Ssl2/TFIIH function in Transcription Start Site Scanning by RNA Polymerase II in Saccharomyces cerevisiae

Tingting Zhao¹,², Irina O. Vvedenskaya³, William K.M. Lai⁴, Shrabani Basu¹, B. Franklin Pugh⁴, Bryce E. Nickels³ and Craig D. Kaplan¹.

¹Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260, USA.
²Current address: Genomics and Bioinformatics Hub, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA 02115, USA.
³Department of Genetics and Waksman Institute, Rutgers University, Piscataway, NJ 08854, USA.
⁴Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA.

ABSTRACT

In Saccharomyces cerevisiae, RNA Polymerase II (Pol II) selects transcription start sites (TSS) by a unidirectional scanning process. During scanning, a preinitiation complex (PIC) assembled at an upstream core promoter initiates at select positions within a window ~40-120 basepairs downstream. Several lines of evidence indicate that Ssl2, the yeast homolog of XPB and an essential and conserved subunit of the general transcription factor (GTF) TFIIH, drives scanning through its DNA-dependent ATPase activity, therefore potentially controlling both scanning rate and scanning extent (processivity). To address questions of how Ssl2 functions in promoter scanning and interacts with other initiation activities, we leveraged distinct initiation-sensitive reporters to identify novel ssl2 alleles. These ssl2 alleles, many of which alter residues conserved from yeast to human, confer either upstream or downstream TSS shifts at the model promoter ADH1 and genome-wide. Specifically, tested ssl2 alleles alter TSS selection by increasing or narrowing the distribution of TSSs used at individual promoters. Genetic interactions of ssl2 alleles with other initiation factors are consistent with ssl2 allele classes functioning through increasing or decreasing scanning processivity but not necessarily scanning rate. These alleles underpin a residue interaction network that likely modulates Ssl2 activity and TFIIH function in promoter scanning. We propose that the outcome of promoter scanning is determined by two functional networks, the first being Pol II activity and factors that modulate it to determine initiation efficiency within a scanning window, and the second being Ssl2/TFIIH and factors that modulate scanning processivity to determine the width of the scanning widow.
INTRODUCTION

Transcription of eukaryotic protein-coding genes is carried out by RNA polymerase II (Pol II) in three sequential steps: initiation, elongation and termination. Accurate initiation requires minimally the assistance of five general transcription factors (GTFs) TFIIB, TFIID, TFIIE, TFIIF and TFIIH, which together with Pol II, comprise the basal transcription machinery. At the beginning of transcription, this machinery assembles at a defined DNA region for each transcript called a promoter, melts the double-stranded DNA yielding a region of unwound DNA forming a “transcription bubble”. Within this bubble a position or positions will be identified to serve as transcription start sites (TSS). Initial promoter melting appears to occur stereotypically ~20-25 nucleotides downstream of promoter elements such as the TATA box across eukaryotes, though most promoters lack a TATA box or other strong sequence signature of promoter elements. During initiation, the process of TSS selection determines the identity and distribution of transcript isoforms that differ by their 5’ ends. Differences in 5’ UTR can alter transcript properties such as translation efficiency or transcript stability through differences in sequence or RNA secondary structure. Furthermore, in conjunction with activators and coactivators, the efficiency of the initiation process will also establish mRNA synthesis rates. How TSS selection is governed by these factors is not well understood for the majority of eukaryotic promoters that utilize multiple TSSs.

Transcription initiation by S. cerevisiae Pol II has been the subject of extensive analysis both in vivo and in vitro, and thus provides a powerful model for system for mechanistic studies of TSS selection. TSS selection by S. cerevisiae Pol II occurs over a range of positions located ~40-120 bp downstream of the core promoter region. Numerous lines of evidence suggest TSS selection by S. cerevisiae Pol II involves a unidirectional scanning mechanism in which the preinitiation complex (PIC) assembles at an upstream core promoter and interrogates consecutive downstream positions for usable TSSs. TFIIH is proposed to drive Pol II scanning through ATP-dependent DNA translocase activity. An optical-tweezer-based single molecule analysis of reconstituted S. cerevisiae PICs indicated that an ATP/dATP-induced activity within the PIC causes shortening of the distance between upstream and downstream DNA. This shortened distance approximates the distance downstream from TATA-elements where TSSs are positioned in yeast (40-120nt) and suggests downstream DNA movement and compaction during promoter scanning by the PIC. Separately, a magnetic-tweezer-based single molecule assay suggested that an initial melted region of 6nt (a “bubble”) is the direct consequence of TFIIH’s ATPase activity. Due to inability of magnetic-tweezers to detect DNA compaction in the particular setup used, how the Pol II machinery reaches downstream TSSs, whether through generation of a large bubble or translocation of a small bubble was not clear. Nevertheless, both studies agree that an ATP-dependent PIC activity for promoter opening is likely Ssl2 within TFIIH, which has been demonstrated as a DNA translocase within purified TFIIH in vitro.

