Supplementary Material for:

Gcn4 binding in coding regions can activate internal and canonical 5' promoters in yeast

Yashpal Rawal[1#], Răzvan V. Chereji[2#], Vishalini Valabhoju[1], Hongfang Qiu[1], Josefina Ocampo[2], David J. Clark[2,†], and Alan G. Hinnebusch[1,†]

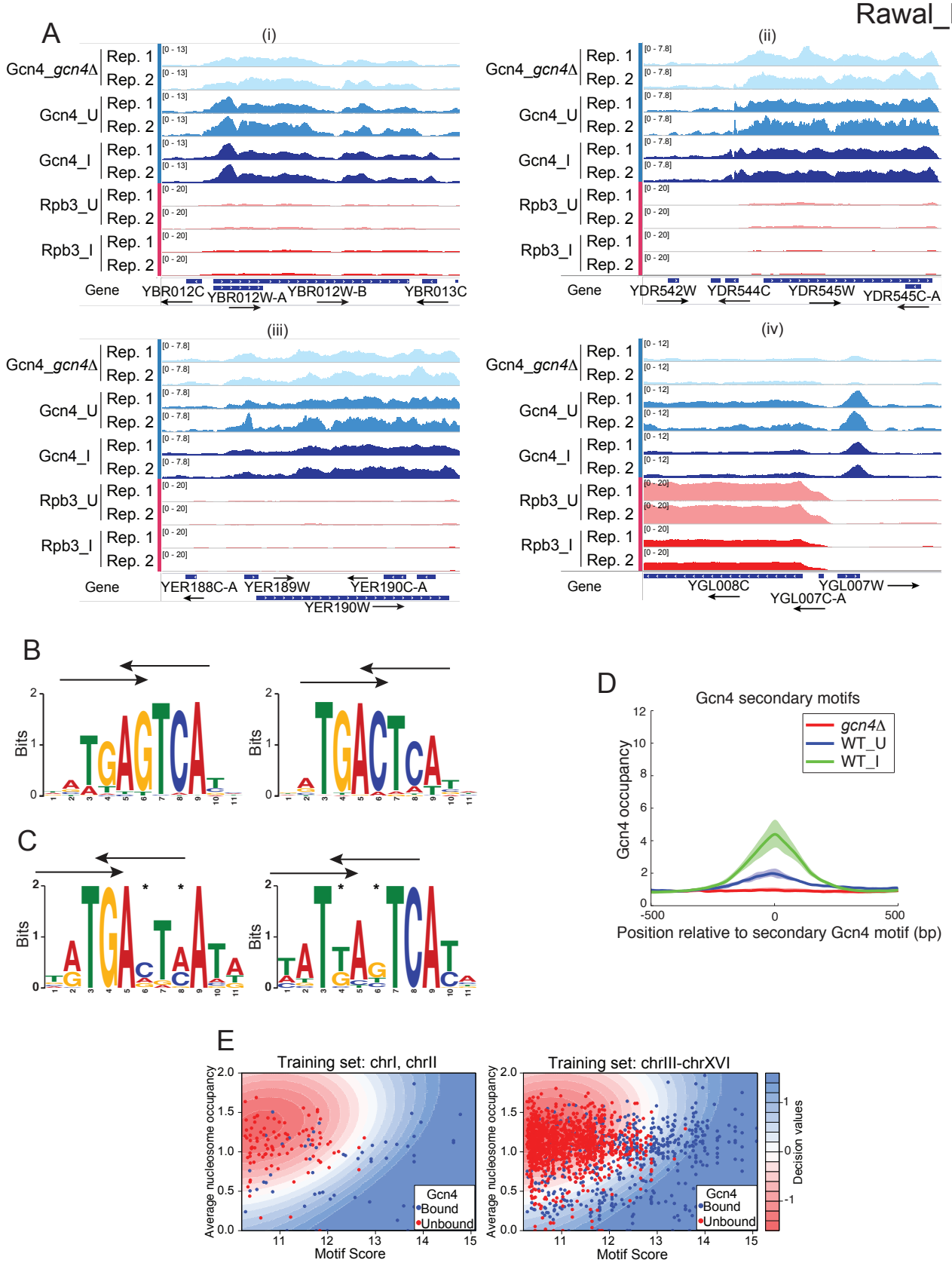

Figure S1. Supporting analysis of ChIP-seq determination of genome-wide Gcn4 occupancies in vivo, related to Figure 1 (See legend on following page).

