Supplementary information

Table of Contents

Figure S1

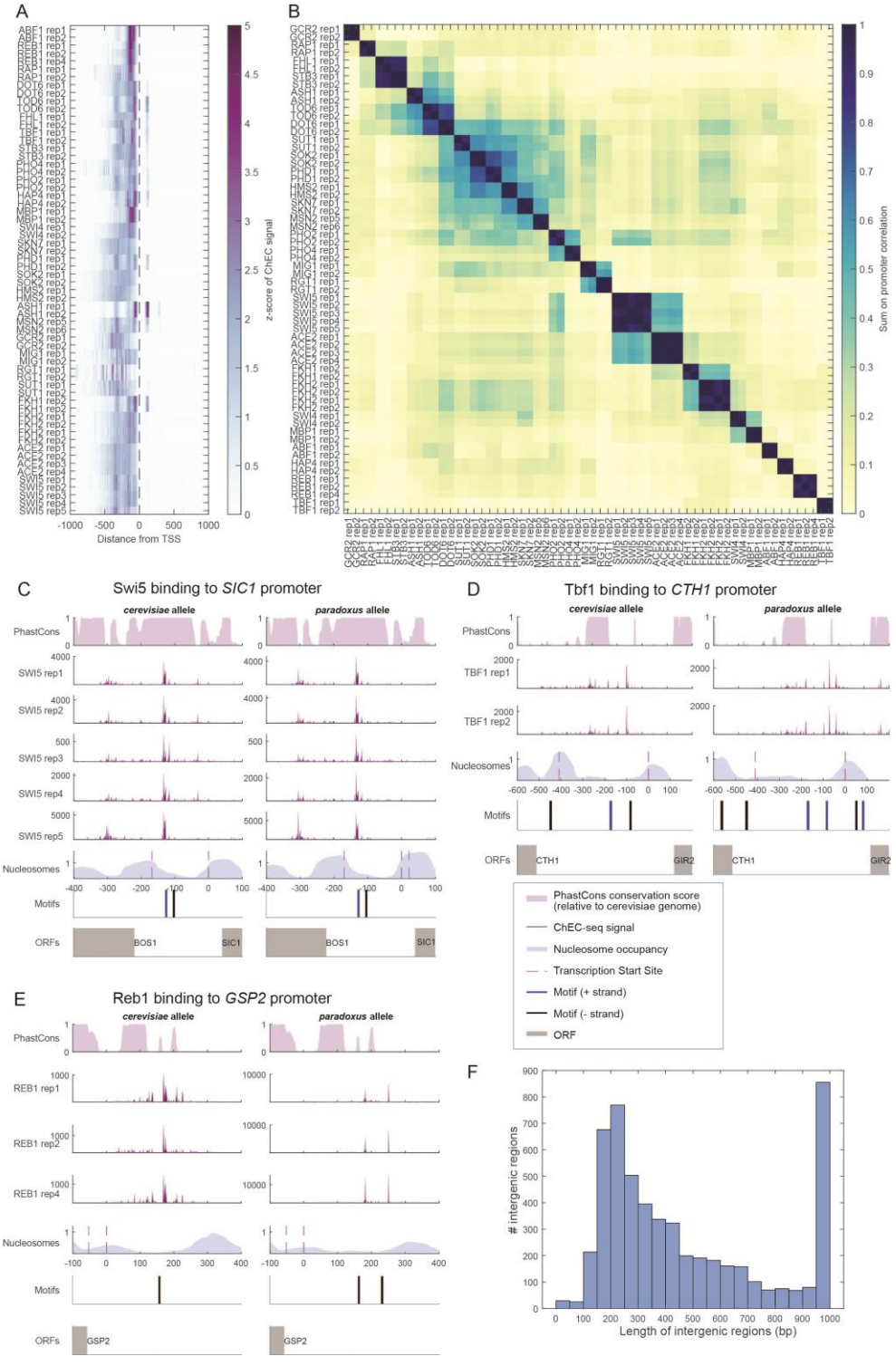


Figure S1: Quality controls for ChEC-seq data. A) ChEC-seq signal is largely restricted to the promoter region. Shown is the metagene: each row is the standardized signal (5' end of reads) around the transcription start site (TSS) over all yeast genes (6701). B) Correlation between ChEC-seq profiles of the different samples. Shown are the Pearson correlation coefficients over sum of signal over gene promoters (6701 genes, sum of 400 bases upstream to TSS), averaged over the two orthologues. C-E) Promoter view of examples in Figure 1 C-E with experimental replicates. F) Length distribution of intergenic regions in *S. cerevisiae*.

Figure S2

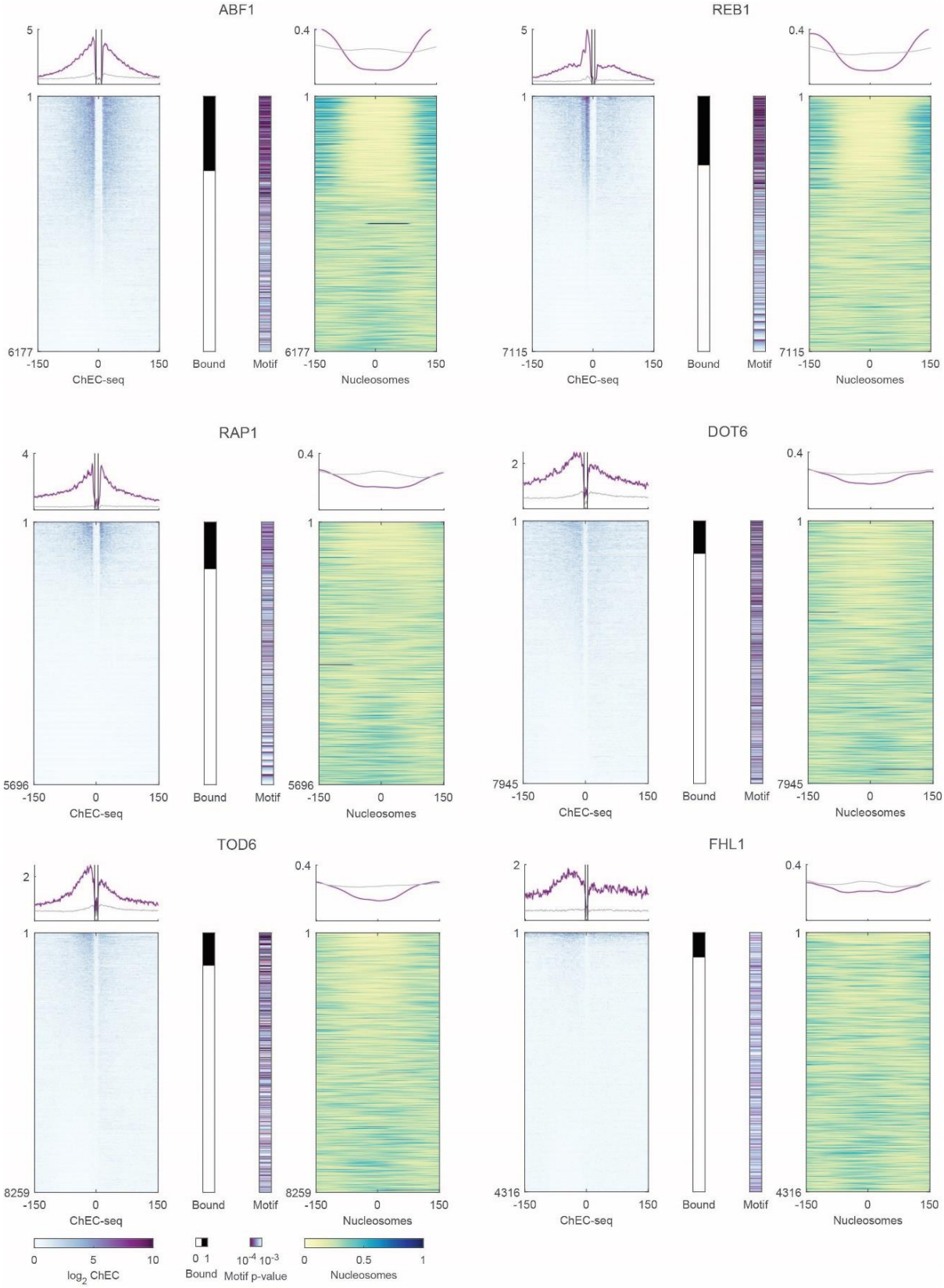

ABF1
REB1
RAP1
DOT6
TOD6
FHL1
ChEC-seq
Bound
Motif
Nucleosomes
log2 ChEC
Bound
Motif p-value
Nucleosomes

Figure S2: page 2

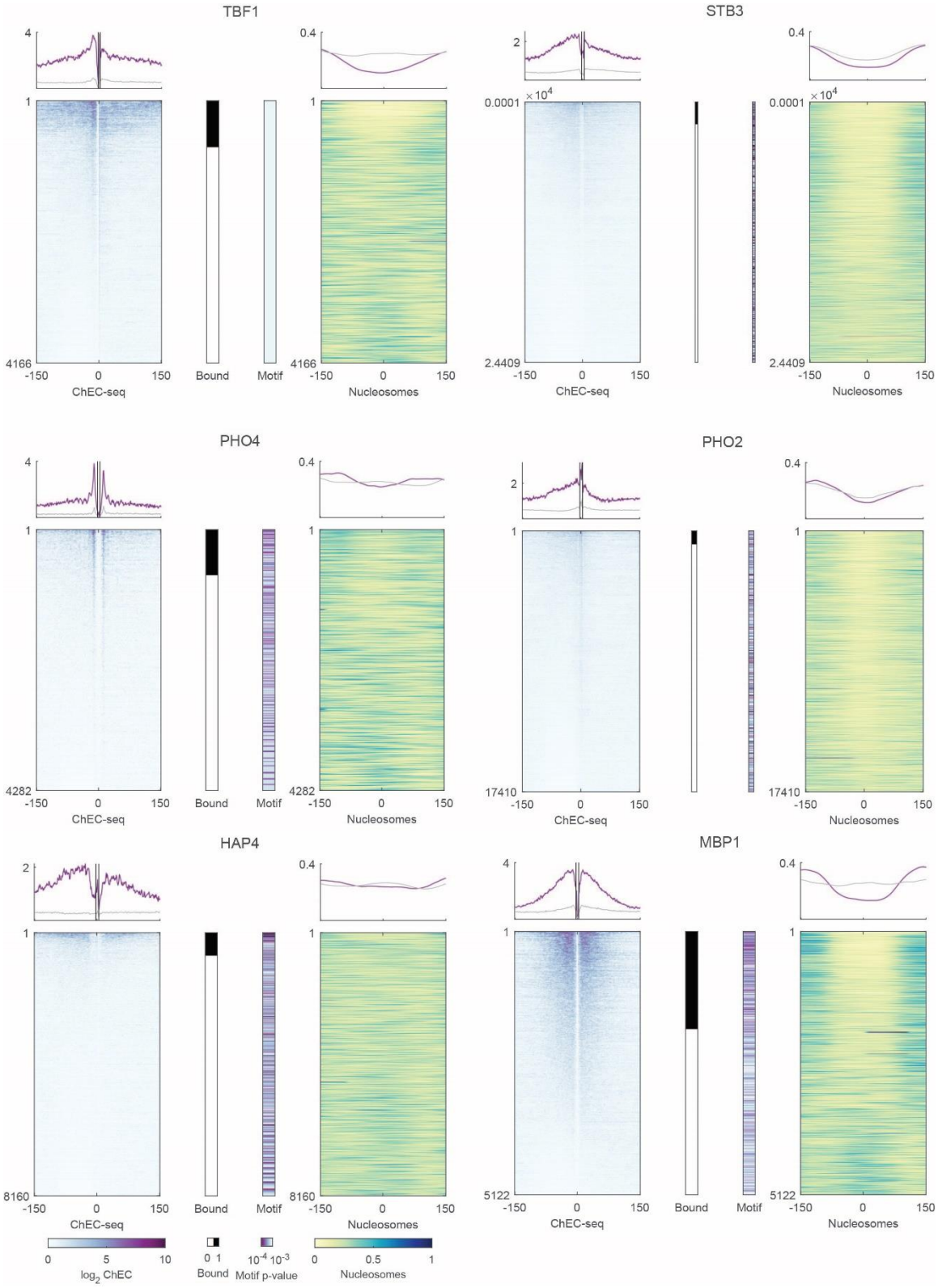


Figure S2: page 3

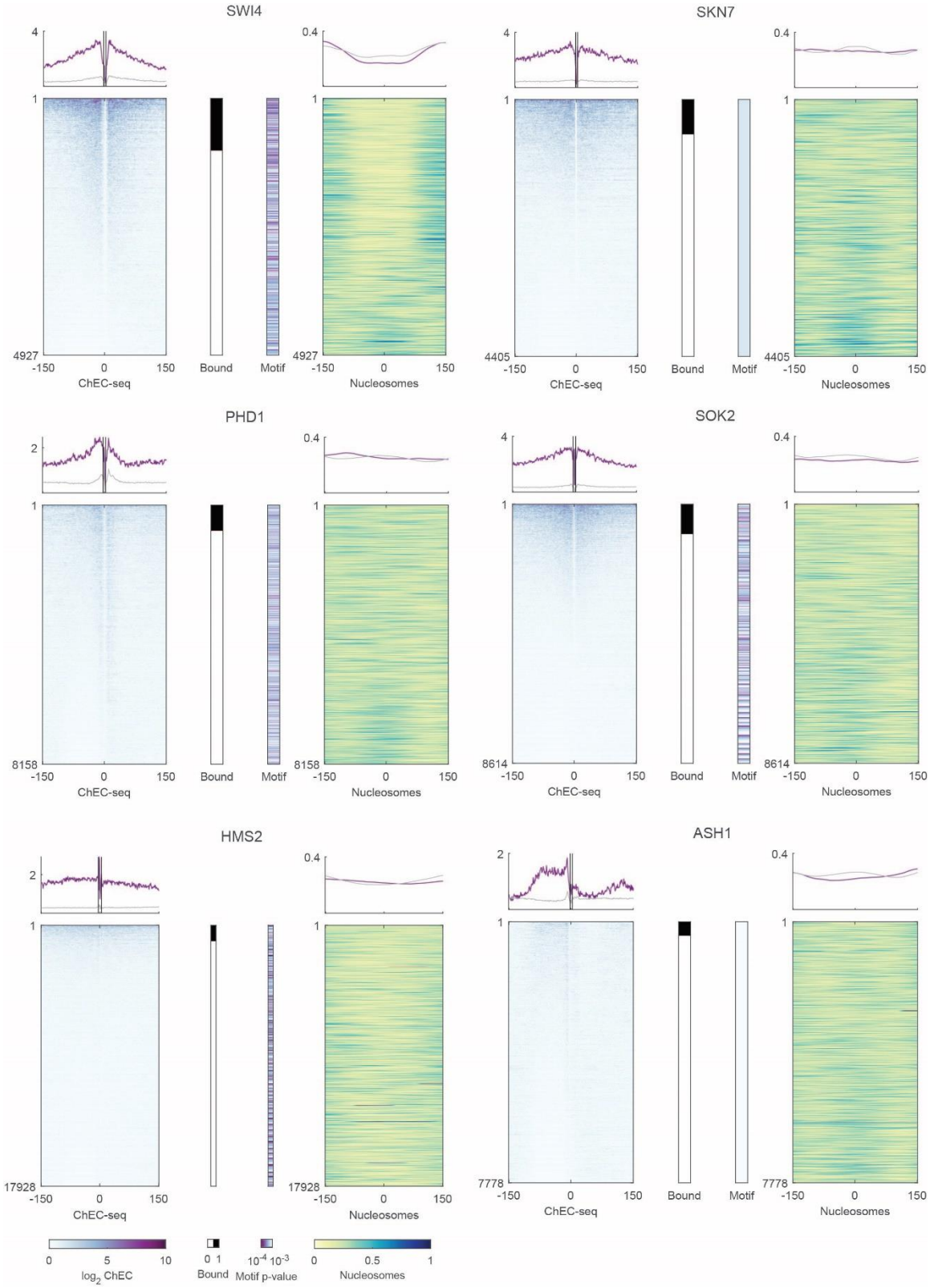
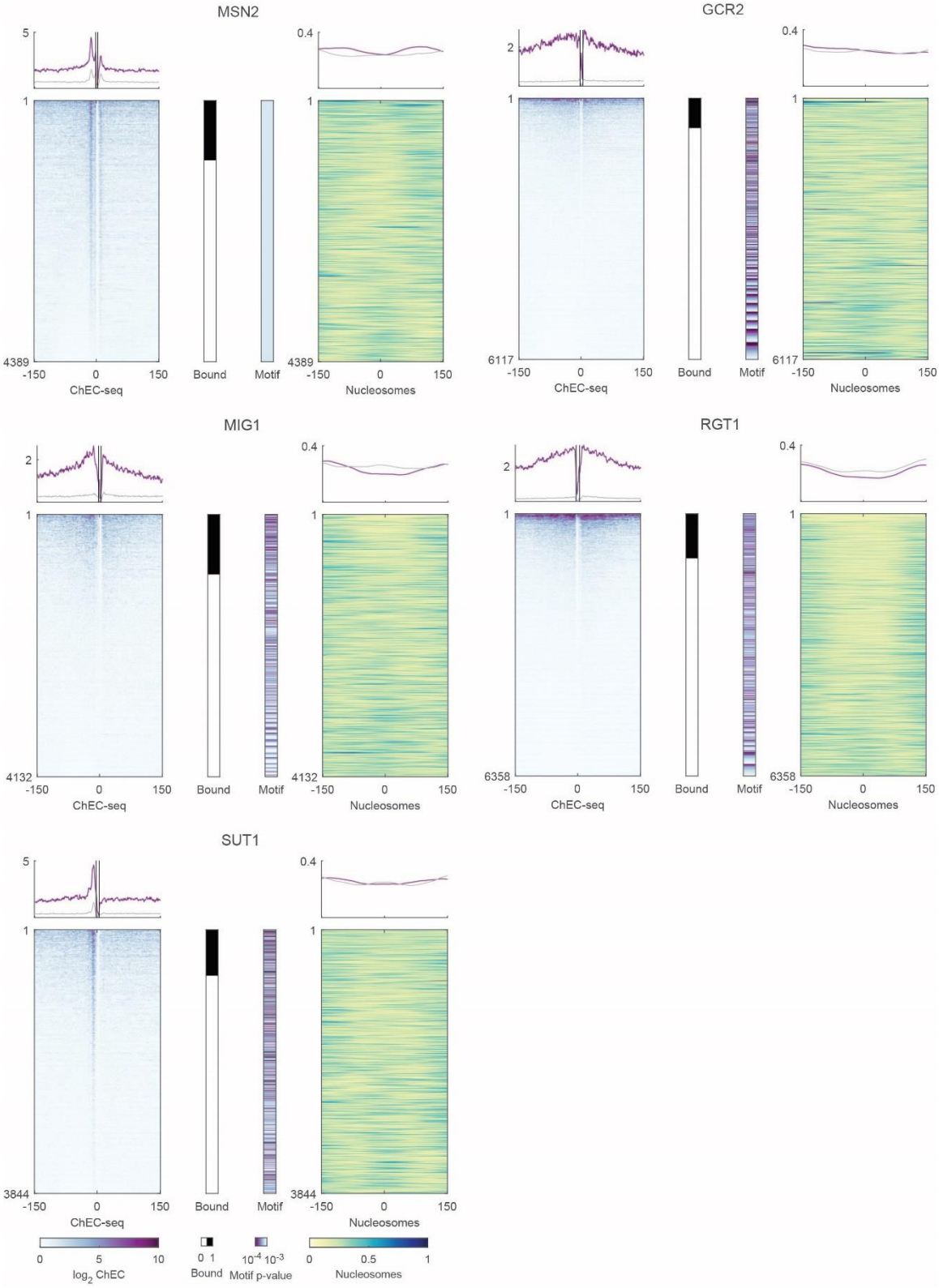


