Supplemental Information for Joo et al.

Supplemental Figures S1-S7

Supplemental References

Supplemental Tables S1-S4

Figure S1, related to Figure 1.

(A) Volcano plot of statistical significance (y-axis, $\log_{10}$ of p-value) versus fold enrichment (x-axis, $\log_2$ of +NTP/-NTP ratio) for proteins associated with the downstream region of the transcription template as illustrated in **Fig 1A**. Data represents three biological replicates each done with two technical replicates of quantitative mass spectrometry analyses, all using naked DNA templates. Each circle represents an individual protein, quantitated as the sum of all identified peptides. Values on x-axis are the average of values from all six experiments. P-value is by t-test of specific protein values against total identified proteins (712 total proteins identified in all six repeats).

(B-C) Plots of change in downstream protein levels (x-axis, $\log_2$) upon NTP (B) or α-amanitin (C) treatment versus the probability of being outside the normal distribution as calculated by the Mixed Model function (normalmixEM) in the R statistic package. Lines are drawn at 0.95 confidence level.

(D) Scatter plot comparing NTP-dependent protein binding to downstream DNA (y-axis, $\log_2$ of +NTP /- NTP ratio from SspI elution) to that of the entire core-promoter and downstream DNA (x-axis, $\log_2$ of + NTP /- NTP ratio by PstI elution) on naked DNA templates. The decrease in total RNApII (PstI elution) contrasts with the increase downstream (SspI elution), indicating that not all upstream RNApII proceeds to elongation or that RNApII in ECs is less stably bound than that in PICs.

(E) Scatter plot comparing NTP-dependent protein binding to downstream DNA (SspI elution) on chromatin templates (y-axis) versus naked templates (x-axis).

Figure S1

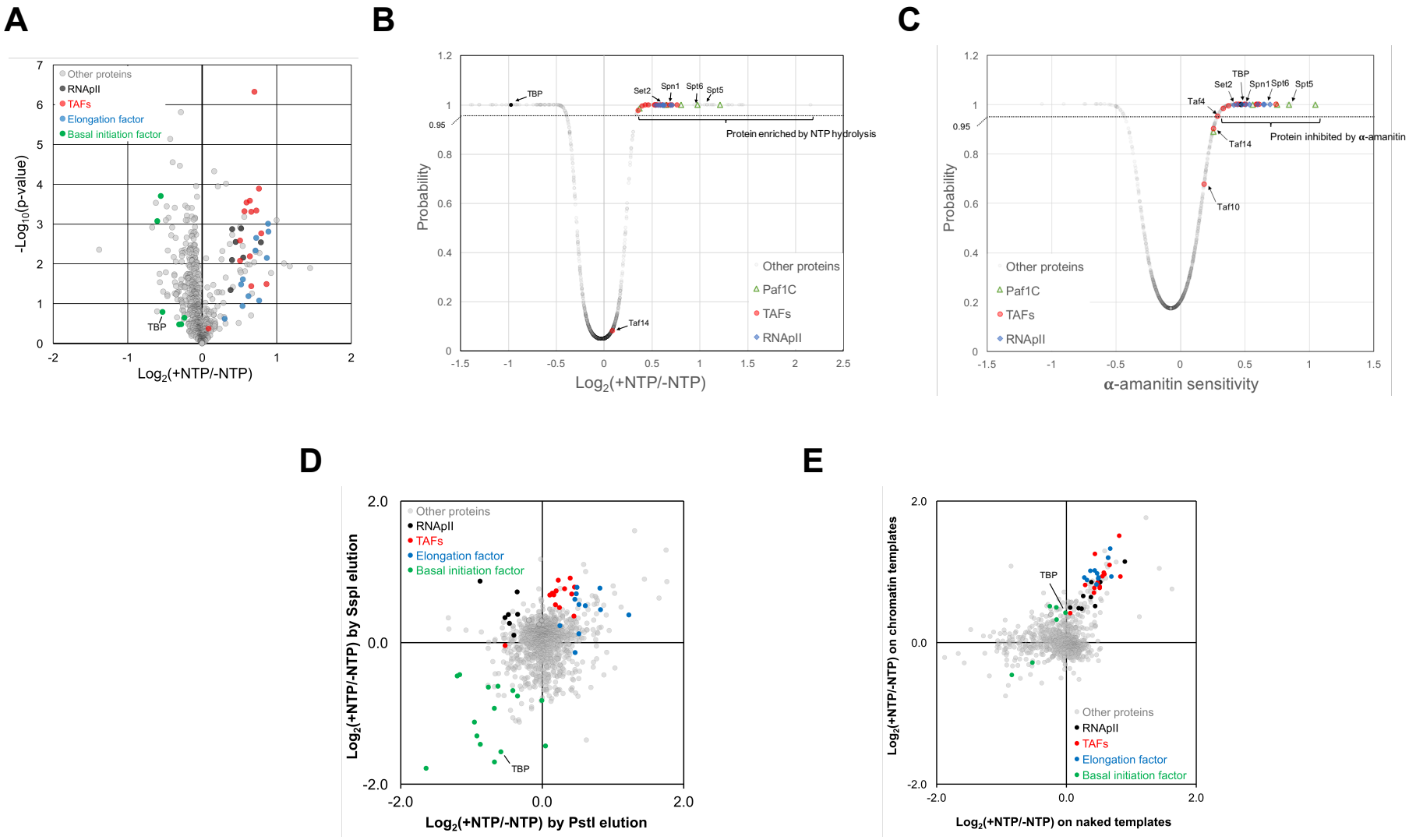

A
B
C
D
E
Other proteins
RNApII
TAFs
Elongation factor
Basal initiation factor
-Log10(p-value)
Log2(+NTP/-NTP)
TBP
Probability
Set2
Spn1
Spt6
Spt5
Taf14
Protein enriched by NTP hydrolysis
Paf1C
α-amanitin sensitivity
Taf4
Taf10
Protein inhibited by α-amanitin
Log2(+NTP/-NTP) by SspI elution
Log2(+NTP/-NTP) by PstI elution
Log2(+NTP/-NTP) on chromatin templates
Log2(+NTP/-NTP) on naked templates

Figure S2, related to Figure 1.