Figure S1. Supporting analysis of ChIP-seq determination of genome-wide Gcn4 occupancies in vivo, related to Figure 1. (A) Examples of hyper-ChIPable regions showing non-specific occupancies in Gcn4 ChIP-seq. Gcn4 occupancies from *gcn4Δ* SM-induced (Gcn4_*gcn4Δ*), WT SM-induced (Gcn4_I) or WT uninduced cells (Gcn4_U) for two biological replicates (Rep.1, Rep.2) (tracks 1-6), or replicates of Rpb3 occupancies from the corresponding WT uninduced (Rpb3_U) or SM-induced (Rpb3_I) cells (tracks 7-10) plotted for the indicated genes using the Integrated Genomics Viewer (IGV). All profiles have been normalized such that the average occupancy for each chromosome equals one. (B) The consensus Gcn4 binding motif was discovered by MEME analysis of 546 sequences occupied by Gcn4 peaks. Forward strand (left); reverse strand (right). (C) Motifs discovered by MEME by analyzing 75 sequences occupied by Gcn4 peaks but lacking a strong match to the consensus sequence in (B): forward (left) and reverse strand (right). (D) Gcn4 occupancies at secondary motifs. Average Gcn4 occupancy for the 40 peaks containing a match to the secondary motif in (C). Different colors indicate the Gcn4 levels in WT cells before (WT_U) and after (WT_I) SM treatment, and in SM-treated *gcn4Δ* cells. The solid lines show the averages of multiple replicates (WT_U – 3 replicates; WT_I – 5 replicates; *gcn4Δ*– 3 replicates), and the shaded areas show the ranges of values for individual replicates. (E) Machine learning, using a type 1 support vector machine (SVM) algorithm, was used to classify the Gcn4 binding status of all predicted Gcn4 motifs, using as classification features the FIMO scores and average nucleosome occupancies of the corresponding genomic loci. All Gcn4 motifs on chromosomes I and II were used as training data, and the motifs from all other chromosomes were used to test the performance of the classification. The results of the training phase are arrayed in a plot of nucleosome occupancy versus FIMO score in which each sector is assigned a color-coded decision value (left). The decision values shown in different levels of red or blue, represent a measure of the distance from a test point to the curve separating the two classes of motifs; a positive value (blue background color) represents a predicted Gcn4-bound motif, while a negative value (red background color) represents a predicted Gcn4-unbound motif. The colors of the dots represent the true status of each Gcn4 motif: bound (blue dot) or unbound (red dot). According to the model, motifs with high FIMO scores are bound by Gcn4 regardless of nucleosome occupancy (far-right sectors), whereas motifs with low scores are bound only if nucleosome-free (lower-left sectors). The model succeeded in predicting the Gcn4 binding status of 83% of the motifs on the remaining 14 chromosomes (right), much greater than the percentage of all 1754 motifs identified experimentally as Gcn4-bound (31%). Thus, whereas a strong match to the consensus motif is the most important determinant of Gcn4 binding, motifs with inferior matches bind Gcn4 if located in sequences relatively depleted of nucleosomes. We attempted to identify other features that would improve the classification accuracy of the Gcn4 motifs, including accessibility to DNase I and occupancies of Abf1, H2A.Z, TBP and subunits of chromatin remodelers, but found only a marginal improvement of the classification accuracy using SVM (see Methods for further details).

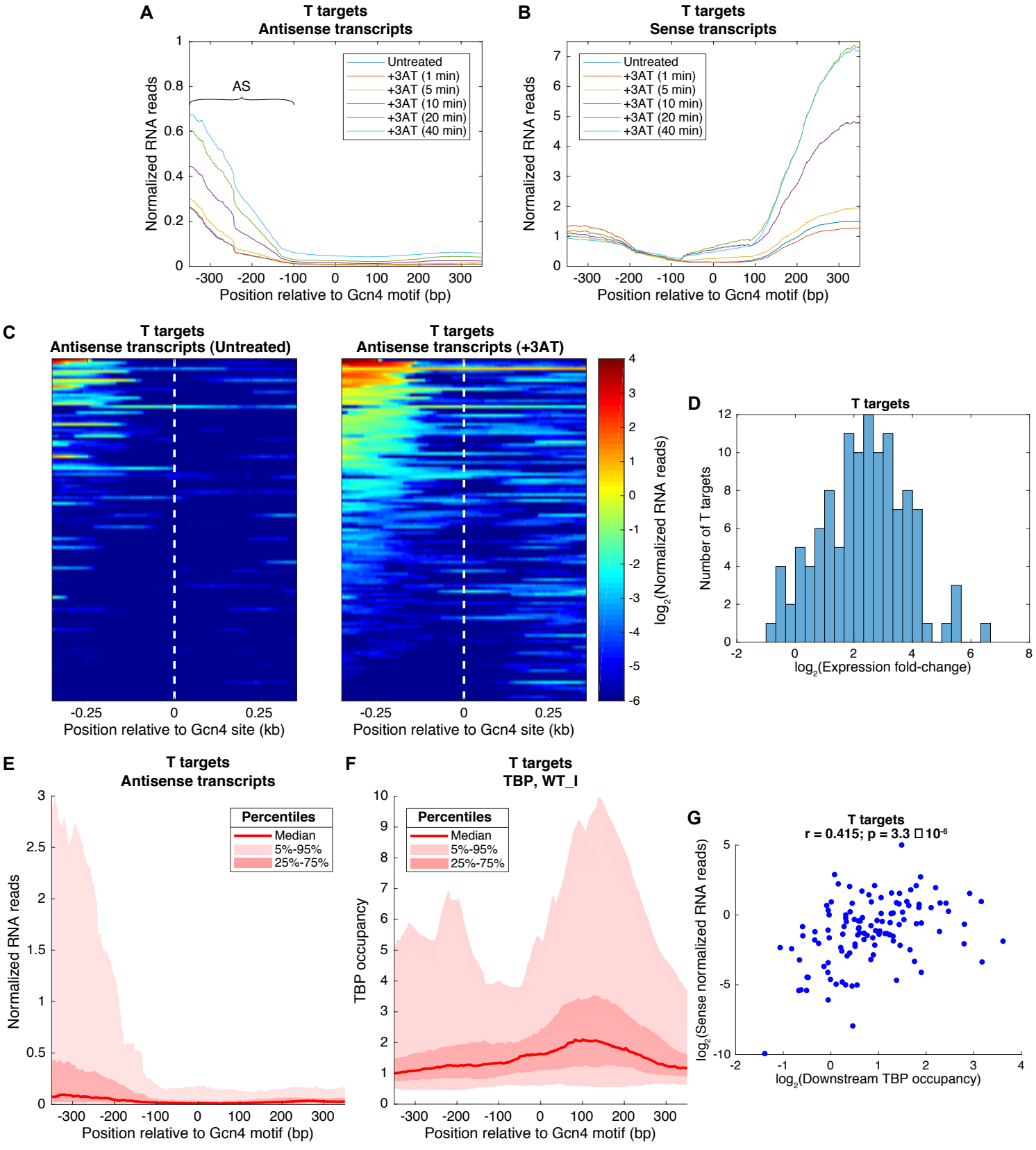

Figure S2. Gcn4 binding to "T" targets also induces bidirectional transcription, related to Figure 5 (See legend on following page).