Figure S2: page 4



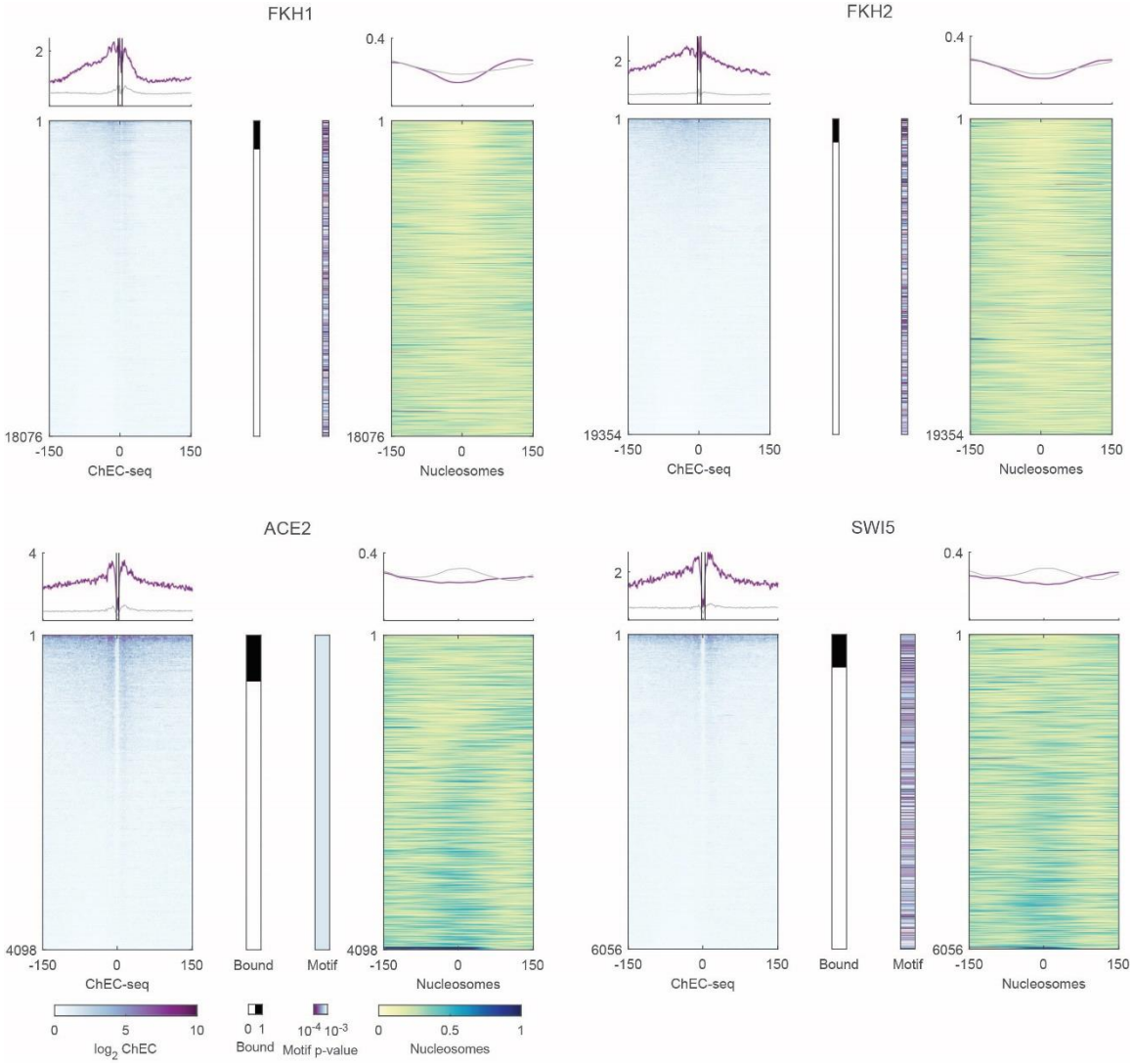


Figure S2: TF binding and nucleosome occupancy at motif-coding sites. Motif-coding sites were defined from the *in vitro* PWM (Table S1) using FIMO (Grant et al., 2011) with p-value < 10^{-3}. For each TF, shown are ChEC-seq profile (5' end of reads, left) and nucleosome occupancy profile (5' end of reads, right) at motif sites, in heatmaps. Nucleosome occupancy profiles were obtained from (Tirosh et al., 2010). Data was sorted by the maximal ChEC-seq signal at 60 bases around motif center. Nucleosome occupancy profiles are aligned to the gene strand, such that the ORF is in the right end of the plot. The mean signal at bound sites (purple) and at non-bound sites (gray) is presented for binding signal (ChEC-seq, left) and for nucleosome occupancy (right) at the top panels. Data points were smoothed with a moving mean of 10 rows.

Figure S3

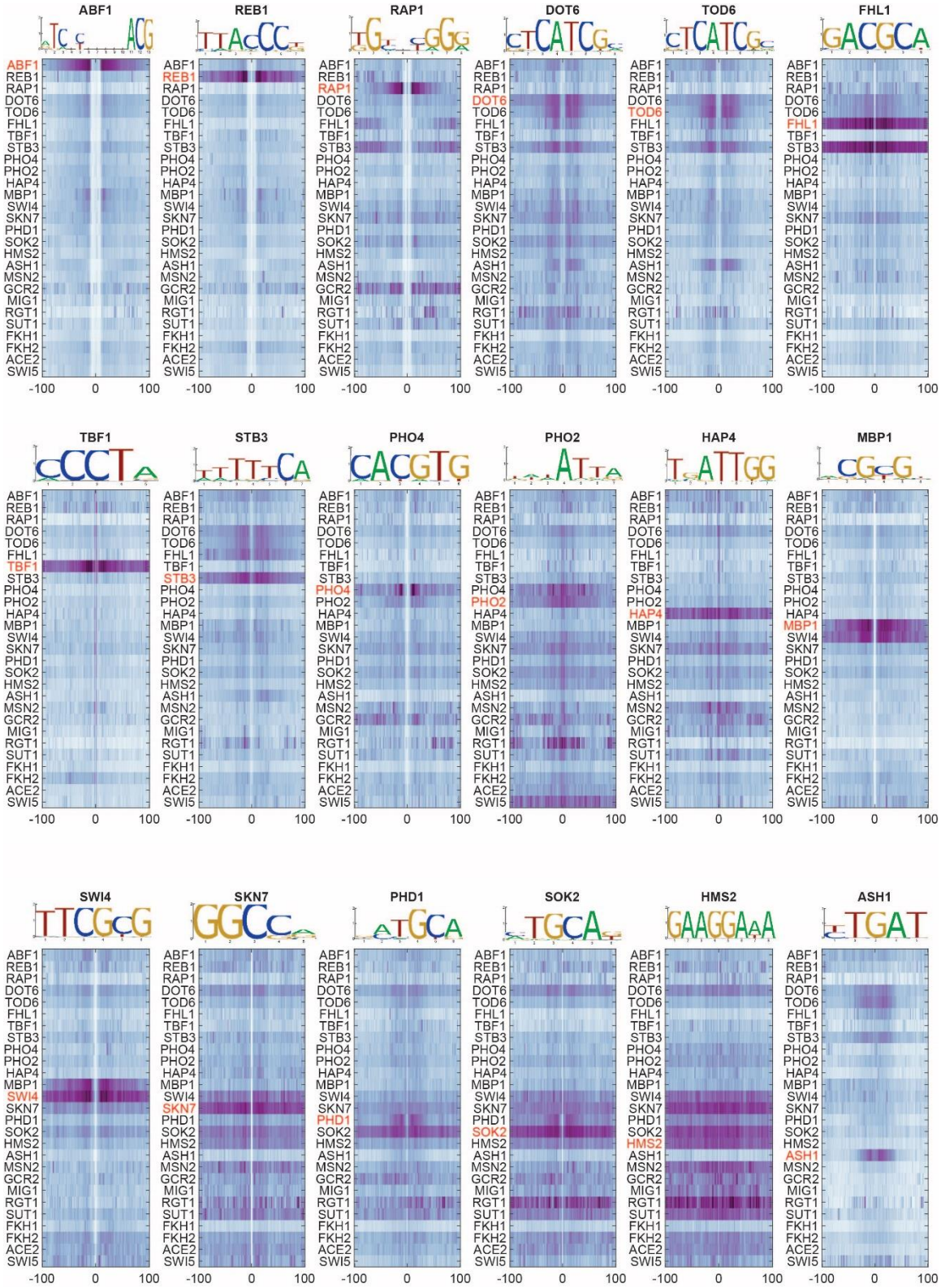

ABF1
REB1
RAP1
DOT6
TOD6
FHL1
TBF1
STB3
PHO4
PHO2
HAP4
MBP1
SWI4
SKN7
PHD1
SOK2
HMS2
ASH1
MSN2
GCR2
MIG1
RGT1
SUT1
FKH1
FKH2
ACE2
SWI5
-100
0
100

Figure S3: page 2

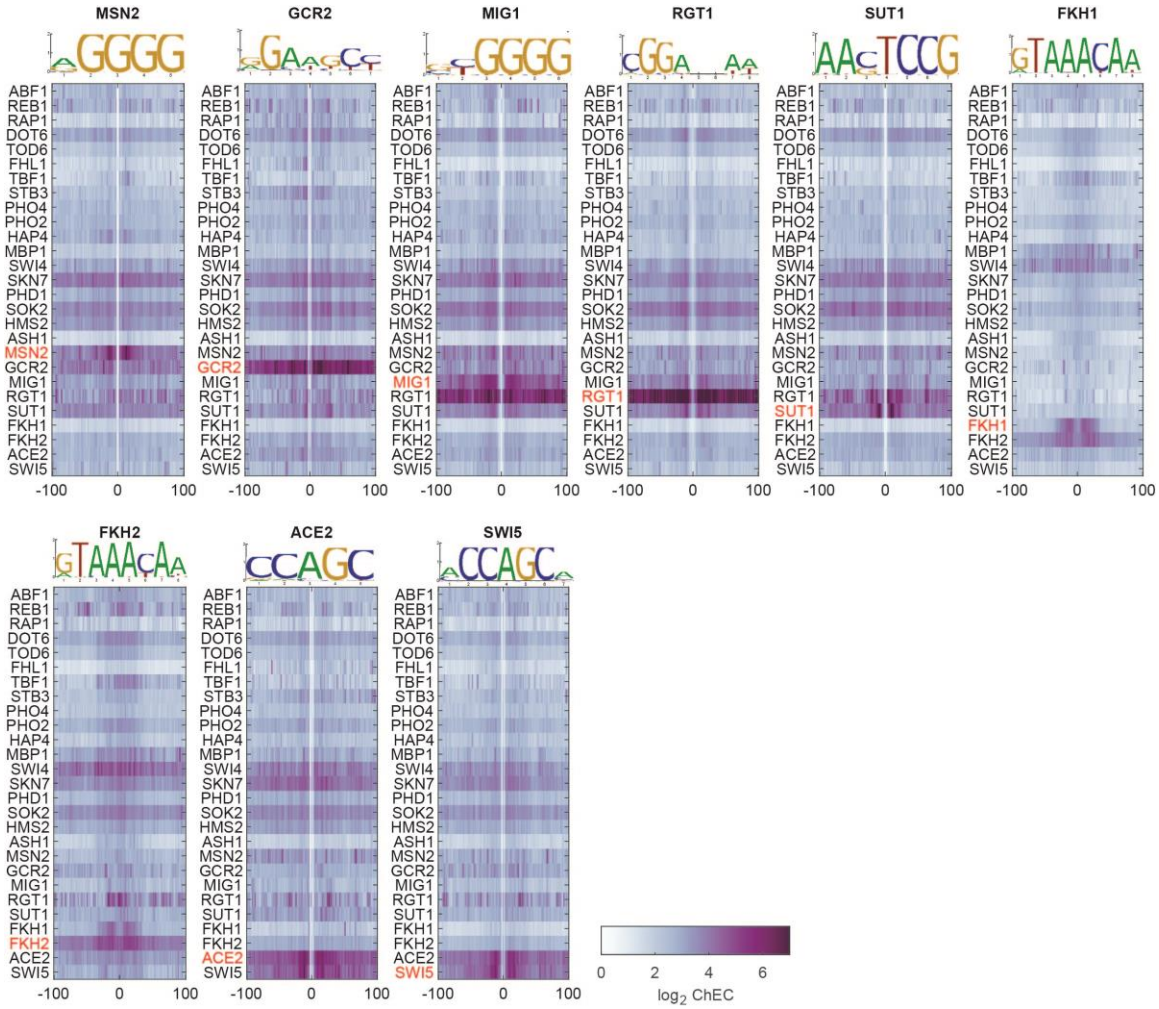


Figure S3: Specificity of TFs to their binding sites. Shown in each panel is the ChEC-seq signal of each TF (rows), averaged over the binding sites of the TF that is indicated in the panel's title. Seqlogo of the *in vitro* motif is presented above each heatmap (source in Table S1). ChEC-seq signal (5' end of reads) is presented in $\log_2$ scale. TF binding sites are peaks associated with a strong *in vitro* motif (FIMO p-value < 0.001). The vertical white line found at the motif center (0 bp) result from the binding of an un-tagged TF to its binding sites. For example, Abf1-bound sites are not digested by a MNase (thus show no read coverage) also in Reb1-MNase cells. The high signal adjacent to motif centers resembles TF specificity to their binding sites.

