

From Structure to Systems: High-Resolution, Quantitative Genetic Analysis of RNA Polymerase II

Hannes Braberg,^{1,4} Huiyan Jin,⁵ Erica A. Moehle,² Yujia A. Chan,⁶ Shuyi Wang,^{1,4} Michael Shales,^{1,4} Joris J. Benschop,⁷ John H. Morris,³ Chenxi Qiu,⁵ Fuqu Hu,⁵ Leung K. Tang,⁵ James S. Fraser,^{1,4} Frank C.P. Holstege,⁷ Philip Hieter,⁶ Christine Guthrie,^{2,*} Craig D. Kaplan,^{5,*} and Nevan J. Krogan^{1,4,8,*}

¹Department of Cellular and Molecular Pharmacology

²Department of Biochemistry and Biophysics

³Resource on Biocomputing, Visualization, and Informatics

University of California, San Francisco, San Francisco, CA 94158, USA

⁴California Institute for Quantitative Biosciences, QB3, San Francisco, CA 94158, USA

⁵Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, USA

⁶Departments of Medical Genetics and Biochemistry, Michael Smith Laboratories, University of British Columbia, Vancouver V6T 1Z4, Canada

⁷Molecular Cancer Research, University Medical Centre Utrecht, 3508 AB Utrecht, the Netherlands

⁸J. David Gladstone Institutes, San Francisco, CA, 94158 USA

*Correspondence: christineguthrie@gmail.com (C.G.), cdkaplan@tamu.edu (C.D.K.), nevan.krogan@ucsf.edu (N.J.K.)

<http://dx.doi.org/10.1016/j.cell.2013.07.033>

SUMMARY

RNA polymerase II (RNAPII) lies at the core of dynamic control of gene expression. Using 53 RNAPII point mutants, we generated a point mutant epistatic miniarray profile (pE-MAP) comprising ~60,000 quantitative genetic interactions in *Saccharomyces cerevisiae*. This analysis enabled functional assignment of RNAPII subdomains and uncovered connections between individual regions and other protein complexes. Using splicing microarrays and mutants that alter elongation rates in vitro, we found an inverse relationship between RNAPII speed and in vivo splicing efficiency. Furthermore, the pE-MAP classified fast and slow mutants that favor upstream and downstream start site selection, respectively. The striking coordination of polymerization rate with transcription initiation and splicing suggests that transcription rate is tuned to regulate multiple gene expression steps. The pE-MAP approach provides a powerful strategy to understand other multifunctional machines at amino acid resolution.

INTRODUCTION

Alterations within a genome often cause specific as well as global phenotypic changes to a cell. Combining two alterations in the same cell allows for measurement of the genetic interaction between them: negative genetic interactions (synthetic sick/lethal) arise when two mutations in combination cause a stronger growth defect than expected from the single mutations. This is often observed for factors participating in redundant path-

ways or as nonessential components of the same essential complex. In contrast, positive interactions occur when the double mutant is either no sicker (epistatic) or healthier (suppressive) than the sickest single mutant (Beltrao et al., 2010) and may indicate that the factors are components of a nonessential complex and/or that the factors perform antagonizing roles in the cell. However, single genetic interactions are often difficult to interpret in isolation; an interaction pattern for a given mutation can be more informative, as it reports on the phenotype in a large number of mutant backgrounds (Schuldiner et al., 2005; Tong et al., 2004). These genetic profiles provide highly specific readouts that can be used to identify genes that are functionally related (Beltrao et al., 2010).

One of the first organisms to be genetically interrogated on a large scale was *Saccharomyces cerevisiae* (*S. cerevisiae*), in which nonquantitative genetic interaction data could be collected using the SGA (synthetic genetic array) (Tong et al., 2004) or dSLAM (heterozygous diploid-based synthetic lethality analysis on microarrays) (Pan et al., 2004) approaches. We developed a technique termed epistatic miniarray profile (E-MAP) (Collins et al., 2010; Schuldiner et al., 2005; Schuldiner et al., 2006), which utilizes the SGA methodology and allows for the quantitative collection of genetic interaction data on functionally related subsets of genes, including those involved in chromatin regulation (Collins et al., 2007b), RNA processing (Wilmes et al., 2008), signaling (Fiedler et al., 2009), or plasma membrane function (Aguilar et al., 2010). However, the vast majority of systematic genetic screening interrogates deletions of nonessential genes or hypomorphic knockdown alleles of essential genes. Because many genes, especially essential ones, are multifunctional, these methods perturb all activities associated with a given gene product.

Here, we describe an important advance of the E-MAP approach, which allows us to address higher levels of complexity

by examining the genetic interaction space of point mutant alleles of multifunctional genes in a technique that we term point mutant E-MAP (pE-MAP). This method greatly increases the resolution achievable by gene function analysis, as it allows assignment of specific genetic relationships to individual residues and domains. In this study, we have used the pE-MAP approach to functionally dissect RNAPII using alteration-of-function alleles in five different subunits of the enzyme. Using the genetic data, we assign transcriptional activity and specific functions to different residues and regions of RNAPII. By examining the relationship between transcription rate and genetic interaction partners, transcription-rate-sensitive factors were revealed. Through the characterization of multiple stages of gene regulation, including start site selection, transcription elongation rate, and mRNA splicing, the pE-MAP technique has provided both global and specific insight into structure-function relationships of RNAPII. We propose this strategy as a useful paradigm for the high-resolution interrogation of any multifunctional protein.

RESULTS

A Set of Alleles for the Functional Dissection of RNAPII

To identify residues that are important for transcriptional regulation *in vivo*, we isolated RNAPII alleles that confer one or more of the following transcription-related phenotypes: suppression of galactose sensitivity in *gal10Δ56* (Gal^R) (Greger and Proudfoot, 1998; Kaplan et al., 2005), the Spt^- phenotype (Winston and Sudarsanam, 1998), or mycophenolic acid (MPA) sensitivity (Shaw and Reines, 2000) (for additional details, see Figure 1A and the Experimental Procedures). Each of these phenotypes relates to a gene-specific transcription defect that can be monitored using plate assays (Figure 1A). Random mutagenesis by PCR was carried out on the entire coding regions of RNAPII subunit genes *RPB2*, *RPB3*, *RPB7*, and *RPB11* and most of *RPO211*/*RPB1* (Experimental Procedures). These genes encode the essential subunits that are unique to RNAPII (Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12 are shared with RNAPI and RNAPIII, and Rpb4 and Rpb9 are nonessential) (Archambault and Friesen, 1993). In total, 53 single point mutants were identified that exhibit at least one of these phenotypes (Kaplan et al., 2012) (Figure 1B and Figure S1 and Table S1 available online).

Analysis of the distribution of phenotypes relative to the RNAPII structure suggested that our alleles might be diverse in their functions. Gal^R and MPA-sensitive mutations were broadly distributed, whereas those with the Spt^- phenotype were less common and more localized (Figure 1B and Table S1). The screens identified mutations in highly conserved residues and structural domains known to be important for RNAPII activity, including the Rpb1 trigger loop, the Rpb1 bridge helix, and the Rpb2 lobe and protrusion (Cramer et al., 2001; Gnatt et al., 2001; Kaplan, 2013) (Table S1). Quantitatively measuring the genetic interactions of specific residues might provide insight into the functions of these RNAPII regions and could therefore also identify protein-protein interaction interfaces.

An RNAPII Point Mutant Epistatic Miniaarray Profile

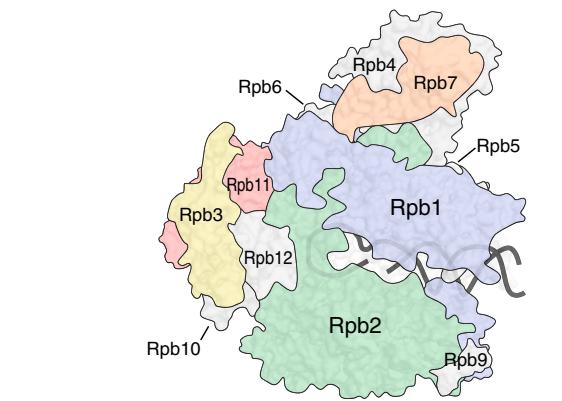
The 53 RNAPII point mutants (Figure 1B) were crossed against ~1,200 deletion and DAmP (decreased abundance by mRNA

perturbation) alleles (Schuldiner et al., 2005) (Table S1), which represent all major biological processes. Thus, a quantitative pE-MAP comprising ~60,000 double mutants was created (Table S2; <http://interactome-cmp.ucsf.edu>). Two-dimensional hierarchical clustering of these data effectively grouped together genes from known complexes and pathways based on their interactions with the point mutants (Figure S2 and Data S1). Previous studies have demonstrated that genes encoding proteins that physically associate often have similar genetic interaction profiles (Collins et al., 2007b; Roguev et al., 2008). The data derived from the point mutants could differ in this respect, as it is based on only five subunits of a single molecular machine. A receiver operating characteristic (ROC) curve was therefore generated to measure how well the genetic profiles of the deletion and DAmP mutants in the pE-MAP predict known physical interactions between their encoded proteins (Experimental Procedures). It was found that the predictive power of the pE-MAP is similar to that of a previously published E-MAP (Collins et al., 2007b), indicating that the genetic interactions of the RNAPII point mutants report on connections among virtually all cellular processes (Figures 2A and S2 and Data S1).

Next, to examine whether the spatial location of a mutated residue is a determinant of its function, we compared the similarity of pairs of RNAPII genetic profiles to the three-dimensional distance between the mutated residues (Wang et al., 2006). We observed a strong correlation ($r = -0.37$, $p < 10^{-22}$) (Table S2), and the trend is significant both for residue pairs residing in the same subunit ($r = -0.25$, $p < 10^{-5}$) and for those in different subunits ($r = -0.28$, $p < 10^{-6}$) (Figure 2B). This suggests that structural proximity correlates with functional similarity and that high-resolution genetic interaction profiling could provide information for targets whose structures have not yet been determined.

Comparison of Genetic and Gene Expression Profiles Derived from the RNAPII Alleles

To determine whether any given genetic interaction might result from the point mutation affecting the expression of the corresponding gene, we subjected 26 of the RNAPII mutants to genome-wide gene expression analysis (Table S2; GEO accession number: GSE47429). We found no correlation ($r = -0.003$) between an RNAPII mutant's genetic interaction score with a gene deletion or DAmP allele and the expression change of that gene due to the RNAPII mutation (Figure 2C). Therefore, connections must be due to more complex relationships between the mutated residues and the library genes. Nonetheless, these data sets allowed us to test whether the clustering of the RNAPII mutants in the pE-MAP (Figure S2 and Data S1) could be recapitulated using their gene expression profiles. We thus assessed pair-wise RNAPII mutant similarity based on genetic and gene expression profiles separately and found that these two measures are highly correlated ($r = 0.71$; Figure 2D). Therefore, these orthogonal data sets provide a common biological framework for functionally organizing the RNAPII mutants, allowing us to study the underlying biology behind these mutants and their phenotypes.

A

Separate mutagenesis of
Rpb1, Rpb2, Rpb3, Rpb7 and Rpb11

Phenotype screening

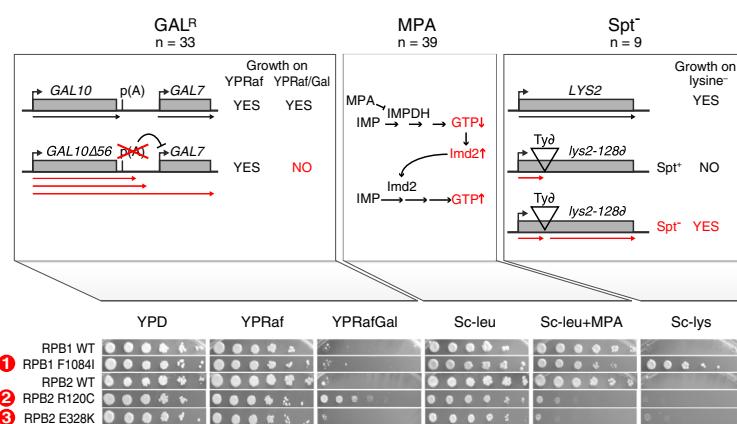
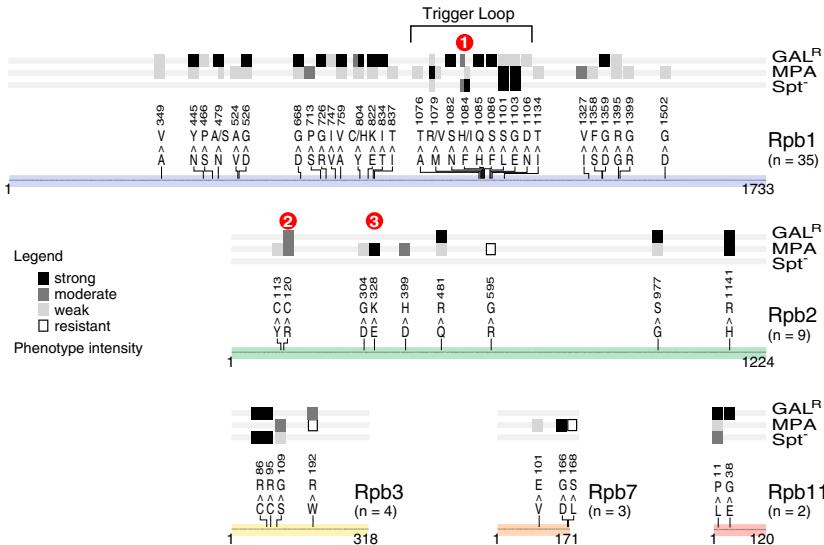
**B**

Figure 1. Generation and Selection of RNAPII Point Mutants

(A) RNAPII point mutants were screened for three transcription-related phenotypes. (Gal^{R} , left) Deletion of the major *GAL10* p(A) site (*gal10Δ56*) results in RNAPII readthrough and interference with initiation at *GAL7*, causing a Gal-sensitive phenotype. Gal^{R} mutants increase *GAL10* 3' end formation/termination, thereby rescuing *GAL7* expression. (MPA, middle) Mycophenolic acid (MPA) inhibits IMP-dehydrogenase (IMPDH)-dependent GTP synthesis but is counteracted by upregulation of an MPA-resistant form of IMPDH, *IMD2*. Transcriptional defects that are sensitive to low GTP levels or reduce *IMD2* expression render cells sensitive to MPA. (Spt^+ , right) Insertion of a Ty retrotransposon into *LYS2* (*lys2-128δ*) results in a lysine auxotrophy due to transcription block. Certain mutants suppress *lys2-128δ* and allow expression of *LYS2* due to activation of an internal promoter. Spot tests to identify each phenotype for three representative mutants are displayed.

(B) Positions, mutations, and phenotypes of the 53 single point mutants analyzed in the pE-MAP. Colored lines represent subunit sequences, with mutations denoted by residue numbers and single letter amino acid codes for WT and mutant. See also Figure S1 and Table S1.

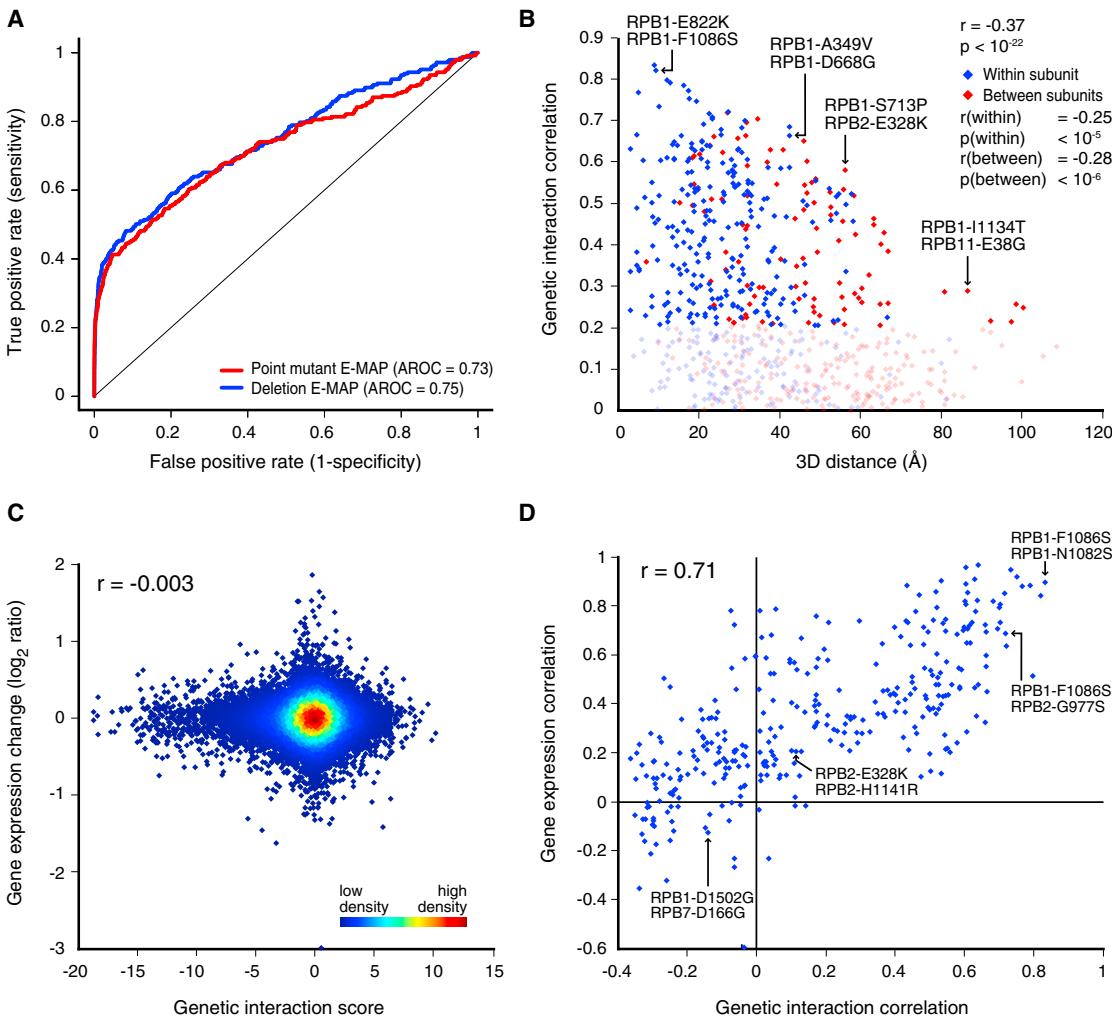


Figure 2. pE-MAP Interactions Span Numerous Biological Processes, Depend on Spatial Location of Mutated Residues, and Are Not Direct Consequences of Changes in Gene Expression

(A) ROC curves comparing the power of genetic profile correlations from the pE-MAP (red) and an E-MAP focused on chromosome biology (blue) to predict physical interactions between pairs of proteins (Collins et al., 2007a; Experimental Procedures). AROC, area under the curve.

(B) Genetic profile correlations between pairs of mutated residues compared to the three-dimensional distance between their α carbons. Blue points denote residue pairs within the same RNAPII subunit; red points represent pairs in different subunits. Negatively correlated residue pairs were excluded, as were four mutants of residues absent from the coordinate file (PDB ID: 2E2H) (Rpb1 D1502, Rpb7 V101, Rpb7 D166, and Rpb7 L168) (Wang et al., 2006). Correlations between 0 and 0.2 are dimmed to highlight trends at higher correlations.

(C) Effect of RNAPII point mutations on gene expression, compared to the corresponding genetic interaction scores between RNAPII mutants and deletion/ΔAmP alleles. All combinations of the 26 RNAPII mutants and 1,192 library genes/mutants that were examined via both pE-MAP and expression analyses are included.

No global changes in gene expression were observed (measured by spike-in control RNA; Experimental Procedures).

(D) Comparison between pairwise RNAPII mutant correlations of genetic interaction profiles and gene expression profiles.

See also Figure S2, Table S1, Table S2, and Data S1.

Functional Associations between RNAPII Residues and Protein Complexes

In an effort to link individual RNAPII mutations to specific cellular functions, data from the pE-MAP were compared to a published genetic interaction data set containing profiles from >4,000 genes (Costanzo et al., 2010). These genes were classified based on complex membership of their encoded proteins (Benschop et al., 2010), and Mann-Whitney U statistics were used to

identify RNAPII mutants with high profile similarity to members of specific complexes (Experimental Procedures, Figure 3A, and Table S3). We uncovered a number of connections, including several point mutants having similar genetic profiles to mutants of components of Mediator and the Rpd3C(L) histone deacetylase complex (Carrozza et al., 2005; Keogh et al., 2005). Unexpectedly, we also observed that several RNAPII mutants are significantly correlated genetically to kinetochore mutants

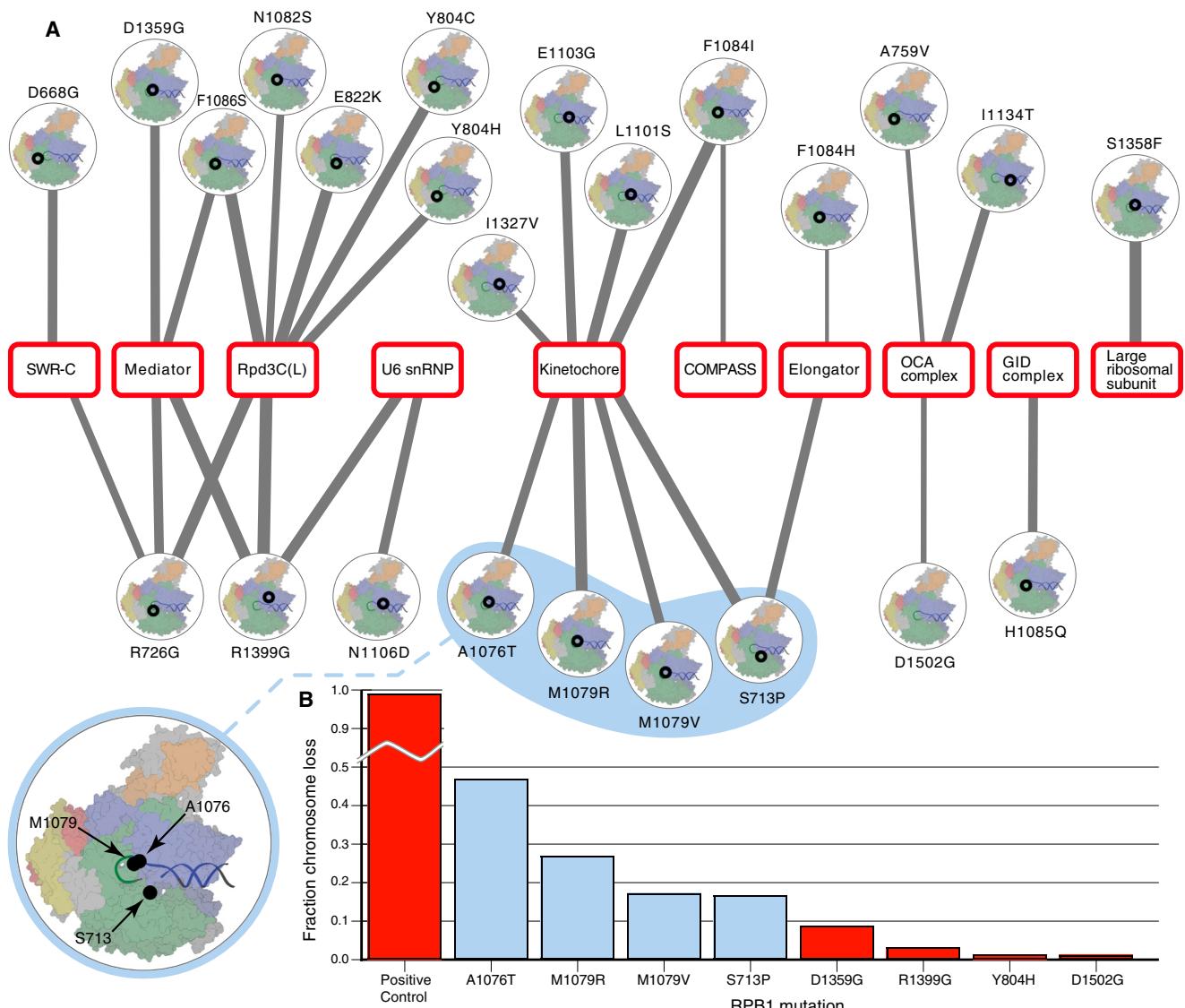


Figure 3. Comparison of the pE-MAP with Previously Collected Genetic Interaction Data Reveals Functional Associations between RNAPII Residues and Protein Complexes

(A) Module map depicting genetic similarity of RNAPII mutants with genes encoding the indicated protein complex subunits (Experimental Procedures). Edge widths correspond to statistical significance of connections. Only *RPB1* edges with a false discovery rate <0.1 are displayed. Four mutated residues linked to the kinetochore are highlighted in blue, and the blow-up indicates their structural locations.

(B) Nineteen mutants were examined using a chromosome transmission fidelity (CTF) assay. The four kinetochore-linked mutants highlighted in (A) exhibit chromosome loss in >15% of their colonies (blue bars), whereas unlinked mutants display no or weak phenotype (red bars, representative set). See also Figure S3 and Table S3.

(Figure 3A). We carried out chromosome transmission fidelity (CTF) assays on 19 of our RNAPII mutants, including those linked specifically to the kinetochore (Experimental Procedures and Table S3) (Spencer et al., 1990). Only four of the tested mutants exhibited chromosome loss in more than 15% of their colonies, and these were genetically linked to the kinetochore in our analysis and had similar genetic profiles to each other in our pE-MAP (Figures 3A and 3B). Recent studies indicate that a certain level of transcription by RNAPII at the centromere is required for

centromere function and high-fidelity chromosome segregation in budding yeast (Ohkuni and Kitagawa, 2011). Using specific constructs designed to ascertain centromere sensitivity to transcriptional readthrough, we did not observe any defects in kinetochore integrity in these RNAPII mutants (data not shown). Ultimately, further work will be required to understand the connection between these RNAPII point mutants and chromosome segregation. A full point mutant module map from alleles of all subunits is presented in Figure S3.

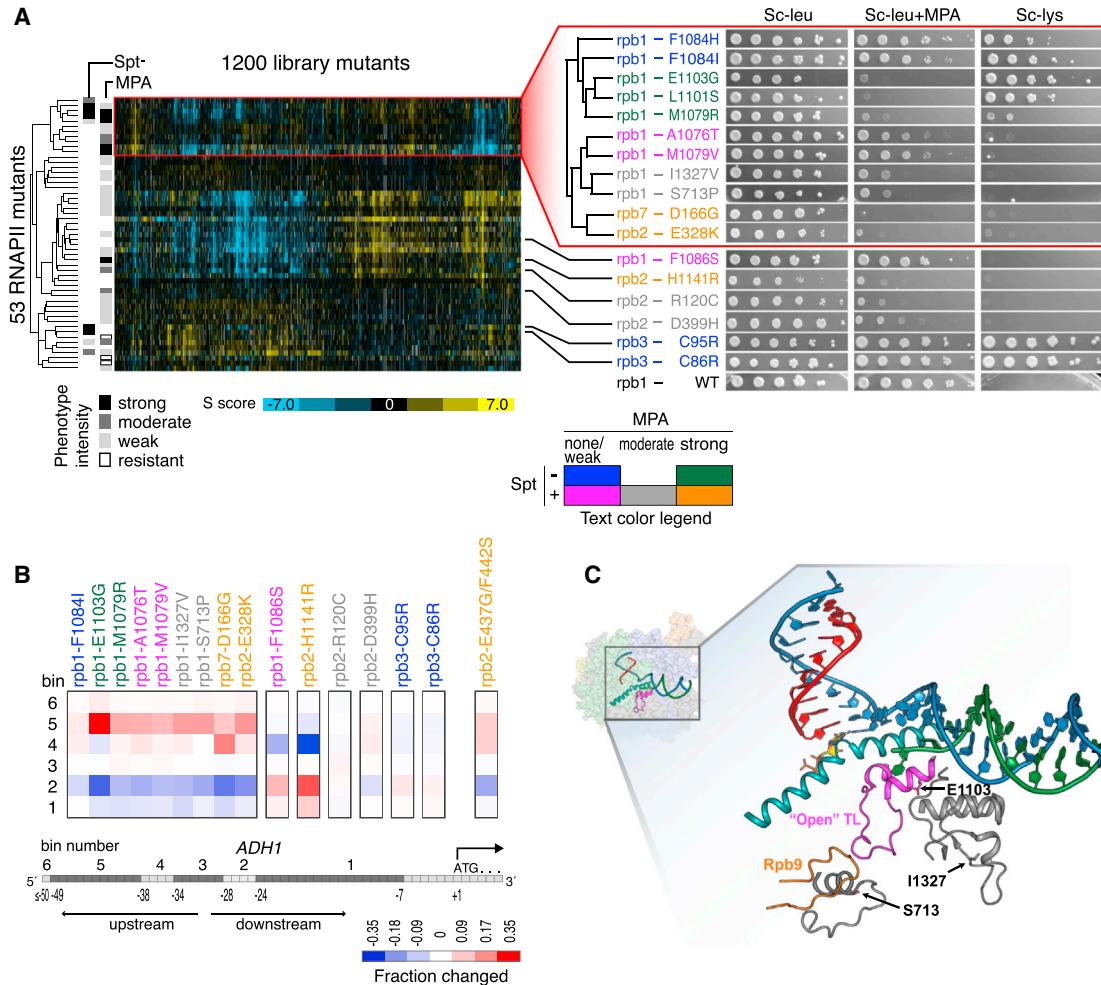


Figure 4. pE-MAP Profiles Differentiate between Subtle Changes in Transcription-Related Phenotypes and Identify RNAPII Mutations that Affect Start Site Selection

(A) pE-MAP clustering in relation to MPA and Spt⁻ phenotypes of alleles. The RNAPII alleles are clustered by pE-MAP profiles, and their colors indicate degrees of MPA and Spt⁻ phenotypes (determined from the spot tests).

(B) Effect of RNAPII mutations on start site selection at *ADH1* determined by primer extension analysis. The heatmap describes the fractional change of start site in each bin of the *ADH1* schematic (bottom).

(C) Rpb1 I1327 and Rpb1 S713 connect to the TL (magenta). Mutations in I1327 could affect the structural region of the TL (E1103) via a network of loops and helices in Rpb1 (gray), and S713 is close to the TL catalytic site in its open conformation, via a Rpb9 loop (orange). In particular, the proline substitution, S713P, could result in structural changes affecting the TL. Coordinates for TL, Rpb9, and the S713 loop are from PDB ID 1Y1V (Kettenberger et al., 2004), and all others are from 2E2H (Wang et al., 2006). The bridge helix is shown in cyan; template DNA, blue; nontemplate DNA, green; and RNA, red. The incoming GTP base is colored by atom.

See also Figure S4, Table S4, and Data S1.

pE-MAP Identifies Alleles Involved in Start Site Selection and Can Finely Distinguish between Different Phenotypic Categories

What other transcription defects might underlie differences in the genetic profiles of specific RNAPII alleles? Recent work has shown that mutations in the Rpb1 trigger loop (TL), a dynamic element in the active site that couples correct NTP substrate recognition with catalysis, can alter transcription start site selection in vivo. For example, *rpb1* E1103G shifts transcription start site selection upstream at *ADH1*, whereas *rpb1* H1085Y shifts distribution of start sites downstream (Kaplan et al., 2012). The

pE-MAP subcluster containing E1103G includes an additional ten mutants in *RPB1*, *RPB2*, and *RPB7* (Figure 4A and Data S1). We examined eight of these for defects in start site selection at *ADH1* by primer extension and found that, like E1103G, all had a preference for upstream start site selection (Figures 4B, S4A, and Table S4), consistent with their clustering with E1103G (Figure 4A). Four of these mutants are also in the TL (*rpb1* F1084I, M1079R, A1076T, and M1079V); however, two are in other regions of Rpb1 (*rpb1* I1327V and S713P). Further inspection reveals that these are in close proximity to the TL (Figure 4C) (Kettenberger et al., 2004; Wang et al., 2006), suggesting that

they too may directly regulate the active site. Interestingly, the other two mutants tested are not close to the TL (*rpb2* E328K and *rpb7* D166G); these mutations may allosterically impact the active site or may function independently of the TL by recruiting other factors to the transcription apparatus. Importantly, *rpb7* D166G is the first identified *RPB7* mutation with a start site defect. We additionally examined start site selection in *rpb4Δ* and *rpb6* Q100R, as both are expected to reduce the association of Rpb4/Rpb7 with RNAPII (Edwards et al., 1991; Tan et al., 2003). Neither altered start site selection at *ADH1* (Figure S4B), indicating that the *rpb7* D166G mutant exerts a unique effect on RNAPII function (see Discussion). These data provide an example of how mechanistic information on structure-function relationships can be extracted from the pE-MAP.

In our screening process, we had also identified an *rpb2* allele mutated at two residues in close proximity (E437G/F442S) within the tip of the Rpb2 protrusion domain, whose genetic profile also clusters with the upstream start site mutants (Figure S4C and Table S2). The Rpb2 lobe and protrusion domains physically contact TFIIF (Chen et al., 2007), and consistent with this, we observed that, similar to what has been reported for TFIIF alleles (Eichner et al., 2010; Ghazy et al., 2004) (Figure S6B), *rpb2* E437G/F442S is sensitive to MPA (Figure S4D) and has a preferential upstream start site selection (Figure 4B) (see Discussion). Additional *rpb2* alleles that confer MPA sensitivity also map near the protrusion (R120C) or to the lobe (D399H); however, these did not alter *ADH1* start site selection and are genetically distinct from *rpb1* E1103G or other *rpb2* alleles, illustrating the fine resolving power of the pE-MAP (Figures 4A and 4B).

Despite its unbiased nature, the pE-MAP precisely grouped the mutants within the upstream start site cluster based on their Spt⁻ and MPA phenotypes in the plate assays (Figure 4A). However, there are other mutants with MPA or Spt⁻ phenotypes that did not exhibit upstream start preference and, notably, cluster apart from the ones that do (Figures 4A, 4B, and S4A and Table S4). These data collectively suggest that the pE-MAP has the resolving power to categorize the mutations causing upstream start site selection and, within this group, can further separate them into specific Spt⁻ and MPA phenotypic categories.

In Vitro Biochemical Activity Correlates with pE-MAP Profiles and Gene Expression

We next focused on the genetic profiles of a series of active-site mutants whose in vitro elongation rates range from <0.1 to >2-fold that of WT RNAPII (Kaplan et al., 2008, 2012; Malagon et al., 2006). This series allowed for addressing questions regarding the in vivo consequences of altered elongation rates. Clustering these mutants based on their pE-MAP profiles yielded two distinct subsets that differ by transcription rate: fast mutants (*rpb1* E1103G, L1101S, and F1084I) and slow mutants (*rpb1* F1086S, N1082S, N479S, and H1085Q) (Figure 5A dendrogram). The gene expression profiles of these mutants also group them into the same two subsets. Thus, pairs of mutants with similar in vitro transcription rates show highly similar genetic and expression profiles (Figure 5B), indicating that altered catalytic activity may likely be a defining feature in vivo.

We reasoned that, because combining two TL mutations that individually increase and decrease elongation rate results in

RNAPII with near-WT elongation rate in vitro and growth rate in vivo (Kaplan et al., 2012) (Figure S5A), these double mutants would also exhibit near-WT pE-MAP and gene expression profiles. Indeed, when we combine *rpb1* E1103G (fast) and F1086S (slow), we see very few high-scoring genetic interactions compared to the single mutants and few genes whose expression changes by more than 1.7-fold (Figures 5C and 5D and Table S2). Similar trends were observed in the genetic and expression profiles of a *rpb1* E1103G/N1082S double mutant (data not shown). However, the pE-MAP and gene expression assays had the resolving power to also identify a double mutant (*rpb1* E1103G/H1085Q) that deviated from this general rule (Figures 5C and 5D and Table S2). Despite exhibiting partially suppressive genetic and expression patterns, there are still significant effects that correlate well with E1103G (Figure S5B), suggesting a more complex genetic relationship with this double mutant. Taken together, our double-mutant analyses are consistent with the notion that the genetic interactions and gene expression changes in fast and slow mutants are, for the most part, defined by the catalytic activity of RNAPII.

Genome-wide and Gene-Specific Effects of Altering Polymerase Speed on In Vivo Splicing Efficiency

Several steps in mRNA processing are coupled to the process of transcription (Bentley, 2005). Based primarily on experiments from metazoa, a kinetic coupling model has been invoked to explain how decreasing polymerase speed affects the usage of alternative splice sites, primarily favoring the recognition of otherwise weak splice sites (de la Mata et al., 2011). We used the RNAPII mutants that differ in their in vitro elongation rates to determine whether this type of coupling exists in *S. cerevisiae*, in which splice sites are of consensus or near-consensus sequence and the sole readout is splicing efficiency. We used splicing-sensitive microarrays to measure the change in total, pre-mRNA, and mature mRNA for each intron-containing gene (Figure 6A, top) in response to mutations in the RNAPII TL. Many genes exhibit the reciprocal pre- and mature mRNA values indicative of changes in splicing (Figure 6A, side panels), and to allow facile comparison across multiple mutants, we calculated pre-mRNA/mature mRNA ratios (Clark et al., 2002) for each gene (Figure 6 and Table S5). Both gene-by-gene (Figure 6A) and global (Figure 6B) measures show that *rpb1* mutants characterized as fast in vitro (E1103G and G1097D) lead to an increase in this ratio for many genes, which indicates a defect in the splicing of those transcripts. Conversely, mutants with slow in vitro elongation (H1085Q and F1086S) decrease this ratio for many genes, a phenotype that is consistent with more efficient splicing. Thus, the global trend that we observe is an anticorrelation between polymerase rate and splicing efficiency ($p < 10^{-5}$ for each mutant in Figure 6A compared to WT; Table S5). Using qPCR, we confirmed this trend at several genes (Figure S6A) and found that these genes exhibit between 1% and 40% unspliced mRNA in WT but up to 70% in G1097D (data not shown).

Given that polymerase rate and splicing efficiency are anticorrelated, it follows that the TL double mutants (*rpb1* E1103G/F1086S and E1103G/H1085Q), which transcribe at near-WT rate in vitro, should also exhibit near-WT splicing. Consistent with this prediction, these strains had very few genes with

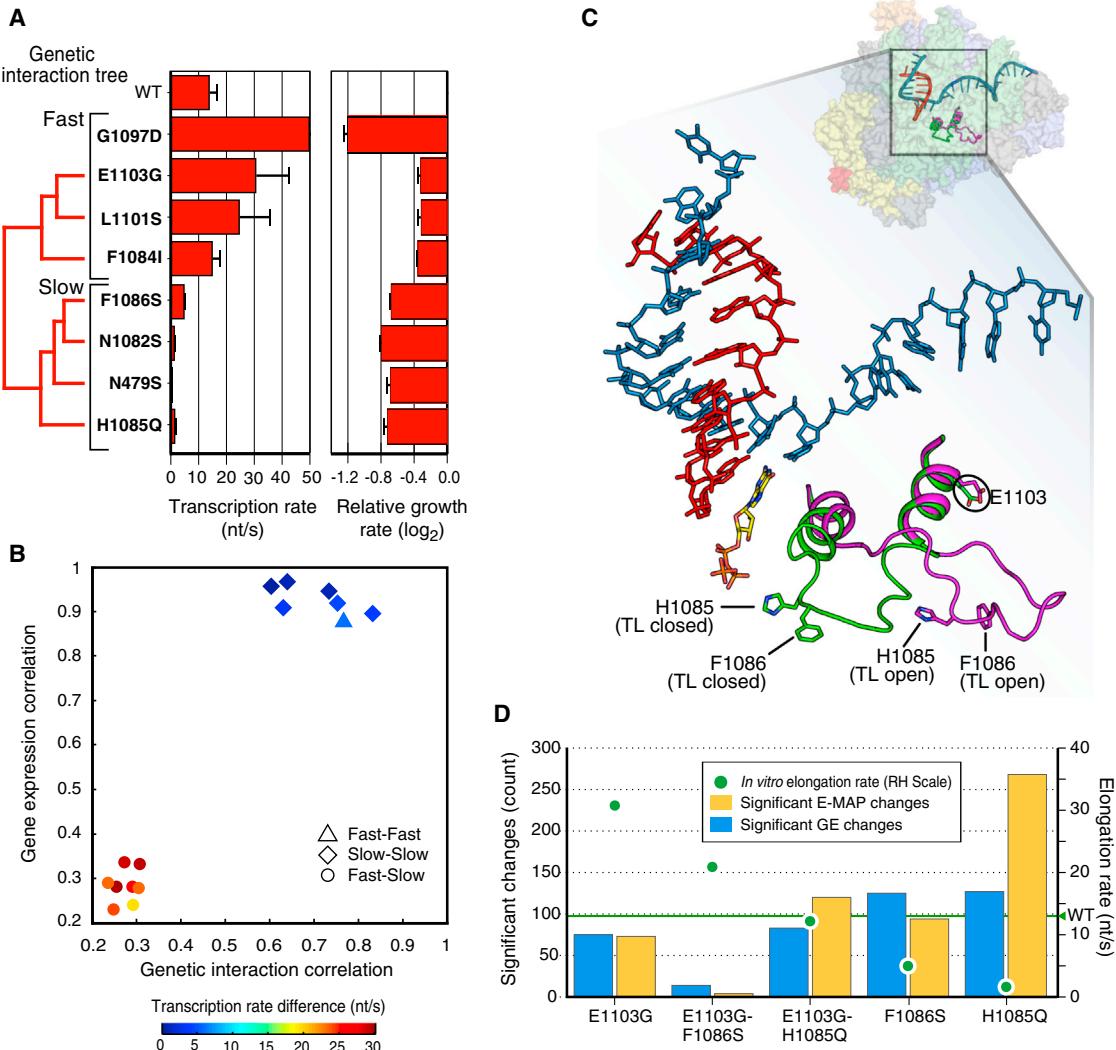


Figure 5. pE-MAP and Expression Profiles Are Indicative of Biochemical Activity

(A) In vitro transcription rates from (Kaplan et al., 2012) and in vivo growth rates relative to WT for RNAPII active-site mutants. The dendrogram was generated via hierarchical clustering of the genetic profiles. Error bars represent 95% confidence intervals for transcription rates and SD for growth rates. Means and SD of growth rates were derived from three technical replicates. Note that *rpb1* G1097D was too sick for reproducible E-MAP analysis.

(B) In vitro transcription rate difference between pairs of active-site mutants in relation to their genetic and expression profile correlations.

(C) Residues H1085 and F1086 reside in the catalytic site of the TL, whereas E1103 is part of the distal flanking α helix that structurally constrains the TL in open conformations. TL is shown in green (closed) and magenta (open); template DNA, blue; and RNA, red. The incoming GTP base is colored by atom. Coordinates for open TL are from PDB ID 1Y1V, and all others are from 2E2H.

(D) Counts of high-scoring interactions (pE-MAP score >3.3 [97.5 percentile] or <-5.1 [2.5 percentile]) in the complete genetic profiles or changes >1.7 -fold in the genome-wide expression profiles of the indicated RNAPII mutants. In vitro transcription rates are indicated on the scale on the right. See also Figure S5.

splicing defects (Figure 6B and Table S5). Thus, suppressive relationships are observed in the double mutants with respect to growth, genetic and expression profiling, in vitro transcription rate, and also mRNA splicing. Furthermore, slowing polymerase by chemical means should phenocopy a genetically slow polymerase mutant. We therefore evaluated splicing in WT cells treated with mycophenolic acid (MPA), which is known to impede transcriptional elongation (Mason and Struhl, 2005). A 10 min MPA treatment resulted in even more genes with improved splicing when compared to a slow polymerase mutant

(Figure 6B, $p < 10^{-15}$ compared to WT; Table S5), possibly because MPA elicits an acute stress. Taken together, these splicing phenotypes are consistent with a direct kinetic coupling between elongation rate and splicing in vivo.

Genetic Interactions between RNAPII Alleles and Other Mutants Reveal Relationships between Transcription Factors and RNAPII Activity

The observation of growth suppression in the TL double mutants (Kaplan et al., 2012) (Figure S5A) suggests that some of the

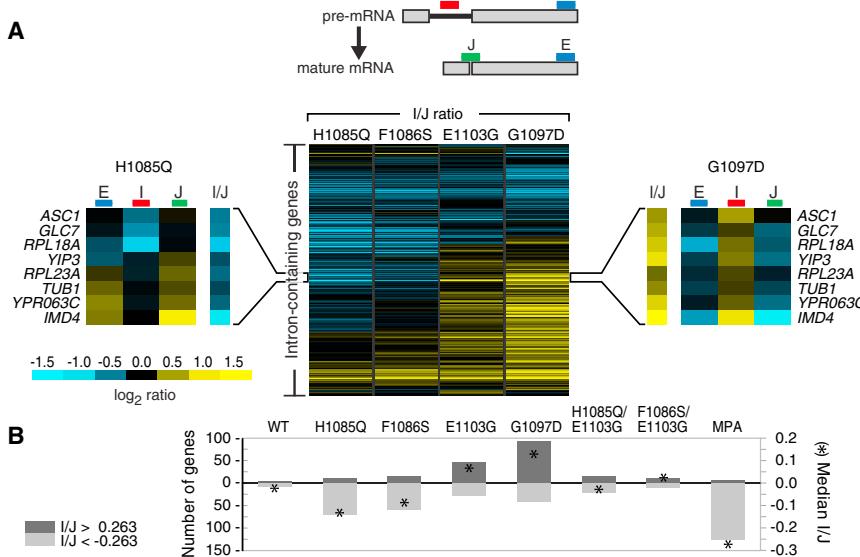


Figure 6. Effects of Altering RNAPII Transcription Rate on In Vivo Splicing Efficiency

(A) (Top) Microarray schematic for each intron-containing gene: probe I (intron) hybridizes to pre-mRNA, J (junction) to mature mRNA, and E (exon) to both. (Center) Heatmap of I/J \log_2 ratios for the slow (*rpb1* H1085Q and F1086S) and fast (*rpb1* E1103G and G1097D) mutants, corresponding to the enrichment of pre-mRNA over mature mRNA. The side panels highlight a subset of genes that behave reciprocally in fast and slow RNAPII backgrounds.

(B) Number of genes exhibiting $>20\%$ change in pre-mRNA-to-mature mRNA ratio (bars, scale on left) and median I/J \log_2 ratio (asterisks, scale on right) across entire array. MPA treatment was 10 min; WT denotes competitive hybridization between two WT cultures. I/J denotes I/J \log_2 ratio.

See also Figure S6 and Table S5.

deletion/DAmP genetic interactors that were detected in the pE-MAP might directly regulate, or collaborate with, RNAPII. We reasoned that disruption of a positively acting transcription factor would result in positive genetic interactions (suppression) with fast RNAPII mutants but negative interactions (synthetic sickness) with slow mutants. Conversely, a negatively acting factor would show opposite genetic trends. To identify these factors, the deletion/DAmP mutants were sorted based on the difference in their average genetic interaction score with fast and slow RNAPII mutants (Experimental Procedures, Figure 7A, and Table S6). We focused on genes that behaved as positively acting factors and observed that *sub1Δ* had the strongest pattern in this regard. Interestingly, previous evidence has implicated Sub1 as a positive factor in in vitro transcription assays (reviewed in Conesa and Acker, 2010), as well as in vivo (García et al., 2012). The genetic relationships from the pE-MAP were confirmed using standard growth assays in which *sub1Δ* exacerbated slow RNAPII alleles (Figure 7B) and partially suppressed fast RNAPII alleles (Figure 7C). Deletion of *SUB1* also exacerbated and suppressed the relevant RNAPII mutant phenotypes (MPA, Spt⁻, and Gal^R) (Figure S7A) (as did other mutants [Figure S7B]). Furthermore, gene expression analysis of *sub1Δ*, *rpb1* E1103G, and *sub1Δ*/E1103G showed an epistatic relationship between the E1103G mutant and *sub1Δ* (Figure 7D), consistent with their positive interaction and suggesting that a fast RNAPII mutant can bypass the requirement for Sub1.

Our recent work has implicated changes in RNAPII activity with the alteration of start site selection in vivo (Kaplan et al., 2012) (Figure 4). Furthermore, Sub1 genetically interacts with TFIIB (*sua7*) (Knaus et al., 1996; Wu et al., 1999), is broadly recruited to RNAPII/III promoters in vivo (Rosonina et al., 2009; Tavenet et al., 2009), and was implicated as a member of the RNAPII preinitiation complex (Sikorski et al., 2011). We therefore sought to determine whether Sub1 might modulate RNAPII start site choice. Notably, primer extension analysis

at *ADH1* revealed that deleting *SUB1* led to a significant downstream shift in start site (Figures 7E and S7C). Slow RNAPII TL mutant *rpb1* F1086S (Figure 4B) and *sua7* alleles (Pinto et al., 1992) (Figure S6B) also initiate downstream and are synthetic sick with *sub1Δ*. These data are consistent with the notion that Sub1 promotes transcription initiation. Double-mutant analysis revealed that *sub1Δ* also exacerbated the downstream start site shift of the slow RNAPII TL allele *rpb1* F1086S and slightly suppressed the *rpb1* E1103G allele (Figure 7E). Because *sub1Δ* has also been linked to another RNA processing step, namely 3' end processing (reviewed in Conesa and Acker, 2010), we examined the effect of *sub1Δ* on splicing and observed a statistically significant increase in splicing efficiency ($p < 10^{-8}$; Table S5), again phenocopying the slow RNAPII mutants (Figure 7F and Table S5).

Based on these data, we propose a model in which transcription start and splicing are intimately coupled with RNAPII elongation: fast RNAPII mutations result in upstream transcription start and diminished splicing, whereas slow mutations or *sub1Δ* give rise to downstream transcription start and enhanced splicing (Figure 7G). Given the possibility of direct coupling between start site selection and downstream mRNA processing, we globally measured splicing defects in *sua7-3* (TFIIB) (Pinto et al., 1994; Wu et al., 1999) and *tgf2Δ261-273* (TFIIF) (Eichner et al., 2010), mutants that exhibit downstream and upstream start site selection, respectively (Figure S6B). If splicing were strictly coupled to start site choice, one would expect these mutants to have similar splicing defects to the slow and fast RNAPII mutants. However, such a correlation was not observed (Figure S6C), suggesting that these processes are, in fact, genetically separable. Taken together, these data support a model in which the catalytic rate of RNAPII has multiple, separable effects on start site selection and mRNA processing and highlight the importance of WT elongation rate for multiple steps in gene expression.

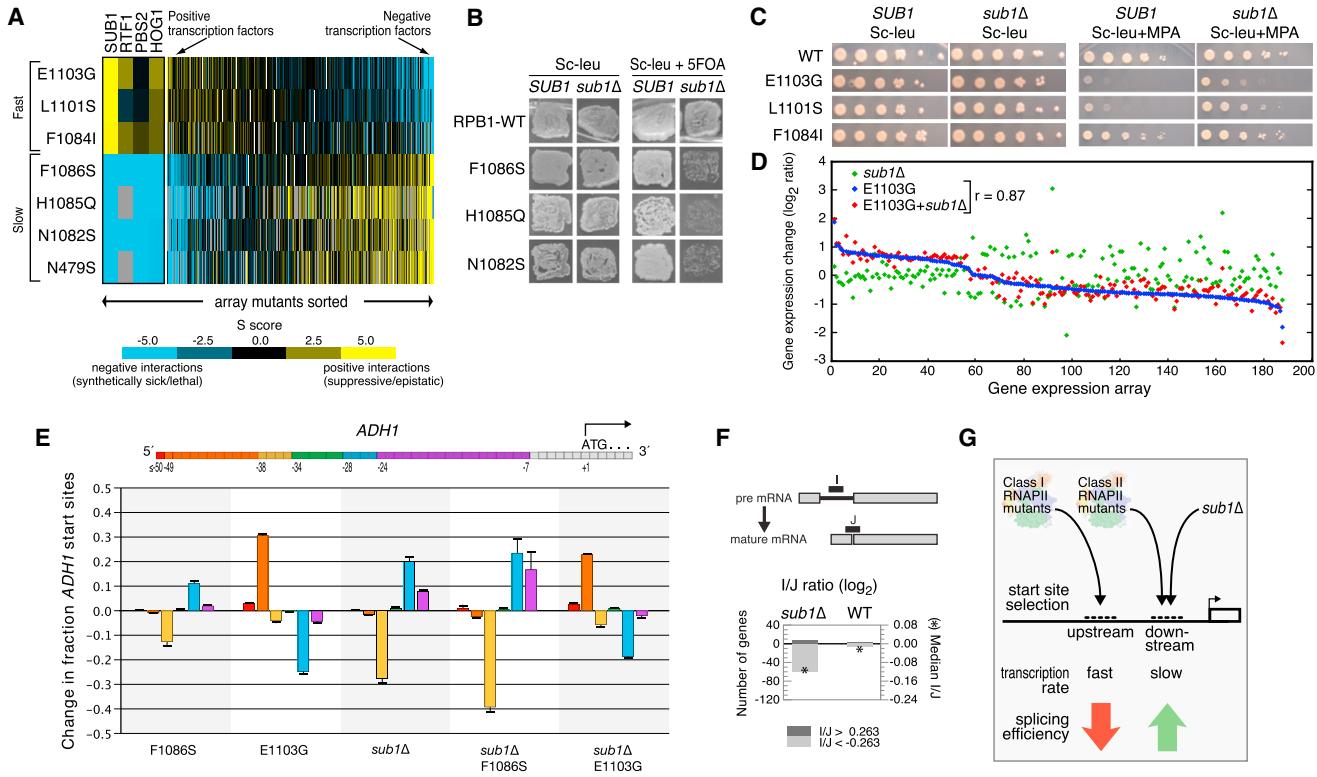


Figure 7. Genetic Interaction Patterns with Fast and Slow RNAPII Mutants Reveal Sub1 as a Transcription Factor that Regulates Start Site Selection and Influences mRNA Splicing

(A) Genetic profiles of library mutants, sorted on the difference between their average interaction with fast and slow RNAPII mutants (Table S6). (B) Patch tests examining the sensitivity of slow TL mutants to *sub1Δ*. WT *RPB1* plasmid covering *rpb1* mutants in left panel is lost in right panel. (C) Spot tests examining the effect of *sub1Δ* on fast mutants in absence (left) and presence (right) of MPA. (D) Comparison of *sub1Δ* effect on gene expression in *rpb1* E1103G (difference between red and blue) and WT (difference between green and y = 0). Included are all array transcripts exhibiting a >1.5-fold expression change in at least one of the three mutants. Transcripts are sorted by expression change in E1103G. (E) Primer extension at *ADH1* to map transcription start sites for *rpb1* F1086S, E1103G, and *sub1Δ* (Figure S7C). Bar colors correspond to sequence windows in the *ADH1* schematic (top), and heights specify the mean fraction change of transcription start in mutant compared to WT. Error bars represent SD. (F) Splicing microarray analysis of *sub1Δ*, as in Figure 6 (Table S5). Number of genes exhibiting >20% change in pre-mRNA-to-mature mRNA ratio (bars, scale on left) and median I/J log₂ ratio (asterisks, scale on right). (G) Model for the effect of Sub1 and RNAPII activity on start site selection and splicing. Fast RNAPII mutations (class I) result in upstream transcription start and diminished splicing efficiency, whereas *sub1Δ* or slow RNAPII mutations (class II) shift transcription start downstream and enhance splicing.

See also Figures S6 and S7 and Tables S4, S5, and S6.

DISCUSSION

In this study, we have described an important extension of our E-MAP genetic interaction mapping strategy, the functional interrogation of individual protein residues. We have first used this approach to genetically dissect RNAPII and demonstrate how the pE-MAP successfully characterized functionally distinct, individual amino acids within this multifunctional complex. This analysis provided not only insight into global structure-function relationships within RNAPII, but also specific details about how RNAPII regulates and is regulated by different factors and processes.

Insights into RNAPII Function Derived from the pE-MAP

Multiple aspects of the genetic interaction map provided insights into transcriptional regulation. First, by comparing the RNAPII genetic interaction profiles with those from previous deletion/

knockdown studies, we could assign function to individual residues, which allowed us to generate a point mutant-protein complex connectivity map (Figures 3 and S3). Furthermore, analyzing double mutants, both within RNAPII itself (Figures 5C and 5D) and between RNAPII and other genes (Figure 7), allowed for a better understanding of TL function, as well as the identification of other factors that directly or indirectly impinge on RNAPII activity. Indeed, the pE-MAP allowed us to identify putative transcription factors (negative and positive), such as Sub1. Although Sub1 was previously implicated at the promoter (Sikorski et al., 2011), we here demonstrated that it positively regulates RNAPII by influencing start site usage (Figure 7). The epistatic relationships between fast RNAPII alleles and *sub1Δ* suggest that Sub1 activity may be bypassed when RNAPII catalytic activity is increased. Collectively, our data indicate that Sub1 plays a direct role in transcriptional initiation and influences mRNA splicing, possibly via its effect on elongation (García et al., 2012).

Finally, structural analysis revealed that mutations close in three-dimensional (3D) space have very similar genetic profiles, including those in different subunits (Figure 2B), suggesting that structural information is ultimately contained within the pE-MAP and can be used to identify specific protein-protein interaction interfaces. For example, the Rpb2 mutation E437G/F442S, which shifts start site selection upstream, is in a domain that contacts TFIIF at a region in which mutations also result in upstream start site shifts (Chen et al., 2007; Eichner et al., 2010). Furthermore, the identification of an Rpb7 mutant (D166G) that alters start site selection is intriguing, as this region interacts with TFIID in a cryo-EM structure of the RNAPII preinitiation complex (He et al., 2013) and TFIID also alters start site selection (Goel et al., 2012). Therefore, the pE-MAP technique could supplement other methods, such as crosslinking and electron microscopy, to identify physically interacting protein regions.

Coordination of Transcriptional Rate with Start Site Selection and mRNA Splicing

The importance of maintaining WT rates of transcription is evidenced by the phenotypic defects observed in the fast and slow mutants (Kaplan et al., 2012) (Figures 5 and 6), as well as the striking mutual suppression seen when combining two mutations that individually make RNAPII slow or fast (Figures 5 and 6). Interestingly, the pE-MAP identified two groups of RNAPII mutants: one that preferentially initiates upstream and exhibits an increased rate of transcription and one that initiates downstream and transcribes slowly (Figure 7) (Kaplan et al., 2012). We propose that both phenotypes are direct consequences of the efficiency of nucleotide selection and incorporation, as the addition of the first nucleotides at initiation is biochemically similar to adding nucleotides during elongation. These data support the model that RNAPII engages in “scanning” during initiation in *S. cerevisiae* (Giardina and Lis, 1993; Kaplan, 2013; Kaplan et al., 2012; Kuehner and Brow, 2006). Whether RNAPII catalysis drives this scanning or whether scanning occurs in the absence of nucleotide incorporation, perhaps driven by TFIID, is unknown. We also note that, because changes in start site selection alter the 5' UTR length and composition, initial transcription decisions may have downstream effects on gene expression such as changes in RNA stability or translational efficiency of the mRNA (Arribere and Gilbert, 2013; Rojas-Duran and Gilbert, 2012).

The pE-MAP has allowed insights into the cotranscriptional process of mRNA splicing. It is now clear that most introns are removed while RNA polymerase is still associated with the DNA template. In metazoans, alternative splicing decisions can be influenced by factors impinging on transcription, including promoter identity and polymerase speed (reviewed in Perales and Bentley, 2009). Slowing the rate of elongation by mutation of RNAPII or chemical means can improve the recognition of splice sites that deviate from consensus signals (de la Mata et al., 2003; Howe et al., 2003; Ip et al., 2011). Because the spliceosome undergoes stepwise assembly on each intron, slowing transcription can afford more time for formation of the catalytically active machine before transcription of a downstream, stronger site. Although budding yeast lack alternative splicing, it nonetheless follows that the efficiency of cotranscriptional

splicing would be favored by allowing sufficient time for spliceosome assembly. Indeed, recent work suggests that RNA polymerase may slow down to favor co- versus posttranscriptional splicing (Aitken et al., 2011; Alexander et al., 2010; Carrillo Oesterreich et al., 2010). Our microarray analyses that directly compare faster and slower RNAPII show a clear trend in which splicing efficiency is anticorrelated with transcription rate (Figure 6); thus, these results satisfy the predictions of kinetic coupling in *S. cerevisiae*.

Interestingly, we observed correlation among start site selection, elongation rate, and splicing efficiency. In fact, promoter-proximal events are known to be able to influence downstream RNA transactions: promoter identity can influence alternative splicing or mRNA stability in other systems (Cramer et al., 1997; Harel-Sharvit et al., 2010; Trcek et al., 2011), and 5' UTR length, determined by start site selection, can strongly alter translation efficiency in budding yeast (Rojas-Duran and Gilbert, 2012). However, when we measured splicing efficiency using mutants in the general transcription factors TFIID and TFIIB that alter start site selection, we found that not all initiation phenotypes are predictive of splicing efficiency (Figure S6C). This suggests that RNAPII catalytic rate has several separable effects on gene expression, a claim supported by recent evidence showing kinetic coupling between RNAPII transcription and Sen1-dependent termination (Hazelbaker et al., 2013).

Taken together, our data highlight the important impact of transcription speed determined by the genetic status of RNAPII and *trans*-acting factors (e.g., Sub1) on start site selection and mRNA splicing. We propose that RNA polymerase may have been evolutionarily tuned to coordinate between multiple steps in gene expression, and we predict that polymerase rate may influence multiple additional cotranscriptional steps in gene expression, including mRNP assembly, 3' end processing, and export.

Future Studies Using the pE-MAP Approach

S. cerevisiae RNAPII provided the groundwork for validating the pE-MAP approach. Application to other molecular machines, including the ribosome, the proteasome, HSP70, HSP90, histones, and DNA polymerases, should prove informative. This analysis can also be carried out in other organisms that are genetically tractable and are amenable to high-throughput genetic interaction mapping, including *S. pombe* (Roguev et al., 2008; Roguev et al., 2007; Ryan et al., 2012) and *E. coli* (Butland et al., 2008; Typas et al., 2008). Furthermore, pE-MAPs could be used to gain structural insight into proteins and complexes that have unknown structures. Finally, as genetic interaction mapping strategies become more prevalent in mammalian cells (Bassik et al., 2013; Laufer et al., 2013; Lin et al., 2012; Roguev et al., 2013) and with the development of genome editing (Gaj et al., 2013), similar work characterizing the function of individual amino acids will have great impact on understanding how point mutations in specific genes result in different disease states.

EXPERIMENTAL PROCEDURES

E-MAP-compatible *MAT α* RNAPII mutant strains carrying a marked *rpb* deletion and mutant *rpb* on a *CEN* plasmid were mated with 1,200 *MAT α*

DAmP/deletion strains by pinning on solid media. Sporulation was induced, and double-mutant *MATa* spores were isolated on selective media. Genetic interactions were scored based on double-mutant colony sizes, which were extracted using automated imaging software.

For gene expression and splicing arrays, total RNA was extracted from mutant and WT log-phase cells. Competitive hybridizations were performed between mutant and WT complementary DNA (cDNA) (splicing) or complementary RNA (cRNA) (gene expression).

CTF assays were carried out by plating strains carrying *ade2-101* and a chromosome VII fragment containing *SUP11* on SC medium with 20% adenine and measuring the fraction of red colonies. The red color caused by *ade2-101* is counteracted by *SUP11*; chromosome fragment loss results in red colonies.

For transcription start site analysis, we performed primer extension from a ^{32}P end-labeled oligo annealed to total RNA. cDNAs were separated by PAGE and bands quantified.

Detailed descriptions of experiments and computational analyses are provided in the [Extended Experimental Procedures](#).

ACCESSION NUMBERS

The GEO accession number for the gene expression profiles reported in this paper is GSE47429.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and six tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2013.07.033>.

ACKNOWLEDGMENTS

The authors wish to thank members of the Krogan, Kaplan, and Guthrie labs; David Agard, Brad Cairns, Ada Cheng, C. James Ingles, Paivand Jalalian, Tanja Kortemme, Andrej Sali, and Keith Yamamoto for helpful discussion; Steve Buratowski, Olga Calvo, and Eva Nogales for sharing results prior to publication; Steve Hahn and Mike Hampsey for providing *tfg2* and *sua7* mutant plasmids, respectively; Colm Ryan for help with statistical analysis; and Ricardo Almeida and Jiewei Xu for assistance with growth assays. This work was supported by grants from QB3 at UCSF and the NIH (R01GM084448, R01GM084279, P50GM081879, and R01GM098101 to N.J.K.; R01GM036659 to Roger D. Kornberg for support of C.D.K. for a portion of this work; R01GM097260 to C.D.K.; R01GM021119 to C.G.; and DP5OD009180 to J.S.F.). C.G. is an ACS Research Professor of Molecular Genetics. E.A.M. was supported by an NSF graduate research fellowship. C.D.K. acknowledges a Helen Hay Whitney Fellowship for the early stages of this work. H.B. was supported by the UCSF Biophysics Graduate Program. N.J.K. is a Searle Scholar and a Keck Young Investigator. The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the acknowledged funding agencies.

H.B., E.A.M., C.G., C.D.K., and N.J.K. designed the research; H.B., H.J., E.A.M., Y.A.C., S.W., J.J.B., C.Q., F.H., and L.K.T. carried out experimental assays; H.B. analyzed data; M.S., J.H.M., and J.S.F. provided computational support; F.C.P.H., P.H., C.G., C.D.K., and N.J.K. supervised the research; H.B., E.A.M., C.G., C.D.K., and N.J.K. wrote the paper. H.J. and E.A.M. contributed equally.

Received: December 16, 2012

Revised: May 16, 2013

Accepted: July 22, 2013

Published: August 8, 2013

REFERENCES

- Aguilar, P.S., Fröhlich, F., Rehman, M., Shales, M., Ulitsky, I., Olivera-Couto, A., Braberg, H., Shamir, R., Walter, P., Mann, M., et al. (2010). A plasma membrane E-MAP reveals links of the eisosome with sphingolipid metabolism and endosomal trafficking. *Nat. Struct. Mol. Biol.* 17, 901–908.
- Aitken, S., Alexander, R.D., and Beggs, J.D. (2011). Modelling reveals kinetic advantages of co-transcriptional splicing. *PLoS Comput. Biol.* 7, e1002215.
- Alexander, R.D., Innocente, S.A., Barras, J.D., and Beggs, J.D. (2010). Splicing-dependent RNA polymerase pausing in yeast. *Mol. Cell* 40, 582–593.
- Archambault, J., and Friesen, J.D. (1993). Genetics of eukaryotic RNA polymerases I, II, and III. *Microbiol. Rev.* 57, 703–724.
- Arribere, J.A., and Gilbert, W.V. (2013). Roles for transcript leaders in translation and mRNA decay revealed by transcript leader sequencing. *Genome Res.* 23, 977–987.
- Bassik, M.C., Kampmann, M., Lebbink, R.J., Wang, S., Hein, M.Y., Poser, I., Weibezahn, J., Horbeck, M.A., Chen, S., Mann, M., et al. (2013). A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility. *Cell* 152, 909–922.
- Beltrao, P., Cagney, G., and Krogan, N.J. (2010). Quantitative genetic interactions reveal biological modularity. *Cell* 141, 739–745.
- Benschop, J.J., Brabers, N., van Leenen, D., Bakker, L.V., van Deutekom, H.W., van Berkum, N.L., Apweiler, E., Lijnzaad, P., Holstege, F.C., and Kemmeren, P. (2010). A consensus of core protein complex compositions for *Saccharomyces cerevisiae*. *Mol. Cell* 38, 916–928.
- Bentley, D.L. (2005). Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr. Opin. Cell Biol.* 17, 251–256.
- Butland, G., Babu, M., Díaz-Mejía, J.J., Bohdane, F., Phanse, S., Gold, B., Yang, W., Li, J., Gagarinova, A.G., Pogoutse, O., et al. (2008). eSGA: *E. coli* synthetic genetic array analysis. *Nat. Methods* 5, 789–795.
- Carrillo Oesterreich, F., Preibisch, S., and Neugebauer, K.M. (2010). Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. *Mol. Cell* 40, 571–581.
- Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P., and Workman, J.L. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* 123, 581–592.
- Chen, H.T., Warfield, L., and Hahn, S. (2007). The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. *Nat. Struct. Mol. Biol.* 14, 696–703.
- Clark, T.A., Sugnet, C.W., and Ares, M., Jr. (2002). Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. *Science* 296, 907–910.
- Collins, S.R., Kemmeren, P., Zhao, X.C., Greenblatt, J.F., Spencer, F., Holstege, F.C., Weissman, J.S., and Krogan, N.J. (2007a). Toward a comprehensive atlas of the physical interactome of *Saccharomyces cerevisiae*. *Mol. Cell. Proteomics* 6, 439–450.
- Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M., et al. (2007b). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* 446, 806–810.
- Collins, S.R., Roguev, A., and Krogan, N.J. (2010). Quantitative genetic interaction mapping using the E-MAP approach. *Methods Enzymol.* 470, 205–231.
- Conesa, C., and Acker, J. (2010). Sub1/PC4 a chromatin associated protein with multiple functions in transcription. *RNA Biol.* 7, 287–290.
- Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L., Toufghi, K., Mostafavi, S., et al. (2010). The genetic landscape of a cell. *Science* 327, 425–431.
- Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* 292, 1863–1876.
- Cramer, P., Pesce, C.G., Baralle, F.E., and Kornblith, A.R. (1997). Functional association between promoter structure and transcript alternative splicing. *Proc. Natl. Acad. Sci. USA* 94, 11456–11460.

- de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D., and Kornblith, A.R. (2003). A slow RNA polymerase II affects alternative splicing in vivo. *Mol. Cell* 12, 525–532.
- de la Mata, M., Muñoz, M.J., Alló, M., Fededa, J.P., Schor, I.E., and Kornblith, A.R. (2011). RNA Polymerase II Elongation at the Crossroads of Transcription and Alternative Splicing. *Genet. Res. Int.* 2011, 309865.
- Edwards, A.M., Kane, C.M., Young, R.A., and Kornberg, R.D. (1991). Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro. *J. Biol. Chem.* 266, 71–75.
- Eichner, J., Chen, H.T., Warfield, L., and Hahn, S. (2010). Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. *EMBO J.* 29, 706–716.
- Fiedler, D., Braberg, H., Mehta, M., Chechik, G., Cagney, G., Mukherjee, P., Silva, A.C., Shales, M., Collins, S.R., van Wageningen, S., et al. (2009). Functional organization of the *S. cerevisiae* phosphorylation network. *Cell* 136, 952–963.
- Gaj, T., Gersbach, C.A., and Barbas, C.F., 3rd. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* 31, 397–405.
- García, A., Collin, A., and Calvo, O. (2012). Sub1 associates with Spt5 and influences RNA polymerase II transcription elongation rate. *Mol. Biol. Cell* 23, 4297–4312.
- Ghazy, M.A., Brodie, S.A., Ammerman, M.L., Ziegler, L.M., and Ponticelli, A.S. (2004). Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II. *Mol. Cell. Biol.* 24, 10975–10985.
- Giardina, C., and Lis, J.T. (1993). DNA melting on yeast RNA polymerase II promoters. *Science* 261, 759–762.
- Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* 292, 1876–1882.
- Goel, S., Krishnamurthy, S., and Hampsey, M. (2012). Mechanism of start site selection by RNA polymerase II: interplay between TFIIB and Ssl2/XPB helicase subunit of TFIIF. *J. Biol. Chem.* 287, 557–567.
- Greger, I.H., and Proudfoot, N.J. (1998). Poly(A) signals control both transcriptional termination and initiation between the tandem GAL10 and GAL7 genes of *Saccharomyces cerevisiae*. *EMBO J.* 17, 4771–4779.
- Harel-Sharvit, L., Eldad, N., Haimovich, G., Barkai, O., Duek, L., and Choder, M. (2010). RNA polymerase II subunits link transcription and mRNA decay to translation. *Cell* 143, 552–563.
- Hazelbaker, D.Z., Marquardt, S., Wlotzka, W., and Buratowski, S. (2013). Kinetic competition between RNA Polymerase II and Sen1-dependent transcription termination. *Mol. Cell* 49, 55–66.
- He, Y., Fang, J., Taatjes, D.J., and Nogales, E. (2013). Structural visualization of key steps in human transcription initiation. *Nature* 495, 481–486.
- Howe, K.J., Kane, C.M., and Ares, M., Jr. (2003). Perturbation of transcription elongation influences the fidelity of internal exon inclusion in *Saccharomyces cerevisiae*. *RNA* 9, 993–1006.
- Ip, J.Y., Schmidt, D., Pan, Q., Ramani, A.K., Fraser, A.G., Odom, D.T., and Blencowe, B.J. (2011). Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. *Genome Res.* 21, 390–401.
- Kaplan, C.D. (2013). Basic mechanisms of RNA polymerase II activity and alteration of gene expression in *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* 1829, 39–54.
- Kaplan, C.D., Holland, M.J., and Winston, F. (2005). Interaction between transcription elongation factors and mRNA 3'-end formation at the *Saccharomyces cerevisiae* GAL10-GAL7 locus. *J. Biol. Chem.* 280, 913–922.
- Kaplan, C.D., Jin, H., Zhang, I.L., and Belyanin, A. (2012). Dissection of Pol II trigger loop function and Pol II activity-dependent control of start site selection in vivo. *PLoS Genet.* 8, e1002627.
- Kaplan, C.D., Larsson, K.M., and Kornberg, R.D. (2008). The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin. *Mol. Cell* 30, 547–556.
- Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J., et al. (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* 123, 593–605.
- Kettenberger, H., Armache, K.J., and Cramer, P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol. Cell* 16, 955–965.
- Knaus, R., Pollock, R., and Guarente, L. (1996). Yeast SUB1 is a suppressor of TFIIB mutations and has homology to the human co-activator PC4. *EMBO J.* 15, 1933–1940.
- Kuehner, J.N., and Brow, D.A. (2006). Quantitative analysis of in vivo initiator selection by yeast RNA polymerase II supports a scanning model. *J. Biol. Chem.* 281, 14119–14128.
- Laufer, C., Fischer, B., Billmann, M., Huber, W., and Boutros, M. (2013). Mapping genetic interactions in human cancer cells with RNAi and multiparametric phenotyping. *Nat. Methods* 10, 427–431.
- Lin, Y.Y., Kiihl, S., Suhail, Y., Liu, S.Y., Chou, Y.H., Kuang, Z., Lu, J.Y., Khor, C.N., Lin, C.L., Bader, J.S., et al. (2012). Functional dissection of lysine deacetylases reveals that HDAC1 and p300 regulate AMPK. *Nature* 482, 251–255.
- Malagon, F., Kireeva, M.L., Shafer, B.K., Lubkowska, L., Kashlev, M., and Strathern, J.N. (2006). Mutations in the *Saccharomyces cerevisiae* RPB1 gene conferring hypersensitivity to 6-azauracil. *Genetics* 172, 2201–2209.
- Mason, P.B., and Struhl, K. (2005). Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo. *Mol. Cell* 17, 831–840.
- Ohkuni, K., and Kitagawa, K. (2011). Endogenous transcription at the centromere facilitates centromere activity in budding yeast. *Curr. Biol.* 21, 1695–1703.
- Pan, X., Yuan, D.S., Xiang, D., Wang, X., Sookhai-Mahadeo, S., Bader, J.S., Hieter, P., Spencer, F., and Boeke, J.D. (2004). A robust toolkit for functional profiling of the yeast genome. *Mol. Cell* 16, 487–496.
- Perales, R., and Bentley, D. (2009). “Cotranscriptionality”: the transcription elongation complex as a nexus for nuclear transactions. *Mol. Cell* 36, 178–191.
- Pinto, I., Ware, D.E., and Hampsey, M. (1992). The yeast SUA7 gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection in vivo. *Cell* 68, 977–988.
- Pinto, I., Wu, W.H., Na, J.G., and Hampsey, M. (1994). Characterization of sua7 mutations defines a domain of TFIIB involved in transcription start site selection in yeast. *J. Biol. Chem.* 269, 30569–30573.
- Roguev, A., Bandyopadhyay, S., Zofall, M., Zhang, K., Fischer, T., Collins, S.R., Qu, H., Shales, M., Park, H.O., Hayles, J., et al. (2008). Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. *Science* 322, 405–410.
- Roguev, A., Talbot, D., Negri, G.L., Shales, M., Cagney, G., Bandyopadhyay, S., Panning, B., and Krogan, N.J. (2013). Quantitative genetic-interaction mapping in mammalian cells. *Nat. Methods* 10, 432–437.
- Roguev, A., WIREN, M., Weissman, J.S., and Krogan, N.J. (2007). High-throughput genetic interaction mapping in the fission yeast *Schizosaccharomyces pombe*. *Nat. Methods* 4, 861–866.
- Rojas-Duran, M.F., and Gilbert, W.V. (2012). Alternative transcription start site selection leads to large differences in translation activity in yeast. *RNA* 18, 2299–2305.
- Rosonina, E., Willis, I.M., and Manley, J.L. (2009). Sub1 functions in osmoregulation and in transcription by both RNA polymerases II and III. *Mol. Cell. Biol.* 29, 2308–2321.
- Ryan, C.J., Roguev, A., Patrick, K., Xu, J., Jahari, H., Tong, Z., Beltrao, P., Shales, M., Qu, H., Collins, S.R., et al. (2012). Hierarchical modularity and the evolution of genetic interactomes across species. *Mol. Cell* 46, 691–704.

- Schuldiner, M., Collins, S.R., Thompson, N.J., Denic, V., Bhamidipati, A., Punna, T., Ihmels, J., Andrews, B., Boone, C., Greenblatt, J.F., et al. (2005). Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniaarray profile. *Cell* 123, 507–519.
- Schuldiner, M., Collins, S.R., Weissman, J.S., and Krogan, N.J. (2006). Quantitative genetic analysis in *Saccharomyces cerevisiae* using epistatic miniaarray profiles (E-MAPs) and its application to chromatin functions. *Methods* 40, 344–352.
- Shaw, R.J., and Reines, D. (2000). *Saccharomyces cerevisiae* transcription elongation mutants are defective in PUR5 induction in response to nucleotide depletion. *Mol. Cell. Biol.* 20, 7427–7437.
- Sikorski, T.W., Ficarro, S.B., Holik, J., Kim, T., Rando, O.J., Marto, J.A., and Buratowski, S. (2011). Sub1 and RPA associate with RNA polymerase II at different stages of transcription. *Mol. Cell* 44, 397–409.
- Spencer, F., Gerring, S.L., Connelly, C., and Hieter, P. (1990). Mitotic chromosome transmission fidelity mutants in *Saccharomyces cerevisiae*. *Genetics* 124, 237–249.
- Tan, Q., Prysak, M.H., and Woychik, N.A. (2003). Loss of the Rpb4/Rpb7 subcomplex in a mutant form of the Rpb6 subunit shared by RNA polymerases I, II, and III. *Mol. Cell. Biol.* 23, 3329–3338.
- Tavenet, A., Suleau, A., Dubreuil, G., Ferrari, R., Ducrot, C., Michaut, M., Aude, J.C., Dieci, G., Lefebvre, O., Conesa, C., and Acker, J. (2009). Genome-wide location analysis reveals a role for Sub1 in RNA polymerase III transcription. *Proc. Natl. Acad. Sci. USA* 106, 14265–14270.
- Tong, A.H., Lesage, G., Bader, G.D., Ding, H., Xu, H., Xin, X., Young, J., Berriz, G.F., Brost, R.L., Chang, M., et al. (2004). Global mapping of the yeast genetic interaction network. *Science* 303, 808–813.
- Trcek, T., Larson, D.R., Moldón, A., Query, C.C., and Singer, R.H. (2011). Single-molecule mRNA decay measurements reveal promoter-regulated mRNA stability in yeast. *Cell* 147, 1484–1497.
- Typas, A., Nichols, R.J., Siegele, D.A., Shales, M., Collins, S.R., Lim, B., Braberg, H., Yamamoto, N., Takeuchi, R., Wanner, B.L., et al. (2008). High-throughput, quantitative analyses of genetic interactions in *E. coli*. *Nat. Methods* 5, 781–787.
- Wang, D., Bushnell, D.A., Westover, K.D., Kaplan, C.D., and Kornberg, R.D. (2006). Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* 127, 941–954.
- Wilmes, G.M., Bergkessel, M., Bandyopadhyay, S., Shales, M., Braberg, H., Cagney, G., Collins, S.R., Whitworth, G.B., Kress, T.L., Weissman, J.S., et al. (2008). A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. *Mol. Cell* 32, 735–746.
- Winston, F., and Sudarsanam, P. (1998). The SAGA of Spt proteins and transcriptional analysis in yeast: past, present, and future. *Cold Spring Harb. Symp. Quant. Biol.* 63, 553–561.
- Wu, W.H., Pinto, I., Chen, B.S., and Hampsey, M. (1999). Mutational analysis of yeast TFIIB. A functional relationship between Ssu72 and Sub1/Tsp1 defined by allele-specific interactions with TFIIB. *Genetics* 153, 643–652.

Supplemental Information

EXTENDED EXPERIMENTAL PROCEDURES

Plate Phenotyping

Plate phenotyping of RNAPII mutants was done as described (Kaplan et al., 2012). Briefly, *CEN LEU2* plasmids containing a mutant *rpb* RNAPII subunit gene were transformed into an appropriate Leu⁺ strain with corresponding endogenous RNAPII subunit gene knocked out but complemented with a *CEN URA3 WT RPB* subunit gene. Leu⁺ transformants were patched on solid medium lacking leucine and replica-plated to medium lacking leucine but containing 5-fluoroorotic acid (5-FOA, Gold Biotechnology) to select against cells maintaining *RPB WT URA3* plasmids. *sub1Δ + RNAPII* mutant strains for direct testing of double mutant phenotypes were constructed based on CKY283 and analyzed in the same fashion as RNAPII single mutants. Yeast media used in phenotyping were prepared as previously described (Amberg et al., 2005; Kaplan et al., 2012). YPD medium contained yeast extract (1% w/v final, BD), peptone (2% w/v final, BD), dextrose (2% final w/v), bacto agar (2%, BD) and supplemental adenine and tryptophan. Alternate carbon source YP media were YP raffinose (2% final w/v, USB) and YP raffinose (2% final w/v, USB) plus galactose (1% final w/v, Sigma-Aldrich), with plates also containing antimycin A (1 μg/ml, Sigma-Aldrich). Synthetic complete medium lacking leucine (SC-Leu) was as described (Kaplan et al., 2012). Mycophenolic acid (MPA, Sigma-Aldrich) was added to SC-Leu at 20 μg/ml final concentration from a 10 mg/ml stock in ethanol.

Galactose sensitivity or resistance (Gal^S or Gal^R, determined based on strain growth on YP raffinose plus galactose) phenotypes are based on an allele of *GAL10*, *gal10Δ56*, that has a deletion of the major *GAL10* polyadenylation signal (Greger et al., 2000; Greger and Proudfoot, 1998; Kaplan et al., 2005). A consequence of this deletion is that *GAL10* transcription reads through into the downstream *GAL7* promoter, interfering with *GAL7* expression. Under conditions where other Gal genes are expressed, specifically Gal1p, Gal7 protein is required to prevent galactose toxicity caused by Galactose catabolism intermediate. Thus defects in *GAL7* due to transcriptional interference from *gal10Δ56* cause the Gal^S phenotype, and resistance to this transcription defect leads to the Gal^R phenotype.

The Spt⁻ phenotype studied here relates to the *LYS2* allele, *lys2-128δ* (Simchen et al., 1984), and suppression of the lysine auxotrophy (Lys⁻ phenotype) conferred by a Ty 0 element insertion into the 5' end of the *LYS2* gene present in this allele. The suppression of Lys⁻ phenotype (growth on medium lacking lysine) is referred to as the Spt⁻ phenotype.

RNAPII Subunit Mutagenesis

PCR-based random mutagenesis was performed to introduce mutations into RNAPII subunits. For each individual RNAPII subunit, one or more screens were performed, mutagenizing different parts of longer subunit genes in different screens (*RPO21/RPB1* or *RPB2*), or the entire open reading frame (*RPB3*, *RPB7*, *RPB11*) for smaller subunit genes. Plasmids containing individual subunits and flanking DNA allowing for expression and termination from native elements were constructed in pRS-based *CEN*, low copy vectors (Sikorski and Hieter, 1989) for complementation of appropriate RNAPII subunit gene deletions. Mutagenesis conditions were under standard PCR conditions (1.5 mM MgCl₂, 200 μM dNTPs) where 6 individual reactions for each mutagenized region were performed for 35 cycles of amplification followed by seeding of 6 additional reactions for 35 more cycles, followed by pooling of second-round reactions for each mutagenized region. Mutagenized PCR products were transformed into yeast along with appropriate gapped plasmid vector lacking most of the region to be mutagenized. Selection of vector marker (*LEU2*) after transformation allowed plasmids putatively created through gap repair recombination between gapped-vector and mutagenized PCR products to be obtained. These transformant populations represented putative mutant libraries that were screened for mutant phenotypes upon standard plasmid shuffling (Boeke et al., 1987) to remove the wild-type *RPB* plasmid, allowing the transformed plasmid to be the only copy of the particular *RPB* subunit in the colony. Following replica-plating to 5FOA, plates were additionally replicated to a number of other media for phenotyping. Candidate mutants were streaked and rescreened for phenotypes followed by either dominance/recessive/plasmid linkage tests or direct plasmid isolation and retransformation to assay plasmid linkage/reproducibility of observed phenotypes. Mutant plasmids were isolated from yeast by standard procedures and portions of insert that had been subjected to PCR in initial mutagenesis were sequenced to identify alterations.

For *RPB3*, two independent replicate screens were performed, HindIII sites present in flanking genomic DNA were used to remove wild-type sequence and gap-repair mutagenize the entirety of the *RPB3* ORF in an individual screen. For *RPB7*, vector *Xhol* and *SacI* sites were used to remove the entire *RPB7* insert to allow mutagenesis of the entire *RPB7* ORF. For *RPB11*, a vector *Xhol* site and a HindIII site present in the 3' end of *RPB11* were used to remove wild-type *RPB11* and mutagenize the majority if not all of the *RPB11* ORF. For *RPB2*, three individual screens were performed: the first utilized *NdeI* and HindIII sites to remove ~700 base pairs of *RPB2* ORF 5' sequence for mutagenesis, the second utilized HindIII and *MscI* to remove a 1.6 kb internal ORF fragment for mutagenesis, and the third used an internal *MscI* and vector *SacI* site to remove a 1.6 kb *RPB2* ORF 3' sequence for mutagenesis. For *RPO21/RPB1*, three screens were performed. For the first, engineered silent *BamHI* and *PstI* restriction sites were utilized for removal of an internal *RPO21/RPB1* ORF fragment for mutagenesis. For the second, a 3' *RPO21/RPB1* ORF fragment up to the CTD was removed with digestion by *XbaI* and *BsiWI* for mutagenesis. For the third, the entire *RPO21/RPB1* ORF 3' region was removed with digestion by *XbaI-SnaBI* for mutagenesis. In general, 3000–5000 colonies were screened per independent screen and mutagenesis levels were approximated by number of 5FOA-sensitive colonies (representing lethal mutations and were between 2%–7% in different screens).

For creation of *URA3*-marked *CEN* plasmids for pE-MAP screens, RNAPII alleles were moved from *LEU2*-marked *CEN* versions into pRS316 or pRS416 by standard cloning procedures. *URA3*-marked RNAPII mutant plasmids were then transformed into pE-MAP *rpbΔ* strains and screened for loss of WT *RPB LEU2* plasmids.

pE-MAP Analysis

E-MAP amenable strains for RNAPII mutants were created by transformation of *MATα* spore selectable marker (*mfa1p::HIS3*) into relevant *MATα rpbΔ* strains (*NAT'*::*rpbΔ* complemented with WT *RPB CEN LEU2* plasmid) via recombination into the *CAN1* locus, generating *mfa1p::HIS3::can1Δ*. Mating type was switched to *MATα* following standard procedures and mutant *rpb* RNAPII subunit genes were introduced on *CEN URA3* plasmids, followed by screening for loss of WT *RPB CEN LEU2* plasmid. Each of these strains was then crossed with 1,200 *MATα* Kan' marked DAmP/deletion strains by pinning on solid media. Sporulation was induced and *MATα* haploid spores were selected by replica plating onto media containing canavanine (selecting *can1Δ* haploids) and lacking histidine (selecting *MATα* spores). Double mutant haploids were isolated on media containing nourseothrycin (selecting *NAT'*::*rpbΔ*) and G418 (selecting Kan' marked deletion/DAmP), and lacking uracil (selecting mutant RNAPII subunit on *CEN URA3* plasmid). Finally, double mutant colony sizes were extracted using imaging software, and genetic interaction scores computed using a statistical scoring scheme described in (Collins et al., 2006). Detailed E-MAP experimental procedures are described in (Collins et al., 2007b; Schuldiner et al., 2006).

ROC Curves

All library deletion mutants that exist in both the RNAPII pE-MAP and the previously published E-MAP (Collins et al., 2007b) were extracted, resulting in 404 mutants used for this analysis. First, pairs of proteins encoded by these genes were defined as physically interacting if listed with PE scores > 2 in (Collins et al., 2007a) and noninteracting if not. Next, Pearson correlation coefficients were computed for all pairwise combinations of the 404 library mutants, based on their pE-MAP profiles. An ROC curve was then generated to determine the power of these correlations to predict physical interactions between encoded proteins. Prior to generating the reference ROC curve from the published E-MAP, query strain profiles with more missing data than the sparsest RNAPII mutant were removed. Additionally, all query mutants that were also represented in the library mutant set were removed. From this filtered E-MAP, 53 rows (query mutant profiles) were randomly selected 1,000 times, and an ROC curve was generated for each run. The median AROC and corresponding curve are reported in Figure 2A.

Genome-wide Gene Expression

Strains were streaked from -80°C stocks onto plates and grown for 3 days. Liquid cultures were inoculated with independent colonies and grown overnight in Synthetic Complete (SC) medium: 2 g/l Drop out mix Complete and 6.71 g/l Yeast Nitrogen Base without AA, Carbohydrate & w/AS (YNB) from US Biological (Swampscott, USA) with 2% D-glucose. Overnight cultures were diluted to OD₆₀₀ = 0.15 in 1.5 ml fresh medium and grown at 30°C in a 24 well plate in a Tecan Infinite F200 under continuous shaking. Growth curves were made for the point mutant cultures (two cultures from two isolates) as well as for two WT inoculates, grown in parallel. Mutant and WT cells were harvested by centrifugation (6100 rpm, 3 min) at midlog phase at OD₆₀₀ = 0.6, and pellets were immediately frozen in liquid nitrogen after removal of supernatant. Up to eleven mutant strains could be grown on a single day. WT cultures were grown parallel to the point mutants to assess day-to-day variance. RNA isolation and purification was performed (Lenstra et al., 2011). In short, total RNA was prepared by phenol extraction and cleaned up using a customized Sciclone ALH 3000 Workstation. For each sample, external control poly-A+ RNAs were added in equimolar amounts to the total RNA to enable monitoring of global changes in mutants (van de Peppel et al., 2003). For each micro-array analysis, 1.25 μg labeled sample cRNA and 1.25 μg reference cRNA was hybridized per slide. Each hybridization performed within this project was subjected to a number of quality controls. Some of these are based on the data from one single hybridization, while others are based on comparing data from one single hybridization against the WT grown in parallel (Lenstra et al., 2011). Two channel microarrays were used. RNA isolated from a large amount of WT yeast from a single culture was used as a common reference. This common reference was used in one of the channels for each hybridization and used in the statistical analysis to obtain an average expression-profile for each mutant relative to the WT. Two independent cultures were hybridized on two separate microarrays. Each gene is represented twice on the microarray, resulting in four measurements per mutant.

Correlations

All correlations are Pearson correlation coefficients, unless otherwise noted. Genetic interaction correlations are based on the complete genetic interaction profiles, and gene expression correlations on the genome-wide expression profiles.

Identification of Functional Links between RNAPII Mutants and Protein Complexes

Published genetic interaction data (Costanzo et al., 2010) were scaled to E-MAP format using a nonlinear scaling method (Ryan et al., 2012), and processed to remove duplicate entries of genes represented by more than one allele (deletion, temperature sensitive or hypomorphic). If the deletion was present it was kept and all other alleles removed, and if no deletion existed the allele to keep was selected randomly. Pearson correlation coefficients were computed for each RNAPII mutant against all genes in this data set, based on the genetic interaction profiles with the intersecting library mutants. The genes were then grouped by membership in physical

complexes of their encoded proteins. Complex definitions were collected from (Benschop et al., 2010), and modified by first removing all binary complexes, and then excluding any gene assigned to more than one complex. To identify RNAPII allele-complex pairs which were significantly correlated, we used a one-sided Mann-Whitney U test to compare the correlations between each RNAPII mutant and the members of each complex to (i) the correlations between the same RNAPII mutant and all genes not in that complex, and to (ii) the correlations between the same complex and all other RNAPII mutants. The highest p value of the comparison to (i) or (ii) was recorded. False discovery rates (FDR) were computed using the method of Benjamini and Hochberg (Benjamini and Hochberg, 1995), and are reported in **Table S3**.

Chromosome Transmission Fidelity Assay

RNA polymerase mutants containing a chromosome VII fragment and *ade2-101* were generated. Red colony color is caused by accumulation of pigment due to a block in adenine production caused by the *ade2-101* (ochre) mutation. The block is relieved in the presence of the *SUP11* gene on the telocentric arm of the chromosome fragment, encoding an ochre-suppressing tRNA^{Tyr}. Cells that lose the chromosome VII fragment develop red color so that colonies exhibiting unstable inheritance of the chromosome fragment develop red sectors. Mutant strains were plated at 100-400 cells a plate on synthetic complete medium with only 20% adenine and left to grow into colonies for 7 days at 30°C. Replicates were counted to obtain the percentage of sectored colonies over total colonies for each mutant strain (Spencer et al., 1990; Yuen et al., 2007).

Primer Extension for Start Site Selection Analysis

Primer extension assays were performed as previously described (Ranish and Hahn, 1991) with the following modifications. Briefly, 30 µg total yeast RNA purified as described (Schmitt et al., 1990) was annealed with ³²P end-labeled oligo priming downstream of start sites in a 15 µl reaction volume instead of 10 µl, followed by reverse transcription reaction supplemented with RNase Inhibitor (Fermentas) by M-MLV Reverse Transcriptase (Fermentas), in a total volume of 45 µl instead of 30 µl. Products were separated on 8% polyacrylamide gels (19:1 acrylamide:bisacrylamide) containing 1 × TBE and 7M urea. *tfg2Δ261-273* mutant plasmids were gifted from the Hahn lab and used to integrate into yeast strain CKY283. Similarly, a *sua7-3* allele plasmid was a gift of the Hampsey lab, and were used to integrate *sua7-3* into CKY283 (**Table S1**). Reported values are means derived from at least three independent experiments.

Growth Rate Assay

Strains were cultured overnight in YPD, then resuspended to OD₆₀₀ = 0.3 in the morning, and grown at 30°C until reaching OD₆₀₀ = 0.7 - 1.0. These cultures were diluted in YPD to OD₆₀₀ = 0.05, 100 µl/well, in a Corning Costar 3631 clear bottom 96-well plate. Each culture was distributed into 3 wells as technical replicates. Plates were sealed with Corning 3930 clear polystyrene lids, and edges taped to avoid evaporation. Growth curves were measured at 30°C in a Tecan Infinite 200 PRO microplate reader, controlled by Tecan Magellan software. OD₆₀₀ measurements were performed every 15 min (top reading, 25 reads per well), with orbital shaking during incubation. Curves were fitted to the measured time series with a model-free spline method from the grofit package (Kahm et al., 2010), in the statistical computing environment R. Exponential growth rates were determined by grofit as the maximum slopes of the spline fits. For each set of technical replicates, the log₂ ratio of mutant to WT growth rate was computed. Reported averages and SD were calculated over the log₂ ratios in the three replicate sets.

Splicing Microarray Assay

Cultures were grown in rich medium according to standard techniques (Guthrie and Fink, 2002). Saturated cultures were diluted to OD₆₀₀ = 0.1 in the morning and allowed to grow at 30°C until reaching midlog phase (OD₆₀₀ = 0.5-0.7). Mutant strains and an isogenic WT were collected by centrifugation and snap frozen in liquid nitrogen. Where indicated, a WT log-phase culture was treated with mycophenolic acid (30 µg/ml) for 10 min, harvested using vacuum filtration and snap frozen in liquid nitrogen. Total cellular RNA was isolated using a hot acid phenol extraction generally as outlined in (Schmitt et al., 1990), but with some specific modifications (Bergkessel et al., 2011). cDNA from each strain was synthesized, and labeled with Cy3 or Cy5 according to the low-throughput sample preparation method (Pleiss et al., 2007). The optimized oligos (Pleiss et al., 2007) were robotically arrayed onto poly-L-lysine coated glass slides (slides from ThermoScientific C40-5257-M20) and slides were processed according to (DeRisi et al., 1997; Pleiss et al., 2007). Microarrays were scanned using Axon Instruments GenePix 4000B at 635 nm and 532 nm wavelengths and image analysis was done using Axon Instruments GenePix Pro version 5.1. Spots were manually removed from analysis if they contained uncharacteristically high background or obvious defects; the ratio of the median intensity for 635 nm and 532 nm was calculated for each remaining spot. Each biological replicate contains 6 technical replicates for each feature as well as dye-flipped replicates, which were combined and normalized (Pleiss et al., 2007). For all mutants, data from at least two biological replicates were used. To emulate a classic splicing measure that compares the levels of pre-mRNA and mature mRNA (Pikielny and Rosbash, 1985), we computed I/J ratios as log₂(Intron/Junction) = log₂(Intron_{mutant}/Intron_{WT})-log₂(Junction_{mutant}/Junction_{WT}) for each gene. Note that all E, I, J and I/J reported in figures and tables are log₂ ratios. The reported p values (**Table S5**) were calculated via one-sided Wilcoxon signed rank tests, comparing each I/J distribution (log₂) to that resulting from direct hybridization of cDNA from two separately grown WT cultures.

qPCR Assay for Splicing Efficiency

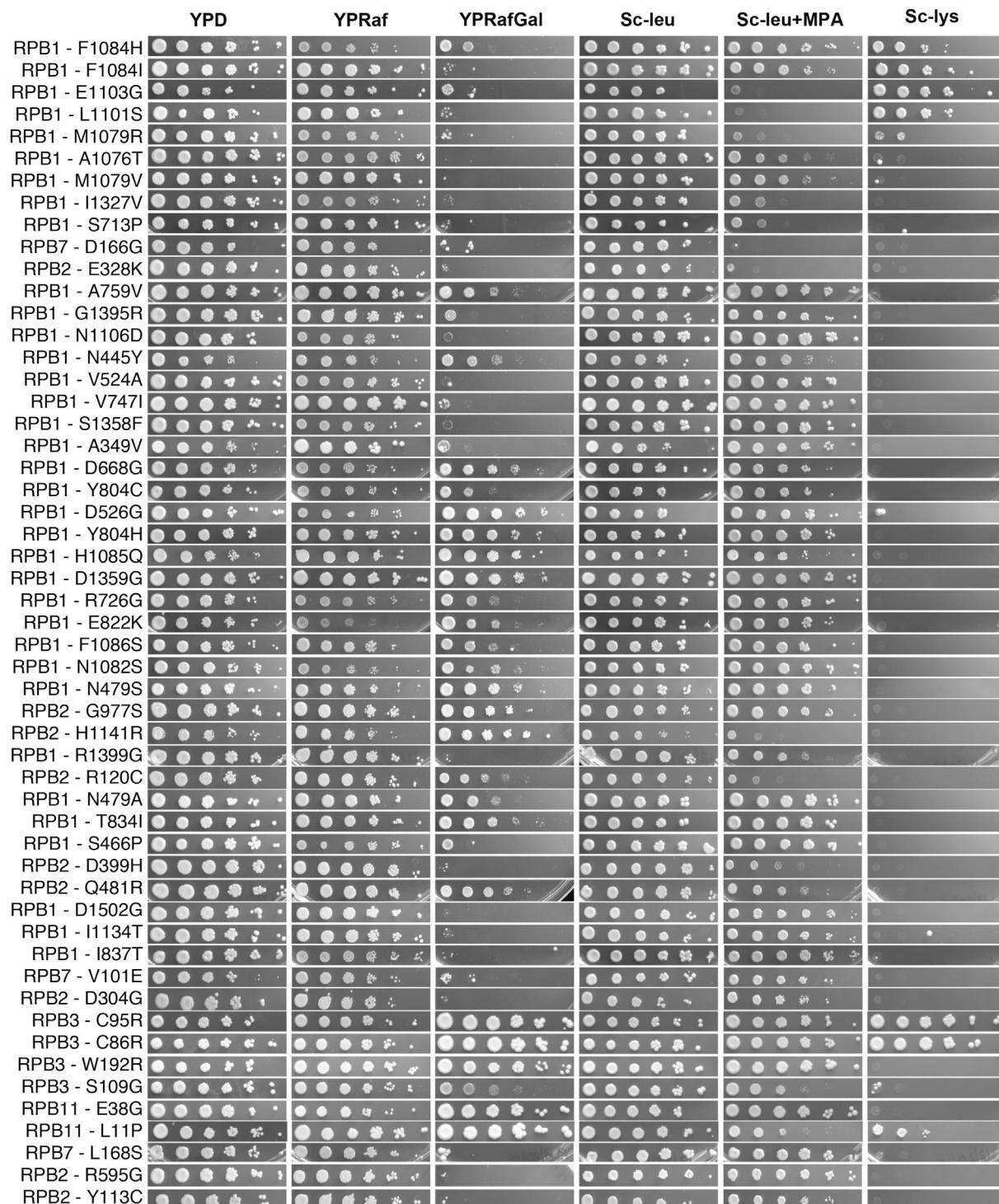
Splicing efficiency was measured by qPCR of cDNA as in (Pleiss et al., 2007) with modifications as described below. Ten µg of RNA were treated with 4U RQ1 DNaseI (Promega) in 16 µL according to manufacturer instruction (20 min 37°C). 4 µL Stop Solution was added, and DNase was inactivated by incubation at 65°C for 10 min. Samples were primed with random 9-mers in 40 µL (50mM Tris-HCl pH8.3, 75mM KCl, 3mM MgCl₂, 10mM DTT, 400ng dN9; 5 min 65°C; 5 min on ice) and divided into equal 40 µL reverse transcription reactions with M-MLV reverse transcriptase (+RT or -RT in 50 mM Tris-HCl pH8.3, 75 mM KCl, 3 mM MgCl₂, 10 mM DTT, 250 µM each dNTP; 2 hr 42°C). qPCRs were performed in 25 µL consisting of 1X Standard Taq Buffer and 1.25U Taq (New England Biolabs), 200 µM each dNTP, 0.0013% SYBR Green I (Sigma), 600nM each primer, DMSO as needed (Table S5), and 25 × - 2500 × dilution of cDNA (-RT samples diluted equivalently to match +RT). Amplification reactions were performed on a C1000 ThermoCycler (BioRad) (94°C 3 min and 35 cycles of 94°C 15 s, 50°C 15 s, 72°C 15 s, plate read). After cycling, each qPCR run was finished with a melt curve to confirm homogeneity of the amplified product. Starting quantity was calculated from a genomic DNA standard curve for each primer set; standard curve reactions not falling in the linear range were removed manually. Two technical replicates apiece were performed for 2–8 biological replicates (F1086S n = 8, H1085Q n = 5, sub1Δ n = 2, MPA n = 3, E1103G n = 6, G1097D n = 5). -RT control samples yielded negligible amplification (data not shown). Each gene was measured using primer sets specific for pre-mRNA and total mRNA (Table S5). To generate the graph in Figure S6A, the pre-mRNA/total mRNA ratio for each mutant was normalized to a within-experiment WT before taking the mean of the biological replicates.

Sorting Array Mutants on Their Interactions with Fast and Slow Mutants

Mean genetic interaction scores with the fast or slow RNAPII groups were calculated for all array mutants that have at least two scores in each group. The 404 mutants with different signs of mean (fast) and mean (slow) were then sorted by the difference of the means.

SUPPLEMENTAL REFERENCES

- Amberg, D.C., Burke, D.J., and Strathern, J.N. (2005). Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual, 2005 Edition, 2005 edn (Cold Spring Harbor, NY: Cold Spring Harbor Press).
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc., B* 57, 289–300.
- Bergkessel, M., Whitworth, G.B., and Guthrie, C. (2011). Diverse environmental stresses elicit distinct responses at the level of pre-mRNA processing in yeast. *RNA* 17, 1461–1478.
- Boeke, J.D., Trueheart, J., Natsoulis, G., and Fink, G.R. (1987). 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. *Methods Enzymol.* 154, 164–175.
- Collins, S.R., Schuldiner, M., Krogan, N.J., and Weissman, J.S. (2006). A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. *Genome Biol.* 7, R63.
- DeRisi, J.L., Iyer, V.R., and Brown, P.O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278, 680–686.
- Greger, I.H., Aranda, A., and Proudfoot, N. (2000). Balancing transcriptional interference and initiation on the GAL7 promoter of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 97, 8415–8420.
- Guthrie, C., and Fink, G.R. (2002). Guide to Yeast Genetics and Molecular and Cell Biology (San Diego, CA: Academic Press).
- Kahm, M., Hasenbrink, G., Lichtenberg-Frathe, H., Ludwig, J., and Kschischko, M. (2010). grofit: Fitting Biological Growth Curves with R. *J. Stat. Softw.* 33, 1–21.
- Lenstra, T.L., Benschop, J.J., Kim, T., Schulze, J.M., Brabers, N.A., Margaritis, T., van de Pasch, L.A., van Heesch, S.A., Brok, M.O., Groot Koerkamp, M.J., et al. (2011). The specificity and topology of chromatin interaction pathways in yeast. *Mol. Cell* 42, 536–549.
- Pikielny, C.W., and Rosbash, M. (1985). mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. *Cell* 41, 119–126.
- Pleiss, J.A., Whitworth, G.B., Bergkessel, M., and Guthrie, C. (2007). Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. *PLoS Biol.* 5, e90.
- Ranish, J.A., and Hahn, S. (1991). The yeast general transcription factor TFIIA is composed of two polypeptide subunits. *J. Biol. Chem.* 266, 19320–19327.
- Schmitt, M.E., Brown, T.A., and Trumper, B.L. (1990). A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 18, 3091–3092.
- Sikorski, R.S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* 122, 19–27.
- Simchen, G., Winston, F., Styles, C.A., and Fink, G.R. (1984). Ty-mediated gene expression of the LYS2 and HIS4 genes of *Saccharomyces cerevisiae* is controlled by the same SPT genes. *Proc. Natl. Acad. Sci. USA* 81, 2431–2434.
- van de Peppel, J., Kemmeren, P., van Bakel, H., Radonjic, M., van Leenen, D., and Holstege, F.C. (2003). Monitoring global messenger RNA changes in externally controlled microarray experiments. *EMBO Rep.* 4, 387–393.
- Yuen, K.W., Warren, C.D., Chen, O., Kwok, T., Hieter, P., and Spencer, F.A. (2007). Systematic genome instability screens in yeast and their potential relevance to cancer. *Proc. Natl. Acad. Sci. USA* 104, 3925–3930.

**Figure S1. Complete Collection of Spot Tests for Identification of Gal^R, MPA Sensitivity, and Spt⁻ Phenotypes, Related to Figure 1**

Spot tests on relevant media were carried out to identify RNAPII mutants that exhibit at least one of the three phenotypes, as described in Figure 1 and Experimental Procedures. The order of RNAPII mutants follows that of the clustered pE-MAP (Figure S2 and Data S1).

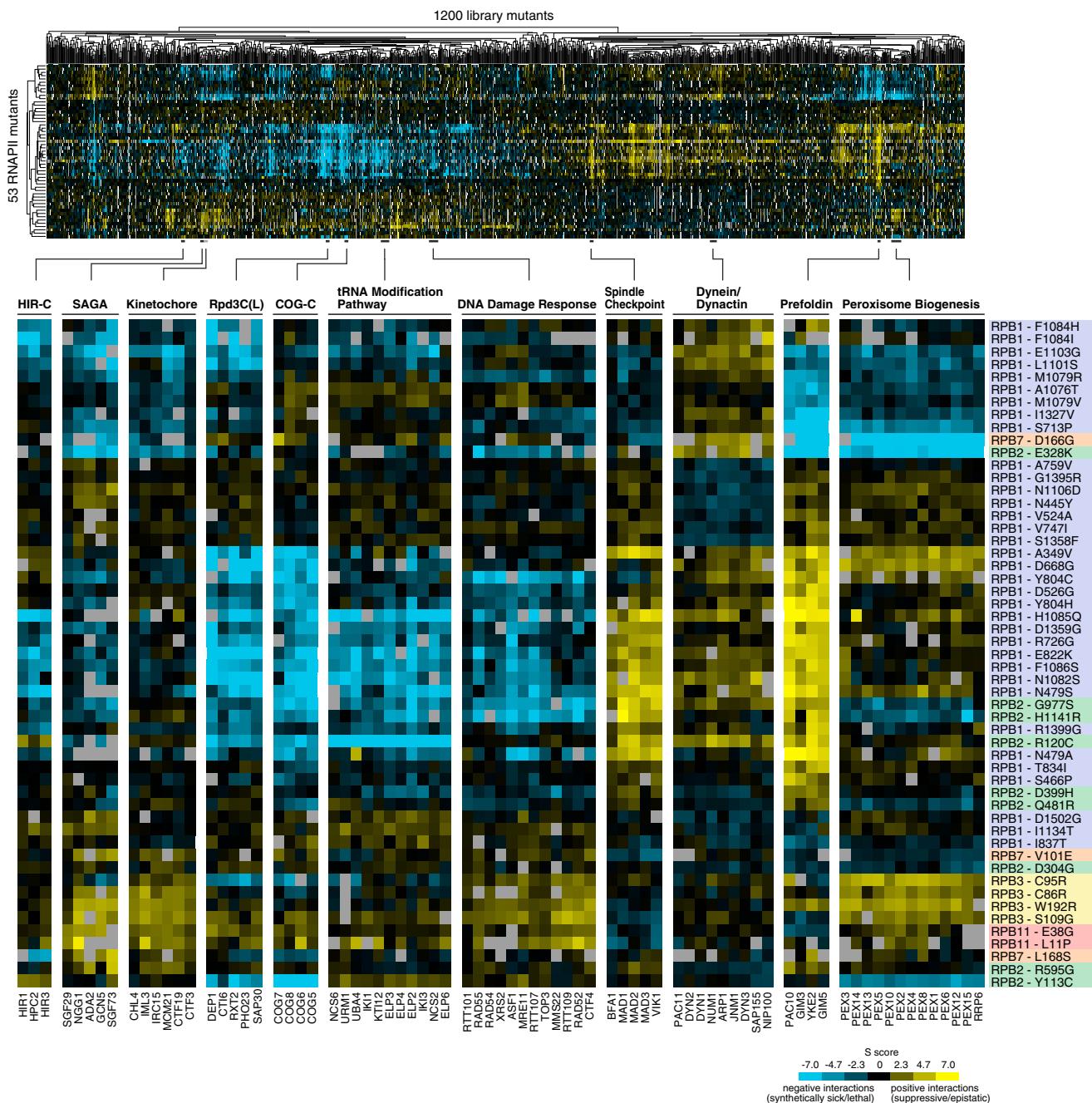
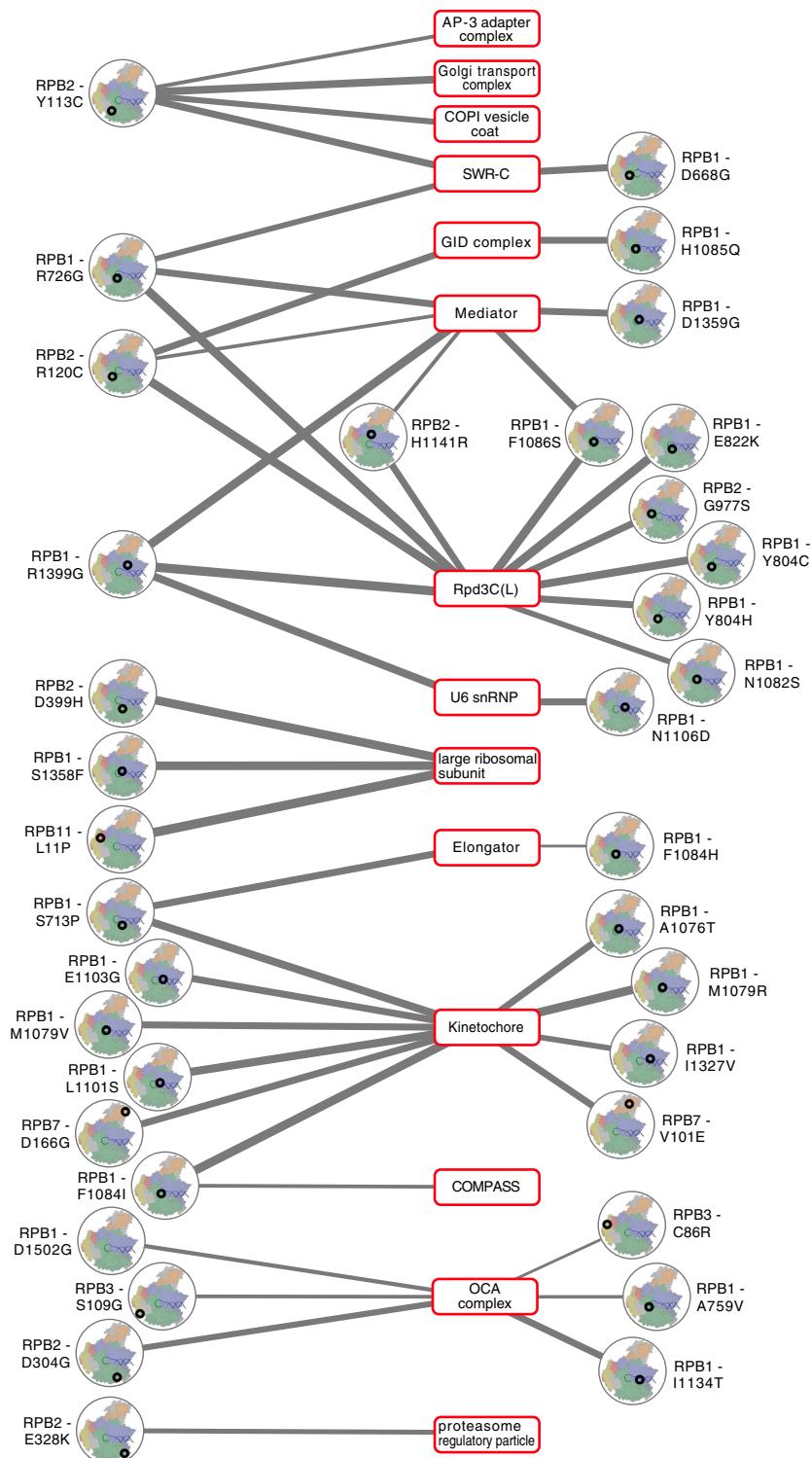


Figure S2. A High-Resolution Genetic Interaction Map of RNAPII Point Mutants, Related to Figure 2

The 53 point mutants described in Figure 1B were genetically screened against a library of 1200 mutants carrying single gene deletions or DAmP alleles (Table S1). The resulting pE-MAP reports on a total of 59,534 genetic interactions between single residue point mutations and deletions of nonessential genes or DAmP alleles of essential genes. The dendograms organize the mutants functionally and were generated by 2-D hierarchical clustering. Several representative clusters of library mutants that belong to the same complex or pathway are highlighted. The background colors of RNAPII mutant labels correspond to subunit and match the color scheme in Figure 1.

**Figure S3. Functional Connections between RNAPII Mutants and Protein Complexes, Related to Figure 3**

Functional associations between RNAPII mutants and protein complexes were determined as described in Figure 3 and Experimental Procedures. Edge widths correspond to the statistical significance of connections, and only edges with a false discovery rate (FDR) < 0.1 are displayed (Table S3).

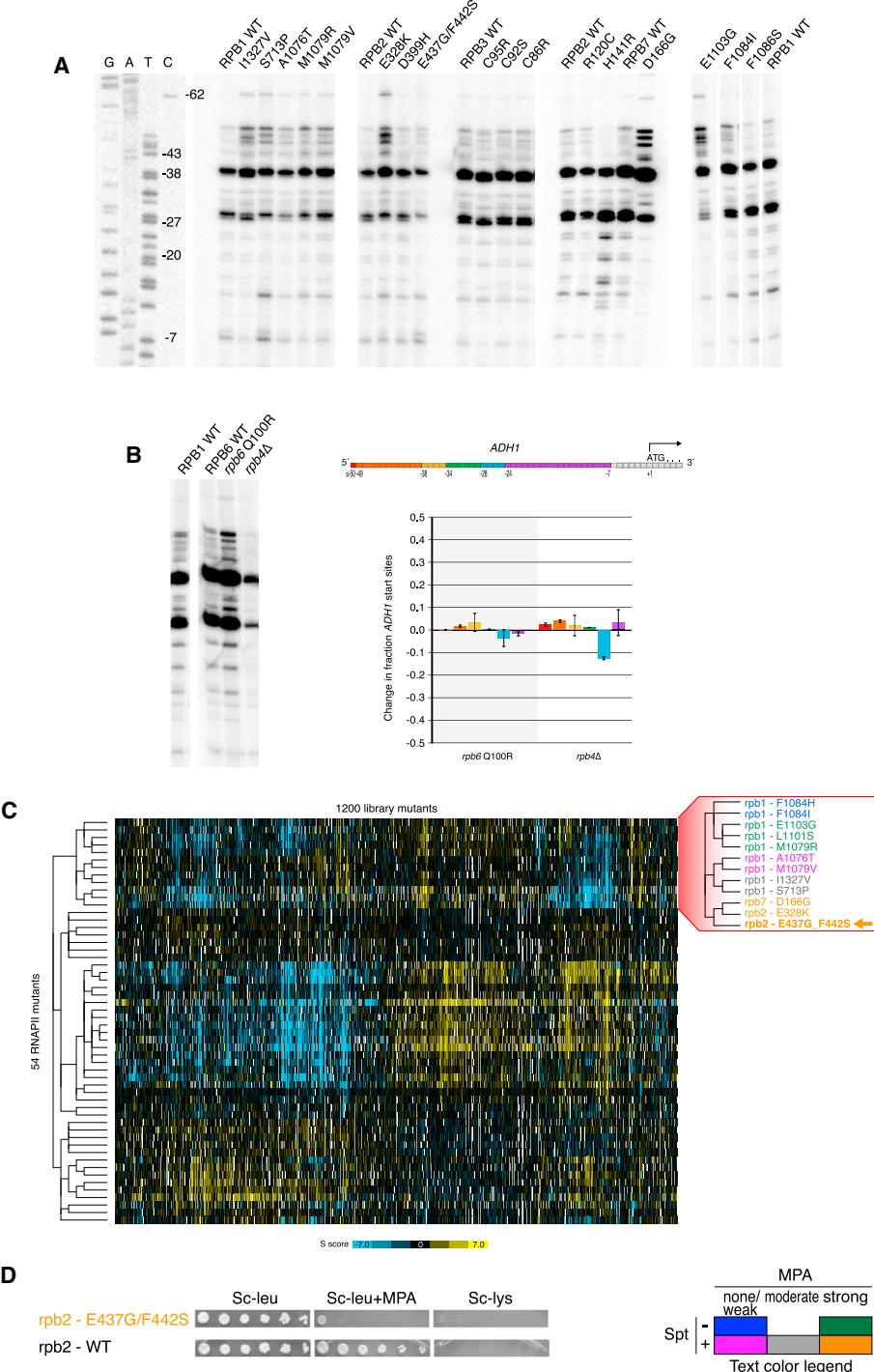


Figure S4. Primer Extension Analysis at *ADH1* to Identify the Effect of RNAPII Mutations on Start Site Selection and Data Relating to *rpb2* E437G/F442S, Related to Figure 4

(A) Sequencing gel for primer extension analysis. Lanes 1-4 correspond to Sanger dideoxy sequencing reactions as reference ladders, and the following lanes carry reverse transcription products from RNAPII alleles as indicated (Table S4).

(B) Primer extension analysis at *ADH1* to map transcription start sites for *rpb4Δ* and *rpb6* Q100R mutations. The colors of the bars correspond to the sequence windows indicated in the *ADH1* schematic (top) and the heights specify the fraction change of transcription start in the mutant compared to WT. Error bars represent SD. Means and SD were derived from three independent experiments.

(C) Clustering of *rpb2* E437G/F442S in pE-MAP.

(D) MPA sensitivity and Spt⁻ phenotypes of *rpb2* E437G/F442S.

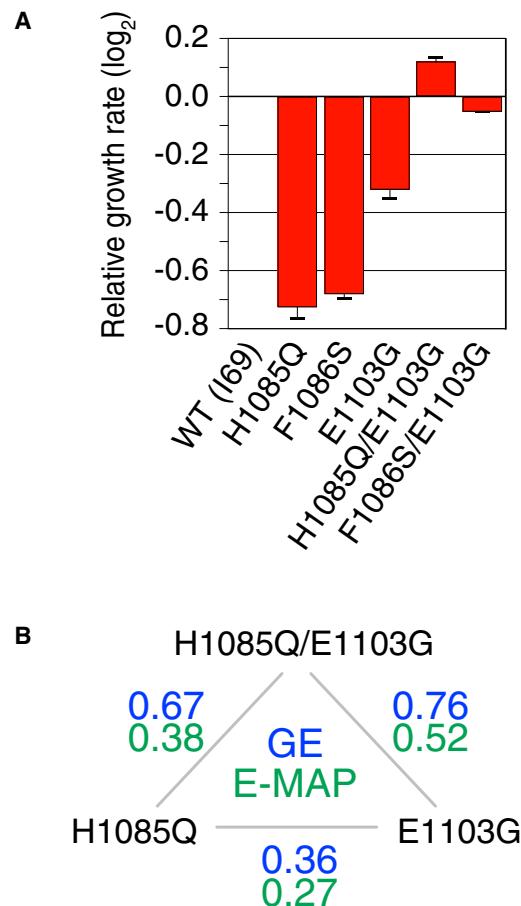


Figure S5. RNAPII Trigger Loop Double- and Single-Mutant Growth Rates and Profile Correlations between *rpb1* E1103G, H1085Q, and E1103G/H1085Q, Related to Figure 5

(A) Growth rates relative to WT of *rpb1* E1103G, F1086S, H1085Q, E1103G/F1086S and E1103G/H1085Q. Error bars represent SD. Means and SD were derived from three technical replicates.

(B) pE-MAP and gene expression (GE) profile correlations between *rpb1* E1103G, H1085Q and E1103G/H1085Q.

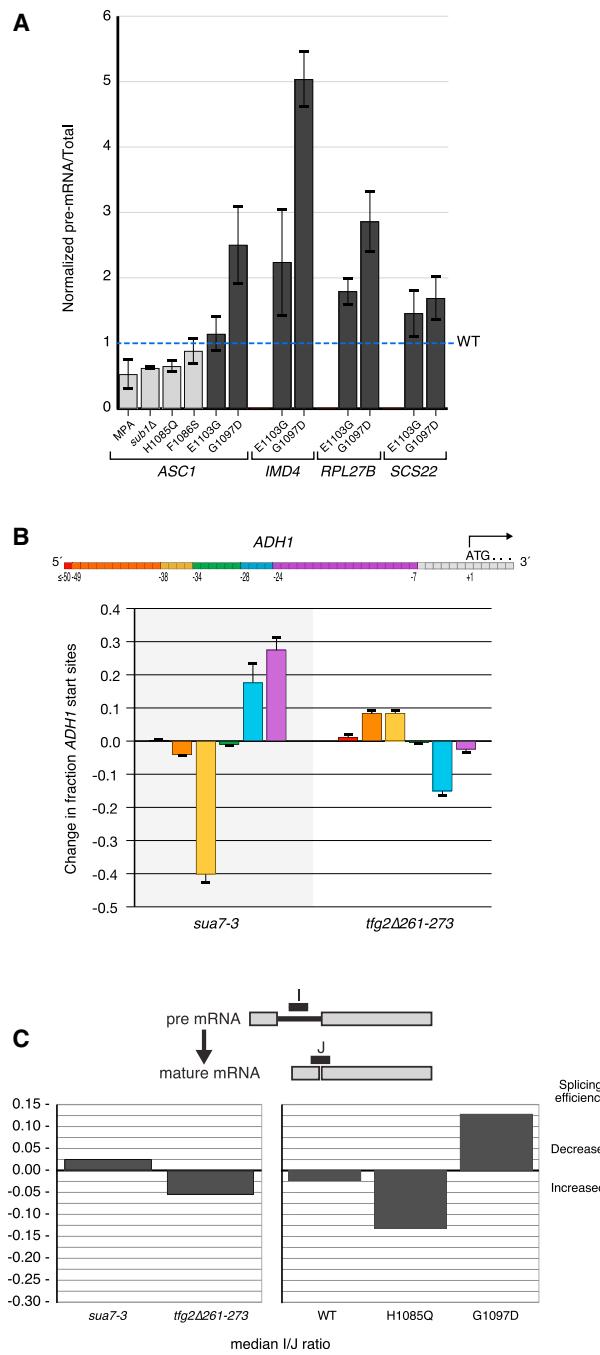


Figure S6. Confirmation of RNAPII Single-Mutant Splicing Phenotypes and Interrogation of Potential Connection between Splicing and Start Site Selection, Related to Figures 6 and 7

(A) Pre-mRNA and total mRNA for transcripts from *ASC1*, *IMD4*, *RPL27B* and *SCS22* were measured in the indicated strains and conditions. Shown are pre-mRNA/total mRNA ratios, normalized to that of an untreated WT. Light gray bars are data points that exhibited improved splicing in the microarray; dark gray bars are data points that exhibited a splicing defect in the microarray. Error bars represent SD of 2–8 biological replicates. “MPA” refers to a 10 min treatment of a WT strain with mycophenolic acid.

(B) Primer extension analysis at *ADH1* to map transcription start sites for mutations in TFIIB (*sua7-3*) and TFIIF (*tfg2Δ261-273*). The colors of the bars correspond to the sequence windows indicated in the *ADH1* schematic (top) and the heights specify the fraction change of transcription start in the mutant compared to WT. Error bars represent SD. Means and SD were derived from three independent experiments.

(C) TFIIB (*sua7-3*) and TFIIF (*tfg2Δ261-273*) mutants were analyzed on splicing microarrays, as described in Figure 6 (Table S5). Median $I/J \log_2$ ratios for the complete profiles of the TFII mutants are shown in the left panel, with WT, a slow and a fast *rpb1* mutant included for reference in the right panel.

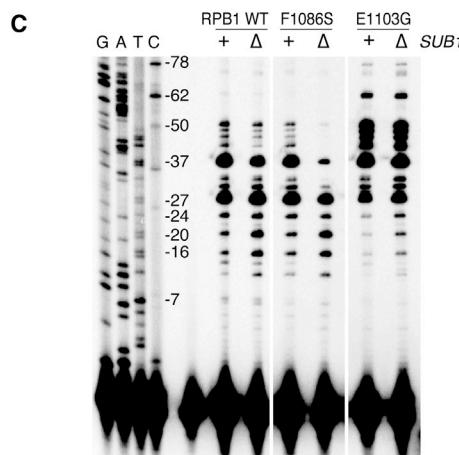
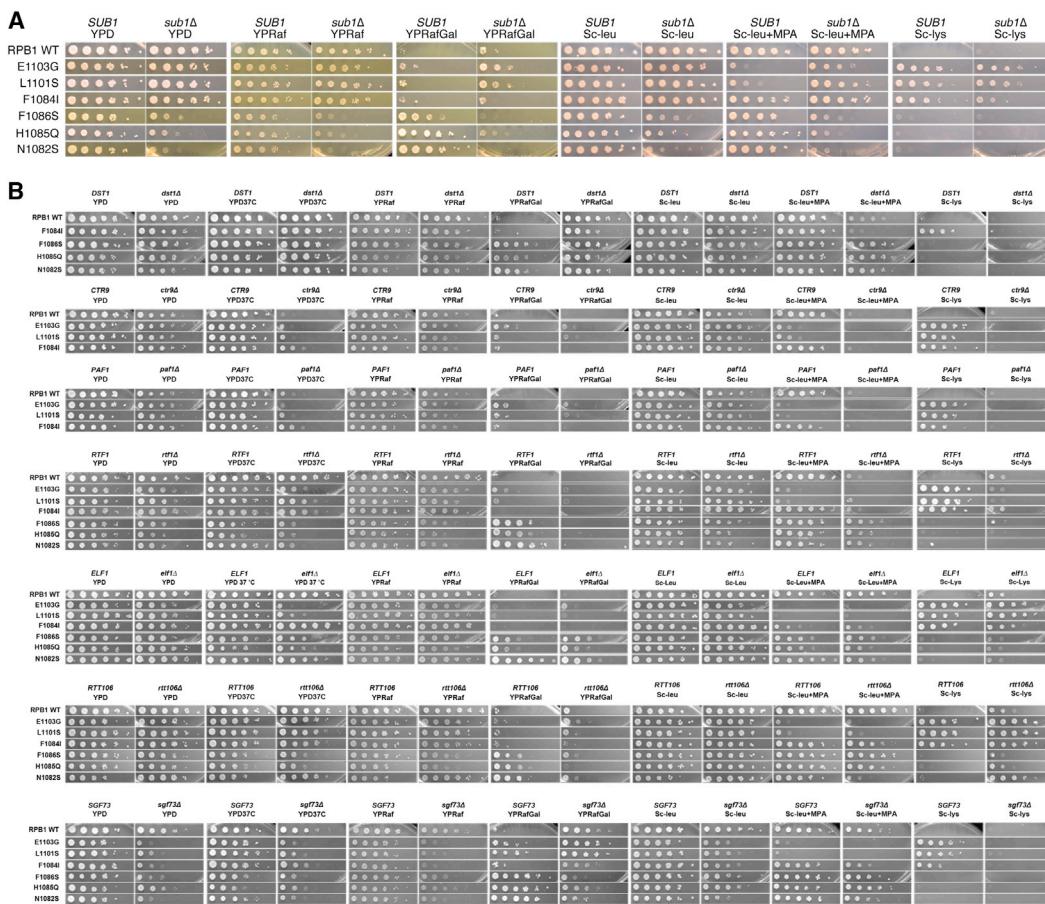


Figure S7. Spot Tests to Determine Gal^R, MPA Sensitivity, and Spt⁻ Phenotypes of *sub1Δ* or Other Deletion Mutants with Fast and Slow RNAPII Alleles and Primer Extension of *sub1Δ* Mutants, Related to Figure 7

(A) Spot tests on relevant media were carried out to determine the effects of *sub1Δ* on the transcription-related phenotypes of fast and slow RNAPII alleles, as described in Figure 1 and Experimental Procedures.

(B) Spot tests on relevant media were carried out to determine the effects of deleting known or predicted transcription factors on the transcription-related phenotypes of fast and slow RNAPII alleles, as described in Figure 1 and Experimental Procedures.

(C) Primer extension analysis at *ADH1* to identify the effects of *sub1Δ* on start site selection. Lanes 1-4 correspond to Sanger dideoxy sequencing reactions as reference ladders and the following lanes carry reverse transcription products from RNAPII mutants in *SUB1* and *sub1Δ* background as indicated (Table S4).