Figure S2. Gcn4 binding to "T" targets also induces bidirectional transcription, related to Figure 5. (A-B) Average abundance of (A) AS and (B) sense transcripts initiated from the "T" targets, before and during a time course of 3AT treatment. Note that AS transcripts are ~10-fold less abundant than induced FL sense transcripts after 40 min of induction (panel B; note different y-axis scales in panels A-B). Low-level AS transcripts also map downstream of the Gcn4 motifs at T genes but they are not induced by 3-AT (A) and presumably derive from constitutive transcription of the non-coding strands of CDSs, possibly driven by transcriptional activators besides Gcn4. There are also uninduced sense transcripts mapping upstream of the Gcn4 motif (B) that most likely derive from constitutive transcription of adjacent upstream genes arranged in the same orientation as the induced T genes. (C) Heat maps showing the AS transcripts originating from all "T" targets, before (left panel) and after (right panel) 40 min. of 3AT treatment. The Gcn4 targets are sorted by the abundance of AS transcripts after induction. (D) Histogram of the AS expression fold-change for the "T" targets. All except seven "T" targets show an increase in antisense transcription ($\log_2$(Fold-change) > 0). (E) Distribution of the abundance of AS transcripts, after 40 min. of 3AT treatment, depicted as in Fig. 5E. Although the median level of transcription (solid red line) is low, the top 25 percentiles (above the upper boundary between the light pink and dark pink shaded area, corresponding to the 75 percentiles) produce AS transcripts at an abundance comparable to the mean level of sense transcription shown in (B). (The RNA reads are normalized such that the genomic average RNA abundance is one). Light pink shaded area – the range between the 5-95 percentiles; dark pink shaded area – the range between the 25-75 percentiles. (F) TBP occupancy distribution near the "T" targets, depicted as in Fig. 5F. Light pink shaded area – the range between the 5-95 percentiles; dark pink shaded area – the range between the 25-75 percentiles; red line – median occupancy. (G) Scatter plot representing the $\log_2$ normalized sense RNA reads in the 350bp regions downstream of the Gcn4 motifs versus the $\log_2$ of the average occupancy of TBP in the same loci. Pearson correlation coefficient: r = 0.415. An F-test for a linear fit gives a p-value = 3.3 x 10-6; null hypothesis: coefficient of proportionality is equal to zero.

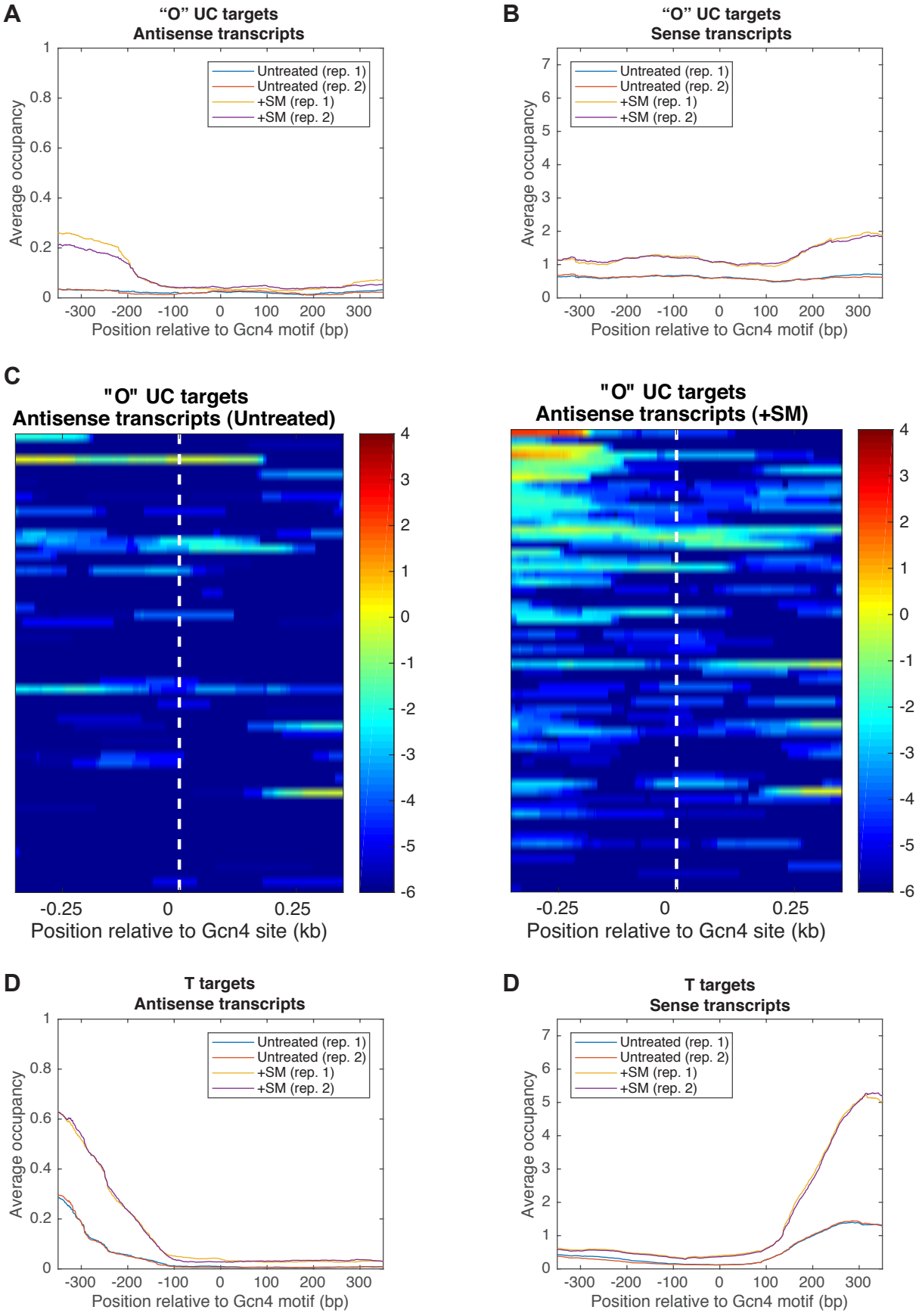

Figure S3. Gcn4 binding on SM-treatment induces bidirectional transcription both within ORFs at "O" target genes and upstream of ORFs at "T" target genes, related to Figure 5 (See legend on following page).