Figure S4

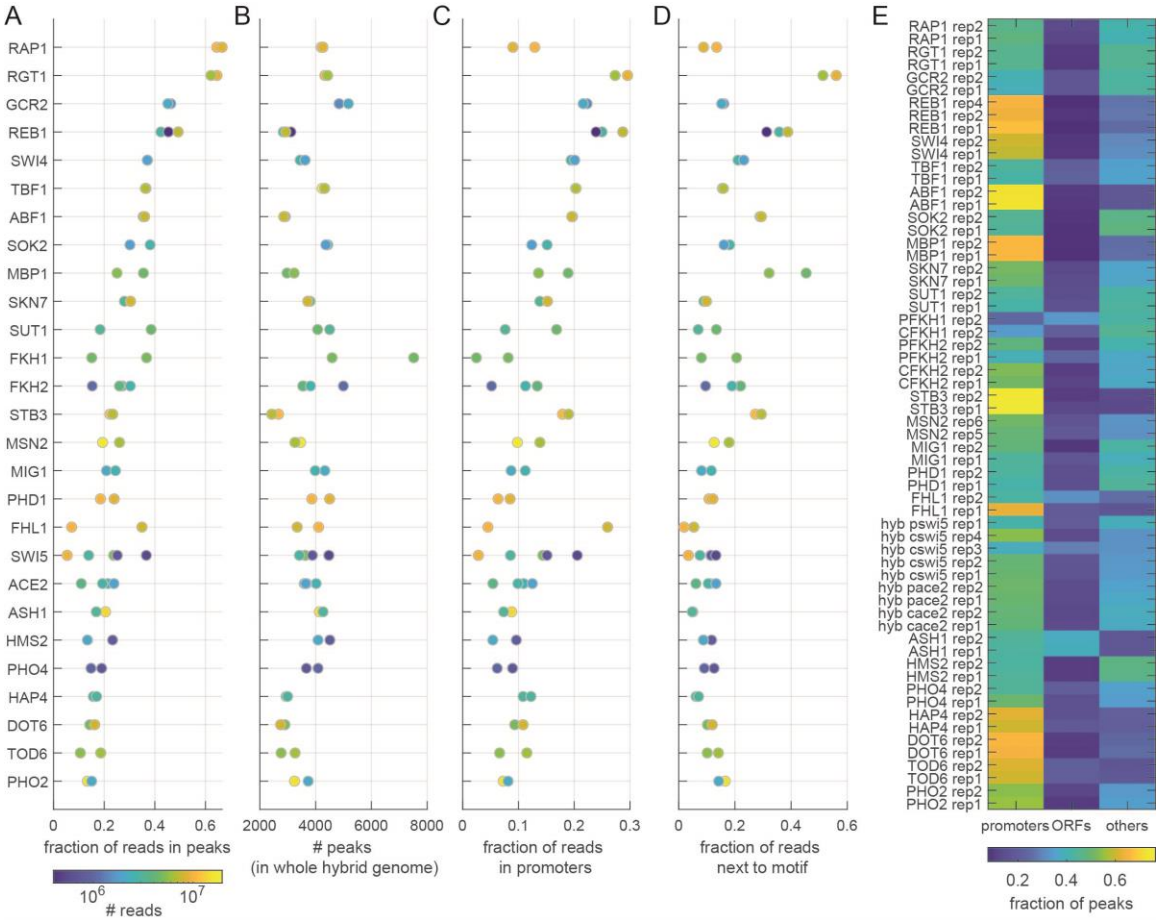


Figure S4: Quality controls for peak calling. Peaks were called using "findpeaks", a MATLAB built-in function (details in Materials and methods). Shown here are the total peaks found on both genomes of the hybrid. The genomic region coding for rDNA on chromosome 12 was removed from all the analyses in this article. A) Fraction of reads in peaks (FRiP) for each experimental replicate of each TF. B) Number of peaks. C) Fraction of reads found in gene promoters. D) Fraction of reads next to a strong motif (FIMO p-value < 0.001), along the full hybrid genome. E) Fraction of peaks found in gene promoters, open reading frames (ORFs) and in other genomic elements.

Figure S5

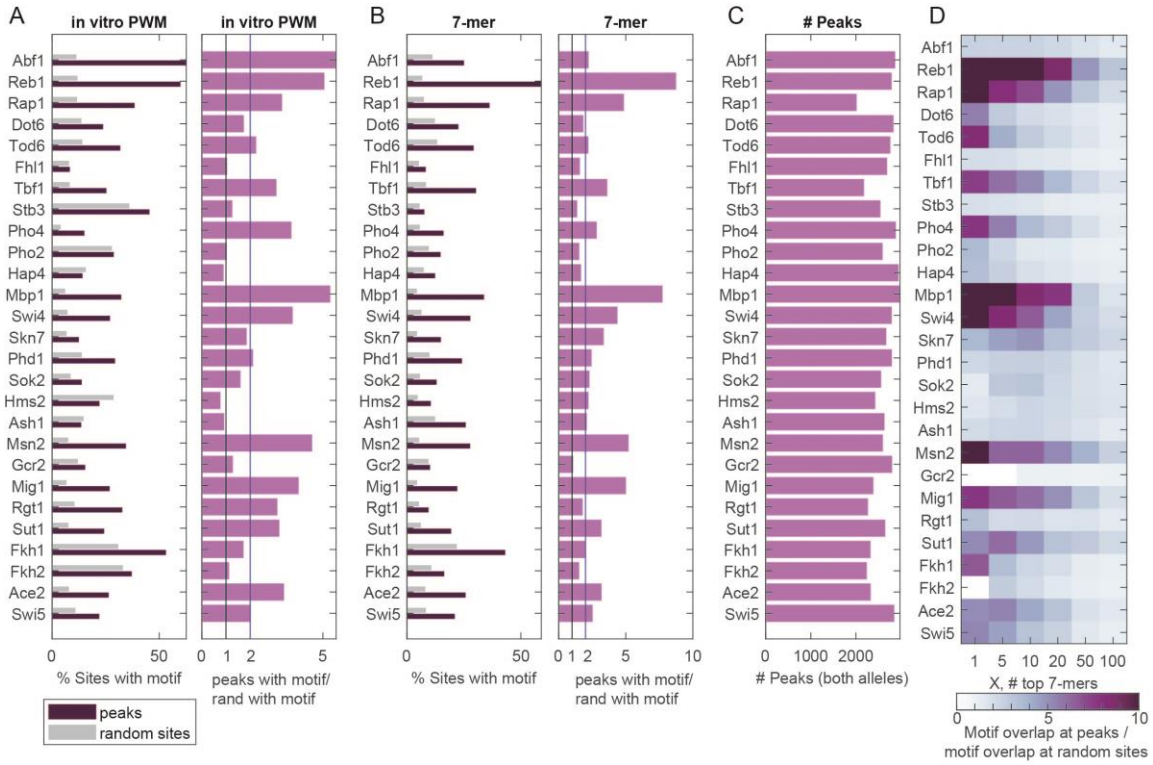


Figure S5: Controls for peak-motif association: Percentage of peaks localized to binding motifs, per TF. Here, we used two measures to define motifs: *in vitro* PWM and 7-mer score.

A) Motifs are the significant motif realizations of *in vitro* PWM from public databases (detailed in Table S1), with FIMO p-value < 0.001. Right: the percentage of peaks (purple) or of random sites in promoters (gray) that reside next to an *in vitro* motif, the maximal distance between peak/random site to motif was 30 bases. Left: The ratio of motif-associated peaks to motif-associated random sites.

B) Motifs are the 20 best sequences of 7-mer enrichment in our hybrid data. Left and right panels as in A.

C) Number of peaks.

D) Different number of top 7-mer sequences and their enrichment within peaks. Here we examined the number of peaks overlapping a certain 7-mer motif versus the number of random sites that overlap this motif (ratio is color-coded). In certain TFs this ratio is high resembling the high correlation between peaks and motifs. In these TFs, the ratio decreases as X (the number of top 7-mers, X-axis) increases. A strong drop in this ratio appears after X = 10 and X = 20, implying that larger X is highly noisy and meaningless.

Figure S6

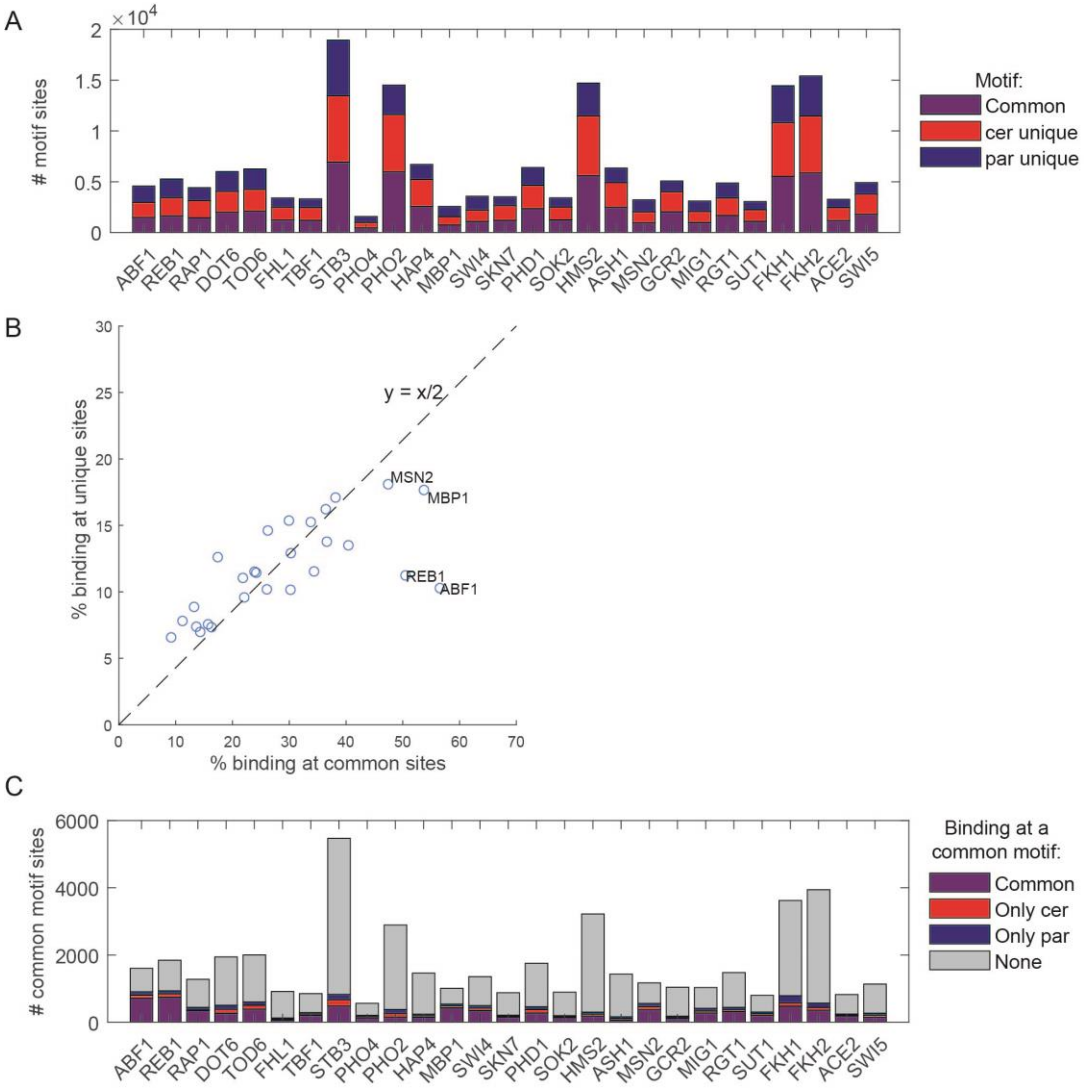


Figure S6: Binding to common and unique motif sites. A) The number of common (conserved), *cerevisiae*-unique and *paradoxus*-unique motif sites in the hybrid genome, per TF. Motif sites are significant realizations of the *in vitro* motif. B) Percentage of bound sites in the hybrid genome. X-axis: percentage of common sites that are bound in both orthologues, Y-axis: percentage of unique-sites that are bound in both orthologues. Dashed line represented x = y*2. C) Proportion of binding to common motif sites, per TF.

Figure S7

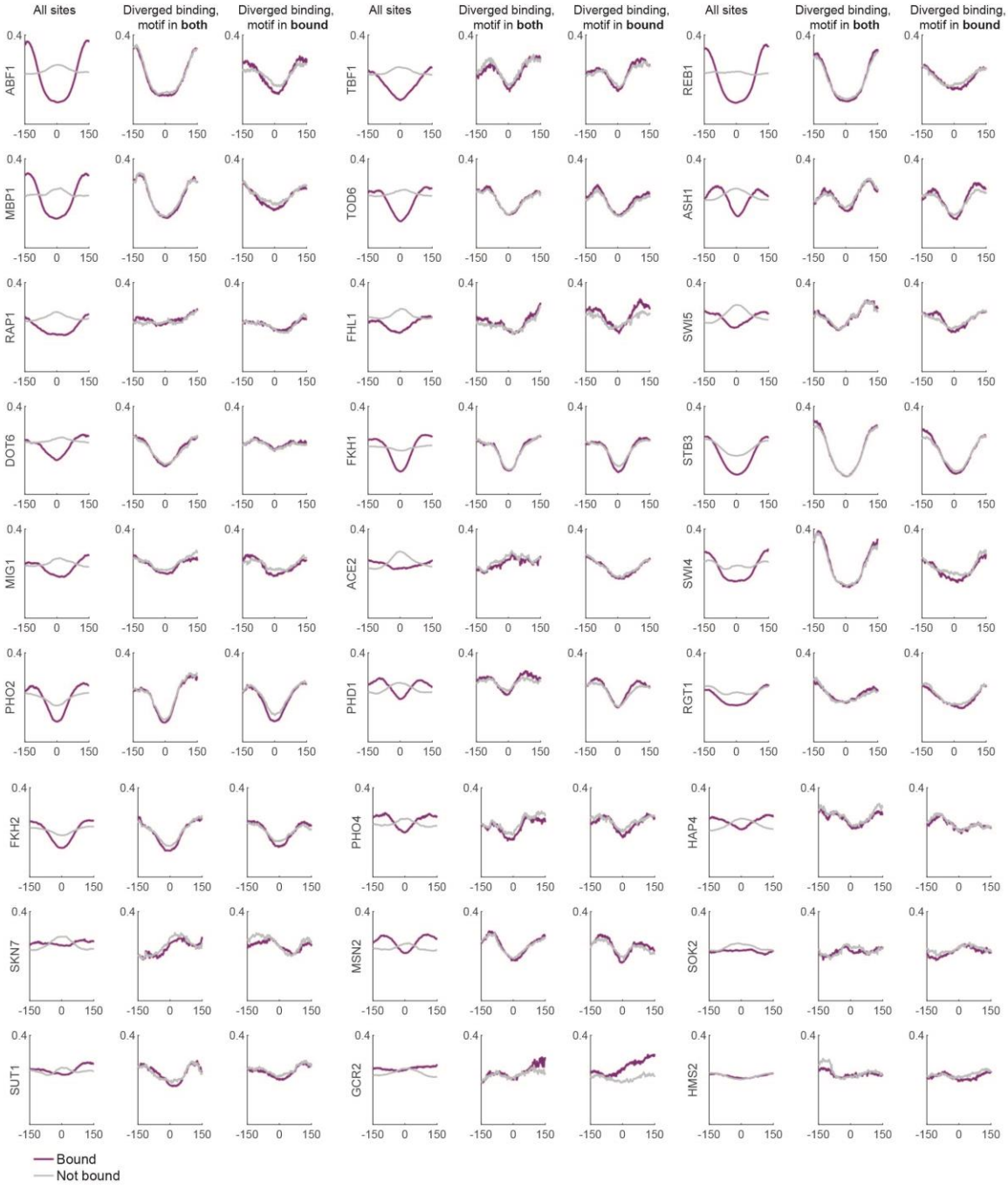


Figure S7: Nucleosome occupancy at motif sites. Shown are three panels per TF.
Left: Motif sites that are either bound or not bound by the TF, regardless of motif conservation or binding conservation.
Middle: Motif sites where only one orthologue is bound, but both orthologues have a motif.
Right: Motif sites of diverged binding, where only the bound orthologue has a motif.
Panels are ordered by the difference in nucleosome occupancy between bound and non-bound at all sites.