(A) TBP and Taf1 binding on different promoter templates before (-NTP) and after (+NTP) incubation with all four NTPs. Upper panel shows schematic of template series. All templates were 5' biotinylated and immoblized on beads, incubated with nuclear extract, and then purified before or after NTP treatment. The lower panels show immunoblots of total bound proteins eluted by boiling. Ratios of factors +/-NTPs is shown between pairs of lanes. All templates contain RNApII promoters except RDN1-1, which is the rDNA promoter template transcribed by RNApI. CYC1p was amplified by PCR from SB649 with primers (O#3151/1477). PGK1 was amplified by PCR from SB1718 with primers (O#3151/3433). HIS4p was amplified by PCR from F529 with primers (O#1105/1104). HIS4 was amplified by PCR from SB1722 with primers (O#1105/3513). RDN1-1 was amplified by PCR from F82 with primers (O#3278/3279).

(B) Comparison of ATP versus NTPs. Immobilized templates used in **Fig 1C** were incubated with nuclear extract, treated without NTPs, with ATP alone, or with three NTPs (A, C, and UTP) plus 3'O-meGTP. Bound proteins were eluted with Ssp I and immunoblotted for Taf1 and TBP.

(C) The promoter is required for downstream TAF recruitment. TBP and Taf1 binding was assayed on different length N-terminal deletion templates before or after incubation with A, C, and UTP plus 3'O-meGTP. Upper panel shows schematic of template series. The lower panels show templates stained with ethidium bromide and immunoblots of proteins eluted by Ssp I.