Figure S3. Gcn4 binding on SM-treatment induces bidirectional transcription both within ORFs at "O" target genes and upstream of ORFs at "T" target genes, related to Figure 5. (A-B) Average RNA read abundance for AS (A) or sense (B) transcripts surrounding Gcn4 motifs at the same group of induced "O" UC target genes analyzed in Fig. 5 for two biological replicates (reps. 1-2) of WT cells untreated or treated with SM. (C) Heat maps showing AS transcript abundance and position relative to Gcn4 motifs at the "O" UC targets, before (left panel) and after (right panel) SM treatment. (D-E) Average RNA read abundance for (D) AS and (E) sense transcripts initiated from the "T" targets, before and after SM treatment. Note that internal AS transcripts are ~5-fold less abundant than the induced FL sense transcripts at "O" UC target genes (A-B), and ~10-fold less abundant than the induced FL sense transcripts at "T" target genes (D-E), after SM treatment (note different y-axis scales in panels A & D vs. B & E).

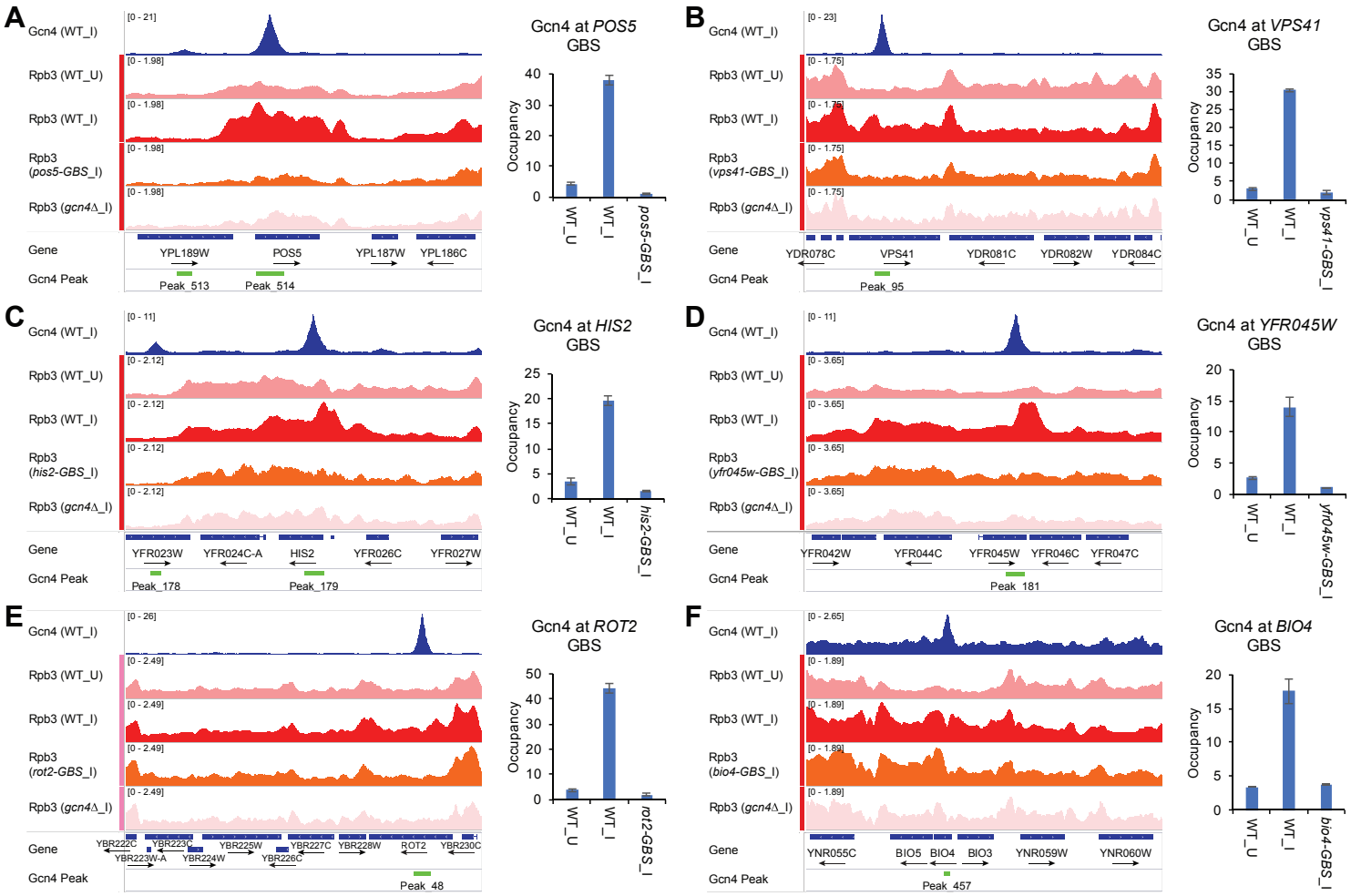

Fig. S4: Eliminating Gcn4 binding sites in CDS reduces Gcn4 binding and diminishes Rpb3 occupancy across that gene but generally not in surrounding genes, related to Figure 7. (A-F) left panels: SM-induced Gcn4 occupancy and Rpb3 occupancies in uninduced (_U) or SM-induced (_I) WT cells, SM-induced GBS mutant, and SM-treated *gcn4Δ* cells (*gcn4Δ*_I), all plotted as in Fig. 1 for the indicated genes. Right panels: Relative Gcn4 occupancy, normalized to POL1 amplicon, determined in uninduced (_U) or SM-induced (_I) WT and SM-induced GBS mutant cells. Mean values (±SD) determined from two biological replicates. Note the exception case in (F) where mutation of the GBS within BIO4 reduced Rpb3 levels in the adjacent gene, BIO3, even though induction of BIO4 itself was not significantly reduced.