Figure S8

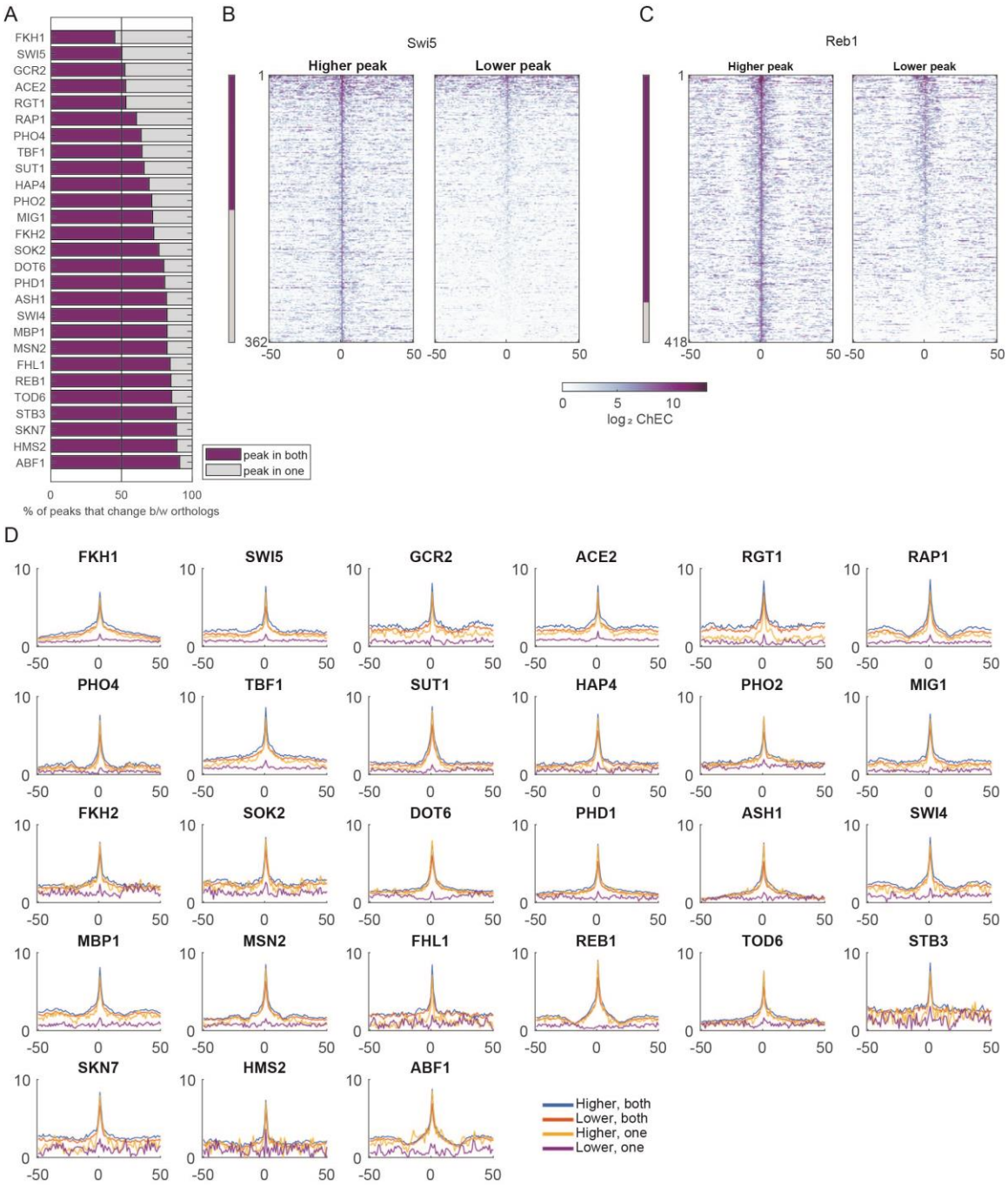


Figure S8: More quantitative than qualitative changes in TF binding peaks.
A) Out of all peaks that change >2 fold between orthologues, presented is the percentage of cases where the difference is quantitative and a peak appears in both orthologues (peak in both, purple) and the percentage of cases where the difference is qualitative (peak in one, gray). Criteria for peak detection are: signal is above the median signal of the specific promoter (above background) and peak has > 20 normalized reads.
B, C) Peak visualization of data presented in A. Shown is ChEC-seq signal at peaks of differential binding (> 2 fold between orthologues), for the orthologue with the higher peak (left) and the orthologue with the lower peak (right). The binary definition of quantitative change (peak in both, purple) or qualitative change (peak in one, gray) is presented on the leftmost panel. Peaks where sorted by the orthologous log ratio. Presented are 50 bp downstream and upstream to the peak center.
D) Meta-peak visualization of the peak categories for each TF, y-axis is in $\log_2$ scale.

Figure S9

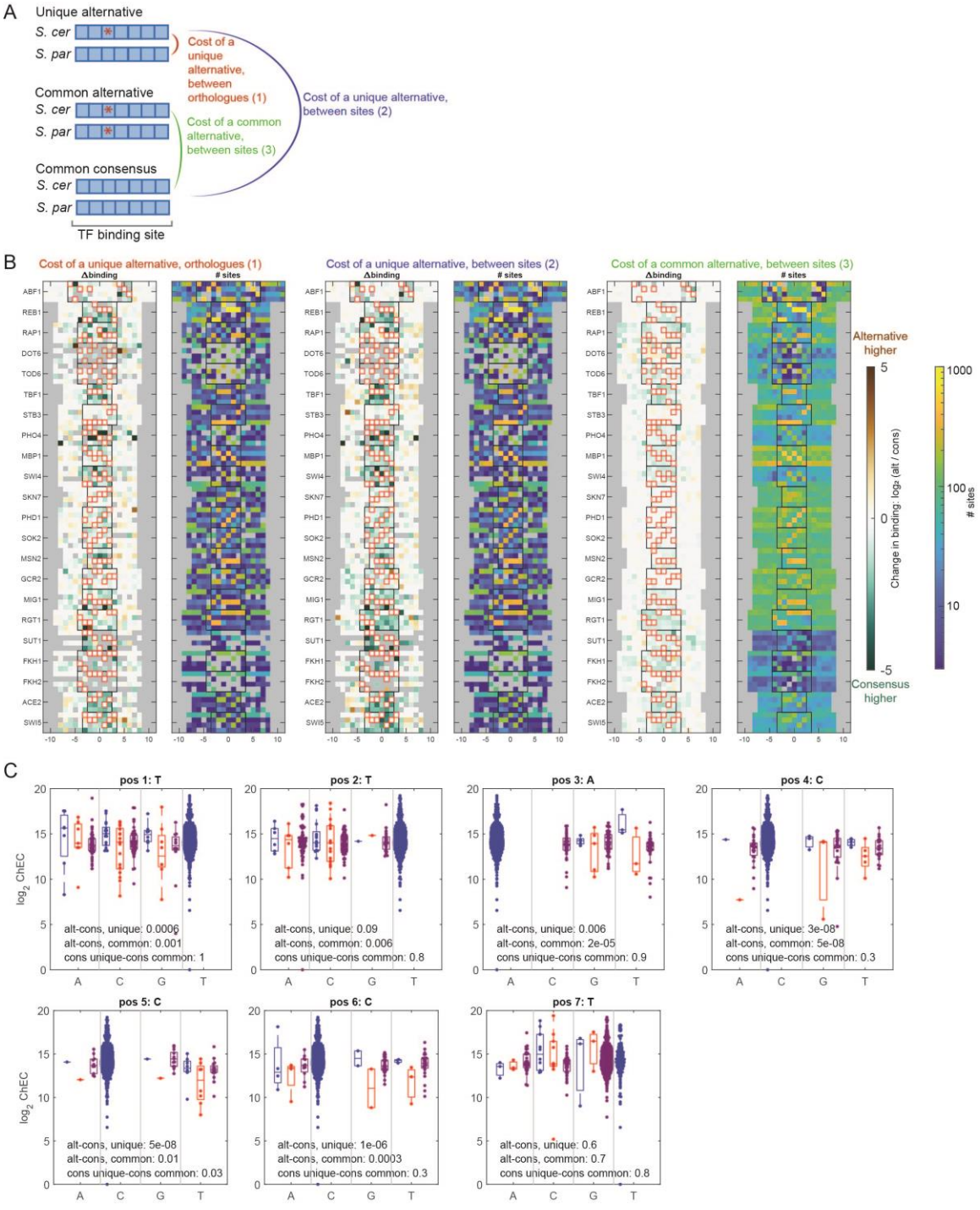


Figure S9: Controls for Figure 4, the cost of cis-regulatory mutations on TF binding.
A) Scheme of the types of mutations examined in the article: unique alternative are cases where one orthologue bears an one-letter variant (alternative allele) and the other has the consensus motif (1). Common alternatives are cases where both orthologues bear the same one-letter variant (3). In addition, we measured the cost of unique alternatives relative to consensus alleles elsewhere in the genome (2).
B) Position-specific mutation cost, of the three types of comparison presented in A. Left: change in binding: color code is $\log_2$ ratio of consensus versus each alternative allele, nucleotide order is: A, C, G, T. Right: number of sites. Gray squares have less than two sites.
C) Binding peak levels per mutation type. Each panel represents a position in the Reb1 motif. Box plots represent the binding levels ($\log_2$ scale) of sites with the consensus (blue), unique alternative (red) and common alternative (purple) alleles. For unique mutation

sites, we plot the binding levels of the consensus allele (blue) and the levels of the alternative allele (red). Shown are t-test p-values for the following comparisons:

- alt-cons, unique: all three alterative alleles at unique sites vs. conserved consensus sites
- alt-cons, common: all three alterative alleles at common sites vs. conserved consensus sites
- cons unique – cons common: consensus alleles at sites of unique mutation vs. conserved consensus sites

T-tests were left tailed, asking if the binding at the conserved consensus sites is higher than binding in the tested sites. The third comparison shows that the consensus allele at unique sites is bound at the same level as sites with a conserved consensus.

Figure S10

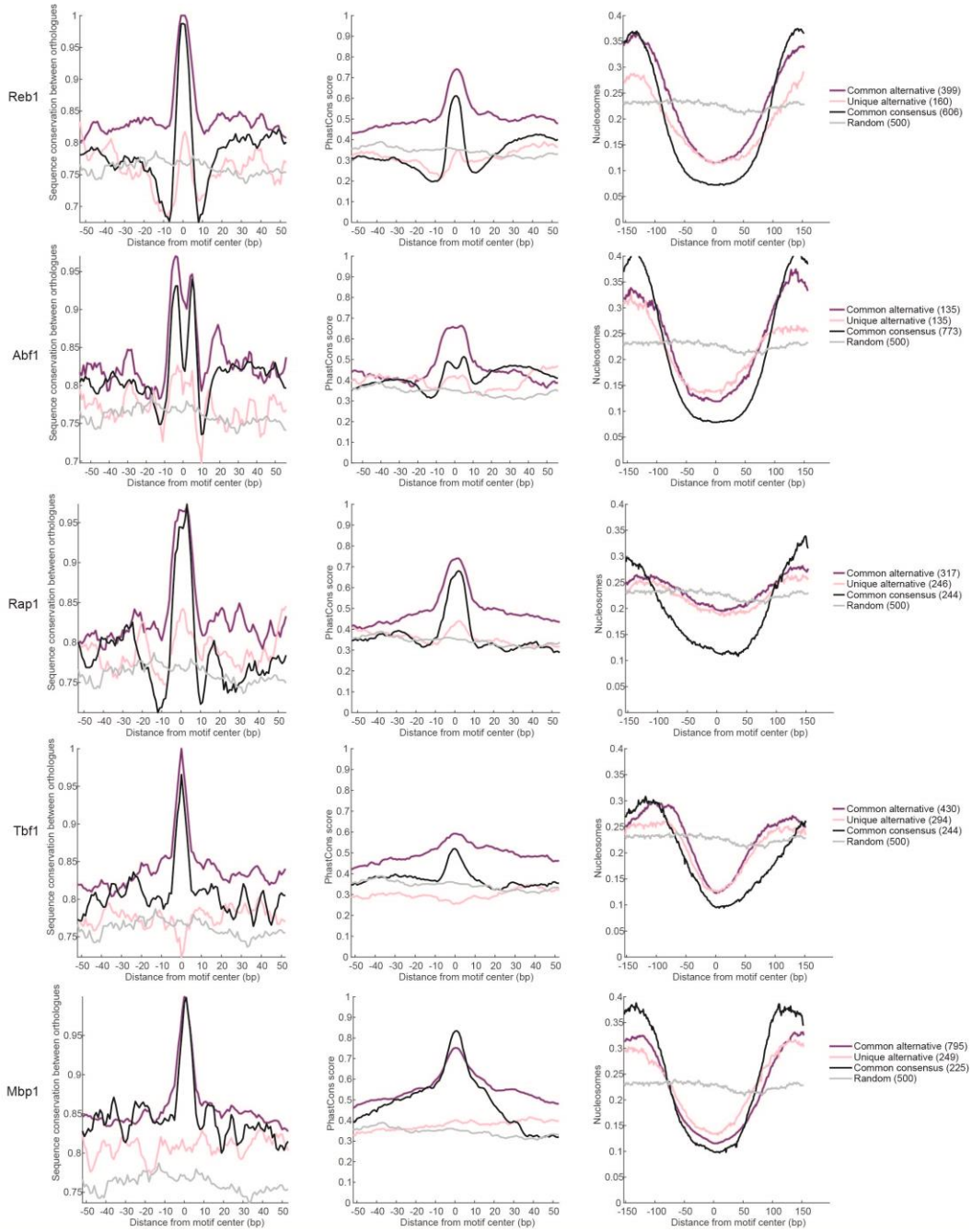


Figure S10: Sequence conservation and nucleosome occupancy at motif regions, distnguished by mutation type. Sequence is more conserved at the motif's vicinity in common alternative sites but not in other sites, for Reb1 (also shown in Figure 4 E, F), Abf1, Rap1, Tbf1 but not for Mbp1 and other factors. Nucleosome depeletion is similar at common alternative and unique alternative sites, but is deeper at common consensus sites, in Reb1, Rap1, Abf1, Tbf1 but not in Mbp1. Shown are sequence conservation between *S. cerevisiae* and *S. paradoxus* orthologues (left panels), PhastCons socre of sequence conservation of seven yeast species (Siepel et al., 2005) (middle panels) and Nucleosome occupancy (right panels). Orthologue conservation profiles were mean-smoothed over five bp. Number of sites per mutation type is indicated in parenthesis. The background conservation at random promoter sites is shown in gray.

Figure S11

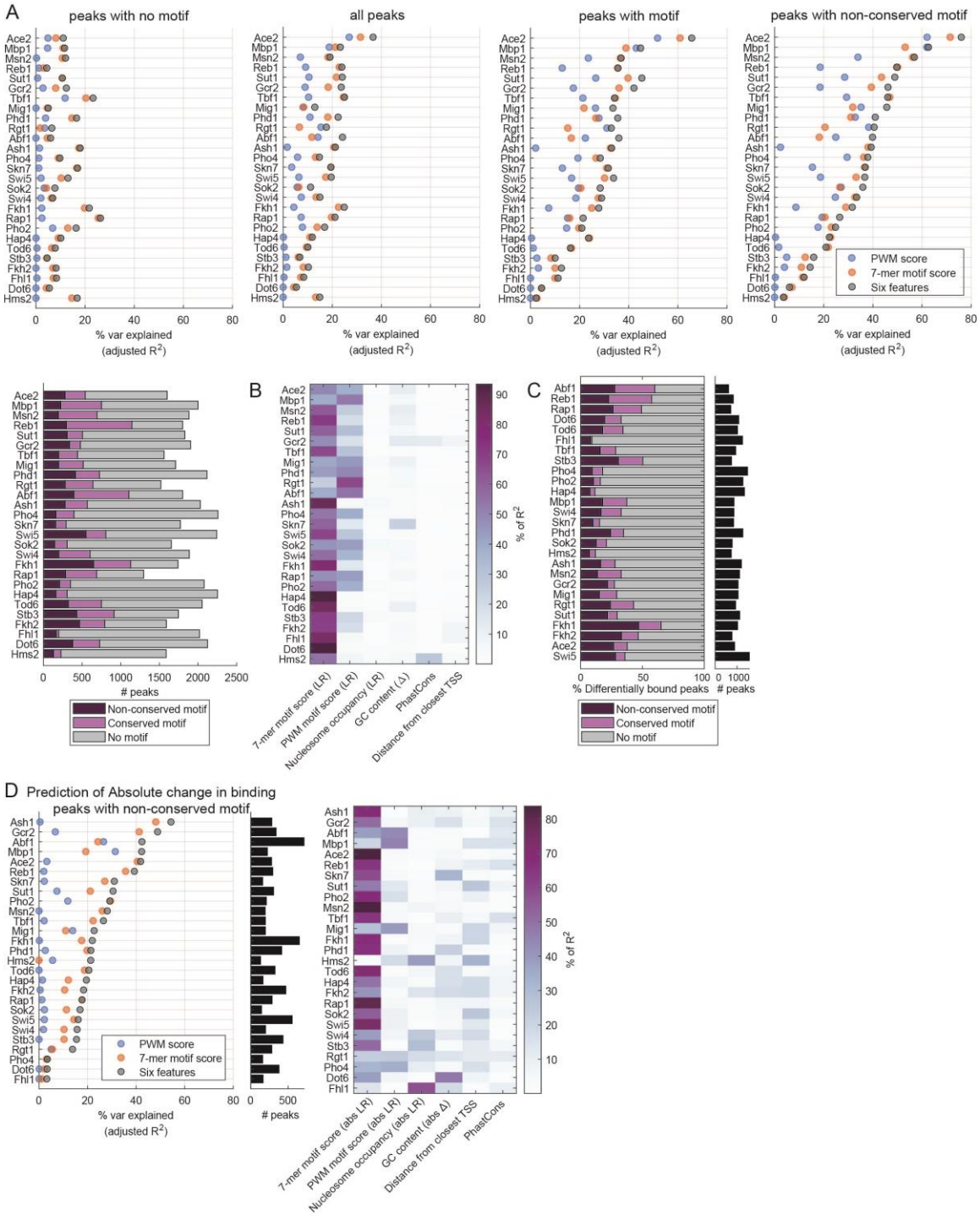


Figure S11: Multi-variate linear models predicts changes in peak binding between orthologues.
A) Prediction was done separately on four peak categories, based on the presence of a strong motif at 60 bases flanking the peak. Strong motifs have FIMO p-value < 0.001. For five TFs we used the 7-mer motif score for strong motif definition (Pho2, Hms2, Ash1, Hap4, Fhl1) because in these TFs, the *in vitro* PWM did not match the motif preference derived from the data.
The number of peaks with no motif, conserved motif or non-conserved motif are presented in the 2nd row. Prediction for peaks with a non-conserved motif (fourth panel) is also presented in Figure 5 B.
B) Relative importance of features in the linear model predictor, using relaImpo R package (Grömping, 2006). Features are: 1. $\log_2$-ratio (LR) of sum of 7-mer motif scores at 60 bases flanking the peak. 2. $\log_2$-ratio of sum of *in vitro* PWM score at 60 bases flanking

the peak. 3. $\log_2$-ratio of sum of nucleosome occupancy signal at 300 bp flanking the peak (nucHybLR). 4. Difference in % GC at 15 bp flanking the motif. 5. Mean phastCons score (Siepel et al., 2005) at 60 bases flanking the peak. 6. Distance of peak from the closest transcription start site.
C) Differentially-bound peaks are associated with non-conserved motifs. The percentage of peaks which are differentially bound between orthologous alleles (fold change > 2), with conserved, non-conserved or no motif are shown.
D) Prediction of the absolute change in binding between orthologues, using the absolute of log-ratio values. Overall prediction power is lower in this case (left panel) and the 7-mer motif score is still the best predictor (right panel).