Figure S2

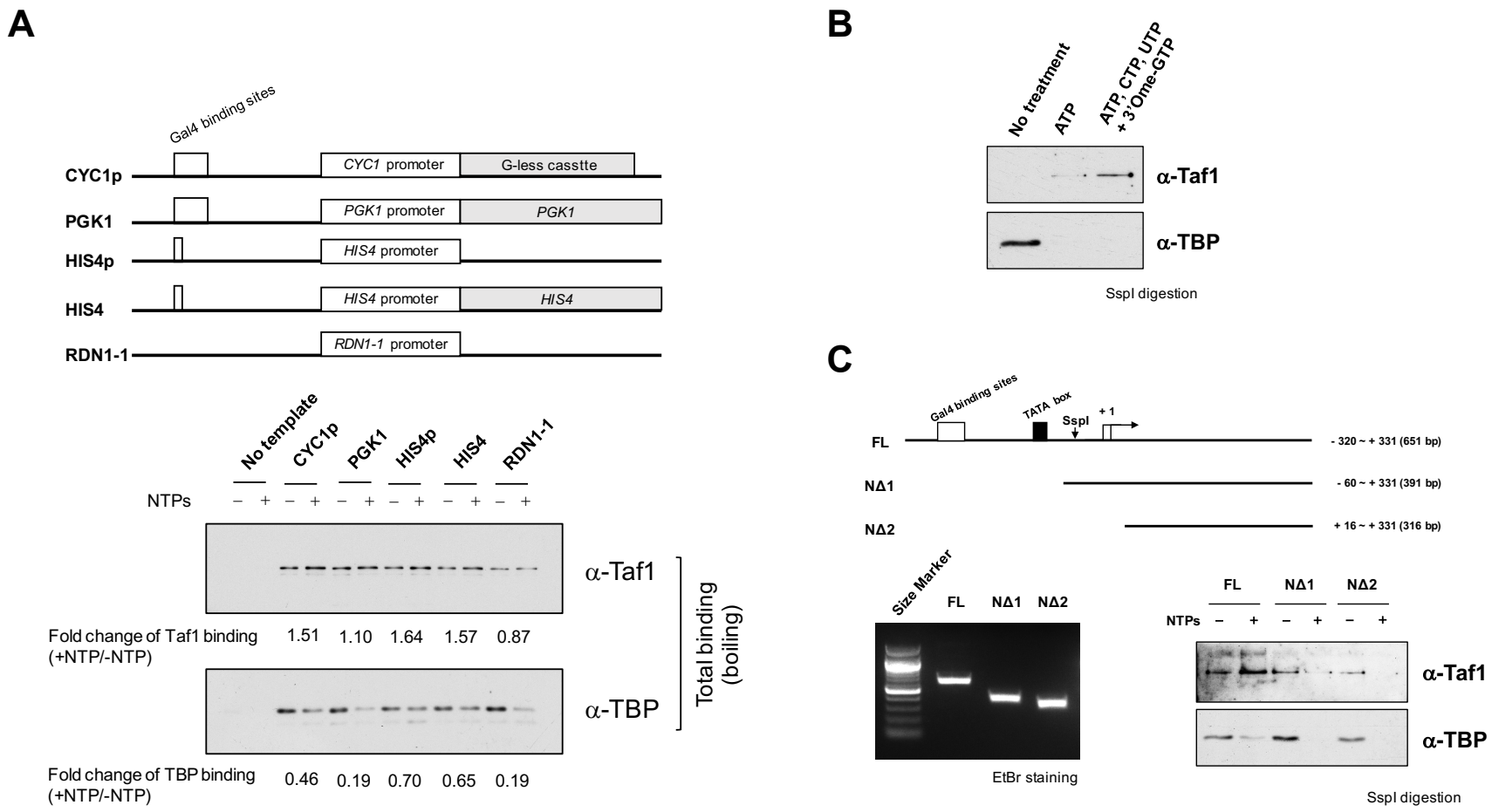

A
Gal4 binding sites
CYC1p
PGK1
HIS4p
HIS4
RDN1-1
CYC1 promoter
G-less casstte
PGK1 promoter
PGK1
HIS4 promoter
HIS4 promoter
HIS4
RDN1-1 promoter
No template
CYC1p
PGK1
HIS4p
HIS4
RDN1-1
NTPs
α-Taf1
Fold change of Taf1 binding (+NTP/-NTP)
1.51 1.10 1.64 1.57 0.87
α-TBP
Total binding (boiling)
Fold change of TBP binding (+NTP/-NTP)
0.46 0.19 0.70 0.65 0.19
B
No treatment
ATP
ATP, CTP, UTP + 3'Ome-GTP
α-Taf1
α-TBP
SspI digestion
C
Gal4 binding sites
TATA box
SspI
+ 1
FL
- 320 ~ + 331 (651 bp)
NΔ1
- 60 ~ + 331 (391 bp)
NΔ2
+ 16 ~ + 331 (316 bp)
Size Marker
FL
NΔ1
NΔ2
EtBr staining
NTPs
α-Taf1
α-TBP
SspI digestion

Figure S3, related to Figure 2.

(A) Comparison of gene expression changes in two different taf1-ts strains (YJS8 and YKH108) from microarray data in (Huisinga and Pugh, 2004). The correlation coefficient between the two *taf1* alleles is 0.89. Dots represent individual genes, color-coded for TAF-dependence as in **Fig 2**. Group High: 825 genes (Fold changes more than one standard deviation (SD = 0.57) below the mean of ratios). Group Medium: 4327 genes (Fold changes between -1 SD and +1 SD from mean of ratios). Group Low: 729 genes (Fold changes more than one SD above the mean of ratios). Group RPG: 137 genes. Average *TAF1*-dependency in High, Medium, Low, and RPG group are -1.73, -0.92, -0.08, and -1.82, respectively.

(B) Metagene anchor-plot shows averaged ChIP-exo occupancy near TSS regions (-200 bp to +400 bp) of Taf1, Taf2, Taf4, Taf8, Bdf1 and TBP adapted and re-analyzed from other studies Reja et al., (2015) and Rhee and Pugh, (2012). Upper panel shows the recently reported cryo-EM structure of TFIID bound to promoter DNA (Louder et al., 2016).

(C) Boxplots showing ratios of downstream (+30 bp ~ +170 bp from TSS) versus upstream (-150 bp ~ -10 from TSS) Taf1 occupancy in different *TAF1*-dependency groups as defined in **Fig 2**. Taf1 occupancy was defined as total ChIP-exo tag counts within the designated region (Rhee and Pugh, 2012). (p-value: **<0.05, ***<0.005)