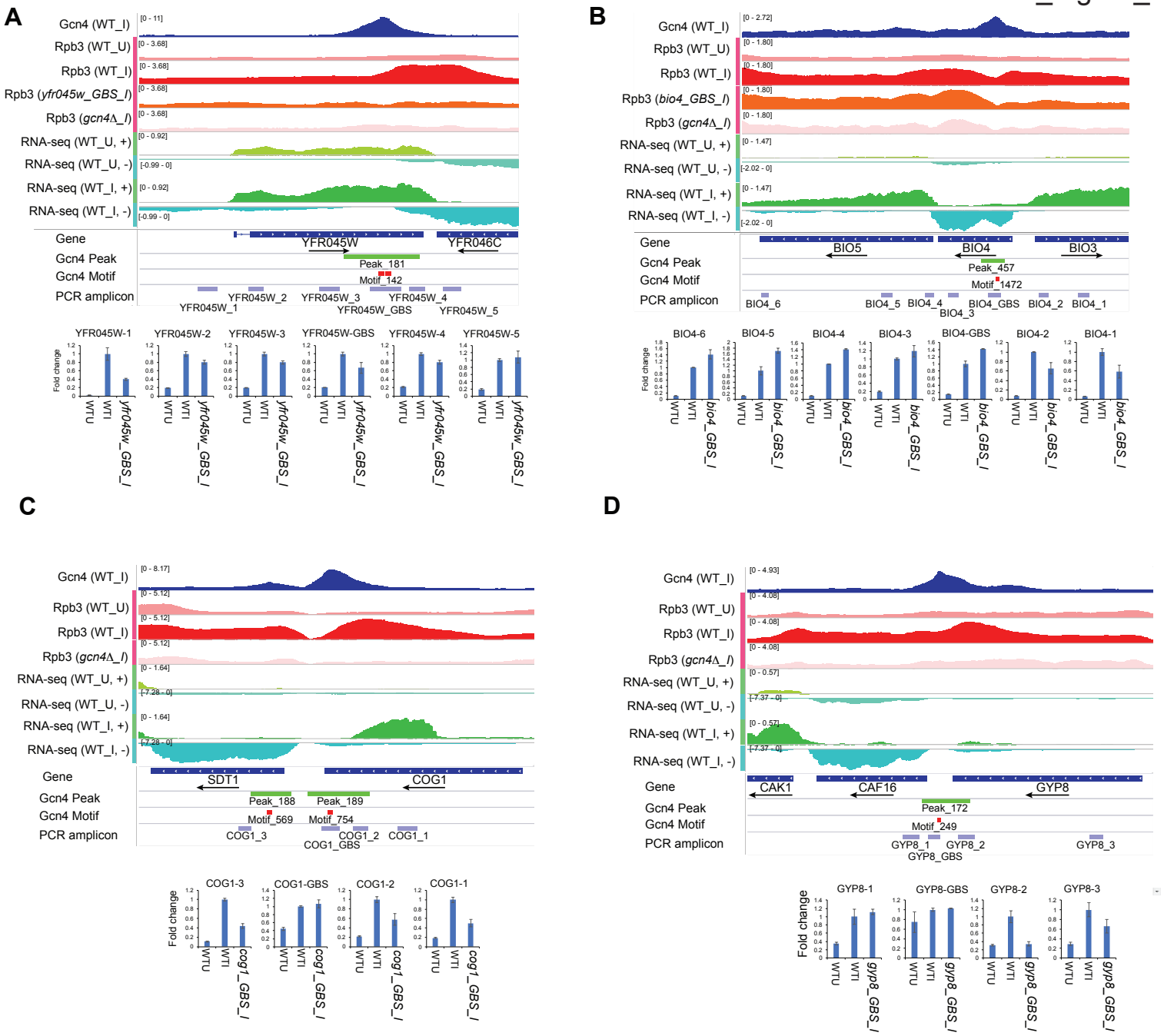

Fig. S5. Elimination of internal Gcn4 binding sites reduces Rpb3_I occupancies of the mutated CDS and SM-induced full-length and sub-genic transcripts at UC genes , related to Figure 1. (A-D) Gcn4 occupancies from SM-induced WT, Rpb3 occupancies from uninduced or SM-induced WT cells, SM-induced GBS mutant cells and *gcn4Δ* and RNA read densities from uninduced or 3AT-induced WT cells, all depicted as in Fig. 7. Below IGV tracks are locations of amplicons produced from total mRNA and quantified by qRT-PCR from uninduced (WT_U) or SM-induced (WT_I) WT cells, or SM-induced GBS mutant cells. Mean (±SD) relative mRNA levels, normalized to actin mRNA, were determined.

Table S1. Multiple internal GBSs function in transcriptional activation, Related to Figure 7.

Table S2. Summary of effects of GBS mutations on full-length, AS, and SGS transcripts, related to Figure 7 and S5.

Table S3. Descriptions of ChIP-seq replicates, related to STAR methods.

Table S4. Yeast strains used in this study, related to STAR methods.