Figure S12

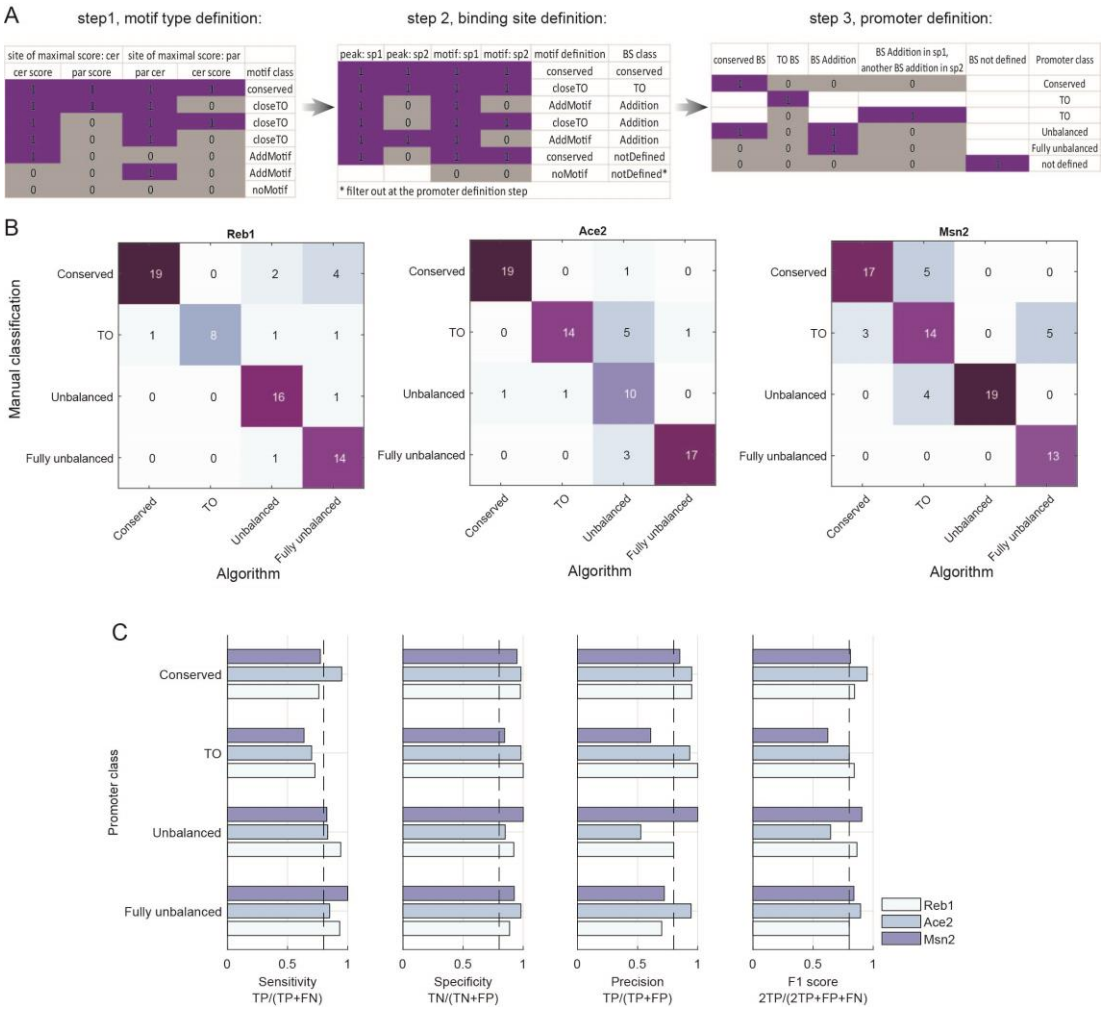


Figure S12: Description and performance of the promoter classification algorithm.

A) Classification algorithm has three steps.

Step1: definition of motifs. The algorithm's input is a list of peaks, and the maximal motif score that reside at 60 bases centered at each peak is examined, per orthologue. Strong motifs were defined using FIMO, with p-value < 0.001. The table in step 1 marks the presence (1) or absence (0) of a strong motif in the max-scoring site of the cerevisiae orthologue (left) or the paradoxus orthologue (right). Conserved sites have a strong motif in cerevisiae and in paradoxus, at the same site. Close turnover (TO) sites have a single peak for both orthologues, but the strong motif is not coded on the same site (4th row of the table), or conserved but another strong motif appears in one of the orthologues (2nd and 3rd rows of the table). Motif addition ("AddMotif") appears in only one of the orthologues (As we compare two species only, no gain/loss events were inferred). These are marked in the 5th and 6th rows of the table.

Step2: Binding site (BS) definition. This step integrates motif definition and the level of binding at peaks, per orthologue. Bound peak (1) have a ChEC-seq signal above threshold (95% of random-site distribution) and above background (as defined in Figure S8). Conserved motifs, that are bound only in one orthologue (6th row on the table) were not defined. Peaks with no motif were not taken for the next step.

Step3: Promoter definition. Takes into account all the BSs at a specific promoter. Here, presence (1) refers to the presence of one or more BS of a specific definition. Conserved promoters contain only conserved binding sites. Turnover promoters can have either a close TO or a far TO BSs (2nd and 3rd rows respectively), where the presence of a conserved or an added BS do not change the TO definition. Unbalanced promoters have at least one conserved BS and at least one added BS. Fully unbalanced promoters are bound, in one or more BS, at only one orthologue.

B) Confusion matrices of manually defined and algorithm-classified promoter classes. For each TF, up to 20 promoters of each class (as defined by the algorithm) were picked randomly. The authors manually visualized the data and defined the promoter classes. Total accuracy percentages are: Reb1: 0.84, Ace2: 0.83, Msn2: 0.79.

C) Classification performance. Shown are the sensitivity, specificity, precision and F1 score, per class, per TF. For each TF,

classification was tested per class versus all other classes. TP = true positives, FN = false negative etc. The mean sensitivity and mean specificity of all TFs and all classes are reported in the main text. Dashed line is at x = 0.8.

Figure S13

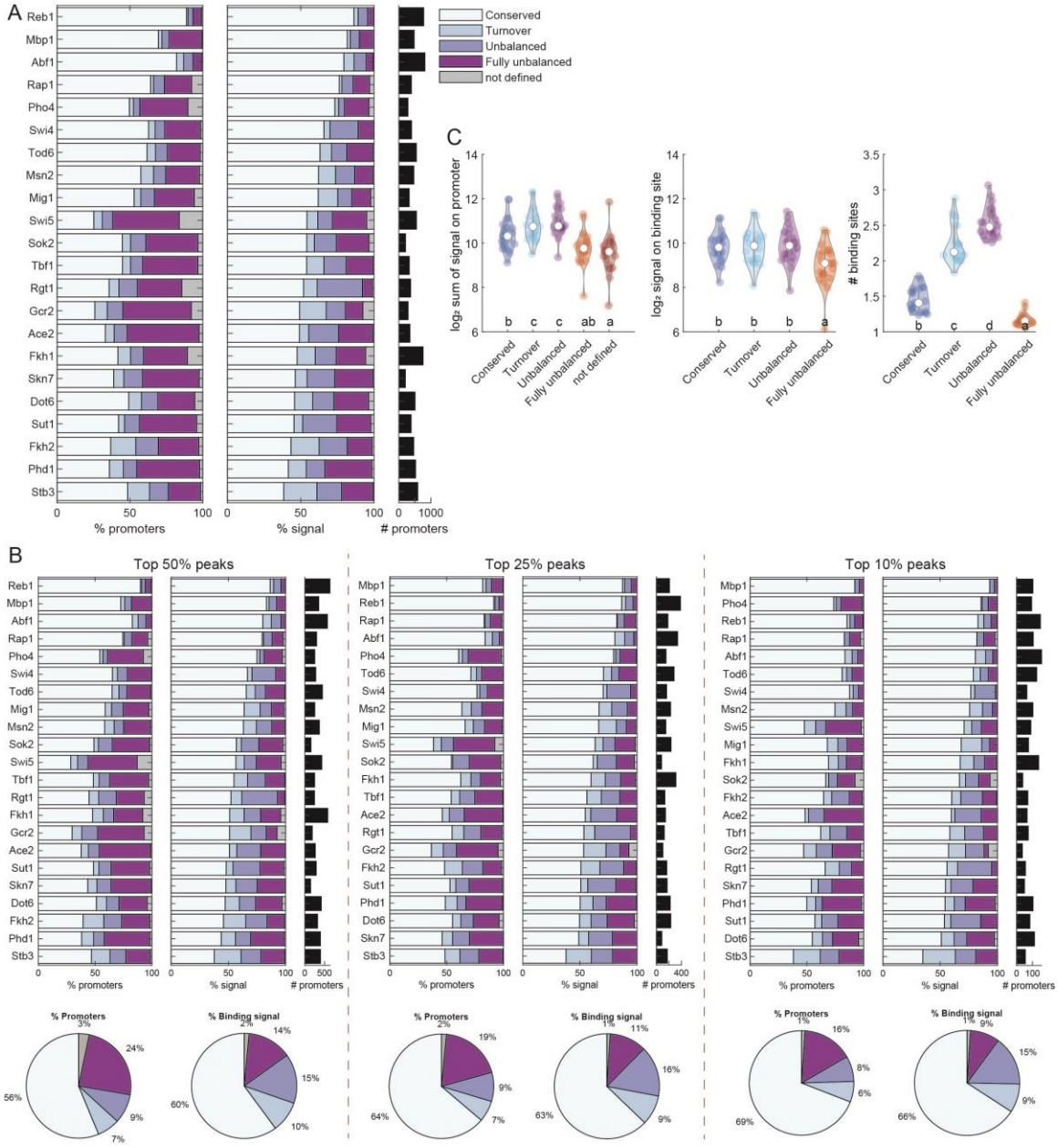


Figure S13: Details of promoter classification.

A) Proportions of promoter classes and binding signal per TF. Left panel: proportion of promoters assigned to each class. Middle panel: proportion of the total binding signal taken by each promoter class. Shown is the binding signal (normalized ChEC-seq reads) of the higher-bound allele. Left panel: Number of promoters in each TF. This analysis includes 22 of the 27 TFs, exceptions are: Pho2, Hms2, Ash1, Hap4 and Fhl1. In these TFs the *in vitro* motif did not match the data-derived motif.

B) Repeating promoter classification with elevated peak threshold. Here we repeated the promoter classification analysis, while restricting peaks to the top 50 %, top 25 % and top 10 % of the peak height distribution. While the proportion of conserved promoters increase with increased threshold, the general trend remains: largest class is Conserved, than Fully unbalanced, Unbalanced and Turnover.

C) Turnover and Unbalanced promoters contain more binding sites than Conserved and Fully unbalanced promoters. Left panel: total promoter signal (log_2 scale) of each promoter class. Shown is the median level per TF. Middle panel: binding signal per binding site (log_2 scale), median level per TF. Right panel: Number of binding sites, mean number per TF. In all panels we plot the median or

mean value per TF, therefore the number of dots in each violin plots equals 22. Letters correspond to statistically distinguished groups after Tukey's-honestly significant difference test.

Figure S14

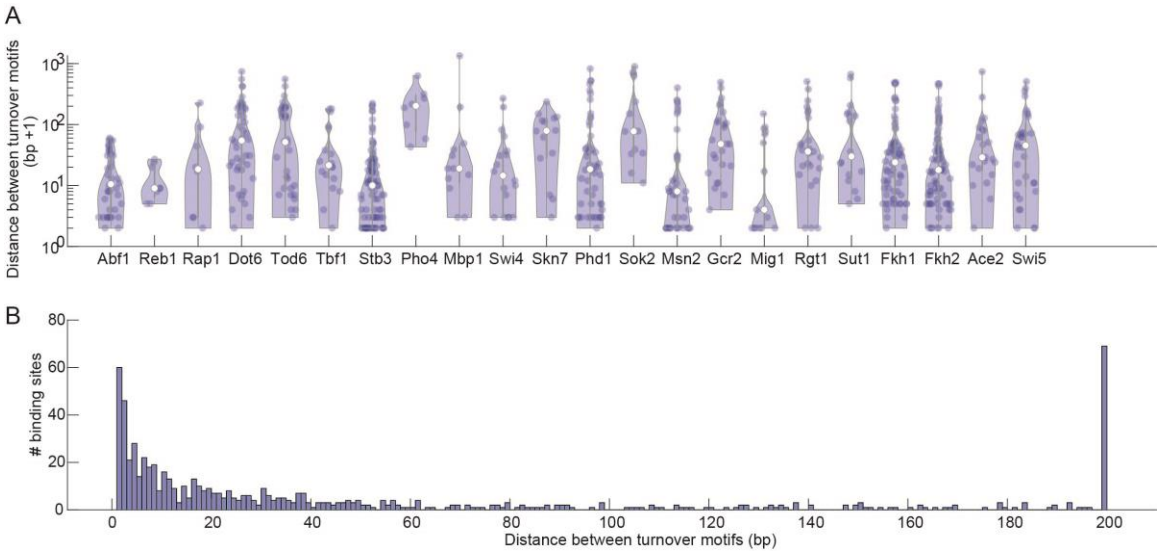


Figure S14: Binding site turnover. A) Distance between turnover motifs per promoter. Upper panel: distance between turnover motifs per promote, per TF. Shown are violin plots in log scale of (distance in bp + 1). B) Histogram of turnover distances in all promoters together (linear scale).