Figure S3

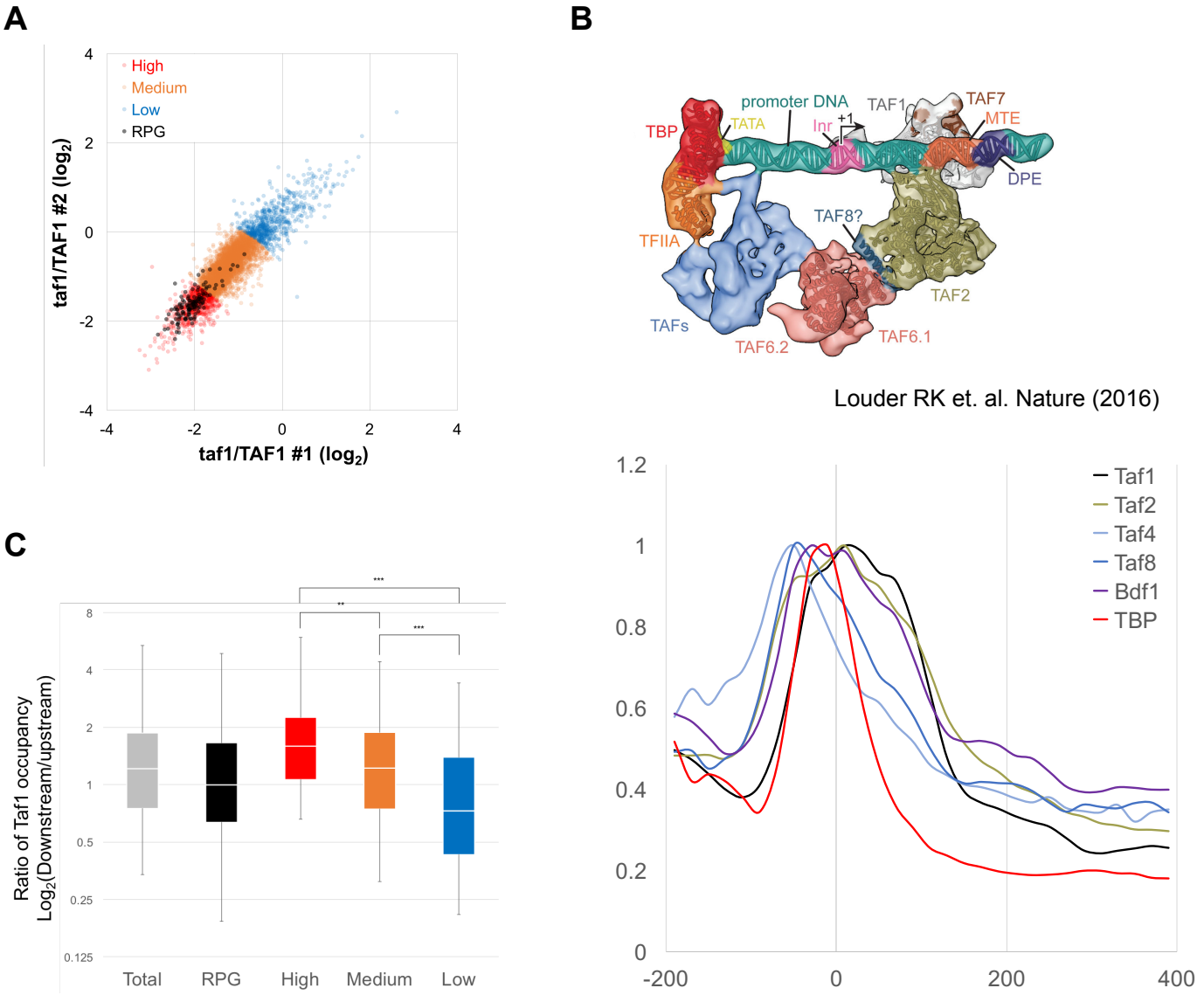

A
High
Medium
Low
RPG
taf1/TAF1 #2 (log2)
taf1/TAF1 #1 (log2)
B
promoter DNA
TAF1
TAF7
TBP
TATA
Inr
+1
MTE
DPE
TAF8?
TFIIA
TAF2
TAFs
TAF6.2
TAF6.1
Louder RK et. al. Nature (2016)
C
Ratio of Taf1 occupancy
Log2(Downstream/upstream)
Total
RPG
High
Medium
Low
Taf1
Taf2
Taf4
Taf8
Bdf1
TBP

Figure S4, related to Figure 3.

(A) Anchor-plot shows averaged ChIP-seq occupancy near TSS (-1 kb ~ +1 kb) of H3 and acetylated H4 (H4ac) for the four Taf1-dependence classes.

(B) Pairwise correlation plots of Taf1(upstream), Taf1(downstream), total Taf1, TBP, Bdf1, Htz1, Swr1, H3, and H4ac levels at individual promoters. Each dot represents a single promoter (total 4720 genes for ChIPexo and 6018 genes for ChIPseq datasets), with colors designating Taf1-dependence class as in **Fig 2**. Occupancies of Taf1(up/downstream), H3, and H4ac were determined as described in **Fig 3**. Occupancy levels of Taf1, TBP, Bdf1, Htz1, and Swr1 were determined as the maximum of ChIPseq reads within a fixed region (-200 bp to +400 bp relative to TSS). Each dot represents one mRNA gene (total 4720 genes with ChIPexo and 6018 genes with ChIPseq results). For each graph, Spearman's rank correlation coefficient (ρ) is shown.