In support of Ssl2/TFIIH’s role in movement of downstream DNA toward the PIC, ssl2 mutants have been identified as altering TSS selection at ADH1 and showed genetic interactions with sua7 (TFIIB) mutants. Specifically, the identified ssl2 mutants shifted TSSs upstream at ADH1. Polar shifts in TSSs distributions have been observed in mutants within Pol II, the GTFs TFIIB and TFIIF, and the PIC cofactor Sub1. In a promoter scanning initiation mechanism, altering initiation efficiency is predicted to alter TSS distributions in a polar fashion when initiation efficiency increases or decreases. We have recently observed polar (directional) shifts at essentially all promoters examined across the genome in yeast for tested Pol II and GTF mutants, as predicted for scanning operating universally across promoter classes. We found that hyperactive Pol II catalytic mutants shifted TSSs upstream at promoters genome-wide, consistent with a higher probability of initiation at every TSS, and thus, initiation happening on average earlier in the scanning process. Conversely, hypoactive Pol II catalytic mutants shift TSS distributions downstream at promoters genome-wide, consistent with initiation happening later in the scanning process. Our previous data on genetic interactions between Pol II and TFIIF or TFIIB support the idea that these mutations altered initiation additively, consistent with their acting in the same pathway during scanning. Tested TFIIB mutants appeared to reduce initiation efficiency while tested TFIIF mutants appeared to increase initiation efficiency. Consistent with...
this idea, a double mutant between TFIIF and a hyperactive Pol II mutant had stronger effects on TSS
shifts than either mutant alone across all promoters. Pol II mutants are proposed to control initiation
efficiency because active site residues important for catalytic activity alter TSS distributions correlating
with the strengths of their catalytic defects in RNA chain elongation. Initiation by promoter scanning
should be dependent both on Pol II catalytic rate together and by whichever factors control the actual
scanning step, i.e. presumptively the rate and processivity of TFIIH as the putative scanning
translocase. Therefore, to understand how promoter scanning works, it is critical to understand how
TFIIH contributes and how its activity is regulated within the PIC.

We have described the scanning process previously using a “Shooting Gallery” analogy\(^\text{13, 40}\). In this
model, initiation is controlled by the rate (TFIIH’s translocase activity) in which a “target” (TSS) passes
the “line of fire” (the Pol II active site) along with the rate of firing (Pol II catalytic activity) and the size of
the target (innate sequence strength). Together, the cooperation and competition between these rates
determines the probability a target is hit (initiation happening). Alteration of enzymatic activities
supporting initiation, either the Pol II active site or TFIIH translocation should have predictable effects
on overall TSS distributions when initiation proceeds by scanning. In addition to the TFIIH translocation
rate, it is predicted that processivity of TFIIH DNA translocation should strongly modulate scanning, but
in ways distinct from controlling innate initiating efficiency. Here, TFIIH processivity would control the
probability that a TSS could be reached during a scanning foray from a core promoter, which appears
to be facilitated by Ssl2’s translocase activity\(^\text{13}\). Optical tweezer experiments are consistent with TFIIH
within the PIC having median processivity on the order of ~90 basepairs, consistent with the average
distance of TSSs downstream of yeast TATA boxes. However, purified holo-TFIIH from yeast has much
reduced measured processivity and human TFIIH has essentially none\(^\text{14}\). Given that TFIIH activity is
predicted to be altered extensively by cofactor interactions in both transcription and nucleotide excision
repair (NER), it will be important to understand TFIIH functions within the PIC in vivo. How alterations to
Ssl2/TFIIH translocase activity control TSS distributions has not been extensively investigated.