Figure S15

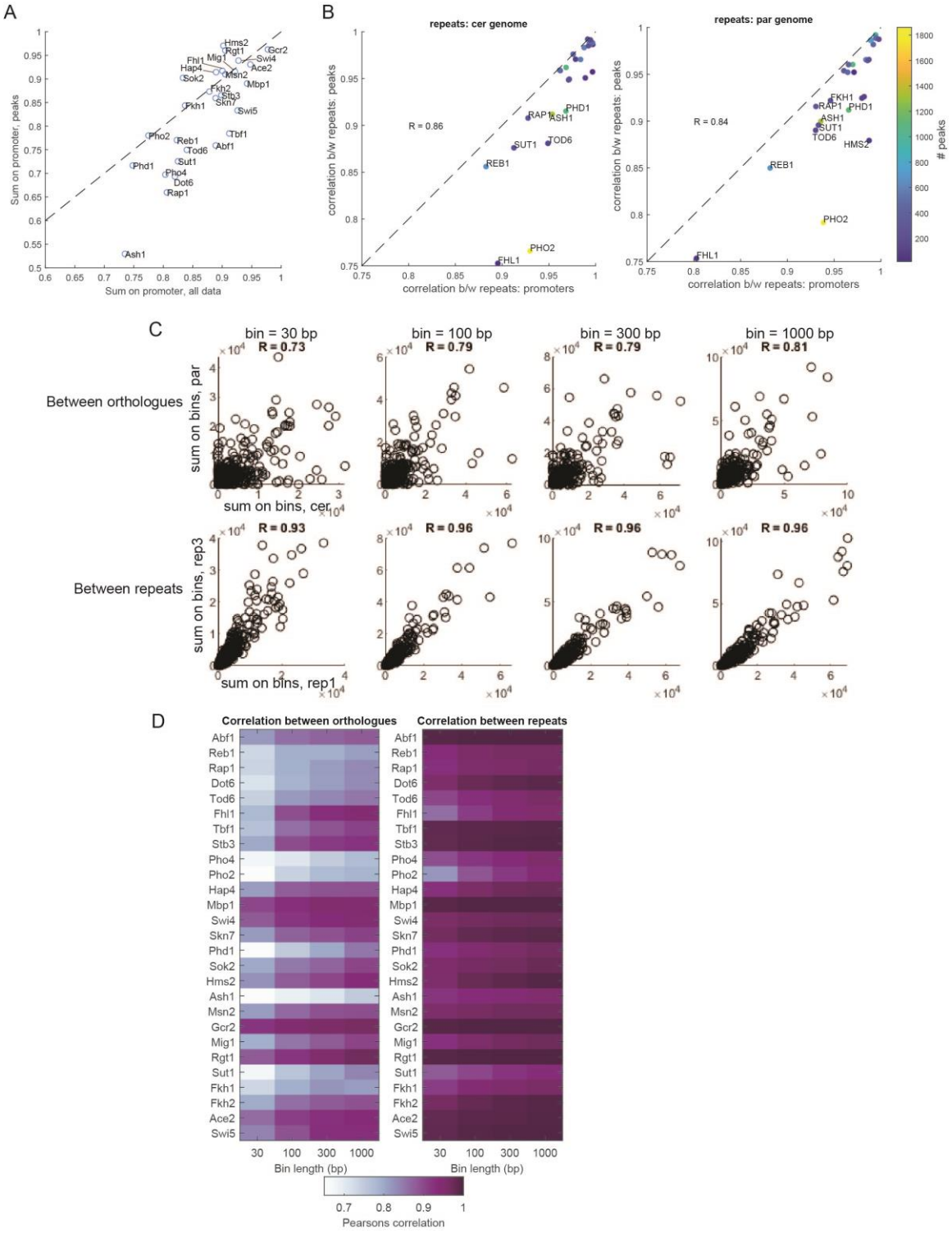


Figure S15: Controls for correlation between orthologues on promoters and peaks.
A) Correlation coefficient are similar when examining the full signal and when restricting to promoters that contain motif-associated peaks. Shown are the correlation coefficients of sum of signal over promoters between orthologues. X-axis: taking all promoters. Y-axis: taking only promoters with motif-associated peaks.
B) Correlation between repeats along promoters and along peaks. Here, we examined only peaks with a strong 7-mer motif (top 20 7-mer score).

C+D) Control for the comparison of correlation between orthologues along peaks and along motifs.
C) We ask whether the observation that peaks correlate less than promoters is a result of the resolution (peaks are ~20 bp wide, promoters are ~350 bp wide) or a result of the functional unit (peak vs. promoter). Here we sum the signal over bins of different length, and compare orthologues (upper panels) and repeats (lower panels) of Reb1 binding profiles. Correlation is lower between orthologues than between repeats. Between orthologues, the large shift in correlation appears between the first (30 bp) and the second bin (100 bp).
D) Repeating the analysis in B on the full set of TFs. In most cases, the large shift appears between the first (30 bp) and the second bin (100 bp).

Table S1

Table S1: List of transcription factors studied in this article, Sequence logos of their known *in vitro*-defined motif from literature, and Seqlogos derived from the data generated in this study, using the top 20 7-mers and via the MEME-ChIP algorithm.

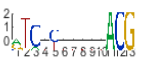
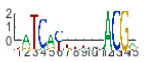
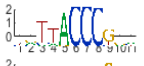
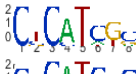
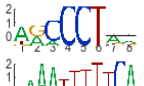
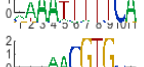
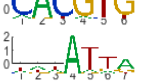
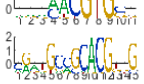
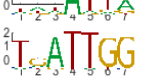
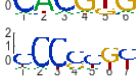
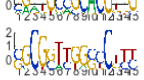
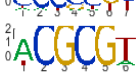
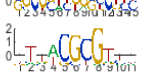
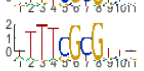
General information		Literature motif			Seqlogo		
Factor	DBD	Database	Method	PMID	Literature	This study (7mer)	This study (MEME-ChIP)
ABF1	ABF1	Yetfasco	PBM	17947326			
REB1	Myb/SANT	Yetfasco	MITOMI	20802496			
RAP1	Myb/SANT	Yetfasco	PBM	17947326			
DOT6	Myb/SANT	Jaspar	PBM	18842628			
TOD6	Myb/SANT	Jaspar	PBM	18842628			
FHL1	Forkhead	Jaspar	PBM	19111667			
TBF1	Myb/SANT	Jaspar	PBM	19111667			
STB3	no DBD, bind Sin3p	Yetfasco	PBM	19997485			
PHO4	bHLH	Jaspar	PBM	18842628			
PHO2	Homeodomain	Yetfasco	PBM	19997485			
HAP4	no DBD (Hap2/3/5: CBF/NF-Y)	Yetfasco	Chip-chip	16522208			
MBP1	APSES	Yetfasco	PBM	19997485			
SWI4	APSES	Yetfasco	PBM	19111667			
SKN7	HSF	Yetfasco	PBM	19111667			
PHD1	APSES	Yetfasco	PBM	19997485			
SOK2	APSES	Yetfasco	PBM	19111667			
HMS2	HSF	Yetfasco	Chip-chip	17500587			
ASH1	GATA	Yetfasco	EMSA	11171979			
MSN2	C2H2 zinc finger	Yetfasco	PBM	19111667			
GCR2	no DBD, co-factor of Gcr1	Yetfasco	PBM	22189060			
MIG1	C2H2 zinc finger	Jaspar	PBM	19111667			
RGT1	Zinc cluster	Yetfasco	PBM	19997485			
SUT1	Zinc cluster	Jaspar	PBM	18842628			
FKH1	Forkhead	Yetfasco	PBM	19158363			
FKH2	Forkhead	Yetfasco	PBM	19158363			
ACE2	C2H2 zinc finger	Yetfasco	MITOMI	20802496			
SWI5	C2H2 zinc finger	Yetfasco	PBM	19111667			

Table S2

Table S2: List of strains used in this study.

Number	strain	source	background strain	ORF-Mnase linker length	Indluded in dataset
2	ABF1-MNase	Bar-Ziv et al.	BY4741	33	yes
3	REB1-MNase	Bar-Ziv et al.	BY4741	33	yes
4	RAP1-MNase	Bar-Ziv et al.	BY4741	33	yes
5	DOT6-MNase	Brodsky et al.	BY4741	33	yes
6	TOD6-MNase	This study	BY4741 C-SWAT	15	yes
8	FHL1-MNase	This study	BY4741 C-SWAT	15	yes
9	TBF1-MNase	This study	BY4741 C-SWAT	15	yes
10	STB3-MNase	This study	BY4741 C-SWAT	15	yes
11	PHO4-MNase	This study	BY4741 C-SWAT	15	yes
12	PHO2-MNase	This study	BY4741 C-SWAT	15	yes
13	HAP4-MNase	This study	BY4741	33	yes
14	MBP1-MNase	This study	BY4741 C-SWAT	15	yes
15	SWI4-MNase	This study	BY4741 C-SWAT	15	yes
16	SKN7-MNase	Brodsky et al.	BY4741	33	yes
18	PHD1-MNase	This study	BY4741 C-SWAT	15	yes
19	SOK2-MNase	This study	BY4741 C-SWAT	15	yes
20	HMS2-MNase	This study	BY4741 C-SWAT	15	yes
21	ASH1-MNase	This study	BY4741 C-SWAT	15	yes
24	MSN2-MNase	Brodsky et al.	BY4741	33	yes
26	GCR2-MNase	This study	BY4741 C-SWAT	15	yes
27	MIG1-MNase	Brodsky et al.	BY4741	33	yes
30	RGT1-MNase	Brodsky et al.	BY4741	33	yes
35	SUT1-MNase	This study	BY4741 C-SWAT	15	yes
40	ACE2-MNase	Lupo et al.	BY4741	33	yes
41	SWI5-MNase	Lupo et al.	BY4741	33	yes
42	FKH1-MNase	Lupo et al.	BY4741	33	yes
43	FKH2-MNase	Lupo et al.	BY4741	33	yes
44	Par alpha HO:Nat-mCherry	Tirosh et al.	OS142 MATα YDL227cΔ::TEF2pr-	-	no
45	free-MNase	This study	BY4741	-	yes

Table S3

Table S3: List of primers that were used for generating the yeast strains for this study.

Primer name	sequence	Use for generating
L3/L1-ATG-MNase Fw	cgtacgctgcaggtcgacggtggcggttctggcggtgg cggatccATGGCCACCAGTACCAAGAAGCT	all strains based on C-SWAT liberary
L4-KanR Rv	atcgatgaattcgagctcgttcaatcgatatctc aaagaagaacttcgatgaattcgagctcgtt	all strains based on C-SWAT liberary
Mnase check Rv	ctggcccttatacatcagcttc	all strains
CLG2_check	CCCAACCTTGGCTGAAGTAG	ABF1
CLG3_check	ATGGTACCTGCTCCATCAGC	REB1
CLG4_check	CCGCTGCTTCCAACTCTTAC	RAP1
CLG5_check	TGAGGTGGAAATTCAATGGAG	DOT6
CLG6_check	AGCCCGTATATGTCACCCAG	TOD6
CLG8_check	GATCCTTCGTCCTTGTCTCG	FHL1
CLG9_check	AACCCCGCTATCACACAATC	TBF1
CLG10_check	ACCTTTCCCCCAATACAACC	STB3
CLG11_check	AACTGCTACAATCAAGCCGC	PHO4
CLG12_check	AAATGACGCATTTGTTGGTG	PHO2
CLG13_check	CCAACAAGTGGACAATGACG	HAP4
CLG14_check	TCATGAACAGCATGACAACG	MBP1
CLG15_check	AGAGGATGCTTTCGATTCCG	SWI4
CLG16_check	ATGCCAAACCTAGATGGTGC	SKN7
CLG18_check	AGCGTCACCAACAGTGACAG	PHD1
CLG19_check	TGTAAGTACAATCGCTGCGG	SOK2
CLG20_check	ACCACCAGTTTCACGCCTAC	HMS2
CLG21_check	TCATCTCCATCTCCCTCCAC	ASH1
CLG24_check	ATTGCGAAAGTGGCGACTAC	MSN2
CLG26_check	ACGACTGCAAACGCTATCAC	GCR2
CLG27_check	TGGTTGGCCAAAGAAATACC	MIG1
CLG30_check	CTATGGGCTCGTCTCCAGTC	RGT1
CLG35_check	TCAATGAAGACGCTAATGCG	SUT1

Supplemental note 1: Comparison of our data to previously published data

We validated the quality of our dataset by comparing it to previously published datasets. In this comparison, we focused on TF binding at the promoter level, locations of binding peaks and motif preference.

Promoter binding signals

As a first test, we asked whether the TF binding signal in our data localize to known target gene promoters. Gene group enrichment is summarized in Table S4, showing that for most TFs, the set of top 100 promoters, bound by each TF in our dataset, is enriched with genes classified to the known functionalities of the respective factor and is further enriched by the known target genes as listed in the Saccharomyces genome database (SGD).

As a next test, we asked whether the promoter binding signal we obtain correlates with the respective binding signals reported in previous genome-wide TF binding mapping experiments. As no respective data for the yeast hybrid is available, we focused this comparison on the *S. cerevisiae* allele within the hybrid, which we assume better approximates the binding of the *S. cerevisiae* parent. The previously reported dataset included in this analysis, and their respective publications, are summarized in Table S5. In the case of Reb1 promoter, our binding is most correlated with the ChEC-seq (Zentner et al., 2015) and Cut & Tag (Brodsky et al., 2020) profiles (R = 0.67, R = 0.52) and moderately correlated with ChIP-exo (Rossi et al., 2021) and ChIP-chip (Venters et al., 2011) profiles (R = 0.4) (Figure S16 A). These correlations are of the same order, or higher than these found when comparing the other datasets between themselves. As can be appreciated, similar results are observed for the majority of other TF, for which previous data is available. As a second way of comparing the different datasets, obtained through different methods for TF-binding measurements, we consider the ranks of matched samples. For example, we asked whether our Reb1 profile is most similar to Reb1 samples of another dataset. In the case of Reb1 as well as for most other TFs, promoter binding matched standard datasets with high specificity (Figure S16 C).

Table S4

Table S4: GO enrichment analysis of top 100 target promoters bound by each TF to the *S. cerevisiae* genome. Functional groups include transcription factors targets (SGD), expression modules (Ihmels et al., 2002), GO slim, KEGG pathways, Glucose signaling (Broach, 2012). Tbf1 target gene list was obtained from Preti et al., 2010. Number of TF targets is indicated in parenthesis.

TF	Description (SGD)	Group	-log10(p-value)
ABF1	DNA binding protein with possible chromatin-reorganizing activity; involved in transcriptional activation, gene silencing, and DNA replication and repair	**ABF1 (328)**	17
		cytoplasm	4
		FKH1 (2782)	4
		cellular respiration	3
		protein modification	2
REB1	DNA binding protein that binds to genes transcribed by both RNA polymerase I and RNA polymerase II	**REB1 (954)**	20
		FKH2 (1602)	7
		FKH1 (2782)	5
		FHL1 (366)	4
		SKN7 (165)	4
RAP1	Essential DNA-binding transcription regulator that binds many loci	**RAP1 (856)**	17
		Protein synthesis	5
		FKH2 (1602)	4
		protein biosynthesis	3
		ribosome	3
DOT6	Protein involved in rRNA and ribosome biogenesis	RAP1 (856)	6
		FKH2 (1602)	6
		TOD6 (8)	6
		FKH1 (2782)	5
		MSN2 (492)	5
TOD6	PAC motif binding protein involved in rRNA and ribosome biogenesis	nucleolus	9
		rRNA processing	9
		ribosome biogenesis and assembly	6
		RNA metabolism	5
		nucleus	4
FHL1	Regulator of ribosomal protein (RP) transcription	**Protein synthesis**	98
		RAP1 (856)	63
		ribosome	56
		structural molecule activity	46
		protein biosynthesis	40
TBF1	Telobox-containing general regulatory factor	**TBF1 (136)**	66
		REB1 (954)	4
		cytoplasm	3
		transcription	3
		motor activity	3
STB3	Ribosomal RNA processing element (RRPE)-binding protein	**Protein synthesis**	81
		RAP1 (856)	50
		ribosome	45
		structural molecule activity	36
		FHL1 (366)	35
PHO4	Basic helix-loop-helix (bHLH) transcription factor of the myc-family; activates transcription cooperatively with Pho2p in response to phosphate limitation	**PHO4 (49)**	14
		RAP1 (856)	12
		FKH2 (1602)	6
		GCR2 (72)	5
		Glycolysis / Gluconeogenesis	5
PHO2	Homeobox transcription factor; regulatory targets include genes involved in phosphate metabolism	**PHO2 (137)**	12
		FKH2 (1602)	8
		plasma membrane	7
		PHO4 (49)	6
		RAP1 (856)	6
HAP4	Subunit of the heme-activated, glucose-repressed Hap2p/3p/4p/5p CCAAT-binding complex, a transcriptional activator and global regulator of respiratory gene expression	**HAP4 (49)**	48
		Oxidative phosphorylation	48
		Repressed in Glu but not in Ras2	40
		mitochondrial membrane	34
		mitochondrion	32
MBP1	involved in regulation of cell cycle progression from G1 to S phase	cell-cycle (G1)	86
		MBP1 (110)	64
		cell cycle	21
		SWI4 (225)	19
		chromosome	17
SWI4	DNA binding component of the SBF complex (Swi4p-Swi6p)	**SWI4 (225)**	64
		MBP1 (110)	32
		cell-cycle (G1)	22
		ASH1 (71)	17
		cell wall	15