Figure S4

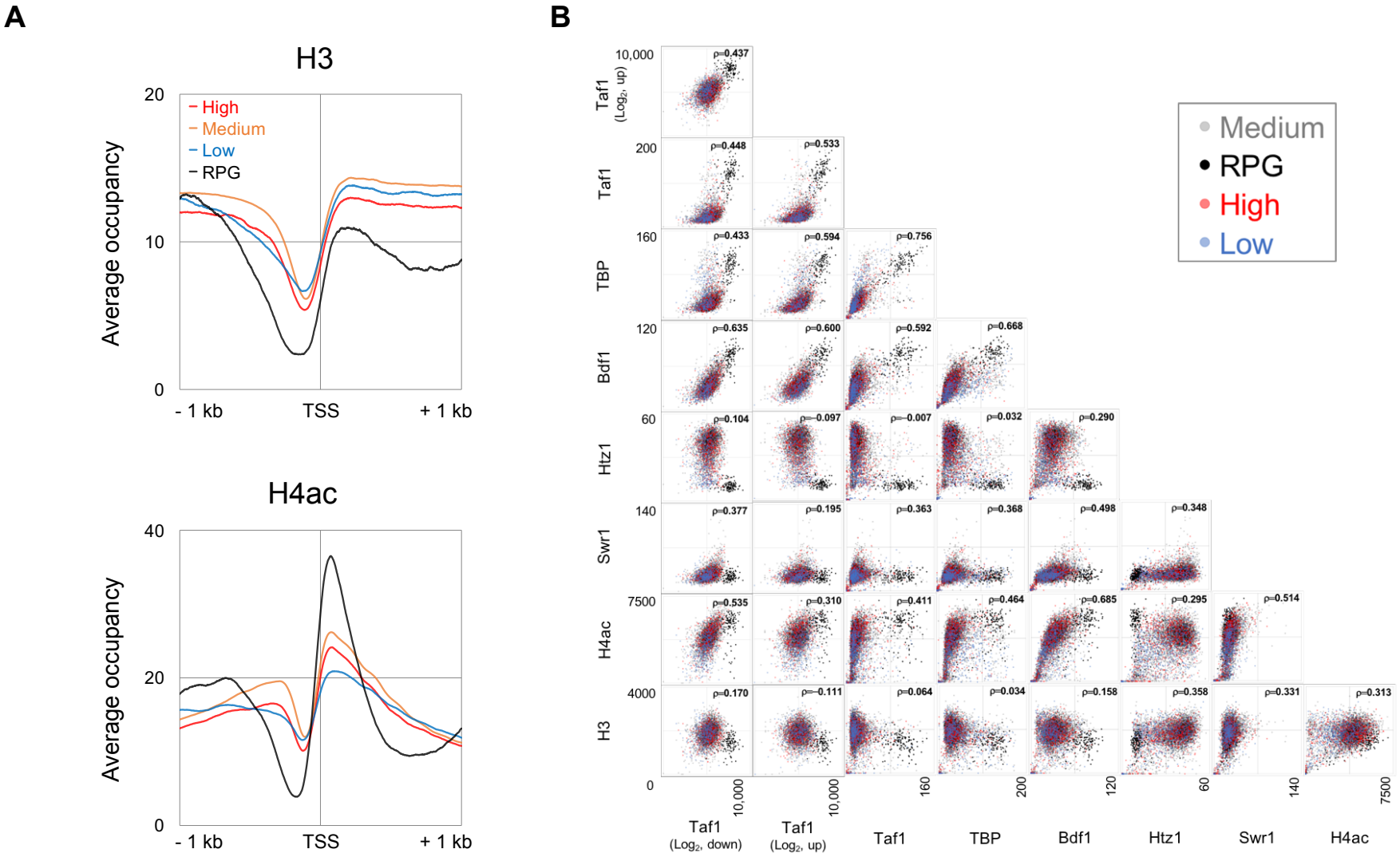

A
H3
Average occupancy
20
10
0
High
Medium
Low
RPG
- 1 kb
TSS
+ 1 kb
H4ac
Average occupancy
40
20
0
- 1 kb
TSS
+ 1 kb
B
Medium
RPG
High
Low
Taf1 (Log2, up)
Taf1
TBP
Bdf1
Htz1
Swr1
H4ac
H3
Taf1 (Log2, down)
Taf1 (Log2, up)
Taf1
TBP
Bdf1
Htz1
Swr1
H4ac

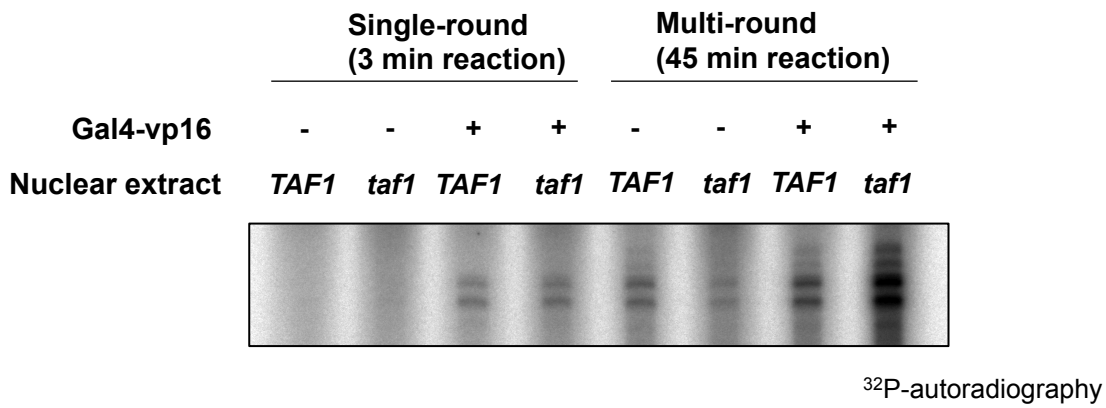
Figure S5, related to Figure 4.

(A) Representative gel for in vitro transcription analysis performed as in **Fig 4** with TAF1 wild type and mutant nuclear extracts. The same amount of reaction mixture was taken at 3 min (Single-round) or 45 min (Multi-round) in the presence or absence of transcription activator Gal4-vp16 as indicated. The transcripts were precipitated and separated on 6% urea-acrylamide gel and detected by autoradiography.

(B) Immobilized HIS4p template used in **Figure S2A** was incubated with nuclear extract for 30 min. After three washes, a run-off reaction was done by adding all four NTPs with ^{32}P-UTP for 5 min in 1X transcription buffer. The run-off transcripts were quantified by gel electrophoresis (left panel) and phosphorimager, normalized, and plotted (right panel).

Figure S5

A


Single-round
(3 min reaction)
Multi-round
(45 min reaction)
Gal4-vp16
- - + + - - + +
Nuclear extract
TAF1 taf1 TAF1 taf1 TAF1 taf1 TAF1 taf1
32P-autoradiography

B

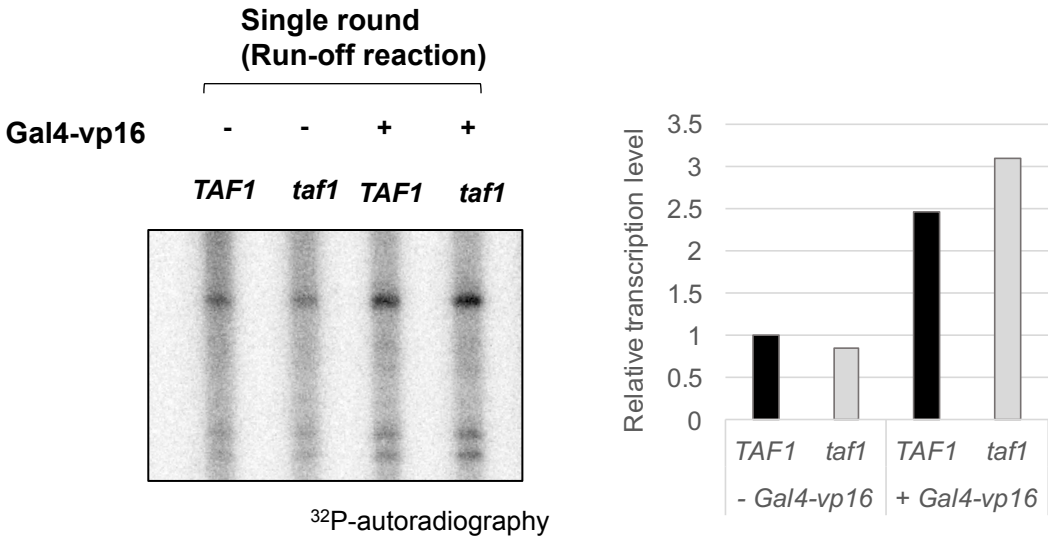

Single round
(Run-off reaction)
Gal4-vp16
- - + +
TAF1 taf1 TAF1 taf1
32P-autoradiography
Relative transcription level
3.5
3
2.5
2
1.5
1
0.5
0
TAF1 taf1 TAF1 taf1
- Gal4-vp16
+ Gal4-vp16

Figure S6, related to Figure 5.