To test if distinct alterations to Ssl2 function have broad effects on promoter scanning and TSS
selection, we used existing and newly identified ssl2 alleles to examine their effect on TSS distributions
genome-wide. Our novel alleles were identified through use of genetic reporters we have developed
and found to be sensitive to different kinds of initiation defects\(^\text{41}\). We have found that ssl2 alleles affect
TSS distributions for the majority of promoters in yeast and for all promoter classes. Furthermore, we
find that ssl2 alleles alter TSS selection distinctly from how changes to the Pol II active site alter TSS
selection, consistent with a distinct role for Ssl2 in promoter scanning. ssl2 alleles appear to extend or
truncate scanning windows at promoters genome-wide, consistent with increase or decrease in the
processivity of scanning. Scanning-window truncating alleles map throughout the Ssl2 structure,
consistent with a hypothetical loss of functions in Ssl2 DNA translocase enzymatic activity that lead to
decreased TFIIH processivity. Conversely, scanning-window extending ssl2 alleles are much more
localized within the Ssl2 N-terminus, including conserved residues within regions that connect Ssl2
helicase domains to the TFIIH component and regulator of Ssl2 activity Tfb2. Our alleles are consistent
with alteration to an Ssl2 regulatory domain resulting in modulated translocase activity or TFIIH
processivity. We further test our model for initiation by promoter scanning through examination of
genetic interactions of initiation-altering Pol II/GTF alleles and ssl2 alleles. The genetic interactions
between Pol II/GTF and Ssl2 alleles support the idea of two major networks controlling TSS selection in
S. cerevisiae. One network shapes TSS distributions through affecting initiation efficiency, represented
by Pol II, TFIIB and TFIIH functions. A second network appears to alter TSS distributions through
regulating TFIIH’s processivity, and includes Ssl2, Sub1, and potentially TFIIF.

RESULTS
Existing SSL2 alleles show transcription-dependent growth phenotypes and distinct TSS usage
patterns
To understand how TSSs are identified by promoter scanning and the potential roles for TFIIH, we first
examined previously identified ssl2 mutants\(^\text{19, 42-44}\) for transcription-related phenotypes that we have
demonstrated are predictive of specific initiation defects (Figure 1A). Two such phenotypes relate to
altered initiation at the IMD2 gene, whose promoter is regulated by a transcription start site switch
We have previously shown that tested mutants that shift TSSs upstream due to altered promoter scanning result in an inability to express a functional IMD2 transcript, causing sensitivity to the IMPDH inhibitor mycophenolic acid (MPA)\(^{13,39,41}\) (Figure 1A, IMD2). In the presence of GTP starvation induced by MPA, WT strains shift start site usage at IMD2 from TATA-proximal GTP-initiating TSSs to a downstream ATP-initiating TSS. This shift results in a functional IMD2 transcript that is required for yeast to survive MPA. Catalytically hyperactive Pol II mutants (termed GOF for “Gain-of-Function”) do not shift TSS usage to the downstream functional IMD2 TSS but instead shift to intermediately positioned non-functional upstream sites, rendering yeast sensitive to MPA (MPA\(^{S}\))\(^{41}\). Pol II GOF mutants with MPA\(^{S}\) phenotypes and a tfg2 MPA\(^{S}\) mutant were additionally found to shift TSS distributions upstream at ADH1 and subsequently genome-wide\(^{13,39,47}\). Correlation between strength of MPA-sensitive phenotypes and quantitative upstream TSS shifts at ADH1 and genome-wide suggest that MPA-sensitivity is a strong predictor for upstream-shifting TSS mutants. Conversely, we found previously that mutants shifting TSSs downstream (reduced catalytic activity loss of function “LOF” Pol II mutants) constitutively express IMD2 in the absence of using MPA as inducer\(^{47}\). Pol II TSS downstream shifting phenotypes at IMD2 can be detected using a reporter allele where IMD2 is replaced by HIS3, placing HIS3 under control of IMD2 promoter and TSS selection (Figure 1A, IMD2\(^\Delta::\text{HIS3}\))\(^{41}\). Indeed, these same LOF Pol II mutants shift TSS distributions downstream at ADH1 and genome-wide\(^{13,39,47}\). These previous results suggest we have phenotypes predictive of alterations to promoter scanning in both directions and can form the basis of a system to characterize mutants across the genome for effects on TSS selection by promoter scanning.