SKN7	Required for optimal induction of heat-shock genes in response to oxidative stress	**SKN7 (165)**	48
		FKH2 (1602)	20
		SWI4 (225)	16
		MSN2 (492)	10
		SOK2 (69)	10
PHD1	Transcriptional activator that enhances pseudohyphal growth	SOK2 (69)	21
		PHD1 (86)	17
		SUT1 (82)	12
		SKN7 (165)	12
		FKH2 (1602)	9
SOK2	Nuclear protein that negatively regulates pseudohyphal differentiation	PHD1 (86)	24
		SOK2 (69)	21
		SUT1 (82)	20
		SKN7 (165)	20
		FKH2 (1602)	15
HMS2	Protein with similarity to heat shock transcription factors	SKN7 (165)	33
		SOK2 (69)	29
		FKH2 (1602)	19
		PHD1 (86)	18
		MSN2 (492)	17
ASH1	Component of the Rpd3L histone deacetylase complex; zinc-finger inhibitor of HO transcription	nucleolus	10
		rRNA processing	9
		ribosome biogenesis and assembly	7
		Purine metabolism	5
		RNA metabolism	5
MSN2	Stress-responsive transcriptional activator	**MSN2 (492)**	21
		stress	13
		SKN7 (165)	13
		carbohydrate metabolism	10
		Starch and sucrose metabolism	10
GCR2	Transcriptional activator of genes involved in glycolysis; interacts and functions with the DNA-binding protein Gcr1p	Glycolysis / Gluconeogenesis	25
		GCR2 (72)	22
		generation of precursor metabolites and energy	17
		carbohydrate metabolism	16
		RAP1 (856)	15
MIG1	Transcription factor involved in glucose repression	Repressed in Glu but not in Ras2	28
		Repressed by glu but not through sch9/pka	18
		Carbohydrate metabolism	13
		Induced by glu through RGT2	12
		MIG1 (44)	11
RGT1	Glucose-responsive transcription factor; regulates expression of several glucose transporter (HXT) genes in response to glucose	**Induced by glu through RGT2**	30
		Induced not through pka/sch9	14
		SUT1 (82)	10
		SKN7 (165)	8
		MSN2 (492)	7
SUT1	positively regulates sterol uptake genes under anaerobic conditions; involved in hypoxic gene expression	**SUT1 (82)**	18
		SKN7 (165)	13
		SOK2 (69)	12
		MSN2 (492)	11
		PHD1 (86)	11
FKH1	rate-limiting replication origin activator; regulates transcription elongation, chromatin silencing at mating loci, expression of G2/M phase genes	cell cycle(G2/M)	16
		FKH2 (1602)	15
		FKH1 (2782)	14
		cell cycle	6
		microtubule organizing center	5
FKH2	rate-limiting activator of replication origins; major role in expression of G2/M phase genes	cell cycle(G2/M)	44
		FKH2 (1602)	42
		FKH1 (2782)	17
		SKN7 (165)	9
		SWI4 (225)	7
ACE2	Transcription factor required for septum destruction after cytokinesis	SWI5 (125)	23
		ACE2 (75)	21
		cell wall	13
		SUT1 (82)	12
		FKH2 (1602)	12
SWI5	Transcription factor that recruits Mediator and Swi/Snf complexes	**SWI5 (125)**	30
		ACE2 (75)	15
		cell wall	13
		SUT1 (82)	7
		bud	6

Location of binding Peaks

We next compared our data to other datasets at the peak level. For this, we considered our peak calling at the *S. cerevisiae* genome, comparing it with publically available peak tables. Of note, different methods were used in each study for peak calling (Table S5). Furthermore, the spatial resolution differs between the datasets, depending on the methodology for binding measurements, influencing the precise peak location. To account for that, we considered regions that are 60 bp wide surrounding each of the reported peak. This resulted in a high similarity between our peaks and previously defined peaks. In the case of Reb1, for example, peak overlap ranged at 40 %– 63% between our dataset and other datasets (Figure S16 B). We noticed that the peak table found in Zentner et al., 2015 supplementary material did not correlate well with our data nor with any of the other datasets, therefore we recalled peaks from that dataset using our method for peak calling (Materials and methods), which led to a higher correlation overall. We compared 25 of the TF we examined to the Rossi et al. ChIP-exo dataset, what resulted in a good match (matching rank of 1 to 3) for 17 TFs (Figure S16 D).

We note that the number of peaks we called for the general TFs (Abf1, Reb1, Rap1) was comparable to that reported in other datasets: 1000-1500 peaks in our data, 300-1500 peaks in Rossi et al, 1000-2000 peaks in Kasinathan et al., 2000-3700 peaks in Zentner et al (considering only the fast-cleaving peaks, as these reported as TF-specific) (Figure S16 E). The number of peaks called in the ChIP-exo dataset (Rossi et al., 2021) was highly variable between TFs and between experimental replicates, with overall smaller number of peaks compared to ours, or other reports (Figure S16 E). We re-called peaks from the ChIP-exo dataset and observed an average of 600 peaks per TF (min = 200, max = 900 peaks) while peak overlap percentages did not change significantly from those called in the original paper (not shown).

To summarize this part, we find our experimental data and peak calling methods highly comparable with standard datasets in the field.

Table S5

Table S5: list of publications used for data comparison.

Promoter-level comparison			
Publication	Method	Number of samples	Number of TFs
Zentner et al., 2015	ChEC-seq	56	4
Rossi et al., 2021	ChIP-exo	1227	791
Harbison et al., 2004	ChIP-chip	351	204
Venters et al., 2011	ChIP-chip	404	202
Brodsky et al., 2020	ChIP	14	11
Brodsky et al., 2020	Cut & Tag	34	14
Peak comparison (peak calling from original publication)			

Zentner et al., 2015	ChEC-seq	4*	4*
Kasinathan et al., 2014	ORGANIC	6	2
Rossi et al., 2021	ChIP-exo	1227 (45**)	791 (25**)

* Including: Abf1, Reb1, Rap1 and free-MNase
** For Rossi et al., 2021 data, peak overlap was computed only for the TFs that were profiled in our study, number indicated in brackets.

Figure S16

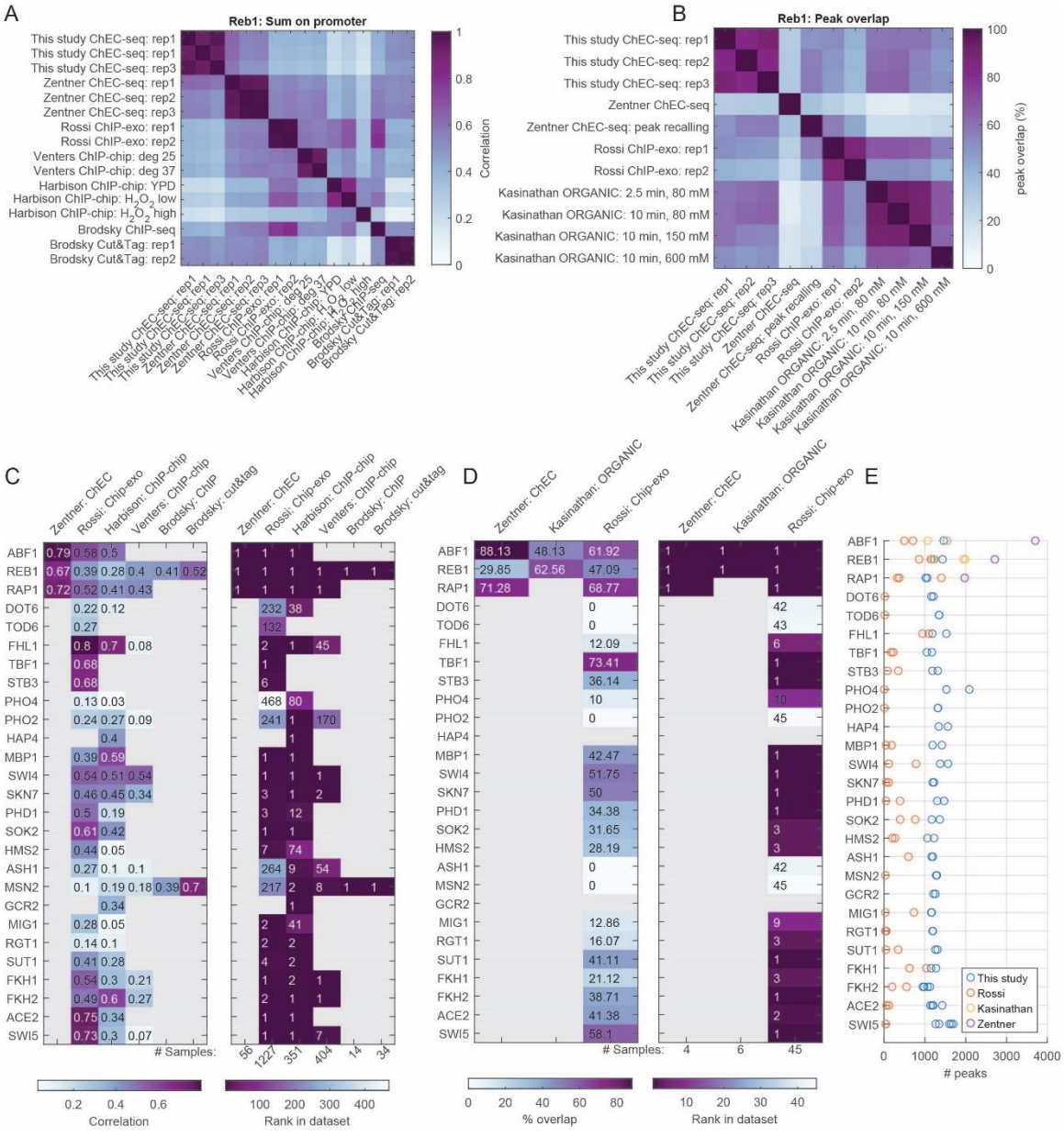


Figure S16: comparison to previously published data. A) Correlation of promoter binding of Reb1 to the *S. cerevisiae* genome between datasets (references in Table S5). Presented is the Pearson correlation coefficient of sum of signal on all 6701 yeast promoters. B) Overlap of Reb1 binding peaks between datasets. As the exact peak location changes between methods, we enlarged the peak width in this comparison to 60 bp. Here we compare the original peak-calling tables from each study with peaks called on

the *S. cerevisiae* genome in the hybrid samples of this study. For Zentner et al., 2015 we considered only the fast-cleaving peaks and in addition we re-called peaks from the raw data (samples taken after 30 seconds in Ca^{2+}). Shown is the number of overlapping peaks divided by the smaller set of peaks. Number of peaks are shown in E. C) Correlation of promoter binding between this study and other datasets, per TF. Left: Pearson correlation coefficient, right: the rank of the matched sample among all samples within each dataset. The number of samples in each dataset is indicated at the bottom of the right panel. D) Peak overlap between this study and other datasets Annotations as in C. E) Number of peaks in each study.

Motifs

Here we refer to *in vitro*-defined motifs as the gold standard of TF binding preference, as these *in vitro* experiments measure the protein's preferential binding and are not affected by cellular factors such as chromatin conformation and interacting proteins. The set of previously published *in vitro* motifs was taken from YetFasco (De Boer and Hughes, 2012) and Jaspar (Sandelin et al., 2004) databases and is described in Table S1. For two TFs in our dataset (Hms2, Hap4), in-vitro motifs were not available, and we therefore used the motifs derived from previous *in vivo* measurements. These published motifs are presented along with motifs derived from our data using two methods: 1) enrichment of 7-mer sequences at locations bound by the TF (7-mers) 2) motif finding algorithm (MEME-ChIP), to which we fed sequences of 60 bp around top-bound peaks as input (Machanick and Bailey, 2011). In both cases, the genomic data used for generating the motifs include both of the hybrid genomes. We find a good correspondence of the motifs derived from our *in vivo* data and the previously reported *in vitro* motif. TFs that show similar motifs include: Abf1, **Reb1**, **Rap1**, **Dot6**, **Tod6**, **Tbf1**, Stb3, **Pho4**, **Mbp1**, **Swi4**, **Skn7**, **Phd1**, **Sok2**, **Hms2**, Ash1, **Msn2**, Gcr2, **Mig1**, **Rgt1**, **Sut1**, **Fkh1**, Fkh2, **Ace2** and **Swi5** (Factors written in bold showed a similar seqlogo for both motif discovery methods: 7-mer and MEME-ChIP, Table S1).

A minority of TFs in our dataset did not show good similarity with the *in vitro* motif. These include Fhl1, Pho2 and Hap4. Fhl1, as well as another TF in our data, Stb3, are activators of ribosomal proteins genes. Both TF tended to localize at 7-mer containing the common sequence 'CCTAG'. However, MEME-ChIP analysis resulted in a different motif preference: Stb3 motif is similar to its in vitro motif ('TTTTTCA') and Fhl1 motif is similar to the PAC motif ('CTCATC') that is recognized by the regulators of ribosomal genes Tod6/Dot6. Fhl1 was shown to bind the 'GACGCA' motif only *in vitro*, with no enrichment for this motif *in vivo* (Gordân et al., 2011). The differences in motifs preferences for Stb3 predicted by different methods (using the same data) might result from the essence of each approach: while the input of MEME-ChIP includes sequences at peaks but not the relative peak level, 7-mer motif score takes into account also the binding strength within a 20 bp window surrounding the respective 7-mer. AT-rich sequences appear at high frequency in nucleosome depleted regions in gene promoters (Kaplan et al., 2009). The high binding to 'TTTTTCA' 7-mer is diluted by many non-bound sites and this 7-mer receives a low score, despite its high frequency in Stb3-bound promoters (Figure S17 A).

Another difference between the two methods for motif finding we employed was observed in paralogous TFs Fkh1 and Fkh2; both of these TFs bind the same motif *in vitro.* In our dataset, however, predictions based on 7-mer enrichment assigned the *in vitro* motif only to Fkh1, while Fkh2 was assigned a different motif. Direct comparison of Fkh1 and Fkh2 motif scores distinguish a common Fkh1 branch ('GTAAACA') from a branch unique to Fkh2 (GC-rich) (Figure

S17 B). This unique Fkh2 branch may result from co-binding of Fkh2 with other TFs, for example Mcm1 or Ndd1 (Pic et al., 2000).

Figure S17

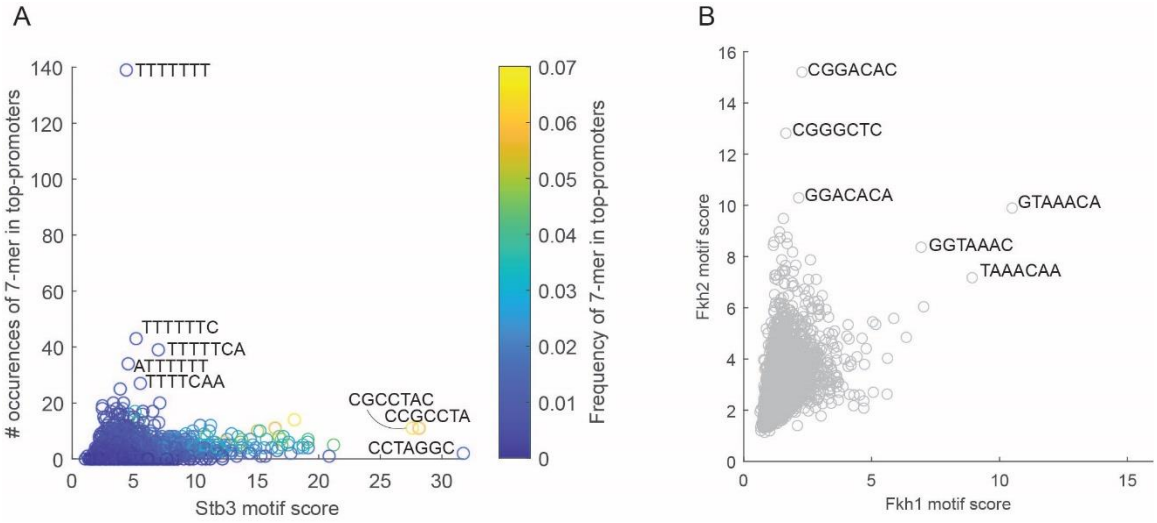


Figure S17: cases of complex motif preference. A) 7-mer motif score of Stb3 does not recognize the 'TTTTTCA' motif although it appears in many Stb3-bound promoters. Shown are the Stb3 7-mer motif score (x-axis), the number of times each 7-mer appears in the top 20 Stb3-bound promoters (y-axis) and their frequency compared to all gene promoters (color axis). B) Fkh2 prefers two types of motifs. Shown are the 7-mer motif score of Fkh1 and Fkh2. Note the two distinct branches, where Fkh1 motifs are also highly-preferred by Fkh2.