Table S1. Multiple internal GBSs function in transcriptional activation

Genes containing GBSs		GBS mutant	Normalized Rpb3 Occupancy						
			WT_I (n=25)	GBS mut (n=2)	P value#	WT_U	*gcn4*Δ_I	GBS-mut_I/WT_I	*gcn4_I*/WT_I
1. YPL188W	*POS5*	*YPL188W(POS5)*	1.31 ±0.27	0.44 ±0.004	<0.001	0.59 ±0.01	0.67 ±0.07	0.335	0.509
2. YDR080W	*VPS41*	*YDR080W(VPS41)*	0.66 ±0.10	0.40 ±0.02	0.002	0.42 ±0.05	0.46 ±0.02	0.609	0.703
3. YFR025C	*HIS2*	*YFR025C(HIS2)*	1.41 ±0.28	0.80 ±0.05	0.010	0.85 ±0.05	0.88 ±0.10	0.568	0.626
4. YFR045W	*YFR045W*	*YFR045W*	1.82 ±0.30	0.90 ±0.07	<0.001	0.67 ±0.02	0.90 ±0.07	0.493	0.497
5. YBR229C	*ROT2*	*YBR229C(ROT2)*	0.93 ±0.12	0.64 ±0.01	0.003	0.61 ±0.01	0.62 ±0.03	0.685	0.666
6. YER037W	*PHM8*	*YER037W(PHM8)*	2.73 ±0.51	2.34 ±0.03	0.332	0.97 ±0.04	2.19 ±0.29	0.855	0.802
7. YBR166C	*TYR1*	*YBR166C(TYR1)*	1.29 ±0.15	0.93 ±0.04	0.009	0.78 ±0.01	0.76 ±0.06	0.723	0.586
8. YNR057C	*BIO4*	*YNR057C(BIO4)*	1.44 ±0.18	1.20 ±0.05	0.062*	0.49 ±0.06	0.79 ±0.16	0.834	0.548
9. YDL223C	*HBT1*	*YDL223C(HBT1)*	1.40 ±0.37	1.17 ±0.22	0.446	0.50 ±0.03	0.73 ±0.14	0.837	0.521
10. YDR380W	*ARO10*	*YDR380W(ARO10)*	4.63 ±1.01	3.91 ±0.15	0.363	0.84 ±0.01	2.45 ±0.41	0.844	0.529
11. YIL112W	*HOS4*	*YIL112W(HOS4)*	1.27 ±0.09	1.13 ±0.03	0.0456*	0.92 ±0.03	1.01 ±0.01	0.894	0.798
12. YOR230W	*WTM1*	*YOR230W(WTM1)*	2.95 ±0.40	2.88 ±0.01	0.782	1.21 ±0.04	2.63 ±0.27	0.977	0.891
Adjacent to Genes w/ GBSs									
13. YFR044C	*YFR044C*	*YFR045W*	1.72 ±0.12	1.54 ±0.07	0.0442	0.79 ±0.03	1.37 ±0.25	0.897	0.794
14. YFR046C	*CNN1*	*YFR045W*	1.49 ±0.19	0.96 ±0.02	0.001	0.71 ±0.09	0.86 ±0.02	0.645	0.573
15. YBR228W	*SLX1*	*YBR229C(ROT2)*	0.90 ±0.08	0.75 ±0.05	0.028	0.72 ±0.03	0.79 ±0.04	0.832	0.879
16. YBR230C	*OM14*	*YBR229C(ROT2)*	2.28 ±0.40	2.15 ±0.08	0.664	1.34 ±0.12	1.38 ±0.12	0.941	0.603

Table S1. (cont'd)

17. YBR165W	*UBS1*	*YBR166C(TYR1)*	1.390 ±0.16	0.913 ±0.02	0.001	0.80 ±0.08	0.91 ±0.09	0.657	0.653
18. YBR167C	*POP7*	*YBR166C(TYR1)*	1.53 ±0.15	1.49 ±0.04	0.707	1.00 ±0.06	1.09 ±0.10	0.972	0.710
19. YNR056C	*BIO5*	*YNR057C(BIO4)*	1.06 ±0.14	1.10 ±0.001	0.713	0.39 ±0.02	0.68 ±0.08	1.038	0.636
20. YNR058W	*BIO3*	*YNR057C(BIO4)*	0.90 ±0.14	0.69 ±0.03	0.029	0.34 ±0.02	0.57 ±0.02	0.762	0.630
21. YDL222C	*FMP45*	*YDL223C(HBT1)*	3.03 ±0.82	2.64 ±0.40	0.550	0.82 ±0.06	1.41 ±0.43	0.872	0.466
22. YDR379C-A	*YDR379C-A*	*YDR380W(ARO10)*	1.94 ±0.26	1.90 ±0.16	0.867	0.93 ±0.05	1.53 ±0.22	0.982	0.788
23. YIL113W	*SDP1*	*YIL112W(HOS4)*	1.23 ±0.20	1.10 ±0.19	0.432	0.60 ±0.06	1.00 ±0.10	0.895	0.813
24. YIL111W	*COX5B*	*YIL112W(HOS4)*	2.95 ±0.30	2.88 ±0.17	0.782	2.16 ±0.18	2.69 ±0.15	0.977	0.913

#Assigned with 2-tailed, unpaired t-test comparing 2 replicate results for each GBS mutant to either 2 replicates each of the other 11 GBS mutants plus 3 WT_I replicates (25 control replicates in total) (*), or to all 27 replicate results including that GBS mutant.

Table S2. Summary of effects of -*GBS* mutations on full-length, AS, and SGS transcripts[1]