We used site-directed mutagenesis and plasmid shuffling to recreate and phenotype five previously described ssl2 mutants, reasoning that this would allow us a first glimpse at the effects and potential diversity present in these classic alleles. These alleles are ssl2-rtt (ssl2 E556K)\(^{44}\), ssl2-DEAD (ssl2 V490A/H491D) and SSL2-1 (ssl2 W427L)\(^{42}\), ssl2-508 (ssl2 H508R)\(^{19}\), and rad25-ts24 (ssl2 V552I/E556K)\(^{43}\) (Figure 1B,1C). Analysis of these five ssl2 mutants showed phenotypes consistent with altered TSS selection (Figure 1C). First, ssl2-DEAD and ssl2-508 exhibited strong and weak MPA\(^{S}\) phenotypes, respectively. These results were consistent with prior analysis of initiation at ADH1 in these mutants\(^{19}\) suggesting that our genetic phenotypes using IMD2 are predictive of potentially global alterations to TSS selection. SSL2-1 exhibited a His\(^{\ast}\) phenotype, which is predictive of downstream shifts in TSS usage and consistent with its identification as a dominant mutant bypassing an inhibitory stem loop in the his4-316 mRNA\(^{42}\). We now can rationalize the original “Suppressor of Stem-loop (SSL) phenotype of SSL2-1 as usage of TSS downstream or within the inhibitory 5’ stem loop his4-316 sequence (though not apparent in Gulyas et al\(^{42}\)). ssl2-rtt and rad25-ts24 show temperature-sensitive phenotypes as expected; however, assayed transcription-related plate phenotypes were not observed. Due to the absence of His\(^{\ast}\) or MPA\(^{S}\) phenotypes, we predicted that ssl2-rtt or rad25-ts24 alleles would not shift TSS usage. Notably, there was no lys2-1280 Spt\(^{\ast}\) phenotype observed among these five existing ssl2 alleles, in contrast to our previous observation of an Spt\(^{\ast}\) phenotype in a subset of MPA\(^{S}\) Pol II TSS alleles\(^{49,50}\), our first indication that ssl2 alleles may alter TSS selection in a distinct fashion from Pol II mutants.

To quantitatively examine TSS usage of these ssl2 mutants, we chose ADH1 initiation. ADH1 has been widely used as a model gene for TSS studies in S. cerevisiae. It contains two major TSSs that are 27 and 37 nucleotides upstream of the start codon (Figure 1D). Using primer extension, transcription products of these two TSSs appear as bands of differing mobility on denaturing polyacrylamide gels (Figure 1 – Figure supplement 1A, left panel WT). Other TSSs positions show minor usage. In most studies, the two major starts are compared qualitatively, but such comparisons miss meaningful alterations that may tell us about initiation mechanisms. To establish that genetic phenotypes using IMD2 correlate with altered TSS selection elsewhere in the genome our quantitative analysis of the ADH1 promoter divides TSSs observed by primer extension into six bins from upstream to downstream, in which bin three and five each contain the TSS for one of the major ADH1 mRNA isoforms (Figure 1D, and Figure 1 – Figure supplement 1A left panel WT). In order to compare a mutant’s TSS distribution to that of WT, distributions are normalized, and the WT distribution is subtracted from tested mutant distributions bin by bin (Figure 1 – Figure supplement 1A, middle and right panel). Negative or positive values indicate that the mutant has relatively lower or higher usage for TSSs in that
We additionally constructed and tested human disease-related XPB mutations in the yeast SSL2 system, together with mutations in the ultra-conserved Arginine-Glutamic Acid-Aspartic Acid (RED) motif. As human disease-related alleles are many times found in conserved residues, we reasoned that some may have effects on Ssl2 biochemistry detectable in our sensitive system. We examined four human disease-related mutants that confer distinct inherited autosomal recessive disorders xeroderma pigmentosum (XP), trichothiodystrophy (TTD), and Cockayne syndrome (CS). Of these Q592*, which is also lethal and confers a C-terminally truncated Ssl2 protein, confers lethality and reduced TSS usage (Figure 1B). T176P confers little if any growth defects and no MPA allele. The SSL2 phenotype shifted downstream at ADH1 (Figure 1G). Mutant Y750*, which mimics a disease-related C-terminally truncated protein, shows a mild to moderate level MPA phenotype and shifts TSS distribution upstream at ADH1 (Figure 1G). The lethal phenotypes of RED motif substitutions in ssl2 revealed their essential roles in S. cerevisiae. These results suggest that a subset of human disease alleles can alter TFIIH functions when placed in the yeast system.

Novel ssl2 mutants with transcriptional defects

Our establishment of a genetic system sensitive to ssl2 mutant mediated initiation defects allowed us to obtain a broad set of alleles for the study of Ssl2 function in promoter scanning by large-scale genetic screens (see Methods). Our genetic screening has identified at least two phenotypically distinguishable classes of ssl2 alleles: the first class is putatively defective for the induction of the IMD2 gene, resulting in sensitivity to MPA; the second class confers constitutive expression of imd2Δ::HIS3, resulting in a His+ phenotype (Figure 1A, Figure 2 – Figure supplement 1). Other transcription-related or conditional phenotypes, Spt+ or suppression of gal10Δ56 (Gal R) or temperature sensitivity (Tsm'), were observed in distinct subsets of alleles of the two major classes (Figure 1A, Figure 2 – Figure supplement 1). The Spt' phenotype reporter used in our strains, lys2-128a, detects activation of a TSS within a Ty1 δ element at the 5' end of the LYS2 gene. Importantly, a subset of TSS-shifting alleles show the Spt' phenotype and it is useful to further classify identified ssl2 alleles. We observed that Spt' and His+ phenotypes were dominant for tested alleles (not shown), suggestive of possible gain of function; in contrast, there were no dominant alleles found among MPA mutants, consistent with either recessive loss-of-function mutations or the nature of the phenotype (sensitivity) or both. We then asked if TSS usage at ADH1 was altered predictably in the mutant classes as we observed for existing ssl2 mutants and previously studied Pol II and other GTF mutants. We find that the two major classes of ssl2 mutants exhibited predicted TSS shifts, with all tested MPA alleles shifting TSS usage upstream and all His alleles shifting TSS usage downstream (Figure 2 – Figure supplement 2) validating our genetic method for identifying ssl2 alleles conferring altered initiation properties.
We examined our substitution mutants in the context of the structure of Ssl2 within the yeast preinitiation complex as determined by the Cramer lab to understand how these alleles relate to PIC architecture and interactions. Substitutions causing MPA\textsuperscript{S} phenotypes and upstream TSS shifts alter amino acids distributed across the protein, with a large number in conserved domains and highly conserved residues (Figure 2A, Figure 2 – Figure supplement 3). In contrast, mutations related to His\textsuperscript{+} phenotypes, while also generally conserved, alter amino acids clustered N-terminally, within a domain predicted and now observed to be homologous to Ssl2’s interaction partner Tfb2 (Figure 2A). This domain was first fully observed in a human core TFIIF structure (PDB 6NMI) but has now been observed in just published higher resolution PIC structures (Figure 2B). This key visualization of the conserved Ssl2 N-terminus allows placement of most ssl2 mutant residues identified as His\textsuperscript{+} or both His\textsuperscript{+} and Spt\textsuperscript{+} (Figure 2C, D). In addition, the Spt\textsuperscript{+} phenotype, not previously observed for classical ssl2 alleles, was found exclusively within a subset of stronger downstream TSS shifting His\textsuperscript{+} mutants (Figure 2A). The coincidence between His\textsuperscript{+} and Spt\textsuperscript{+} phenotypes for ssl2 mutants is in contrast to what is observed for Pol II mutants. In our previous studies of Pol II mutants, the Spt\textsuperscript{+} phenotype was observed in Pol II catalytic center substitutions, overlapping with MPA\textsuperscript{S} to a large extent, and tightly linked with increased Pol II activity. Pol II alleles with both Spt\textsuperscript{+} and MPA\textsuperscript{S} phenotypes increased the efficiency of TSS usage resulting in upstream shifts to TSS distributions across promoters in vivo. However, in our identified ssl2 alleles, none of the Spt\textsuperscript{+} mutants also conferred MPA\textsuperscript{S} (Figure 2A). These observations together are consistent with distinct effects on structure and function in ssl2 mutant classes.

Some TSS shifting substitutions are located on the Ssl2 surface (Figure 2B) and a subset (e.g. D610) are located proximal to DNA (Figure 2C). D610A confers an upstream TSS shift. In contrast, R636C, which is also close to DNA, confers a downstream TSS shift. F498, which is located in the groove of Ssl2 lobe 1 and facing DNA, caused an upstream TSS shift when substituted with leucine. Additionally, a small patch of residues with many TSS shifting substitutions is found on the Ssl2 surface proximal to DNA (Figure 2A). These substitutions are from the helicase domain 1 and shift TSSs upstream, including D522V, K523N, N528I, F529L and E537G. In addition, substitutions I383T and K372E are in residues proximal to this small patch but shift TSS downstream. Intriguingly a number of alleles of both classes are found in the Ssl2 N-terminal domain homologous to Tfb2 (TFB2C-like or “Clutch”) that forms interaction with Tfb2 and bridges Tfb2 with Ssl2 helicase domain 1 (Figure 2A-D). As Tfb2/p52 recruits Ssl2/XPB into TFIIF and stimulates Ssl2/XPB catalytic activity, this interface is of special interest for how other factors might communicate with Ssl2/XBP function in both transcription and NER.

Highlighting the uniqueness and potential plasticity of this region, we identified multiple substitutions at the conserved N230 in this domain, with N230D/S conferring MPA\textsuperscript{S} and an upstream TSS shifts, while N230I conferred both His\textsuperscript{+} and Spt\textsuperscript{+} phenotypes and a downstream TSS shift (Figure 2A-D). These results suggest that altered Ssl2 DNA or intradomain interactions alter Ssl2 function in TSS selection in distinct ways, likely through distinct effects on Ssl2 biochemical activity, discussed below and in Discussion.

**ssl2 alleles behave differently from Pol II and other GTF alleles for TSS usage**

We highlight two alleles as representative of distinct ssl2 allele classes in comparison to examples of the two Pol II allele classes: ssl2 N230D, which is MPA sensitive and appears to reduce ability to use downstream TSSs at ADH1, and ssl2 N230I, which shows both His\textsuperscript{+} and Spt\textsuperscript{+} phenotypes and shifts TSSs downstream relative to WT at ADH1 (Figure 2E, F rbp1 alleles, 2G, H ssl2 alleles). As we have shown previously, rbp1 mutants also fall into two major classes regarding transcription phenotypes and shifts. As a comparison, rbp1 E1103G is MPA\textsuperscript{S} and shifts TSS usage upstream while rbp1 H1085Y is His\textsuperscript{+} and shifts TSS usage downstream. MPA-sensitivities for both rbp1 E1103G and ssl2 N230D correlate with an upstream TSS shift at ADH1, similarly to our observation for existing ssl2 MPA\textsuperscript{S} alleles, ssl2-DEAD and ssl2-508. However, all ssl2 MPA\textsuperscript{S} alleles examined appear to shift TSS distribution upstream by limiting or truncating TSS usage at downstream sites and not by activating lowly used upstream sites as rbp1 E1103G does and as known TFIIF alleles do (Figure 2F, H). This behavior suggests that rbp1 and ssl2 alleles may alter TSS usage by distinct mechanisms.
Transcription start site sequencing (TSS-seq) identifies global effects on TSSs in *ssl2* alleles in *S. cerevisiae*

To gain an insight into the impact of TFIIH’s activity on TSSs genome-wide, we have examined 5’ ends of RNA transcripts for these two classes of *ssl2* allele in *S. cerevisiae* by performing TSS-seq\(^\text{13,67}\) (Figure 3A). In total, six *ssl2* alleles were analyzed along with a WT control, including three His\(^+\) alleles (L225P, N230I and R636C) that shift TSSs downstream at *ADH1*, and three MPA\(^5\) alleles (N230D, D522V and Y750*) that shift TSSs upstream at *ADH1*. Furthermore, we performed TSS-seq on previously analyzed Pol II WT, E1103G and H1085Y for direct comparison purposes using our updated protocol (Methods). The positions and counts of the 5’ ends of uniquely mapped reads were extracted to evaluate correlation and assess the reproducibility between biological replicates (Figure 3B, Figure 3 – Figure supplement 1A). We previously found that clustering of correlation coefficients among libraries could distinguish Pol II mutants into GOF and LOF groups\(^13\). Here, we also found that mutant phenotypic classes were distinguished by this clustering with *ssl2* and Pol II alleles separated suggesting effects observed at individual promoters are predictive of effects observed across the genome (Figure 3C, Figure 3 – Figure supplement 1B).

**ssl2** alleles shift transcription start site distribution genome-wide

As in Qiu, Jin et al\(^13\), we focused on initiation within a cohort of 5979 promoters for a large number of mRNA genes and non-coding RNAs. These promoters are separated into Taf1-enriched or -depleted subclasses as a proxy for the two primary promoter classes in yeast\(^68\). We took a simple approach to examine how mutants affect observed TSS distributions using a few key metrics, for example the position in a promoter window containing the median read in the distribution as a measure of where the TSS distribution is in a particular promoter window (Figure 4A). We observed decreased TSS signal downstream of the WT median TSS signal in *ssl2* N230D and other *ssl2* MPA\(^5\) alleles (Figure 4B, Figure 4 – Figure supplement 1). In contrast, substantially increased TSS usage was observed at downstream sites in *ssl2* N230I and other *ssl2* His\(^+\) alleles (Figure 4B, Figure 4 – Figure supplement 1). To quantify the change of TSS usage, we measured the TSS shift between WT and mutant strains by subtracting the WT median TSS position within each promoter window from the mutant median TSS position (Figure 4C). For each mutant, we show the median TSS shifts across promoters in both heatmap and boxplot format (Figure 4D-E, Figure 4 – Figure supplement 2A). As predicted from our model, *ssl2* MPA\(^5\) alleles shift the median TSS positions upstream at most promoters, showing a similar profile to Pol II GOF mutant (Figure 4D, Figure 4 – Figure supplement 2A). Also as predicted from our model, *ssl2* His\(^+\) alleles shift median TSS positions downstream at the majority of promoters and show a similar profile to a Pol II LOF mutant\(^13\) (Figure 4D, Figure 4 – Figure supplement 2A). Additionally, mutants were clustered into upstream and downstream classes based on shifts in promoter median TSS position (Figure 4D, Figure 4 – Figure supplement 2A). Principle Component Analysis (PCA) of TSS shifts distinguishes *ssl2* and Pol II mutants into four major classes, namely Pol II LOF (downstream shifting) and GOF (upstream shifting), and *ssl2* upstream and downstream shifting mutants (Figure 4 – Figure supplement 2B). We observed that the magnitude of TSS shift was consistent with the strengths of putative TSS-shift dependent growth phenotypes in *ssl2* upstream shifting mutants (Figure 4E). For example, alleles of N230D, D522V and Y750*, from the left to right, show weaker MPA\(^5\) phenotypes in genetic tests (Figure 2A), while also showing a gradient in TSS shift magnitudes across promoters (Figure 4E). Notably, the extents of TSS shifts in *ssl2* alleles are less than for Pol II activity mutants, indicating a more dramatic effect of Pol II’s catalytic activity on TSS distributions (Figure 4E). These results are consistent with phenotypes of mutants being predictive of global TSS defects among mutants for a particular gene but not necessarily between genes, which we discuss later as indicative of different mechanisms for Pol II and Ssl2 effects on TSS selection.

**Distinct alterations to TSS distribution in *ssl2* mutants**

To evaluate the effects of *ssl2* and Pol II alleles on scanning length, the width between positions of the 10th and 90th percentiles of the TSS signal at each promoter window was determined (TSS “spread”) (Figure 5A). We observed obvious narrowing of TSS spreads in *ssl2* upstream shifting mutants and widening of TSS spreads in *ssl2* downstream shifting mutants (Figure 5B-D). The profiles of the TSS spread difference between mutant and WT at each individual promoter (Figure 5A) also differentiates mutants into clear shift classes (Figure 5C; Figure 5 – Figure supplement 1A, B). Changes in TSS
spread for ssl2 alleles are distinct from how Pol II mutants alter TSS spread (Figure 5D). Both classes of ssl2 mutant show large bias in direction of shift in spread, relative to WT across a number of types of assessment (Figure 5 – Figure supplement 2D-F). Consistently, MPA\textsuperscript{3} ssl2 alleles (those that shift TSSs upstream) showed narrowing in TSS spread at the majority of promoters while His+ ssl2 alleles (those that shift TSSs downstream) showed widening in TSS spread at the majority of promoters. These results extend the idea that while classes of initiation allele may be easily identified for Pol II or Ssl2 using the same genetic phenotypes, their effects on TSSs at individual promoters are quantitatively and likely qualitatively distinct.

**Genetic interactions between initiation factors and ssl2 alleles suggest distinct roles for Ssl2 and other factors in TSS scanning**

To explain the observed quantitative differences between Pol II initiation efficiency alleles and ssl2 alleles, we hypothesize that ssl2 alleles that narrow TSS spreads (ssl2 N230D and similar alleles), resulting in upstream shifts in TSS distributions, are defective in scanning processivity due to decreased Ssl2/TFIIF translocase activity. In contrast, ssl2 alleles that show increased TSS spreads and usage at downstream sites (N230I and similar alleles) behave as increased scanning processivity (GOF) alleles, due to an increase in Ssl2/TFIIF activity. Alternatively, ssl2 N230I might instead be a scanning rate GOF allele that decreases initiation efficiency across TSSs by decreasing the exposure time of each TSS within Pol II active site. As a consequence, there would be, hypothetically, increased TSS usage at downstream sites due to increased Pol II flux reaching those positions, similar to Pol II LOF efficiency alleles. To probe mechanisms of Ssl2 function we designed ssl2 genetic interaction studies for which we have specific predictions based on their possible roles (Figure 6 – Figure supplement 1). These studies emerge from our prior observations of genetic interactions from three angles between initiation factors such as Pol II activity mutants and alleles of GTFs\textsuperscript{39}. First, we have examined general effects on growth between classes of mutant, which can manifest as synthetic sickness or lethality, suppression, or lack of effects on growth. Second, we have examined suppression or enhancement of transcription-related phenotypes in double mutants, allowing detection of additive or epistatic interactions based on putative transcription defects at the genetic reporter loci. Finally, we have quantitatively examined effects of double mutants on TSS distributions at ADH1, allowing additive or epistatic interactions to be observed at the individual promoter level. In our previous studies, we found that Pol II activity mutants and alleles of GTFs (sua7/TFIIB and tfg2/TFIIF) showed additivity/suppression for transcription-dependent phenotypes as well as additivity/suppression for alterations to TSS distributions at ADH1. These studies suggested these GTF alleles were functioning in same pathway as Pol II catalytic mutants, namely controlling the efficiency of initiation across individual TSSs. In order to understand how ssl2 alleles interact with other initiation factors and how Ssl2 functions within the network of activities that are essential for TSS selection, we generated double mutants among a collection of ssl2 alleles, Pol II activity-altering alleles, sua7-1, tfg2Δ146-180, and sub1Δ. To streamline display a large number of genetic interactions and growth phenotypes, we have converted general growth on plates and growth on phenotype-specific media to qualitative scores (Figure 6D-F, Figure 6 – Figure supplement 2) and these are shown as heat maps (Figures 6, 7) with primary data shown in Figure 6 – Figure supplement 3, Figure 7 – Figure supplement 1 and Figure 7 – Figure supplement 2.

Our detailed studies are summarized as follows (for detailed discussion, see version one of the pre-print of this work\textsuperscript{39}). In contrast to the suppressive/additive interactions that were broadly observed between Pol II and TFIIB/TFIIF alleles\textsuperscript{39}, we observe primarily epistatic effects between Pol II and ssl2 alleles. First, the broad synthetic lethality or enhancement/additivity of transcriptional phenotypes through combining Pol II and TFIIB/TFIIF alleles that shift TSS distributions in the same direction were not observed between Pol II and ssl2 alleles (Figure 6). Examples of epistatic interactions, where double mutants between Pol II and ssl2 alleles have phenotypes of either the Pol II single mutant or the ssl2 single mutant, were found in a number of cases. Each case supports a model where ssl2 alleles are functioning through scanning processivity and not initiation efficiency directly. This epistasis is best reflected by nearly complete absence of synthetic lethality between ssl2 downstream shifting alleles and Pol II downstream shifting alleles (Figure 6A-C), in contrast to interactions between all other classes of downstream shifting allele (e.g. Pol II, sua7/TFIIB, sub1Δ\textsuperscript{39}). Epistasis for both transcription
phenotypes and **ADH1** TSS shifts was also observed between Pol II upstream shifting alleles and both classes of **ssl2** allele, meaning that double mutants had phenotypes of Pol II single mutants (**Figure 6A-K**, **Figure 6 – Figure supplement 3**). For each of these cases, results support a model where if initiation is efficient enough or early enough in a scanning window, i.e. due to increased Pol II initiation activity, then increase in scanning processivity (e.g. **ssl2** N230I) loses ability to alter TSS distributions, while a decrease in scanning processivity is buffered against due to high enough gain in transcription efficiency in tested Pol II alleles. Similarly, both classes of **ssl2** alleles appeared epistatic or non-additive with Pol II downstream shifting alleles (**Figure 6A-C**, **Figure 6 – Figure supplement 3**), also consistent with determination of scanning window by **ssl2** activity to be upstream of ability of Pol II mutants to alter TSS distributions through altered initiation efficiency.

Interactions between **ssl2** alleles and other GTFs or **sub1Δ** reveal complexities that are of special note as they suggest non-obvious roles/interactions between these factors and Ssl2 function (**Figure 7**, **Figure 7 – Figure supplements 1,2**). We anticipated that **sua7-1** and **tfg2Δ146-180**, encoding mutant forms of TFIIB and TFIIF respectively, would behave strictly as TSS efficiency alleles due to their additive behavior with Pol II alleles, and therefore would similarly show epistatic effects with **ssl2** alleles. Notably, lethal phenotypes were observed between individual **ssl2** alleles and **sua7-1** or **tfg2Δ146-180** alleles for combinations between single mutants that alter TSS distributions in the same direction, distinct from their interactions with Pol II alleles (**Figure 7A**). We suggest two possibilities for this observation: first, **sua7-1** and **tfg2Δ146-180** could confer additional defects causing increased sensitivity to **ssl2** defects, for example, altered PIC integrity; second, **sua7-1** and **tfg2Δ146-180** might be sensitized to increased Ssl2 processivity (for **sua7-1**) or decreased Ssl2 processivity (for **tfg2Δ146-180**) in addition to their altered TSS efficiency effects (see Discussion). When we combined alleles of **sua7-1** or **tfg2Δ146-180** with **ssl2** alleles that shift TSS distributions in opposite directions, interactions were complex but significant epistasis was observed. