Individual letters of the RNA polymerase II CTD code govern distinct gene expression programs in fission yeast

Beate Schwer[a,1], Danny Asher Bitton[b], Ana M. Sanchez[a], Jürg Bähler[b], and Stewart Shuman[c,1]

[a]Department of Microbiology and Immunology, Weill Cornell Medical College, New York, NY 10065; [b]Department of Genetics, Evolution and Environment and UCL Cancer Institute, University College London, London WC1E 6BT, United Kingdom; and [c]Molecular Biology Program, Sloan–Kettering Institute, New York, NY 10065

Edited by Fred M. Winston, Harvard Medical School, Boston, MA, and approved February 7, 2014 (received for review November 21, 2013)

The primary structure and phosphorylation pattern of the tandem $Y^1S^2P^3T^4S^5P^6S^7$ repeats of the RNA polymerase II carboxyl-terminal domain (CTD) comprise an informational code that coordinates transcription, chromatin modification, and RNA processing. To gauge the contributions of individual CTD coding "letters" to gene expression, we analyzed the poly(A)$^+$ transcriptomes of fission yeast mutants that lack each of the four inessential CTD phosphoacceptors: Tyr1, Ser2, Thr4, and Ser7. There was a hierarchy of CTD mutational effects with respect to the number of dysregulated protein-coding RNAs, with *S2A* (n = 227) >> *Y1F* (n = 71) > *S7A* (n = 58) >> *T4A* (n = 7). The majority of the protein-coding RNAs affected in *Y1F* cells were coordinately affected by *S2A*, suggesting that Tyr1-Ser2 constitutes a two-letter code "word." *Y1F* and *S2A* elicited increased expression of genes encoding proteins involved in iron uptake (Frp1, Fip1, Fio1, Str3, Str1, Sib1), without affecting the expression of the genes that repress the iron regulon, implying that Tyr1-Ser2 transduces a repressive signal. *Y1F* and *S2A* cells had increased levels of ferric reductase activity and were hypersensitive to phleomycin, indicative of elevated intracellular iron. The *T4A* and *S7A* mutations had opposing effects on the phosphate response pathway. *T4A* reduced the expression of two genes encoding proteins involved in phosphate acquisition (the Pho1 acid phosphatase and the phosphate transporter SPBC8E4.01c), without affecting the expression of known genes that regulate the phosphate response pathway, whereas *S7A* increased *pho1*$^+$ expression. These results highlight specific cellular gene expression programs that are responsive to distinct CTD cues.

iron homeostasis | phosphate homeostasis | transcription profiling

The carboxyl-terminal domain (CTD) of the Rpb1 subunit of RNA polymerase II (Pol II) consists of tandemly repeated heptapeptides of consensus sequence $Y^1S^2P^3T^4S^5P^6S^7$. The inherently plastic CTD structure is modulated by phosphorylation of the Tyr1, Ser2, Thr4, Ser5, and Ser7 residues and by *cis–trans* isomerization of the prolines (1, 2). With as many as 128^n potential CTD primary structures (where n is the number of heptads), the CTD provides information about the state of the transcription machinery—a CTD code—that is "read" by CTD receptor proteins that control transcription, modify chromatin structure, and catalyze or regulate mRNA capping, splicing, and polyadenylation (1, 2).

Basic informational rules that govern the CTD code have been elucidated by genetically manipulating the composition and structure of the Rpb1 CTD in the fission yeast *Schizosaccharomyces pombe* (3–7). By introducing alanines and conservative mutations in lieu of Tyr1, Ser2, Pro3, Thr4, Ser5, Pro6, and Ser7 of every consensus heptad of a fully functional Rpb1 CTD array (comprising 14 consensus heptad repeats linked to the body of Rpb1 by a "rump" consisting of four degenerate heptads), we determined that: (*i*) Tyr1, Pro3, Ser5, and Pro6 are essential for viability, by the criterion that alanine substitution is lethal, whereas Ser2, Thr4, and Ser7 are not essential; and (*ii*) *Y1F*, *S2A*•*S7A* and *T4A*•*S7A* mutants are viable, signifying that phenylalanine is functional in lieu of Tyr1 and that Ser5 is the only strictly essential phosphorylation site in fission yeast (6).

Further interrogation of the grammar and punctuation of the CTD code established that: (*i*) all essential coding letters and "words" are encompassed within a diheptad repeat; (*ii*) the minimal fission yeast CTD coding unit is a decapeptide $Y^1S^2P^3T^4S^5P^6S^7Y^1S^2P^3$ and the spacing between coding units is flexible; (*iii*) Tyr1 must be present in consecutive heptads and proper spacing between consecutive tyrosines is important for CTD function; (*iv*) Pro3, Ser5, and Pro6 need not be present in adjacent heptads; and (*v*) Ser5(P)-Pro6 comprises an essential two-letter code word that is read chiefly by the mRNA capping apparatus (7). Querying how perturbations of CTD primary structure affect CTD serine phosphorylation patterns in vivo, as gauged by Rpb1 reactivity with phosphospecific antibodies, showed that Ser2 phosphorylation does not rely on Ser5, Pro6, Ser7, or Thr4, whereas Ser5 phosphorylation does not depend on Ser2, Thr4, or Ser7.

The fact that four of the five phosphoacceptor coding letters of the CTD heptad are not essential in fission yeast raises important questions as to whether and how these phosphate marks impact gene expression, the extreme situations being that absence of a particular CTD-PO_4 mark has little or no effect or that loss of a coding cue does exert significant effects, albeit on the expression of genes that are not essential under the laboratory

Significance

The carboxyl-terminal domain (CTD) code encrypted within the YSPTSPS heptad repeats of RNA polymerase II is deeply rooted in eukaryal biology. This paper exploits fission yeast CTD mutants in which single coding cues (or combinations of cues) are subtracted from all consensus heptads to gauge the roles of the Tyr1, Ser2, Thr4, and Ser7 phosphoacceptors in gene expression. Transcriptome profiling revealed distinctive signatures for the CTD phosphoacceptors that were verified by cell-based activity assays. The results highlight (*i*) the role of Tyr1 and Ser2 in repressing the iron uptake regulon and (*ii*) the distinctive roles of Thr4 and Ser7 in the phosphate starvation response, whereby Thr4 facilitates and Ser7 represses the induction of Pho1 acid phosphatase.

Author contributions: B.S., D.A.B., J.B., and S.S. designed research; B.S., D.A.B., A.M.S., and S.S. performed research; B.S., D.A.B., A.M.S., J.B., and S.S. analyzed data; and B.S., D.A.B., and S.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE52370).

[1]To whom correspondence may be addressed. E-mail: bschwer@med.cornell.edu or s-shuman@ski.mskcc.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1321842111/-/DCSupplemental.

conditions surveyed. The latter scenario applies to the Ser2 coding letter, insofar as a heterothallic haploid fission yeast *rpb1-S2A* mutant can grow vegetatively, but is unable to mate with heterothallic WT cells of either mating type, i.e., *S2A* is sterile. Sterility results from the inability of *S2A* cells to induce expression of the Ste11 transcription factor that triggers the meiotic sexual differentiation pathway (5, 6, 8). Microarray analyses of the transcriptome of *S2A* cells vs. WT highlighted ≥1.5-fold changes in the expression level of many meiotically regulated genes (5, 9). The Ser2 requirement for transcription during sexual differentiation was bypassed by subtracting Ser7, suggesting that imbalance in the phosphorylation array, not absence of a CTD-PO_4 cue, underlies a CTD-associated pathology (6).

In the present study, we apply high-throughput RNA sequencing methods (RNA-seq) to gauge globally the impact of the loss of each of the four inessential CTD phosphoacceptors on gene expression. By comparing the poly(A)$^+$ transcriptomes of four fission yeast mutants, *rpb1-Y1F*, *-S2A*, *-T4A*, and *-S7A*, we illuminate how individual letters of the Pol II CTD code affect the expression of distinct sets of genes. Validation of selected transcriptional programs by functional tests identified Tyr1 and Ser2 as a two-letter coding word that governs an iron-uptake regulon (10) and Thr4 and Ser7 as components of the phosphate response pathway that regulates Pho1 expression.

Results

RNA-Seq Analysis of *rpb1-CTD* Mutants. RNA-seq affords a detailed strand-specific annotated map of the protein-coding and noncoding fission yeast transcripts vs. their genomic DNA loci (11–13). Here we performed RNA-seq on poly(A)$^+$ RNA isolated from cells expressing Rpb1 with CTDs composed of the rump plus 14 consensus YSPTSPS heptads or an identical number of Y1F, S2A, T4A, or S7A mutant heptads (6). We also analyzed CTD double mutants *Y1F•S7A*, *S2A•S7A*, and *T4A•S7A*. cDNAs obtained from two nonsimultaneous biological replicates (using RNA from cells grown to midlog phase at 30 °C) were sequenced for each strain. In all of the datasets, ∼90% of the reads were mapped to unique genomic loci (Fig. S1), and expression levels were highly reproducible between biological replicates (Fig. S2). As an internal control, we queried the cDNA sequences encoding the Rpb1 CTD itself and found that between 99.7% and 100% of the CTD mRNA reads were correct with respect to the WT or relevant mutant CTD, with the rare exceptions corresponding to other CTD variants that were bar-coded in the pooled sequence run (Fig. S1). A cutoff of plus or minus twofold change in normalized transcript read level and a corrected *P* value of <0.05 were the criteria applied to derive lists of differentially expressed annotated loci for each *rpb1-CTD* single-mutant vs. the WT CTD strain.

The results are presented as heat maps for the ensemble of significantly changing protein-coding and noncoding RNAs (Fig. S3) and as four separate Venn diagrams in Fig. 1, showing protein-coding transcripts that were down-regulated (Fig. 1*A*) or up-regulated (Fig. 1*B*) and noncoding transcripts that were down-regulated (Fig. 1*C*) or up-regulated (Fig. 1*D*) in the *Y1F*, *S2A*, *T4A*, and *S7A* strains. There was a clear hierarchy of CTD mutational effects on the number of annotated coding plus noncoding RNAs dysregulated, whereby *S2A* (n = 481) had greater impact than *Y1F* (n = 107) or *S7A* (n = 97) and *T4A* (n = 15) had only scant effects on gene expression. Of the 5,133 annotated protein-coding transcripts in fission yeast, 157 (3.1%) were down-regulated in *S2A*; 54 (1.1%) in *Y1F*; 25 (0.5%) in *S7A*; and only 4 (0.08%) in *T4A*. To the extent that the at least twofold decrement in individual transcript levels in these mutant strains reflects a contribution of the missing CTD coding letter or phosphorylation mark for achieving normal levels of transcription or processing of the affected mRNAs, these results are remarkable for how few fission yeast genes rely on these CTD cues. However, the modest effects of the *rpb1-CTD Y1F*, *T4A*, and *S7A* mutations on mRNA levels were in keeping with the fact that these strains grow as well as WT cells in YES medium at 30 °C (6). The *S2A* strain, in which more mRNAs are down-regulated, grows slightly slower than WT at 30 °C; this is the case even when the length of the CTD S2A array appended to the rump is increased from 14 to 25 heptads (Fig. S4).

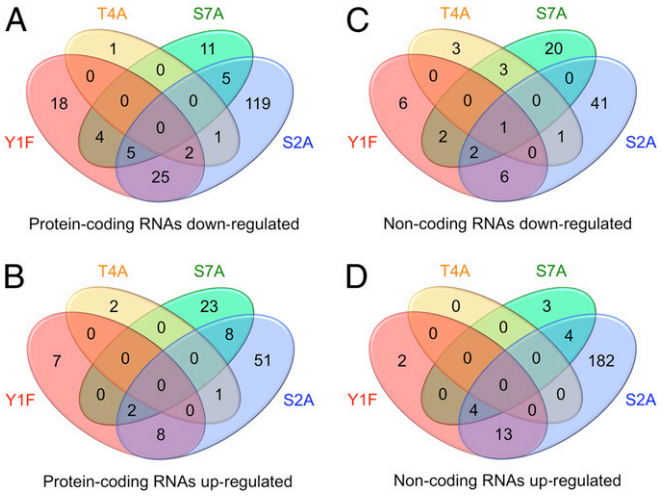

Fig. 1. Effect of CTD mutations on the fission yeast transcriptome. Four Venn diagrams specify the numbers of protein-coding (*A* and *B*) and noncoding (*C* and *D*) poly(A)$^+$ transcripts that were significantly down-regulated by at least twofold (*A* and *C*) or significantly up-regulated by at least twofold (*B* and *D*) in the indicated *rpb1-CTD* strains relative to WT strain.

Even fewer protein-coding transcripts were up-regulated in response to CTD mutations: 70 (1.4%) in *S2A*; 33 (0.6%) in *S7A*; 17 (0.3%) in *Y1F*; and 3 (0.06%) in *T4A*. We surmise that the pertinent CTD coding letters or phosphorylation marks negatively regulate the transcription or processing of a limited set of fission yeast mRNAs.

The situation was quite different with respect to the effect of CTD mutations on the noncoding poly(A)$^+$ transcriptome, whereby *S2A* elicited up-regulation of 203 (11%) of the 1,889 annotated noncoding RNAs. By contrast, 19, 11, and 0 noncoding RNAs were up-regulated in *Y1F*, *S7A*, and *T4A* cells, respectively (Fig. 1*D*). These findings implicate the Ser2 hydroxyl and/or the Ser2-PO_4 mark as a negative regulator of noncoding RNA biogenesis. Relatively few noncoding RNAs were down-regulated in the CTD mutants: 51 (2.7%) in *S2A*; 28 (1.5%) in *S7A*; 17 (0.9%) in *Y1F*; and 8 (0.4%) in *T4A* (Fig. 1*C*).

RNAs Dysregulated by *Y1F* Are Coordinately Dysregulated by *S2A*. From this point forward, we will focus our analysis on the extent to which the CTD mutations exert their effects on distinct vs. overlapping sets of RNAs, as a means of gleaning clues to the syntax of the CTD code. The rationale is that, if mutations in two coding letters coordinately affect the expression of a common set of genes, the two letters might well comprise a coding word read by proteins that control expression of the shared target gene set. Inspection of the Venn diagrams in Fig. 1 highlights a consistent overlap pattern, whereby the majority of the RNAs that were dysregulated by *Y1F* were coordinately dysregulated by *S2A*. For example, 32 of the 54 protein-coding RNAs down-regulated in *Y1F* cells were also repressed in *S2A* cells ($P < 5.8 \times 10^{-36}$, hypergeometric test). The shared target set for *Y1F* and *S2A* down-regulation (Fig. S5, *Upper*) is rich in genes encoding enzymes of nucleotide, sugar, and amino acid metabolism. Similarly, 9 of the 17 noncoding RNAs down-regulated by *Y1F* were also down-regulated by *S2A* ($P < 7.6 \times 10^{-11}$; Fig. 1).

The coregulation theme extends to the RNAs that increase in response to CTD mutations. Ten of the 17 protein-coding RNAs that were up-regulated by *Y1F* were coordinately up-regulated by *S2A* ($P < 2.1 \times 10^{-15}$). The shared set for up-regulation (Fig. S5, *Lower*) comprises mRNAs encoding proteins involved in iron uptake and/or regulated by iron (Str3, Frp1, Fip1, Fio1, Str1, Srx1)

and proteins induced during meiosis (Ppk31, Mde5, Mug136, Meu14). Seventeen of the 19 noncoding RNAs up-regulated by *Y1F* were also up-regulated by *S2A* ($P < 3.1 \times 10^{-17}$). No comparable extent of overlap was evident for any other pair of CTD mutants. These results suggest that Tyr1 and Ser2 are a two-letter CTD coding word that governs the expression of a discrete subset of fission yeast genes.

Tyr1 and Ser2 Transduce Repression of an Iron Uptake Regulon. Fission yeast has a finely tuned system of iron homeostasis entailing regulated expression of genes encoding proteins involved in iron uptake (10). Key players are the cell surface ferric reductase Frp1 that converts extracellular ferric ion to ferrous ion, which is then transported across the cell membrane by the oxidase–permease complex composed of Fio1 and Fip1 proteins. A parallel pathway of iron acquisition entails the action of a ferrochrome siderophore, synthesized by Sib1 and Sib2, and cell surface siderophore transporters Str1, Str2, and Str3. The mRNAs encoding Frp1, Fio1, Fip1, Str1, and Str3 were all up-regulated in the *Y1F* strain versus the isogenic WT strain, by 16-fold, 15-fold, 14-fold, 10-fold, and 98-fold, respectively (Fig. 2). These five RNAs were also up-regulated in *S2A* cells, as was the RNA encoding Sib1. Moreover, the mRNAs encoding Frp1, Fip1, Fio1, Str3, and Str1 were up-regulated in the CTD double-mutants *Y1F•S7A* and *S2A•S7A* (Fig. 2), fortifying the inference that CTD coding letters Tyr1 and Ser2 act to down-regulate iron uptake genes. The RNA-seq results were verified by reverse transcriptase (RT)-qPCR analyses, which showed that the levels of mRNAs encoding Fio1, Fip1, and Frp1 were up-regulated in the *Y1F*, *S2A*, *Y1F•S7A*, and *S2A•S7A* mutants, but not in the *S7A* mutant (Fig. S6 *A–C*).

The GATA-family transcription factor Fep1 represses expression of the aforementioned iron uptake genes under iron-replete growth conditions (10, 14). Fep1 binds in an iron-dependent manner to a GATA sequence element [5′-ATC(A/T)GATA(A/T)] in the regulatory regions upstream of the repressible iron uptake genes. Additional iron/Fep1-regulated genes identified by expression profiling with DNA microarrays include those encoding sulfiredoxin Srx1, Fe–S cluster assembly protein Isu1, and heme utilization protein Shu1 (15). The Srx1 mRNA was up-regulated in *Y1F*, *S2A*, *Y1F•S7A*, and *S2A•S7A* cells (Fig. 2). The Shu1 and Isu1 transcripts were up-regulated in *Y1F•S7A* and *S2A•S7A* cells (Fig. 2). (They were also up-regulated in *Y1F* cells, albeit not at the *P* value cutoff of 0.05). Thus, the loss of the Tyr1 and Ser2 CTD cues specifically derepresses the regulon governed by Fep1.

A trivial explanation for this derepression would be that *Y1F* and *S2A* down-regulate the expression of the proteins that repress the iron regulon. These include, in addition to Fep1, the corepressors Tup11 and Tup12 (10, 14). However, this scenario was vitiated by the findings that the Fep1, Tup11, and Tup12 mRNA levels, gauged by RNA-seq, were not affected significantly by the *Y1F*, *S2A*, *Y1F•S7A*, or *S2A•S7A* mutations (Fig. 2). RT-qPCR analysis verified that the level of mRNA encoding the transcriptional repressor Fep1 was unaffected by these CTD mutations (Fig. S6*D*). Another possibility would be that *Y1F* and *S2A* prompt overexpression of a factor that derepresses the iron regulon, e.g., by negatively regulating Fep1. Grx4 is implicated as a negative regulator of Fep1 (10), but Grx4 expression was not affected significantly in *Y1F*, *S2A*, *Y1F•S7A*, or *S2A•S7A* cells (Fig. 2). Therefore, we infer that the CTD Tyr1-Ser2 coding word is needed to transduce a repressive signal from Fep1 to the transcription apparatus.