Supplemental note 2: free-MNase as a negative control

Our analysis relies on comparing binding peaks between the *S. cerevisiae* and the *S. paradoxus* genomes. It is therefore sensitive to false-positive peaks, namely peaks that do not represent actual binding events. As we only use peaks reproducible in multiple repeats, false positives, if exist, do not come from random measurement errors. However, they could still arise from systematic biases of the method. To control for such biases, we used a negative control in which we consider cells expressing an MNase not fused to any TF (free-MNase). This free MNase therefore provides the cleavage profile in the absence of TF-dependent DNA localization. We therefore mapped the binding profile of this free-MNase expressed in our hybrid cells, using short incubation time in calcium (30 seconds), following the same procedure used for our TF. In this, we followed previous practice (Zentner et al., 2015, 2021). Further, to gain a high signal, we over-expressed the free-MNase under the *TDH3* promoter.

We first examined the data by plotting the metagene profile, namely, the average signal of all genes (6701) centered at the transcription start site (TSS). This profile was low and flat, being highly distinguished from these obtained for TF-fused MNase in our dataset, see e.g. Reb1-MNase (Figure S18 A). In this sense our free-MNase control was similar to the respective profile described by Zentner et al., and was different from their long calcium incubations that peak at nucleosome free regions and linkers (Zentner et al., 2021).

We next used the free-MNase profile to assess the false positive rate of TF-MNase peaks in our dataset. Binding peaks that are common to both TF-MNase and free-MNase are not TF-specific, and therefore represent false positives. We plotted the overlap percentage of peaks called from our data (*S. cerevisiae* genome of the hybrid, labels in red) and peaks recalled from genomic profiles of the original ChEC-seq paper (labels in black, Figure S18 B). Notably the free-MNase samples taken at short calcium incubations showed little overlap between repeats (2% overlap) and were fully distinct from all TF profiles in our data. Of note, free-MNase at longer incubation times did generate a typical cleavage pattern as previously discussed (Mittal et al., 2021; Zentner et al., 2021). The number of peaks was comparable between the different samples: ranging between 500 - 1000 peaks for the free-MNase samples, and 700 – 1500 peaks for the TF-MNase samples.

Next, we examined the overall peak overlap between samples of our dataset. Experimental replicates were highly similar in terms of peak overlap and peak profiles were highly distinguishable between TFs but were similar between related TFs such as the duplicated TFs Ace2 and Swi5 (Figure S18 C). We find a small peak overlap (10-20%) between different, unrelated, TFs (Figure S18 C). Furthermore, when considering only motif-associated peaks, which are the focus of the analysis in this paper, background overlap was reduced to 2-4 % (Figure S18 D). Experimental replicates of the same TF showed high peak overlap, between 64%-97% with median of 87%.

Finally, to assess the false-positive rate of peak calling for each TF we averaged the genomic signal of experimental replicates and called peaks (Figure S18 E). We find that only 2.7%– 4.4% of TF peaks overlapped with free-MNase peaks (corresponding to 71/2547 and 118/2670 of the peaks in Stb3 and Skn7, representing the lowest and the highest overlap values). These few overlapping peaks were not enriched for- nor depleted from- motifs, as their fraction of motif

overlap matched the overall peak fraction (Figure S18 F). See figure S5 A for the fraction of peaks that reside next to a motif compared with the fraction of random sites that reside next to a motif.

We conclude that our data and peak finding procedure result in highly specific TF binding peaks with low (2.5%-4%) false positive rate.

Figure S18

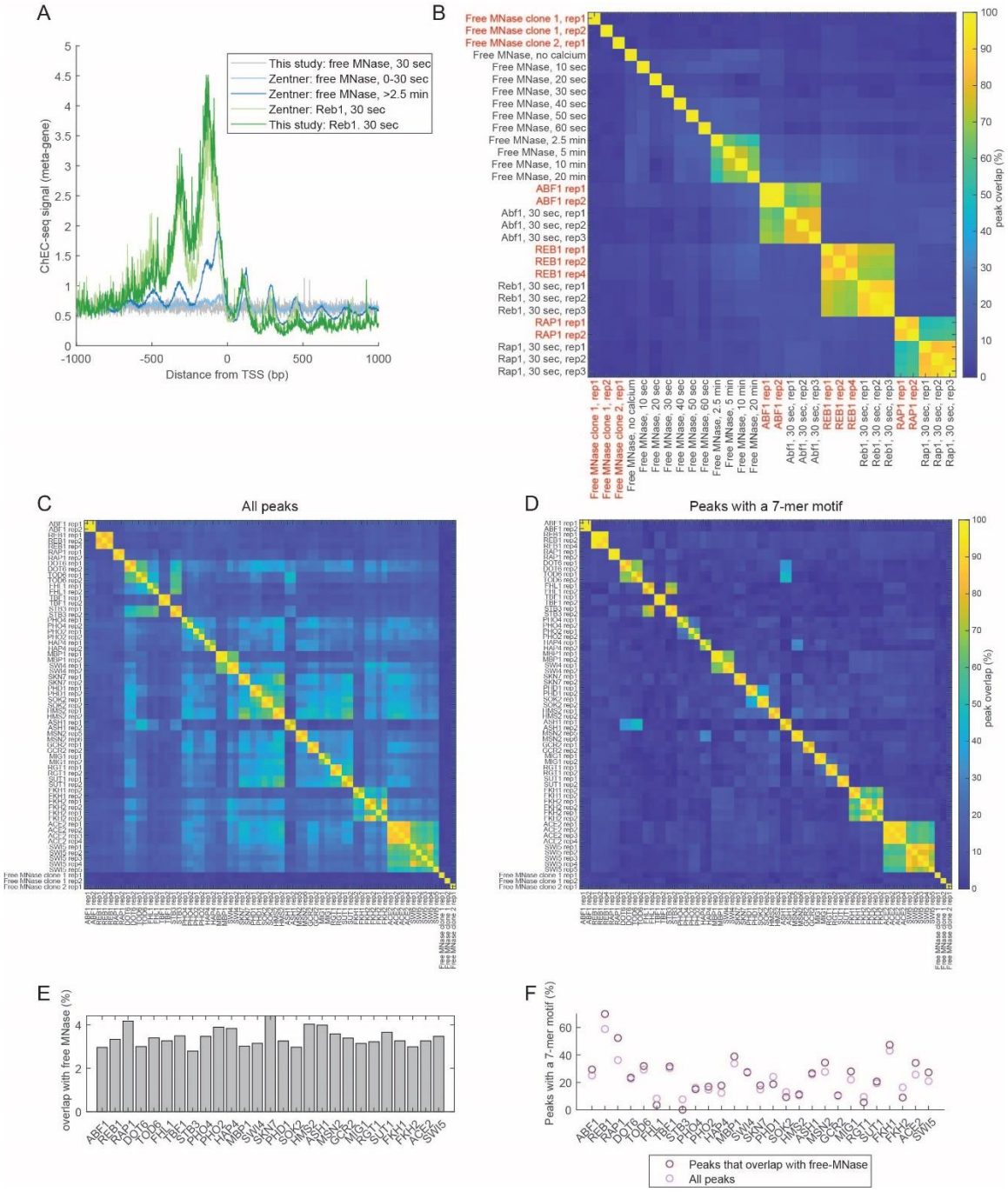


Figure S18: Free-MNase as a negative control. A) Free-MNase profiles of short calcium incubations are highly distinguished from a TF-MNase signal. Shown are metagene profiles centered at the transcription start site of all yeast genes (6701). B) Free-MNase peaks

do not overlap with TF-MNase peaks. Red labels indicate samples from this study, black labels indicate samples from the original ChEC-seq paper (Zentner et al., 2015). Shown is the number of peaks that overlap divided by the smaller list of peaks, where the number of peaks is overall comparable (see text). Note that long calcium incubations result in a typical peak profile, and were avoided in this study and in previously published studies from our lab (Bar-Ziv et al., 2020; Brodsky et al., 2020; Lupo et al., 2021). C) Peak overlap between all experimental samples of our dataset, using all detected peaks. D) Peak overlap between all experimental samples, restricting only to 7-mer motif-associated peaks. E) Low false positive rate of peak calling, defined as the percentage of TF-MNase peaks that overlap free-MNase peaks. Presented data is the average of all TF-MNase and free-MNase replicates. F) Peaks that are common to TF-MNase and free-MNase are not depleted of motifs. Shown is the percentage of peaks associated with a motif, for peaks that are common to TF-MNase and free-MNase and for all peaks.

Supplemental note 3

As most TFs bind a small number of target promoters at high level, we found the Pearson correlation coefficient most appropriate to describe the similarity in TF binding between two samples or two orthologous alleles. However, we also observed the data through different correlation measurements and transformations. In the following figure (Figure S19) we present: Pearson correlation on the linear data (as we show in the manuscript), Spearman correlation on linear data, Pearson correlation on $\log_2$-transformed data and Pearson correlation on $\log_2$(sum on promoter + 700). In the last measure, we added a minimal promoter binding level. Spearman coefficient mostly agrees with the log-transformed measures. However, to our understanding, as most TFs bind a small number of targets, the Pearson linear correlation is the most appropriate one.

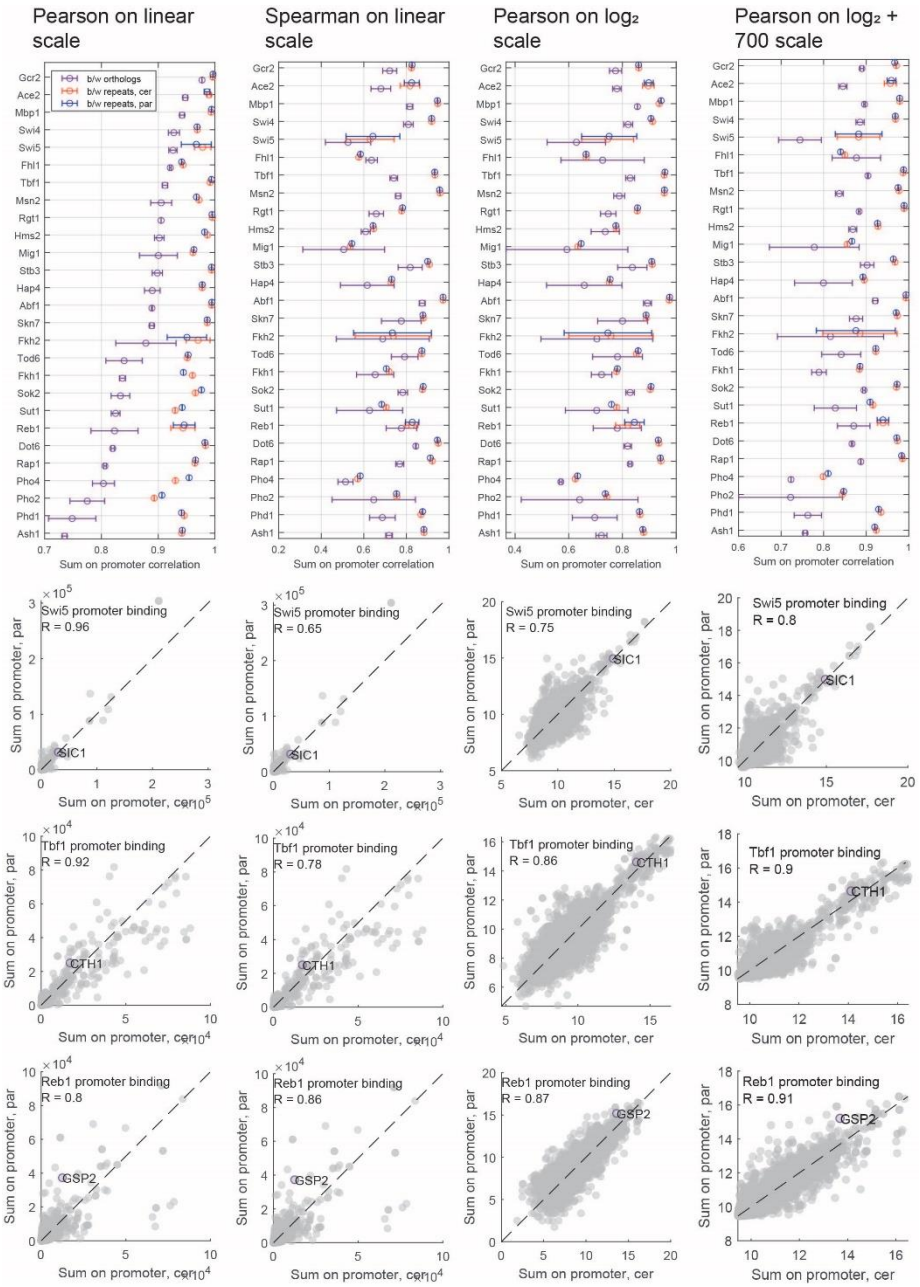


Figure S19: Comparison of measurements for correlation of promoter binding.

References

Bar-Ziv, R., Brodsky, S., Chapal, M., and Barkai, N. (2020). Transcription Factor Binding to Replicated DNA. Cell Rep. *30*, 3989-3995.e4.

Broach, J.R. (2012). Nutritional Control of Growth and Development in Yeast. Genetics *192*, 73–105.

Brodsky, S., Jana, T., Mittelman, K., Chapal, M., Kumar, D.K., Carmi, M., and Barkai, N. (2020). Intrinsically Disordered Regions Direct Transcription Factor In Vivo Binding Specificity. Mol. Cell

79, 459-471.e4.

Gordân, R., Murphy, K.F., McCord, R.P., Zhu, C., Vedenko, A., and Bulyk, M.L. (2011). Curated collection of yeast transcription factor DNA binding specificity data reveals novel structural and gene regulatory insights. Genome Biol. *12*.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017–1018.

Grömping, U. (2006). Relative importance for linear regression in R: the package relaimpo. J. Stat. Softw. *17*, 1–27.

Harbison, C.T., Gordon, D.B., Lee, T.I., Rinaldi, N.J., Macisaac, K.D., Danford, T.W., Hannett, N.M., Tagne, J.B., Reynolds, D.B., Yoo, J., et al. (2004). Transcriptional regulatory code of a eukaryotic genome. Nature *431*, 1–5.

Ihmels, J., Friedlander, G., Bergmann, S., Sarig, O., Ziv, Y., and Barkai, N. (2002). Revealing modular organization in the yeast transcriptional network. Nat. Genet. *31*, 370–377.

Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J., et al. (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. Nature *458*, 362–366.

Kasinathan, S., Orsi, G.A., Zentner, G.E., Ahmad, K., and Henikoff, S. (2014). High-resolution mapping of transcription factor binding sites on native chromatin. Nat. Methods *11*, 203–209.

Lupo, O., Krieger, G., Jonas, F., and Barkai, N. (2021). Accumulation of cis- and trans-regulatory variations is associated with phenotypic divergence of a complex trait between yeast species. G3 Genes|Genomes|Genetics *11*, jkab016.

Machanick, P., and Bailey, T.L. (2011). MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics *27*, 1696–1697.

Mittal, C., Rossi, M.J., and Pugh, B.F. (2021). High similarity among ChEC-seq datasets. BioRxiv 2021.02.04.429774.

Pic, A., Lim, F.L., Ross, S.J., Veal, E.A., Johnson, A.L., Sultan, M.R., West, A.G., Johnston, L.H., Sharrocks, A.D., and Morgan, B.A. (2000). The forkhead protein Fkh2 is a component of the yeast cell cycle transcription factor SFF. EMBO J. *19*, 3750–3761.

Preti, M., Ribeyre, C., Pascali, C., Bosio, M.C., Cortelazzi, B., Rougemont, J., Guarnera, E., Naef, F., Shore, D., and Dieci, G. (2010). The Telomere-Binding Protein Tbf1 Demarcates snoRNA Gene Promoters in Saccharomyces cerevisiae. Mol. Cell *38*, 614–620.

Rossi, M.J., Kuntala, P.K., Lai, W.K.M., Yamada, N., Badjatia, N., Mittal, C., Kuzu, G., Bocklund, K., Farrell, N.P., Blanda, T.R., et al. (2021). A high-resolution protein architecture of the budding yeast genome. Nature.

Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., and Richards, S. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. *15*, 1034–1050.

Tirosh, I., Sigal, N., and Barkai, N. (2010). Divergence of nucleosome positioning between two closely related yeast species: genetic basis and functional consequences. Mol. Syst. Biol. *6*, 365.

Venters, B.J., Wachi, S., Mavrich, T.N., Andersen, B.E., Jena, P., Sinnamon, A.J., Jain, P., Rolleri, N.S., Jiang, C., Hemeryck-Walsh, C., et al. (2011). A comprehensive genomic binding map of gene and chromatin regulatory proteins in Saccharomyces. Mol. Cell *41*, 480–492.

Zentner, G.E., Kasinathan, S., Xin, B., Rohs, R., and Henikoff, S. (2015). ChEC-seq kinetics discriminates transcription factor binding sites by DNA sequence and shape in vivo. Nat. Commun. *6*, 8733.

Zentner, G.E., Policastro, R.A., and Henikoff, S. (2021). ChEC-seq produces robust and specific maps of transcriptional regulators. BioRxiv 2021.02.11.430831.