	Name/Location of Amplicons							
	Spanning GBS	**Upstream of the GBS**			**Downstream of the GBS**			
Amplicons	POS5_GBS	POS5_1			POS5_2			
% WT_I-WT_U	28.02 ±1.92 **	1.21 ±0.03 ***			23.8 ±0.30 **			
Amplicons	HIS2_GBS	HIS2_1			HIS2_2	HiS2_3		
% WT_I-WT_U	51.05 ±8.35 *	12.45 ±2.14 **			66.34 ±0 *	55.38 ±3.25 **		
Amplicons	SPO21_GBS	SPO21_1			SPO21_2	SPO21_3		
% WT_I-WT_U	90.77 ±22.87	8.8 ±6.16 ***			7.069 ±1.68 ***	26.58±10.68 ***		
Amplicons	COG1-GBS	COG1-1	COG1-2		COG1-3			
% WT_I-WT_U	112.35 ±17.49	38.55 ±12.78 ***	45.94 ±12.78 ***		36.83 ±4.95 ***			
Amplicons	GYP8-GBS	GYP8-1	GYP8-2		GYP8-3			
% WT_I-WT_U	112.95 ±1.98	116.26 ±13.09	4.152 ±7.84 ***		54.51 ±28.22 *			
Amplicons	SOL1-GBS	SOL1-1	SOL1-2		SOL1-3			
% WT_I-WT_U	46.73 ±11.01 ***	14.29 ±4 ***	48.98 ±36.75 *		56.79 ±21.69 **			
Amplicons	VPS41_GBS	VPS41_1	VPS41_2		VPS41_3	VPS41_4		
% WT_I-WT_U	60.67 ±0.54 **	39.32 ±0.62 **	26.93 ±5.94 **		38.29 ±11.80 **	60.81 ±12.80 *		
Amplicons	HMG2-GBS	HMG2-1	HMG2-2	HMG2-3	HMG2-4			
% WT_I-WT_U	79.035 ±20.3	78.66 ±6.48 ***	74.70 ±19.17	76.31 ±35.16	56.016 ±22.99**			
Amplicons	YFR045w_GBS	YFR045w_1	YFR045w_2	YFR045w_3	YFR045w_4	YFR045w_5		
% WT_I-WT_U	58.82 ±16.23 *	38.71 ±2.7 *	75.55 ±5.89 *	75.11 ±5.12 *	75.26 ±5.32 *	109.82 ±20.19		
Amplicons	ROT2_GBS	ROT2_1	ROT2_2	ROT2_3	ROT2_4	ROT2_5		
% WT_I-WT_U	35.94 ±25.32 *	69.37 ±21.43 **	38.63 ±8.67	-37.23 ±9.45 **	50.46 ±35.18	39.71 ±35.78		
Amplicons	BIO4-GBS	BIO4 - 1	BIO4-2		BIO4-3	BIO4-4	BIO4-5	BIO4-6
% WT_I-WT_U	149.87 ±0.41 **	52.78 ±23.6 **	62.8 ±27.99 *		123.57 ±16.9	146.02 ±1.15 ***	178.9 ±10.3 *	146.86 ±15.95*

[1]Summary of qRT-PCR analysis of mRNA expression changes conferred by the indicated GBS mutations. For each -*GBS* mutant allele, 3 or more amplicons quantified by qRT-PCR to probe transcription (i) spanning the GBS, which should quantify primarily FL transcripts, (ii) upstream of the GBS, to quantify AS transcripts; (iii) downstream of the GBS, to quantify SGS transcripts. The cell below the label of each amplicon gives the percentage of the difference in mRNA expression between induced and uninduced conditions in the -*GBS* mutant versus WT, expressed as a percentage of the WT difference, calculated as follows: (i) A = (Mean_WT_I) - (Mean_WT_U), calculated from the mean values determined from two or more biological replicates of WT_I and WT_U total RNA samples; (ii) B_n = $(GBS\_I)_n$-Mean_WT_U (calculated for biological replicates n = 1, 2,..4 of the -*GBS* mutant; (iii) $C_n = (B_n/A)X100$; (iv) %(WT_I-WT_U) = Mean C_n (±Average deviation). A two-tailed Student's t-test was conducted to determine whether the mean of Cn values calculated for the -*GBS* mutant from n replicates differs significantly from the mean of Cn values calculated from n replicates of the WT strain.
* P≤ 0.05, ** P≤ 0.01 and *** P≤ 0.001.

Table S3. Descriptions of ChIP-seq replicates

Genotype	ChIP	Sample Name	PE reads	PE rmdup reads	Pearson correlation between replicates			
					AGH24_05	AGH24_06	AGH32_86	AGH32_87
BY4741_I	Gcn4_I	AGH24_04	13142345	501413	0.9775	0.9796	0.9424	0.9452
BY4741_I	Gcn4_I	AGH24_05	13309102	443492		0.9764	0.937	0.9412
BY4741_I	Gcn4_I	AGH24_06	15050690	446177			0.9427	0.9458
BY4741_I	Gcn4_I	AGH32_86	11022951	1462391				0.9959
BY4741_I	Gcn4_I	AGH32_87	10329576	1086737				
					AGH32_HQ46		AGH32_HQ47	
BY4741_U	Gcn4_U	AGH32_HQ45	9434957	834910	0.9972		0.9967	
BY4741_U	Gcn4_U	AGH32_HQ46	8293362	352621			0.9958	
BY4741_U	Gcn4_U	AGH32_HQ47	6312086	223598				
					AGH24_02		AGH24_03	
gcn4Δ_I	Gcn4_I	AGH24_01	9956992	807437	0.9771		0.9753	
gcn4Δ_I	Gcn4_I	AGH24_02	12785177	522922			0.969	
gcn4Δ_I	Gcn4_I	AGH24_03	8469126	235539				
					AGH12-02		AGH23-11	
gcn4Δ_I	Rpb3_I-1	AGH12-01	13200360	8404587	0.9911		0.8895	
gcn4Δ_I	Rpb3_I-2	AGH12-02	12475932	7843968			0.8891	
gcn4Δ_I	Rpb3_I-3	AGH23-11	8667515	1858940				
					Replicate 2			
rot2-GBS_I	Rpb3_I-1	AGH85-01	21939943	8660775	0.9871			
rot2-GBS_I	Rpb3_I-2	AGH85-02	23462617	5818055				
hbt1-GBS_I	Rpb3_I-1	AGH85-03	12628893	3455163	0.9878			
hbt1-GBS_I	Rpb3_I-2	AGH85-04	21413927	6530496				
vps41-GBS_I	Rpb3_I-1	AGH85-05	14653698	4922416	0.989			
vps41-GBS_I	Rpb3_I-2	AGH85-06	24355045	7729944				
aro10-GBS_I	Rpb3_I-1	AGH85-07	19695500	3631346	0.9798			
aro10-GBS_I	Rpb3_I-2	AGH85-08	14365583	6765744				
phm8-GBS_I	Rpb3_I-1	AGH85-09	21902229	3327726	0.9753			
phm8-GBS_I	Rpb3_I-2	AGH85-10	30995106	3954872				
his2-GBS_I	Rpb3_I-1	AGH85-11	24430264	2285433	0.9686			
his2-GBS_I	Rpb3_I-2	AGH85-12	9198411	1134697				
yfr045w-GBS_I	Rpb3_I-1	AGH86-01	10550130	2197196	0.9814			
yfr045w-GBS_I	Rpb3_I-2	AGH86-02	14987193	3271705				
hos4-GBS_I	Rpb3_I-1	AGH86-03	14196735	2976920	0.9717			
hos4-GBS_I	Rpb3_I-2	AGH86-04	9939528	4868827				
wtm1-GBS_I	Rpb3_I-1	AGH86-07	13315018	4018608	0.9762			
wtm1-GBS_I	Rpb3_I-2	AGH86-08	6491683	3617616				
pos5-GBS_I	Rpb3_I-1	AGH86-09	11600154	2483310	0.9526			
pos5-GBS_I	Rpb3_I-2	AGH86-10	11911824	3778767				
tyr1-GBS_I	Rpb3_I-1	AGH94-05	22507027	17095399	0.9934			
tyr1-GBS_I	Rpb3_I-2	AGH94-06	26417337	20048682				
bio4-GBS_I	Rpb3_I-1	AGH94-07	25093873	19596202	0.9929			
bio4-GBS_I	Rpb3_I-2	AGH94-08	21617804	16770972				
TBP-myc$_{13}$_U	Myc_U-1	AGH84-01	23867690	17101713	0.998			
TBP-myc$_{13}$_U	Myc_U-2	AGH84-02	29431993	22193596				
TBP-myc$_{13}$_I	Myc_I-1	AGH84-05	26549545	19424748	0.9995			
TBP-myc$_{13}$_U	Myc_I-2	AGH84-06	28039770	19832697				
BY4741_U	Mnase-H3_U-1	AGH68-01	21155789	16387173	0.9862			
BY4741_U	Mnase-H3_U-2	AGH73-02	33931620	20843860				
BY4741_I	Mnase-H3_I-1	AGH68-03	22985503	17204691	0.9866			
BY4741_I	Mnase-H3_I-2	AGH68-04	19959153	15376225				