(A) Sequential transcription analysis was performed as in **Fig 5A,** except for the inclusion of Gal4-vp16 (200 ng) throughout the reaction. The same amount of reaction mixture was taken at four different time points (2, 4, 8, and 16 min) and transcripts were precipitated and separated on 6% urea-acrylamide gel. Relative levels of Transcript 1 over time are displayed below each lane after normalizing to Transcript 2 levels and setting the 16 minute maximum to 1.00.

(B) Sequential transcription analysis was performed with ATP in pre-incubation. The same amount of reaction mixture was taken at four different time points (2, 4, 8, and 16 min) and transcripts were analyzed the same as in panel (A).

(C) Sequential transcription assay was performed with wild-type (*TAF1*) and two taf mutants (*taf1* and *taf11*) nuclear extracts in the first pre-incubation reaction. After washes, the second transcription reaction was performed with wild type nuclear extract. Quantitation from three or more independent replicates of each extract is shown in **Fig 5D**.

Figure S6

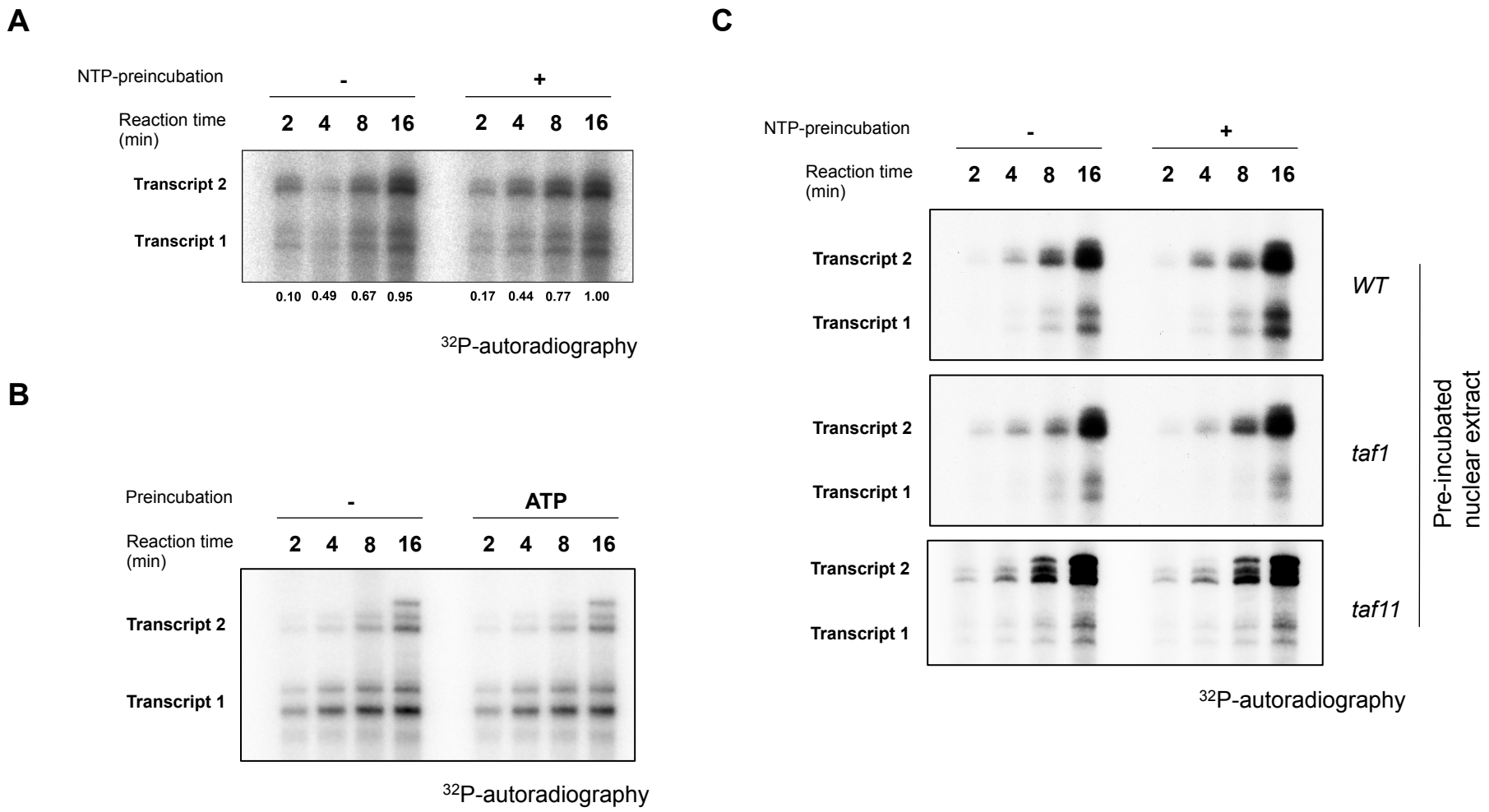

A
NTP-preincubation
-
+
Reaction time (min)
2 4 8 16 2 4 8 16
Transcript 2
Transcript 1
0.10 0.49 0.67 0.95 0.17 0.44 0.77 1.00
32P-autoradiography
B
Preincubation
-
ATP
Reaction time (min)
2 4 8 16 2 4 8 16
Transcript 2
Transcript 1
32P-autoradiography
C
NTP-preincubation
-
+
Reaction time (min)
2 4 8 16 2 4 8 16
Transcript 2
Transcript 1
WT
taf1
taf11
Pre-incubated nuclear extract
32P-autoradiography