		RPKM: expt. 1 / 2 [$\log_2$ fold change]				
		WT	Y1F	S2A	Y1F•S7A	S2A•S7A
Frp1	ferric reductase	0.79 / 0.31	4.44 / 4.69 [4.01]	2.68 / 1.85 [1.44]	5.06 / 3.88 [3.94]	4.05 / 4.15 [3.34]
Fip1	iron permease	4.58 / 4.96	8.36 / 8.97 [3.93]	7.54 / 7.57 [2.50]	8.40 / 7.98 [3.36]	8.04 / 8.39 [3.25]
Fio1	iron transport oxidase	1.08 / 0.97	4.24 / 5.28 [3.83]	3.95 / 4.79 [3.15]	3.91 / 4.14 [2.94]	4.30 / 4.61 [3.24]
Str3	siderophore-iron transporter	−5.50 / −2.38	3.59 / 3.02 [6.61]	0.82 / −0.13 [3.36]	3.38 / 3.44 [6.59]	3.03 / 4.40 [6.89]
Str1	siderophore-iron transporter	3.11 / 3.67	6.51 / 6.88 [3.31]	5.54 / 4.86 [1.53]	6.68 / 5.87 [2.84]	6.16 / 6.35 [2.65]
Sib1	ferrichrome synthetase	1.53 / 1.57	1.97 / 1.33 0.16	3.38 / 2.58 [1.18]	2.30 / 1.59 0.36	2.80 / 2.45 0.91
Srx1	sulfiredoxin	3.74 / 3.57	5.80 / 5.53 [2.03]	5.32 / 4.94 [1.19]	5.33 / 5.05 [1.47]	4.71 / 5.40 [1.23]
Shu1	heme utilization	−2.04 / −16.6	4.25 / 2.38 *[6.63]*	1.30 / −0.35 *[3.38]*	4.26 / 2.89 [6.66]	3.89 / 4.39 [6.96]
Isu1	Fe-S cluster assembly	6.74 / 6.74	7.68 / 8.31 *[1.31]*	7.20 / 7.53 0.37	7.83 / 7.93 [1.08]	7.84 / 8.06 [1.02]
Fep1	repressor of iron acquisition	4.65 / 4.67	4.33 / 4.54 −0.19	4.37 / 4.58 −0.44	4.18 / 4.34 −0.44	4.32 / 4.52 -0.41
Tup11	co-repressor	5.68 / 5.66	5.65 / 5.45 −0.10	5.85 / 5.81 −0.12	5.72 / 5.48 −0.13	5.70 / 5.61 −0.20
Tup12	co-repressor	6.45 / 6.66	6.37 / 6.47 −0.12	6.31 / 6.62 −0.35	6.53 / 6.78 0.05	6.72 / 6.97 0.09
Grx4	de-repressor (glutaredoxin)	7.61 / 7.59	7.98 / 8.07 0.44	8.04 / 7.78 0.03	8.17 / 7.89 0.37	7.96 / 8.07 0.23

Fig. 2. CTD mutations *Y1F* and *S2A* derepress an iron uptake regulon. The expression levels (reads per kilobase per million reads; RPKM) of genes involved in iron uptake and utilization are shown for the biological replicates of each of the strains specified. The $\log_2$ fold changes vs. WT are denoted below the expression levels. Values meeting the dual cutoff criteria of at least twofold change and $P < 0.05$ are highlighted in brackets in plain font. Values of at least twofold change with $P > 0.05$ are italicized and in brackets. Other $\log_2$ fold change values not meeting either criterion are in plain font. The genes derepressed by CTD mutations are listed above the spacer row shaded gray. The transcriptional regulatory genes that repress the iron regulon are listed below the spacer row.

***Y1F* and *S2A* Affect Iron Homeostasis.** If the derepression of the iron regulon at the mRNA level is physiologically important, we would expect to see increased activity of Fep1 regulated proteins in cells bearing the *Y1F* and *S2A* CTD mutations. We assayed ferric reductase Frp1 activity by overlaying cells grown on iron-replete medium with agar containing the ferric reductase substrate triphenyltetrazolium chloride, which is reduced by cell surface-associated Frp1 to yield a red pigmented product (14). This assay distinguished the pale WT, *S7A*, and *T4A* colonies from the red *Y1F*, *Y1F•S7A*, and *S2A•S7A* colonies (Fig. 3, *Right*). The *S2A* strain grew more slowly under these conditions, and hence there were fewer cells on the plate, but it was still apparent that the *S2A* cells generated red pigment. Thus, the effect of CTD mutations on the activity of the Frp1 enzyme in vivo correlated with the mRNA levels gauged by RNA-seq.

We surveyed the effects of the CTD mutations on the sensitivity of fission yeast to phleomycin, a glycopeptide antibiotic that cleaves the phosphodiester backbone of nucleic acids in the presence of iron and oxygen. Derepression of the Fep1-regulated iron uptake regulon causes hypersensitivity to phleomycin toxicity (14) by virtue of increasing the available intracellular iron. The salient findings were that *Y1F*, *S2A*, *Y1F•S7A*, and *S2A•S7A* strains failed to grow in the presence of 10 μg/mL phleomycin, a concentration of drug that did not prevent growth of the WT, *S7A*, and *T4A* strains (Fig. 3, *Left*). We conclude from these experiments that CTD Tyr1 and Ser2 are critical informational cues for proper iron homeostasis in fission yeast.

Thr4 and Ser7 Govern Expression of Acid Phosphatase Pho1 and Pho1 Induction by Phosphate Starvation. Although the *T4A* CTD mutation resulted in decreased expression of merely four mRNAs, two of them encode proteins involved in phosphate homeostasis: the secreted acid phosphatase Pho1 and the inorganic phosphate transporter SPBC8E4.01c. The fission yeast phosphate transporter and Pho1 are encoded by adjacent cooriented genes on chromosome II. The Pho1 and phosphate transporter mRNAs were down-regulated by 7.5-fold and fourfold, respectively, in *T4A* cells and by twofold and 2.7-fold, respectively, in the *S2A* mutant (Fig. 4*A*). By contrast, Pho1 was up-regulated twofold in response to the *S7A* mutation (Fig. 4*A*). An instructive finding was that Pho1 was down-regulated 12-fold in the *T4A•S7A* double mutant (Fig. 4*A*), signifying that the decrement in

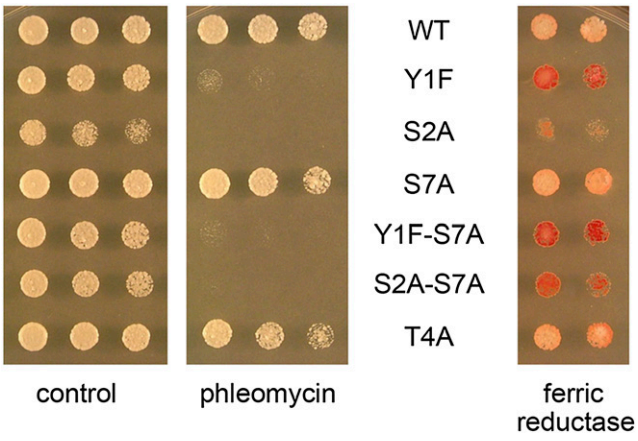

Fig. 3. *Y1F* and *S2A* derepress ferric reductase and sensitize to phleomycin. (*Left*) Phleomycin sensitivity. Midlog phase cultures of the indicated strains grown in yeast extract with supplements (YES) broth at 30 °C were adjusted to 6×10^4 cells/mL and aliquots (5 μL) of serial twofold dilutions were spotted on YES agar medium with no addition (control) or containing 10 μg/mL phleomycin. The plates were incubated at 30 °C and photographed after 3 d (control) or 4 d (phleomycin). (*Right*) Ferric reductase activity. Cells as specified were spotted on YES agar medium containing 50 μg/mL $FeCl_3$ and grown for 2 d at 30 °C. The colonies were overlaid with 1% Bacto agar containing 0.1% TTC and 0.067 M phosphate (pH 7). The plate was photographed after incubation for 45 min at 22 °C.

expression caused by loss of the Thr4 cue was dominant over the gain seen with *S7A*. The phosphate transporter mRNA was also down-regulated in *T4A•S7A* cells (Fig. 4*A*).

The RNA-seq findings that *pho1* mRNA was down-regulated in *T4A* and *T4A•S7A* cells and up-regulated in *S7A* cells were verified by two independent methods: RT-qPCR (Fig. S7*D*) and primer extension (Fig. S7 *A* and *B*). Reverse transcriptase extension of a ^{32}P-labeled DNA primer complementary to nucleotides +33 to +51 of the *pho1* ORF revealed a unique product of identical length (~100 nt) in each of the yeast strains, albeit differing in abundance relative to an internal *act1* primer extension control (Fig. S7*A*). By analyzing the *pho1* primer extension product in parallel with a chain-terminated sequencing ladder generated by DNA polymerase-catalyzed extension of the same ^{32}P-labeled primer annealed to a *pho1* DNA template (Fig. S7*C*), we mapped the transcription start site to the G position 51 nt upstream of the *pho1* AUG start codon.

To gauge the effect of the CTD mutations on acid phosphatase activity, we plated serial dilutions of cells on pombe minimal glutamate (PMG) agar medium (which contains 15.5 mM inorganic phosphate), allowed the colonies to attain comparable cell densities, and then overlaid the plate with phosphatase substrate α-naphthyl phosphate, Fast Blue Salt B stain, and 0.1 M sodium acetate (pH 4.2). The secreted acid phosphatase causes formation of a red pigment and relative color intensity provides a semiquantitative assay of Pho1 activity (16). The *T4A* and *T4A•S7A* colonies were clearly paler than the WT, whereas the *S7A* colonies were more red than the WT (Fig. 4*B*), signifying that the amount of active phosphatase was regulated down or up by CTD mutations in step with the levels of *pho1* mRNA. Similar trends in phosphatase staining were seen when the yeast colonies were grown on YES agar medium (which contains ~2 mM phosphate), albeit with a higher basal level of phosphatase activity (Fig. 4*B*). We surmise that the variations in phosphatase activity reflect effects on Pho1 abundance, insofar as the RNA-seq analysis showed that the level of mRNA encoding a second fission yeast phosphatase Pho4 was unaffected by the CTD mutations (Fig. S8).

The RNA-seq analysis was performed on yeast cells grown in the presence of phosphate. To determine the impact of the CTD mutations on the induction of Pho1 under conditions of phosphate starvation, cells grown in YES medium were washed and then incubated for 3 h in synthetic medium containing 15.5 mM phosphate, or in medium lacking exogenous phosphate to elicit the starvation response. Acid phosphatase activity was quantified in a solution assay by incubating suspensions of equal numbers of the phosphate-replete or phosphate-starved cells for 5 min with *p*-nitrophenylphosphate and scoring colorimetrically the formation of *p*-nitrophenol. The basal phosphatase activity of WT cells was increased sixfold by phosphate starvation (Fig. 4*C*). Whereas

A

		$\log_2$ fold change			
		S2A	T4A	S7A	T4A•S7A
Pho1	Acid phosphatase	−1.06	−2.91	1.41	−3.57
SPBC8E4.01c	Phosphate transporter	−1.42	−2.02	0.65	−2.22

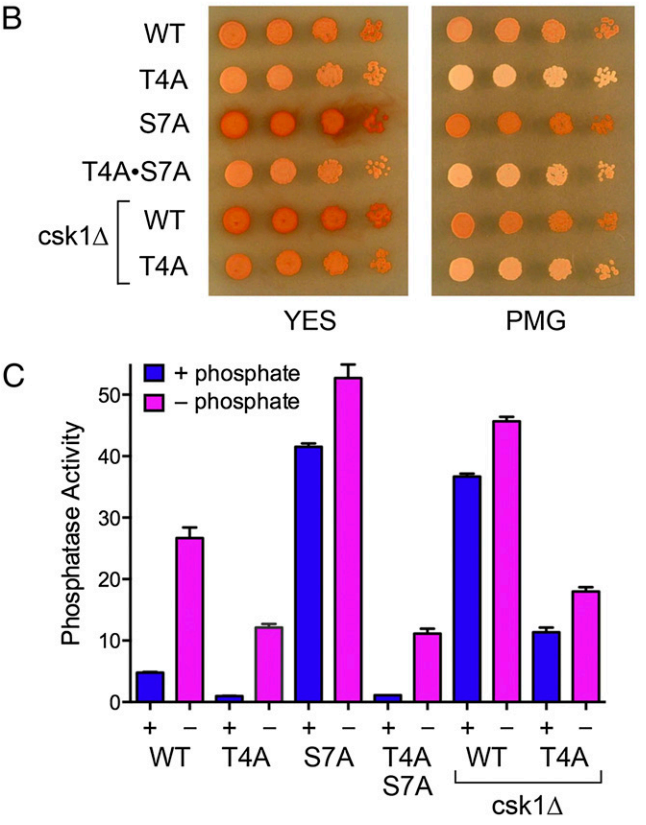

Fig. 4. CTD mutations affect expression of acid phosphatase Pho1. (*A*) Changes in transcript levels for *pho1*$^+$ and SPBC8E4.01c are shown for the indicated CTD mutant compared with WT cells. (*B*) Detection of acid phosphatase activity. Exponentially growing cultures of the indicated strains were adjusted to A_{600} of 0.1 and aliquots (3 μL) of serial fivefold dilutions were spotted on YES agar or PMG agar medium. After growth for 3 d at 30 °C, the colonies were overlaid with 1% Bacto agar containing 0.015% α-naphthyl phosphate and 0.15% Fast Blue Salt in 0.1 M sodium acetate (pH 4.2). The plates were photographed after incubation for 5 min at 22 °C. (*C*) Induction of acid phosphatase activity after phosphate starvation. Aliquots of exponentially growing cultures in YES medium were harvested, the cells were washed in water and adjusted to A_{600} of ~0.3 in PMG medium (with 5 μg/mL thiamine), either with 15.5 mM phosphate (+) or without phosphate (–) to induce Pho1. After incubation for 3 h at 30 °C, cells were harvested, washed, and suspended in water to attain an A_{600} of 1.25. To quantify acid phosphatase activity, reaction mixtures (200 μL) containing 100 mM sodium acetate (pH 4.2), 10 mM *p*-nitrophenylphosphate, and 0.025, 0.05, or 0.1 A_{600} unit of cells were incubated for 5 min at 30 °C. The reactions were quenched by adding 1 mL of 1 M sodium carbonate, the cells were removed by centrifugation, and the absorbance of the supernatant at 410 nm was measured. Acid phosphatase activity is expressed as the ratio of A_{410} (*p*-nitrophenol production) to A_{600} (cells). Each datum in the bar graph is the average of three assays by using cells from three independent cultures ± SEM.

the basal phosphatase activity of *T4A* cells was fivefold lower than WT, the basal activity of *S7A* cells was ninefold higher than WT. *T4A* cells responded to phosphate starvation, but the level of induced phosphatase activity was only half that of WT cells. The constitutively high level of phosphatase activity in S7A cells was increased further during phosphate starvation, to a level twofold higher than in starved WT cells (Fig. 4*C*). The superinduction of Pho1 activity by *S7A* was effaced completely in the *T4A•S7A* strain, which displayed the same diminished basal and induced phosphatase activities seen in the *T4A* single mutant (Fig. 4*C*). We conclude that the Thr4 mark normally exerts a positive influence on the phosphate response, whereas the Ser7 mark exerts a strong repressive effect.

Pho1 and phosphate transporter expression are induced in response to phosphate starvation via the action of the Zn-finger transcription factor Pho7 (17, 18). The protein kinase Csk1 is a negative regulator of Pho1 and phosphate transporter expression, insofar as both are constitutively induced in *csk1*Δ cells. The fact that constitutive expression of Pho1 is abolished in a *csk1*Δ *pho7*Δ double mutant implies that Pho7 is essential for Pho1 transcription and acts downstream of Csk1 (17, 18). The RNA-seq data indicate that the decrements in Pho1 and phosphate transporter expression in *T4A* and *T4A•S7A* cells are not a consequence of decreased expression of the positive regulator Pho7 or increased expression of the negative regulator Csk1, because *pho1* and *csk1* mRNA levels were not affected significantly by the CTD mutations (Fig. S8).

The in vivo chromogenic colony assay indicated that *csk1*Δ cells with a WT CTD elaborated increased phosphatase activity compared with *csk1*$^+$ cells (Fig. 4*B*). Quantitative assay of acid phosphatase showed that basal of *csk1*Δ cells was eightfold higher than *csk1*$^+$ cells (Fig. 4*C*). The increase in basal Pho1 activity was attenuated in a *csk1*Δ *T4A* strain, as was the induction by phosphate starvation (Fig. 4*C*), suggesting that Csk1 acts upstream of the CTD in modulating the phosphate response.

It is worth noting that the *T4A* and *T4A•S7A* mutations did not significantly affect expression of the SPBC1271.09 gene that encodes a glycerophosphodiester transporter (Fig. S8), even though SPBC1271.09 is part of the phosphate response and coordinately regulated with Pho1 and the phosphate transporter by Pho7 and Csk1 (19). On the contrary, SPBC1271.09 was up-regulated along with Pho1 by the *S7A* mutation (Fig. S8). These findings highlight the selective impact of *T4A* and *T4A•S7A* on expression of the tandemly arranged phosphate transporter and acid phosphatase genes.

***S7A* Up-Regulates the Expression of Many Meiotic Genes.** The *S7A* CTD mutation resulted in the up-regulation of 33 protein-coding mRNAs; 16 of these transcripts encode proteins with known roles in fission yeast meiosis or are RNAs known to be up-regulated during meiosis (Fig. S9), accounting for their naming as meiosis up-regulated genes (Mug) or Mei4-dependent (Mde) genes (20–22). Mei4 is a meiosis-specific forkhead transcription factor that up-regulates the expression of multiple meiotic genes, either during sexual differentiation and sporulation or when Mei4 is ectopically expressed in vegetative cells (20–22); these include the genes encoding Spo6, Spn5, Spn6, Mfr1, Ght5, Mde3, Mde7, Mde1, Mei2, Mug130, Mug136, Mug147, and SPBC725.10 that were up-regulated in vegetative *S7A* cells. Because *mei4* was itself up-regulated by *S7A*, it is likely that Mei4 drives the observed up-regulation of several of its target genes in *S7A* cells. Our results implicate the CTD Ser7 mark as a negative influence on the expression of Mei4 and other meiotic genes.

The up-regulation elicited by *S7A* was largely effaced in the *Y1F•S7A* and *T4A•S7A* double mutants (Venn diagrams in Fig. S10). Whereas the majority of mRNAs up-regulated in *Y1F* cells remained up-regulated in *Y1F•S7A* cells (11 of 17), this was not the case for the set of mRNAs up-regulated by *S7A*, of which only 6 of 33 were up-regulated in *Y1F•S7A* cells. Similarly, 2 of 3 mRNAs up-regulated by *T4A* remained up-regulated in *T4A•S7A* cells, but only 8 of 33 mRNAs up-regulated by *S7A* were up-regulated in *T4A•S7A* cells (Fig. S10). With respect to the 16 meiotic genes up-regulated by *S7A*, 11 and 12 were not up-regulated in the *Y1F•S7A* and *T4A•S7A* mutants, respectively. These findings suggest that the biological output of a CTD coding letter may be contingent on the state of neighboring coding cues. This point is underscored by the observations that the *Y1F•S7A* and *T4A•S7A* double mutations elicited up-regulation of distinctive sets of coding mRNAs that were not up-regulated by the component single mutations (Fig. S10).

Discussion

The CTD code is deeply rooted in eukaryal biology, yet its complexity and output can vary among taxa. The trend, albeit oversimplified, is that the number of heptad repeats comprising the CTD, the minimal number of repeats required for viability, and the prevalence of nonconsensus coding letters increases as one moves up the evolutionary ladder from unicellular to multicellular eukarya, as a means to accommodate and orchestrate the actions of increasing numbers of CTD receptors that regulate gene expression (1, 2). It is reasonable to posit a core CTD code that governs events common to most eukarya; for example, the essential function of the Ser5-PO_4–Pro6 coding word in recruiting mRNA capping enzymes to the Pol II elongation complex, which is conserved in budding yeast, fission yeast, and mammals, notwithstanding major differences in genetic and physical organization of the capping apparatus in these taxa (6, 7, 19, 23, 24). The code is plastic (by virtue of the 128 potential structures each heptad can adopt), and can easily expand its functional capacity as new receptors emerge to read the code letters and words. Accordingly, the necessity for specific CTD cues and their impact on cellular gene expression may differ among taxa. The Ser2, Thr4, Ser7, and Tyr1 hydroxyl groups are inessential for vegetative growth of fission yeast (6). In metazoa, Ser2, Thr4, and Ser7 are critical for cell viability and/or specific steps of transcription or RNA processing, e.g., because these coding cues are recognized by components of the processing machinery, some of which are unique to metazoa (25–27).

The homogeneous primary structure of the fission yeast CTD and the availability of viable *rpb1* mutants in which single coding cues (or combinations of cues) are subtracted from all the consensus heptads makes for an attractive experimental system to survey the roles of the Tyr1, Ser2, Thr4, and Ser7 phosphoacceptors in gene expression. As reported here, it was remarkable how few mRNAs were significantly dysregulated at least twofold by the CTD mutations *S2A*, *Y1F*, *S7A*, and *T4A*, which affected only 4.4%, 1.4%, 1.2%, and 0.14% of the annotated fission yeast protein-coding RNAs, respectively. We note that the method of RNA-seq analysis used here, whereby equal amounts of RNA from different fission yeast strains are subjected to high-throughput sequencing, assumes that the fission yeast strains analyzed produce similar levels of RNA per cell; we cannot exclude the possibility that a CTD mutation might elicit a global increase or decrement in the steady-state levels of all, or most, RNAs per yeast cell, an issue that has been discussed in relation to gene expression analyses in metazoan cells (28). Nonetheless, the analysis is highly instructive with respect to the selective effect of CTD mutations.

RNA-seq revealed distinctive gene expression signatures for the CTD phosphoacceptors that were verified by cell-based activity assays. Two examples stand out. First, our studies highlight the role of Tyr1 and Ser2 in repressing the expression of multiple unlinked genes involved in iron uptake that comprise a Fep1-repressed regulon. Although the molecular details remain to be elucidated, we speculate that Fep1 bound to DNA flanking the target gene promoters might interact, through its C-terminal repressor domain (10) or corepressors bound to the repressor domain, with the Pol II CTD to exert a gene-specific repressive effect on transcription (or cotranscriptional RNA processing) and that this interaction depends on Tyr1 and Ser2.

Second, we implicated Thr4 as an activator of the expression of two adjacent genes encoding proteins involved in phosphate homeostasis: a phosphate transporter and the acid

phosphatase Pho1. The CTD *T4A* mutation coordinately reduced *pho1* RNA and Pho1 activity at the cell surface. *T4A* blunted the increase in Pho1 activity that resulted from elimination of the repressive kinase Csk1, thereby placing Thr4 downstream of Csk1, perhaps acting in tandem with the transcriptional activator Pho7. However, the observation that an unlinked third Pho7-responsive gene in the phosphate pathway was not down-regulated by *T4A* raises the prospect that a particular feature of the transporter-phosphatase locus dictates its reliance on the Thr4 CTD cue. The CTD Ser7 residue was found to exert a negative effect on Pho1 expression, whereby *S7A* coordinately increased *pho1* RNA and Pho1 activity; when *S7A* was combined with *T4A*, the latter held sway and Pho1 was down-regulated.

In conclusion, the marriage of CTD genetics and transcriptome profiling provides unique insights to the output of the CTD code, especially by Tyr1, Thr4, and Ser7. That said, we cannot at present distinguish whether the narrow but specific gene expression dyscrasias elicited by CTD mutations *Y1F*, *T4A*, and *S7A* reflect the lack of phosphorylation of Tyr1, Thr4, or Ser7 or, rather, interference with phosphorylation-independent functions of the respective hydroxyl groups. Making this distinction will be contingent on identifying and genetically controlling the protein kinases responsible for Tyr1, Thr4, and Ser7 phosphorylation in fission yeast, an area that is presently *tabula rasa*.

Methods

The isogenic heterothallic *S. pombe rpb1-CTD* strains studied herein express Rpb1 variants in which 14 consensus YSPTSPS WT heptad repeats or 14 heptad repeats with single amino acid substitutions (Y1F, S2A, T4A, S7A) or double substitutions (Y1F•S7A, S2A•S7A, T4A•S7A) in each of the heptads are appended to amino acid 1577 of Rpb1 (6). The mating type, determined by crossing with tester strains, is h– for all strains studied, except for the sterile *S2A* mutant, which is h+ (as determined by diagnostic PCR using mat1-P– and mat1-M–specific primers). RNA was isolated from yeast cells grown in liquid culture in YES medium at 30 °C to an A_{600} of 0.4–0.7. Cells were harvested by centrifugation and total RNA was extracted by using the Qiagen RNeasy kit. The integrity and quantity of total RNA was gauged with an Agilent Technologies 2100 Bioanalyzer. Poly(A)$^+$ RNA was purified from 1 μg of total RNA by using the Illumina TruSeq RNA sample preparation kit. A TruSeq stranded total RNA sample preparation kit was used for subsequent steps of poly(A)$^+$ RNA fragmentation, strand-specific cDNA synthesis, indexing, and amplification. Indexed libraries were normalized and pooled (n = 8 samples per lane) for paired-end sequencing performed by using an Illumina HiSeq2000/1000 system. Paired-end reads of length 51 bases originating from each sample were aligned with Bowtie 0.12.7 (29) to the *S. pombe* genome sequence (Ensembl *S. pombe*, Build EF1, version 13) as well as to its corresponding exon–exon junction database (only the second part of the paired-end reads was considered). As many as three mismatches were allowed. Reads that matched multiple loci were removed from further analysis, and the resulting alignment files were processed to generate "pile-ups" against each chromosome. Differential expression between samples was determined using the DESeq Bioconductor package (30). A cutoff of plus or minus twofold change and corrected P value <0.05 were applied to derive a list of differentially expressed RNAs. The data analysis reported here was limited to the 7,022 annotated protein-coding and noncoding RNAs. Full details of the RNA-seq data processing and normalization of expression levels are provided in *SI Methods*.

ACKNOWLEDGMENTS. This research was supported by National Institutes of Health Grant GM52470 (to B.S. and S.S.) a Wellcome Trust Senior Investigator Award (to D.A.B. and J.B.), and a Royal Society Research Merit Award (to J.B.).

1. Eick D, Geyer M (2013) The RNA polymerase II carboxy-terminal domain (CTD) code. *Chem Rev* 113(11):8456–8490.
2. Corden JL (2013) RNA polymerase II C-terminal domain: Tethering transcription to transcript and template. *Chem Rev* 113(11):8423–8455.
3. Karagiannis J, Balasubramanian MK (2007) A cyclin-dependent kinase that promotes cytokinesis through modulating phosphorylation of the carboxy terminal domain of the RNA Pol II Rpb1p sub-unit. *PLoS ONE* 2(5):e433.
4. Schneider S, Pei Y, Shuman S, Schwer B (2010) Separable functions of the fission yeast Spt5 carboxyl-terminal domain (CTD) in capping enzyme binding and transcription elongation overlap with those of the RNA polymerase II CTD. *Mol Cell Biol* 30(10): 2353–2364.
5. Coudreuse D, et al. (2010) A gene-specific requirement of RNA polymerase II CTD phosphorylation for sexual differentiation in *S. pombe*. *Curr Biol* 20(12):1053–1064.
6. Schwer B, Shuman S (2011) Deciphering the RNA polymerase II CTD code in fission yeast. *Mol Cell* 43(2):311–318.
7. Schwer B, Sanchez AM, Shuman S (2012) Punctuation and syntax of the RNA polymerase II CTD code in fission yeast. *Proc Natl Acad Sci USA* 109(44):18024–18029.
8. Sukegawa Y, Yamashita A, Yamamoto M (2011) The fission yeast stress-responsive MAPK pathway promotes meiosis via the phosphorylation of Pol II CTD in response to environmental and feedback cues. *PloS Genet* 7(12):e1002387.
9. Saberianfar R, Cunningham-Dunlop S, Karagiannis J (2011) Global gene expression analysis of fission yeast mutants impaired in Ser-2 phosphorylation of the RNA pol II carboxy terminal domain. *PloS ONE* 6(9):e24694.
10. Labbé S, Khan MG, Jacques JF (2013) Iron uptake and regulation in *Schizosaccharomyces pombe*. *Curr Opin Microbiol* 16(6):669–676.
11. Wilhelm BT, et al. (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* 453(7199):1239–1243.
12. Rhind N, et al. (2011) Comparative functional genomics of the fission yeasts. *Science* 332(6032):930–936.
13. Marguerat S, et al. (2012) Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. *Cell* 151(3):671–683.
14. Pelletier B, Beaudoin J, Mukai Y, Labbé S (2002) Fep1, an iron sensor regulating iron transporter gene expression in *Schizosaccharomyces pombe*. *J Biol Chem* 277(25): 22950–22958.
15. Rustici G, et al. (2007) Global transcriptional responses of fission and budding yeast to changes in copper and iron levels: a comparative study. *Genome Biol* 8(5):R73.
16. Schweingruber ME, Schweingruber AM, Schüpbach ME (1982) Isolation and characterization of acid phosphatase mutants in *Schizosaccharomycespombe*. *Curr Genet* 5(2):109–117.
17. Henry TC, et al. (2011) Systematic screen of *Schizosaccharomyces pombe* deletion collection uncovers parallel evolution of the phosphate signal pathways in yeasts. *Euk. Cell* 10:198–206.
18. Carter-O'Connell I, Peel MT, Wykoff DD, O'Shea EK (2012) Genome-wide characterization of the phosphate starvation response in *Schizosaccharomyces pombe*. *BMC Genomics* 13:697.
19. Fabrega C, Shen V, Shuman S, Lima CD (2003) Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Mol Cell* 11(6):1549–1561.
20. Abe H, Shimoda C (2000) Autoregulated expression of *Schizosaccharomyces pombe* meiosis-specific transcription factor Mei4 and a genome-wide search for its target genes. *Genetics* 154(4):1497–1508.
21. Horie S, et al. (1998) The *Schizosaccharomyces pombe mei4*$^+$ gene encodes a meiosis-specific transcription factor containing a forkhead DNA-binding domain. *Mol Cell Biol* 18(4):2118–2129.
22. Mata J, Wilbrey A, Bähler J (2007) Transcriptional regulatory network for sexual differentiation in fission yeast. *Genome Biol* 8(10):R217.
23. Pei Y, Hausmann S, Ho CK, Schwer B, Shuman S (2001) The length, phosphorylation state, and primary structure of the RNA polymerase II carboxyl-terminal domain dictate interactions with mRNA capping enzymes. *J Biol Chem* 276(30):28075–28082.
24. Ghosh A, Shuman S, Lima CD (2011) Structural insights to how mammalian capping enzyme reads the CTD code. *Mol Cell* 43(2):299–310.
25. Hsin JP, Sheth A, Manley JL (2011) RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3′ end processing. *Science* 334(6056):683–686.
26. Hintermair C, et al. (2012) Threonine-4 of mammalian RNA polymerase II CTD is targeted by Polo-like kinase 3 and required for transcriptional elongation. *EMBO J* 31(12):2784–2797.
27. Egloff S, et al. (2007) Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* 318(5857):1777–1779.
28. Lovén J, et al. (2012) Revisiting global gene expression analysis. *Cell* 151(3):476–482.
29. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10(3):R25.
30. Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11(10):R106.

Supporting Information

Schwer et al. 10.1073/pnas.1321842111

SI Methods

Genome Level Alignments and Annotation. Paired-end reads of length 51 bases (each part) originating from each sample were aligned using Bowtie 0.12.7 (1) to the *Schizosaccharomyces pombe* genome sequence (Ensembl *S. pombe*, Build EF1, version 13) as well as to its corresponding exon–exon junction database (only the second part of the paired-end reads was considered). As many as three mismatches were allowed. Reads that matched multiple loci were removed from further analysis and the resultant alignment files were processed to generate "pile-ups" against each chromosome. The numbers of total reads and "mappable" reads assigned to single loci (~90% of total) for each sample are listed in Fig. S1.

Confirmation of *rpb1-CTD* Sample Purity. The eight strains subjected to multiplex sequencing of bar-coded cDNA libraries differed with respect to the DNA sequence encoding the carboxyl-terminal domain (CTD) segment of Rpb1. As an internal control to gauge the "purity" of the strain-specific reads, the WT and mutant CTD-coding sequences were compiled into a single database against which each of the samples were aligned using Bowtie, but without tolerating mismatches. The numbers of "correct" and total CTD reads for each sample are listed in Fig. S1.

Exon–Exon Junctions. Searches were performed against the genome sequence combined with a dataset of known exon–exon junctions as defined by Ensembl *S. pombe* release 13. To ensure that a 51-base read mapped to a splice junction, only the last 45 bases of the first exon and the first 45 bases of the second exon were considered (if the exon length exceeded 45 nt). In this way, reads that overlapped a junction by less than 6 nt were excluded. Reads that matched to more than one junction or elsewhere in the genome were also discarded.

Defining Intronic, Antisense, and Intergenic Regions. The known annotated set of *S. pombe* loci ($N = 7{,}022$; Ensembl version 13) was used to define unambiguous antisense transcripts (i.e., those that exactly mirror known annotated transcripts without overlapping nearby genes), and unique accessions were accordingly assigned ("anti_xxx"; $n = 3{,}097$). By using this augmented annotation, intergenic regions were defined, i.e., regions between known annotated and antisense transcripts on each strand ("inter_xxx"; $n = 8810$ accessions were assigned). Unique accessions were also assigned to all known introns ($N = 5{,}361$); thus, in total, 24,290 regions were interrogated across the 16 samples. To allow accurate gauge of library size, read summaries were computed for all regions and analyzed by DESEq (2), but only the analyses of the 7,022 annotated protein-coding and noncoding RNA regions are reported here.

High-Throughput RNA Sequencing Expression Level. Normalized expression levels (E) for individual exons and introns were calculated by using the formula $E = \log_2[(C \times R_i)/(T_i \times L)]$ as described previously (3) to obtain reads per kilobase per million reads (RPKM) values. In brief, the number of reads (R) detected across a given region in a given sample (i) was multiplied by a constant ($C = 1 \times 10^9$) and divided by the total number of reads in that sample (Ti) multiplied by the region's length (L). A small constant (10^{-5}) was added to all expression values to avoid taking logs of zero. Gene expression level values were determined by using exon data. Reproducibility of expression levels in nonsimultaneous biological replicates for all 7,022 annotated genes is reported in Fig. S2, with correlation coefficient values of ≥0.975 in every case.

Sample-specific expression levels for all annotated loci interrogated in this study are provided in Dataset S1. Fold changes and differential expression between samples were determined using the DESeq Bioconductor package (1). A cutoff of plus or minus twofold change and corrected P value <0.05 were applied to derive lists of differentially expressed loci.

Reverse Transcriptase Quantitative PCR Analysis. Total RNA was prepared from exponentially growing cells (two independent cultures for each yeast strain analyzed) with a Qiagen RNeasy kit. The RNAs were treated with DNase I, extracted serially with phenol:chloroform and chloroform, and then precipitated with ethanol. The RNAs were resuspended in 10 mM Tris·HCl (pH 6.8) and 1 mM EDTA and adjusted to a concentration of 600 ng/μL. Reverse transcription was performed with 2 μg of this RNA template plus oligo$(dT)_{18}$ and random hexamer primers by using the Maxima First Strand cDNA synthesis kit (Thermo Scientific). After cDNA synthesis for 30 min at 55 °C, the reverse transcription reaction mixtures were diluted 10-fold with water. Aliquots (2 μL) were used as templates for gene-specific quantitative PCR (qPCR) reactions directed by the sense and antisense primers listed in Table S1. The qPCR reactions were constituted with the Maxima SYBR Green/ROX master mix (Thermo Scientific) and monitored with an ABI Prism 7900HT Sequence Detection System. The qPCR reactions were performed in triplicate for each cDNA population. The level of individual cDNAs was calculated relative to that of *act1* cDNA by the comparative Ct method (4).

Primer Extension Analysis. Total RNA was prepared from exponentially growing cells (two independent cultures for each yeast strain analyzed) with a Qiagen RNeasy kit. The RNAs were treated with DNase I, extracted serially with phenol:chloroform and chloroform, precipitated with ethanol, and resuspended in 10 mM Tris·HCl (pH 6.8) and 1 mM EDTA and adjusted to a concentration of 1 μg/μL. Aliquots (10 μg) of this RNA were used as templates for M-MuLV reverse transcriptase-catalyzed extension of ^{32}P-5′-end labeled oligodeoxynucleotide primers complementary to the *pho1* ORF (nucleotides +33 to +51) or *act1* ORF (nucleotides +55 to +73). The primer extension reactions were performed as described previously (5) and analyzed by electrophoresis of the reaction mixtures through a 22-cm 8% (wt/vol) polyacrylamide gel containing 7 M urea in 80 mM Tris·borate, 1 mM EDTA. The ^{32}P-labeled primer extension products were visualized by autoradiography of the dried gel and quantified by scanning the gel with a Typhoon PhosphorImager (GE Healthcare). The *pho1* primer extension signals were normalized to those of *act1* for each RNA sample.

Primer Extension Mapping of the 5′ End of *pho1* mRNA at Single-Nucleotide Resolution. To precisely map the 5′ end of the *pho1* mRNA, we analyzed the reverse transcriptase primer extension product in parallel with a series of DNA-directed primer extension reactions that contained mixtures of standard and chain-terminating nucleotides. The DNA template fragment containing the first 51 nt of the *pho1* ORF and 1,150 nt of 5′ flanking genomic DNA was annealed to the ^{32}P-5′-end labeled primer complementary to the *pho1* ORF (nucleotides +33 to +51). Primer extension by Therminator DNA polymerase (NEB) in reaction mixtures (50 μL) containing 50 μM dNTPs and 2 μM of acyGTP, acyATP acyTTP, or acyCTP (NEB) were performed

according to the NEB protocol. The primer extension products were analyzed by electrophoresis through a 42-cm denaturing 8% polyacrylamide gel and visualized by autoradiography of the dried gel.

1. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10(3):R25.
2. Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11(10):R106.
3. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5(7):621–628.
4. Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative $C_{(T)}$ method. *Nat Protoc* 3(6):1101–1108.
5. Schwer B, Mao X, Shuman S (1998) Accelerated mRNA decay in conditional mutants of yeast mRNA capping enzyme. *Nucleic Acids Res* 26(9):2050–2057.

Sample	Total Reads	Mappable Reads	CTD Reads (correct / total)	Outliers
WT 1	23891119	21609959	1396 / 1400	1 S2A, 1 T4A, 2 S7A
WT 2	24553450	22284397	3436 / 3440	2 S2A, 1 T4A, 1 S7A
Y1F 1	24008502	21695617	1840 / 1842	1 S2A, 1 S7A
Y1F 2	26192539	23565149	3741 / 3745	3 S2A, 1 S2A•S7A
S2A 1	28679106	26157568	484 / 484	
S2A 2	24421156	21839717	429 / 429	
T4A 1	33318692	30146189	1548 / 1551	1 S2A, 2 Y1F•S7A
T4A 2	23966944	21507147	1206 / 1206	
S7A 1	29902020	27041449	1096 / 1098	1 T4A, 1 S2A•S7A
S7A 2	23368553	21079905	1432 / 1434	1 WT, 1 Y1F
Y1F•S7A 1	33937891	30754787	2044 / 2045	1 T4A•S7A
Y1F•S7A 2	22428372	20264796	2840 / 2840	
S2A•S7A 1	30496499	27442991	1654 / 1661	3 S2A, 2 T4A, 1 S7A, 1 Y1F•S7A
S2A•S7A 2	20978819	19175660	1338 / 1341	1 S2A, 1 T4A, 1 S7A
T4A•S7A 1	34979892	31586495	668 / 668	
T4A•S7A 2	22779273	20732602	1136 / 1136	

Fig. S1. Total and uniquely mappable read counts for high-throughput RNA sequencing analysis and confirmation of sample purity via *rpb1-CTD* sequence reads. Biological replicates 1 and 2 are listed separately for each *rpb1-CTD* strain. The numbers of "correct" and total CTD-coding reads for each sample are listed, along with the number and identity of the outlier CTD reads.

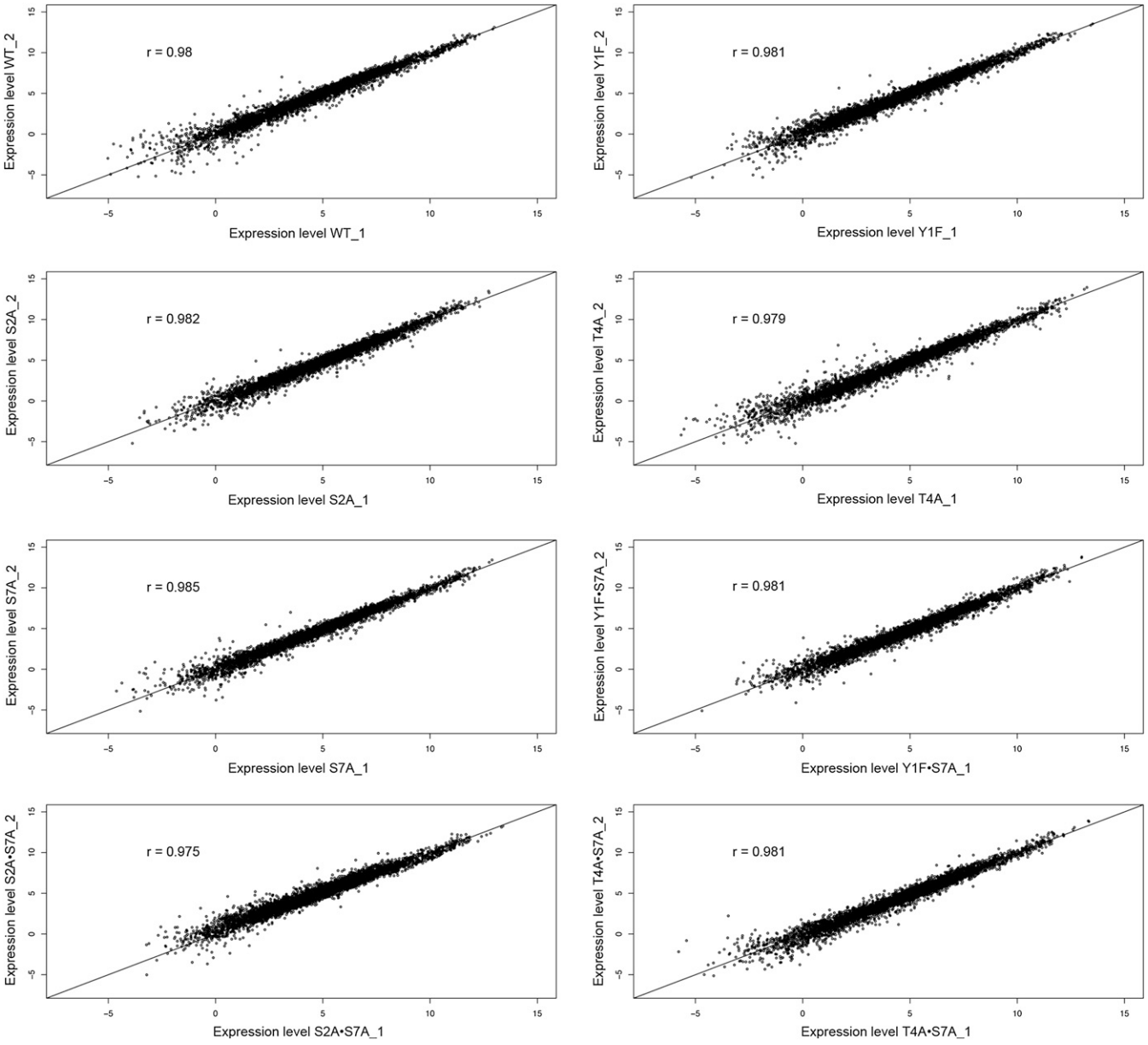

Fig. S2. Reproducibility of expression levels in nonsimultaneous biological replicates. Each point represents an annotated gene (7,022 in total; Pearson correlation coefficient *r*). Expression levels for each gene in each sample were calculated as $\log_2$ of the number of sequencing reads starting within that locus, normalized both by the length of the region and the total number of sequence reads within the sample (i.e., as RPKM).

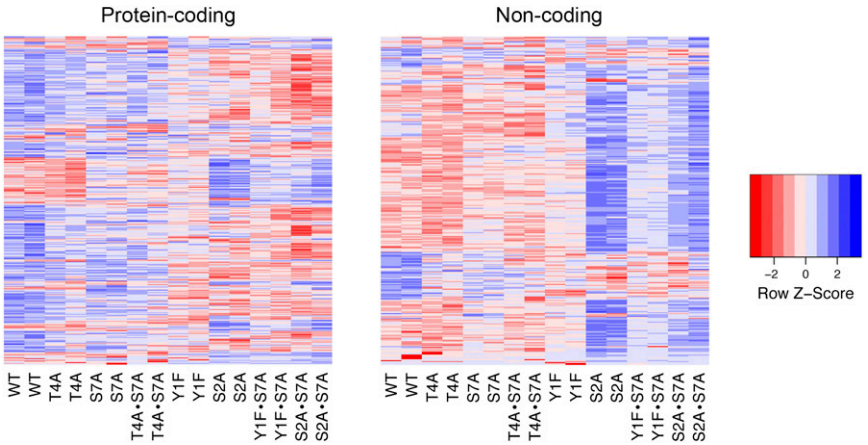

Fig. S3. Protein-coding and noncoding transcripts exhibit distinctive expression profiles in the transcriptomes of CTD mutants. Heat maps showing the expression profiles of 421 significantly changing protein-coding RNAs (*Left*) and 349 significantly changing noncoding RNAs (*Right*) in at least one mutant compared with WT. Expression levels for each locus in each sample were calculated as $\log_2$ of the number of sequencing reads starting within that locus, normalized both by the length of the region and the total number of sequence reads within the sample (i.e., RPKM). Maximum distance between rows was determined using the "dist" function in R (method "maximum") followed by hierarchical clustering using the "hclust" function. Row *z*-score: RPKM values in each row were scaled by subtracting the mean of the row from each value, followed by division of the resultant values by the SD of the row, i.e., (RPKM value – row mean)/row SD.

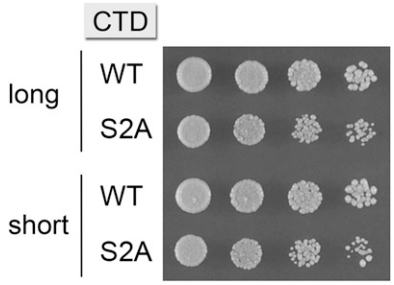

Fig. S4. Growth of fission yeast *rpb1-S2A* "short" and "long" variants. Isogenic strains harbored the indicated *rpb1-CTD* alleles at the chromosomal *rpb1* locus flanked by a 3′ *natMX* marker. The short versions of WT and S2A CTDs consist of the "rump" plus 14 consensus WT heptads or S2A heptads. The long WT is the native WT CTD consisting of the rump plus 25 heptads. The long version of the S2A CTD consists of an S2A mutant rump plus 25 S2A heptads. This gene segment was PCR-amplified from genomic DNA of the $S2A_{29}$ strain obtained from Damien Hermand (1) and inserted into our *rpb1-CTD-natMX* integration cassette, which was then used for allelic replacement of *rpb1* in a diploid, followed by sporulation and selection for nourseothricin-resistant "S2A-long" haploid progeny. (The S2A-long genotype was confirmed by diagnostic Southern blotting and sequencing of the *rpb1-CTD-S2A* locus.) The indicated strains were grown in YES medium at 30 °C. Exponentially growing cultures were adjusted to A_{600} of 0.1, and aliquots (3 μL) of serial fivefold dilutions were spotted to YES agar. The plate was photographed after incubation for 3 d at 30 °C.

1. Cassart C, Drogat J, Migeot V, Hermand D (2012) Distinct requirement of RNA polymerase II CTD phosphorylations in budding and fission yeast. *Transcription* 3(5):231–234.

Gene	Function	[log_2 fold change]	
		Y1F	S2A
SPBC1683.04	glycosyl hydrolase family 3	-4.64	–2.06
Pfl7	DIPSY cell surface glycoprotein	-4.44	–2.42
SPAC11D3.11c	Zn finger transcription factor	–3.13	–1.36
SPAC11D3.10	cysteine desulfurase	–3.06	–1.28
Mal1	maltase alpha-glucosidase	–2.65	–1.64
SPBC1683.06c	uridine ribohydrolase	–2.55	–1.64
Urg2	uracil phosphoribosyltransferase	–2.23	–2.45
SPAC11D3.14c	5-oxoprolinase	–2.22	–1.17
SPBC1773.12	Zn finger transcription factor	–2.12	–1.35
Psf3	GINS complex subunit	–1.88	–2.15
Atd2	aldehyde dehydrogenase	–1.80	–1.06
SPBC1703.08c	5-formyltetrahydrofolate cyclo-ligase	–1.80	–1.78
Gep4	phosphatidylglycerophosphatase	–1.79	–2.39
SPAC1039.08	serine O-acetyltransferase	–1.71	–2.36
SPBC1683.02	adenine/adenosine deaminase	–1.70	–1.31
SPCC1884.01	–	–1.69	–1.17
SPAC1399.04c	uracil phosphoribosyltransferase	–1.60	–1.68
SPBC1709.14	peptide N-glycanase	–1.59	–1.68
Ght4	hexose transporter	–1.57	–1.16
Set7	histone lysine methyltransferase	–1.47	–1.78
Jmj4	histone demethylase	–1.35	–1.31
Set10	ribosomal lysine methyltransferase	–1.35	–1.10
SPBC418.02	N-acetyltransferase subunit	–1.24	–1.65
SPAC18G6.12c	–	–1.21	–1.40
Klf1	Zn finger transcription factor	–1.18	–1.58
Tpr1	Paf1 complex subunit	–1.17	–1.36
SPBC8E4.01c	inorganic phosphate transporter	–1.17	–1.41
SPBC1683.05	uracil/uridine permease	–1.16	–1.27
Rtt109	histone lysine acetyltransferase	–1.14	–1.08
SPAC1039.11c	alpha-glucosidase	–1.13	–1.26
Aah2	alpha-amylase	–1.07	–1.39
SPAC977.17	transmembrane water channel	–1.03	–1.68

Gene	Function	[log_2 fold change]	
		Y1F	S2A
Str3	siderophore-iron transporter	6.61	3.36
Frp1	ferric-chelate reductase	4.01	1.44
Fip1	iron permease	3.93	2.50
Fio1	iron transport multicopper oxidase	3.83	3.15
Str1	siderophore-iron transporter	3.31	1.53
Srx1	sulfiredoxin	2.03	1.19
Ppk31	serine/threonine protein kinase; meiotically upregulated gene	1.95	2.34
Mde5	alpha amylase homolog; meiotically upregulated gene	1.73	3.23
Mug136	acetylglucosaminyltransferase; meiotically upregulated gene	1.54	1.43
Meu14	sporulation specific PIL domain protein	1.48	2.42

Fig. S5. RNAs dysregulated by *Y1F* are coordinately dysregulated by *S2A*. The transcripts that were coordinately down-regulated (*Upper*) and up-regulated (*Lower*) in *Y1F* and *S2A* cells are listed in descending order of the magnitude of the decrease or increase in expression level in *Y1F* vs. WT cells.

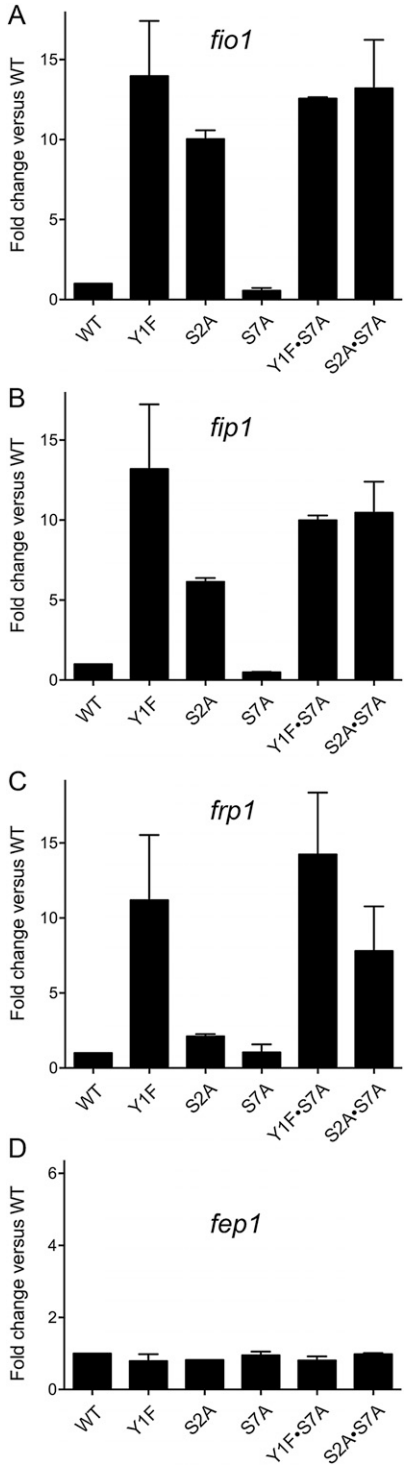

Fig. S6. RT-qPCR analysis of the effects of CTD mutations on the expression of iron homeostasis genes. The analysis of mRNA levels for *fio1* (*A*), *fip1* (*B*), *frp1* (*C*), and *fep1* (*D*) in *rpb1-CTD WT*, *Y1F*, *S2A*, *S7A*, *Y1F•S7A*, and *S2A•S7A* yeast strains was performed as described in *SI Methods*. The level of each transcript was normalized to that of *act1* measured for the same RNA sample. The bar graph shows fold change in mRNA levels for the *Y1F, S2A, S7A, Y1F•S7A*, and *S2A•S7A* mutants relative to the WT control strain (defined as 1.0). Each datum in the bar graph is the average of values from RT-PCR analyses of RNAs from two independent yeast cultures; the error bars denote the range.

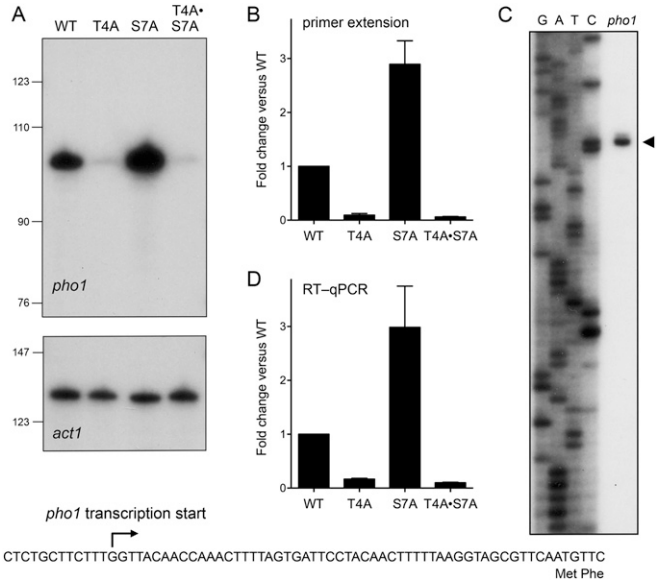

Fig. S7. Effect of *T4A* and *S7A* on *pho1* expression. (*A*) Primer extension analyses were performed as described in *SI Methods*. The ^{32}P-labeled oligonucleotide primers complementary to *pho1* mRNA (nucleotides +33 to +51 of the ORF) or *act1* RNA (nucleotides +55 to +73 of the ORF) were annealed to total RNA from the indicated *rpb1-CTD* strains and extended with reverse transcriptase. The reaction products were analyzed by denaturing PAGE and visualized by autoradiography. The sizes (in nucleotides) of ^{32}P-labeled marker DNAs are indicated (*Left*). (*B*) The *pho1* primer extension product was quantified and was normalized to that of *act1* measured for the same RNA sample. The bar graph shows the fold change in *pho1* for the *T4A*, *S7A*, and *T4A•S7A* mutants relative to the WT control (defined as 1.0). Each datum in the bar graph is the average of values from primer extension analyses of RNAs from two independent yeast cultures; the error bars denote the range. (*C*) Mapping of the *pho1* transcription start site. The *pho1* primer extension product was analyzed in parallel with a chain-terminated sequencing ladder generated by DNA polymerase-catalyzed extension of the same ^{32}P-labeled primer annealed to a *pho1* DNA template. An autoradiogram of the dried gel is shown. The sense strand sequence of the *pho1* locus upstream of the translation start site is shown at bottom; the transcription start site at the G position 51 nt upstream of the AUG start codon is denoted by the arrowhead. (*D*) RT-qPCR analysis of mRNA levels for *pho1* was performed as described in *SI Methods*. The *pho1* transcript level was normalized to that of *act1* measured for the same RNA sample. The bar graph shows the fold change in *pho1* mRNA levels for the *T4A*, *S7A*, and *T4A•S7A* mutants relative to WT control (defined as 1.0). Each datum in the bar graph is the average of values from RT-qPCR analyses of RNAs from two independent yeast cultures; the error bars denote the range.

		RPKM: expt. 1 / 2 [$\log_2$ fold change]				
		WT	**S2A**	**T4A**	**S7A**	**T4A•S7A**
Pho1	Acid phosphatase	9.08 / 8.87	8.26 / 8.15 [–1.06]	6.12 / 5.62 [–2.91]	10.39 / 10.36 [1.41]	5.46 / 5.18 [–3.57]
SPBC8E4.01c	Phosphate transporter	8.58 / 8.28	7.54 / 7.06 [–1.42]	6.63 / 5.73 [–2.02]	9.30 / 8.82 0.65	6.49 / 5.69 [–2.22]
Pho7	transcription activator	5.56 / 6.00	6.03 / 6.12 0.004	5.93 / 6.16 0.43	5.59 / 5.76 –0.06	5.97 / 5.97 0.27
Csk1	repressor of *pho1*	5.31 / 5.42	4.88 / 5.08 –0.65	5.63 / 5.62 0.44	4.92 / 5.20 –0.29	5.61 / 5.60 0.33
SPBC1271.09	glycerophosphodiester transporter	4.04 / 3.47	4.27 / 3.93 0.03	3.40 / 2.80 –0.48	5.79 / 5.85 [2.04]	3.32 / 3.09 –0.49
Pho4	thiamine-repressible acid phosphatase	6.85 / 6.68	6.70 / 6.60 –0.40	6.86 / 6.61 0.15	6.77 / 6.49 –0.13	7.00 / 6.53 0.09

Fig. S8. CTD mutations affect expression of phosphate homeostasis genes. The expression levels (RPKM) of genes involved in phosphate homeostasis are shown for the biological replicates of each of the strains specified. The $\log_2$ fold changes vs. WT are denoted below the expression levels. Values meeting the dual cutoff criteria of at least twofold change and $P < 0.05$ are highlighted in brackets. Other $\log_2$ fold change values not meeting either criterion are in plain font. The acid phosphatase and phosphate transporter genes affected by CTD mutations are listed above the spacer row shaded gray. Other phosphate acquisition genes and the transcription factors that control the phosphate response, which were not affected by CTD mutations, are listed below the spacer row.

meiotic genes upregulated by *S7A*		[log_2 fold change]
Mfr1	fizzy-related protein	3.86
Mde3	serine/threonine kinase	3.82
Mde7	RNA binding protein	3.21
Spn6	septin	2.85
Ght5	hexose transporter	2.74
Spn5	septin	2.45
Mug130		2.42
Ppk31	serine/threonine kinase	2.27
Spo6	Spo4-Spo6 kinase regulatory subunit	2.20
Mei4	forkhead transcription factor	1.74
Rec8	meiotic cohesin complex subunit	1.55
Mde1		1.48
Mei2	RNA binding protein	1.44
Mug136	acetylglucosaminyltransferase	1.41
Mug147		1.32
SPBC725.10	tspO homolog	1.13

Fig. S9. Meiotic genes up-regulated by *S7A*. The affected transcripts and their protein products are listed in descending order of magnitude of the increased expression level in *S7A* vs. WT cells.

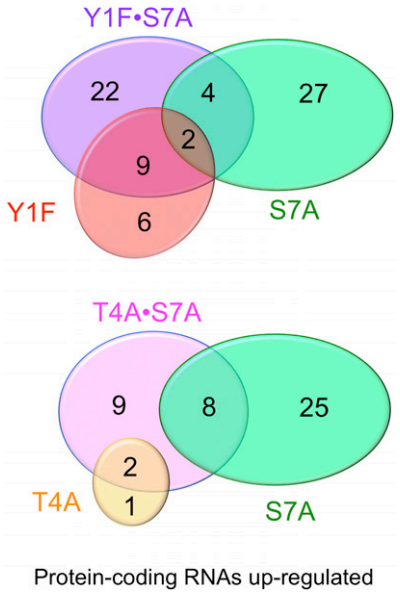

Fig. S10. *Y1F* and *T4A* attenuate the up-regulation of gene expression by *S7A*. Three-way Venn diagrams comparing the numbers of protein-coding genes significantly up-regulated at least twofold in the *Y1F* and *S7A* single mutants vs. the *Y1F•S7A* double mutant (*Upper*) and the *T4A* and *S7A* single mutants vs. the *T4A•S7A* double mutant (*Lower*).

Table S1. Oligonucleotide primers used for qPCR analyses

Gene	Strand	Sequence
act1	Sense	5′-AAGTACCCCATTGAGCACGG
	Antisense	5′-CAGTCAACAAGCAAGGGTGC
fep1	Sense	5′-CTGTTGGCGAGTCAGTCTGT
	Antisense	5′-ATTTGAGGGTGGCTCTTCGG
fio1	Sense	5′-AGTGTTGGAGCGTGGAGAAG
	Antisense	5′-CTGGGTATCATCTCCGGTGC
fip1	Sense	5′-GTGAGACCCTTGAGGCTTCC
	Antisense	5′-CGATAAAGCCACCGCCAATG
frp1	Sense	5′-TTACGGCCCAGTCTCAAACC
	Antisense	5′-AAAGCCGAAGAACGAGCAGA
pho1	Sense	5′-GCAGCGCGTTGTAGATTCTG
	Antisense	5′-TCGAGAGCAGGGGTAGTGAA

Other Supporting Information Files

Dataset S1 (XLS)