Table S4. Yeast strains used in this study

Strain Name	Parent	Genotype	Motif sequences replaced	Source
F729/ BY4741	NA	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	NA	Research genetics
F731	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 gcn4Δ::kanMX4*	NA	Research genetics
HQY366/ H3256	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 SPT15-myc$_{13}$::HIS3MX6*	NA	Qiu et al., 2004 MCB
YDC111	NA	*MATa ade2–1 can1-100 leu2-3,112 trp1-1 ura3-1*	NA	Kim et al., 2006
YR201	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 tyr1-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	TCTGAGTCATT	This study
YR202	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rot2-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	GATGACTCATT	This study
YR203	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 hbt1-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	AATGACTCACG	This study
YR204	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 vps41-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	TATGAGTCATT	This study
YR205	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 aro10-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	GATGAGTCAAA	This study
YR206	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 phm8-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	TATGAGTCAGA	This study
YR207	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 his2-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	CATGAGTCATG	This study
YR208	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 yfr045w-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	[1]**AATGACTCAGC**CCATT GACGTCGTAAAAACAA GGAT**GATGAGTCAAA**	This study
YR209	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 hos4-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	GCTGACTCACC	This study
YR211	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 bio4-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	ATTGAGTCAGA	This study
YR212	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 wtm1-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	TGTGACTCACA	This study
YR214	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 pos5-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	CATGAGTCATA	This study
VV 001	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 hmg2-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	TATGACTCACAAC	This study
VV 002	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 gyp8-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	GATGACTCAAA	This study
VV 003	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 cog1-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	CAATTAGTCATC	This study
VV 004	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 sol1-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	GATGAGTCATT	This study
VV 005	BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 spo21-GBS::P$_{GAL1}$SCE1-hyg-KlURA3*	AATGAGTCAT	This study
YR216	YR201	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 tyr1-GBS::BamHI*	TCTGAGTCATT	This study
YR217	YR202	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rot2-GBS::BamHI*	GATGACTCATT	This study
YR218	YR203	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 hbt1-GBS::BamHI*	AATGACTCACG	This study
YR219	YR204	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 vps41-GBS::BamHI*	TATGAGTCATT	This study
YR220	YR205	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 aro10-GBS::BamHI*	GATGAGTCAAA	This study
YR221	YR206	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 phm8-GBS::BamHI*	TATGAGTCAGA	This study
YR222	YR207	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 his2-GBS::BamHI*	CATGAGTCATG	This study
YR223	YR208	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 yfr045w-GBS::BamHI*	[1]**AATGACTCAGC**CCATT GACGTCGTAAAAACAA GGAT**GATGAGTCAAA**	This study
YR224	YR209	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 hos4-GBS::BamHI*	GCTGACTCACC	This study
YR226	YR211	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 bio4-GBS::BamHI*	ATTGAGTCAGA	This study
YR227	YR212	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 wtm1-GBS::BamHI*	TGTGACTCACA	This study
YR229	YR214	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 pos5-GBS::BamHI*	CATGAGTCATA	This study
VV 006	VV 001	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 hmg2-GBS::BamHI*	TATGACTCACAAC	This study

VV 007	VV 002	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 gyp8-GBS::BamHI*	GATGACTCAAA	This study
VV 008	VV 003	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 cog1-GBS::BamHI*	CAATTAGTCATC	This study
VV 009	VV 004	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 sol1-GBS::BamHI*	GATGAGTCATT	This study
VV 010	VV 005	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 spo21-GBS::BamHI*	AATGAGTCAT	This study

[1] Two motifs highlighted in bold.