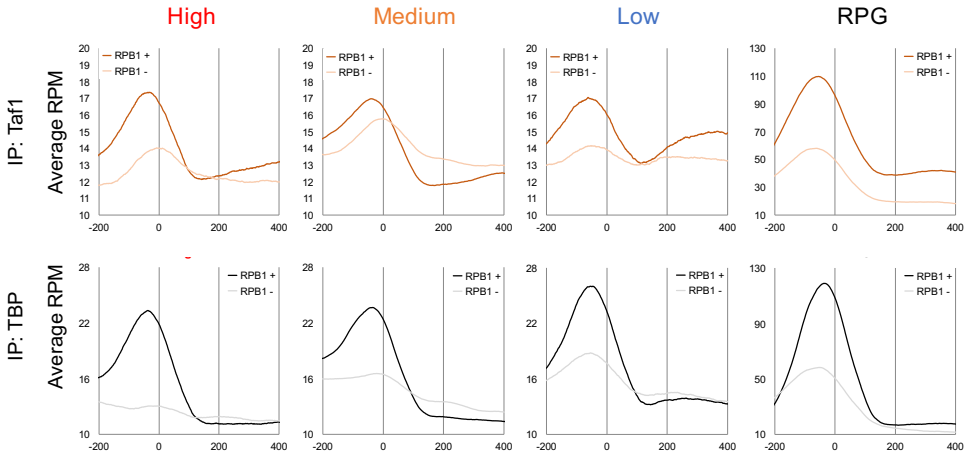
Figure S7, related to Figure 6.

(A) Metagene anchor-plots with ChIPseq analysis show average Taf1 and TBP binding near TSS (-200 bp to +400 bp) after RPB1-depletion for TAF-dependency groups defined in Fig 2.

(B) Same as in A, except Bdf1 was depleted.

Figure S7

A


High
Medium
Low
RPG
IP: Taf1
IP: TBP
Average RPM
RPB1 +
RPB1 -

B

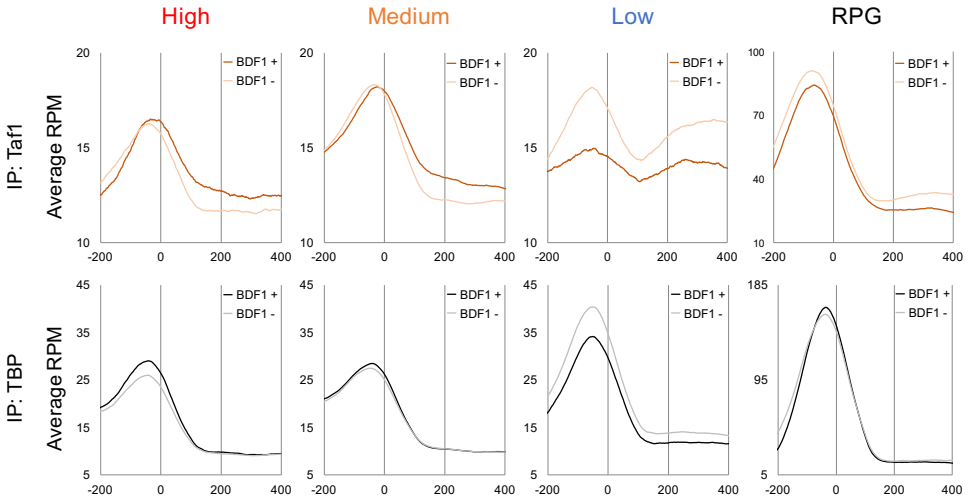

High
Medium
Low
RPG
IP: Taf1
IP: TBP
Average RPM
BDF1 +
BDF1 -

Table S1. Enrichment values (log_2) from quantitative mass-spectrometry analysis of proteins showing transcription-dependent association with downstream promoter DNA.

Complex	Protein	NTP enrichment w/o α-amanitin	NTP enrichment w/ α-amanitin	α-amanitin sensitivity
RNApII	RPB1	0.588	0.082	0.506
	RPB2	0.565	0.122	0.443
	RPB3	0.611	0.130	0.481
	RPB4	0.616	-0.039	0.654
	RPB6	0.613	0.197	0.415
	RPB7	0.711	0.066	0.646
	RPB9	0.695	0.094	0.601
	RPB10	0.631	0.097	0.534
	RPB12	0.538	-0.158	0.696
TAFs	BDF1	0.552	0.179	0.373
	BDF2	0.568	0.180	0.388
	TAF1	0.688	0.095	0.593
	TAF2	0.707	0.106	0.602
	TAF3	0.397	0.061	0.336
	TAF4	0.356	0.063	0.293
	TAF5	0.466	0.013	0.452
	TAF6	0.532	0.032	0.500
	TAF7	0.761	0.015	0.746
	TAF9	0.610	0.011	0.598
	TAF8	0.585	0.150	0.435
	TAF11	0.527	0.018	0.509
	TAF12	0.541	0.074	0.468
	TAF13	0.655	0.044	0.612
PAF1	CTR9	1.213	0.166	1.046
	PAF1	0.973	0.130	0.843
	CDC73	0.803	0.052	0.751
	RTF1	0.650	0.088	0.562
Elongation factor	SPT5	1.154	0.290	0.863
	SPT6	0.945	0.267	0.679
	SPN1	0.721	0.219	0.502
SWI/SNF	SWI1	1.415	0.799	0.616
	SNF2	1.309	1.000	0.309
	SWI3	1.081	0.773	0.307
	SNF6	1.053	0.719	0.334
Others	ABD1	0.355	-0.048	0.403
	EAF3	0.439	0.072	0.367
	HEK2	0.764	-0.069	0.833
	RAD26	0.668	0.303	0.365
	RPH1	0.364	-0.031	0.395
	SET2	0.668	0.267	0.401
	SSD1	0.606	-0.060	0.666
	SWC3	1.295	0.956	0.339
	YJR039W	0.980	0.602	0.378

All numbers are expressed as log_2 values. α-amanitin sensitivity is the difference between the -α-amanitin and +α-amanitin enrichment values.

Yeast strains used in this study

Name	Stock#	Genotype	Source
BY4741	YF336	*MATa, ura3Δ0, leu2Δ0, his3Δ1, met15Δ0*	Winzeler et al. (1999) Scie
BJ2168	YF4	*MATa, ura3-52, leu2-, trp1-, prb1-1122, pep4-3, prc1-407, gal2-*	E. Jones (1991) Methods I
Rpb1-FRB	YSB3202	*MATα, ura3-, leu2-3,112, trp1-1, his3-11,15, fpr1::NAT, RPL13A-2xFKBP12::TRP1, tor1-1, ade2-1, can1-100, RPB1-FRB::HIS3MX*	In this study
Bdf1-FRB	YSB3323	*MATα, ura3-, leu2-3,112, trp1-1, his3-11,15, fpr1::NAT, RPL13A-2xFKBP12::TRP1, tor1-1, ade2-1, can1-100, BDF1-FRB::HIS3MX*	In this study
YSW87	YF157	*MATa, ura3Δ99, leuΔ1, his3Δ200, lys2-801, ade2-101, taf1Δ::LEU2, [pSW104 = TAF1, HIS3, CEN/ARS]*	Walker et al. (1996) Nature
YSW90 (*taf1*)	YF158	*MATa, ura3Δ99, leuΔ1, his3Δ200, lys2-801, ade2-101, taf1Δ::LEU2, [pRS313-TAF145^{ts} = taf1 ts allele ts-1, HIS3, CEN/ARS]*	Walker et al. (1996) Nature
taf11	YSB1732	*MATa, ura3-52, leu2, trp1Δ63, his3Δ200, taf11Δ::LEU2, [pRS314-FLAG-taf40-3100 = taf11-3100, TRP1, CEN/ARS]*	Komarnitsky, Michel, and B

Oligonucleotides used in this study

Purpose	Name	Strand	Sequence (5' → 3')	stock(O#)
G-less template (CYC1p)	Biotin - universal	Forward	5Biosg/TTGGGTAACGCCAGGGT	3151
	M13 -21 Reverse (FL)	Reverse	GGAAACAGCTATGACCATG	1477
	G-LESS PRIMER (+128)	Reverse	AGAGTGAATGATGATAGATTTGGGAAA	301
	mid G-less-runoff-R (+90)	Reverse	GAGAGAAGAGGAGAAGATAATA	3157
	Mid G-less-d3_rev (+16)	Reverse	GGAGGTATAGAAGTATAGTAATTTATG	3898
	Mid G-less-d4_rev (-16)	Reverse	AAAAATAATGTATAAGGAAAGAATATTTAGAG	3899
	Bio_CYC1pro_for (NΔ1)	Forward	5Biosg/TTCTTCTTTTCTCTAAATATTCTTTCC	3923
	Bio_Gless_for (NΔ2)	Forward	5Biosg/ATTTTTTTTTTTTAATATTATACCCTTCCTCCATCTATAC	3922
HIS4p template	p965	Forward	5Biosg/TAATGCAGCTGGCACGACAGG	1105
	pNot	Reverse	GGCCGCTCTAGCTGCATTAATG	1104
HIS4 template	HIS4 5' ORF rev	Reverse	GGTACATTCAATTGTTCAACCA	3513
PGK1 template	PGK1 5' rev	Reverse	GCAACAATTCTTTGGTTAGAAG	3433
RDN1-1 template	Bio_pGEM_for	Forward	5Biosg/GATATAGGCGCCAGCAAC	3278
	Pol I pro_rev	Reverse	AACGAACGACAAGCCTACT	3279

Plasmids used in this study

Plasmid name	Stock #	Features
pUC18-G_5CYC1 G-	SB649	5 copies of GAL4 binding sites upstream of *CYC1* promoter driving short G-less cassette (250, 277 nt transcript), Amp^R
pG5CG-D2	F916	5 copies of GAL4 binding sites upstream of *CYC1* promoter fused to a 377 bp G-less cassette with an introduced DraI site at bp 307, A
pUC18-G5 PGK1	SB1718	5 copies of GAL4 binding site upstream of PGK1 promoter and ORF, Amp^R
pSH515-HIS4	SB1722	wild type HIS4 transcription unit (5' core promoter, ORF, and 3' UTR) with a single Gal4 binding site, Amp^R
pSH515	F529	wild type HIS4 core promoter region with a single Gal4 binding site, Amp^R
pYrIIA	F82	yeast rDNA promoter, Amp^R

References

Huisinga, K.L., and Pugh, B.F. (2004). A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in Saccharomyces cerevisiae. Mol Cell *13*, 573-585.

A. Johnson, G. Li, T.W. Sikorski, S. Buratowski, C.L. Woodcock, and D. Moazed (2009).Reconstitution of heterochromatin-dependent transcriptional gene silencing. Mol. Cell *35*, 769-781.

Jones, E.W. (1991). Tackling the protease problem in Saccharomyces cerevisiae. Methods Enzymol *194*, 428-53.

Komarnitsky, P.B., Michel, B., and Buratowski, S. (1999). TFIID-specific yeast TAF40 is essential for the majority of RNA polymerase II-mediated transcription in vivo. Genes Dev *13*, 2484–2489.

Louder, R.K., He, Y., Lopez-Blanco, J.R., Fang, J., Chacon, P., and Nogales, E. (2016). Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. Nature *531*, 604-609.

Pardee, T.S., Ghazy, M.A., and Ponticelli, A.S. (2003). Yeast and Human RNA polymerase II elongation complexes: evidence for functional differences and postinitiation recruitment of factors. Eukaryotic Cell *2*, 318–327.

Reja, R., Vinayachandran, V., Ghosh, S., and Pugh, B.F. (2015). Molecular mechanisms of ribosomal protein gene coregulation. Genes Dev *29*, 1942-1954.
Rhee, H.S., and Pugh, B.F. (2012). Genome-wide structure and organization of eukaryotic pre-initiation complexes. Nature *483*, 295-301.

Winzeler, E.A. et al. (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science *285*, 901-6.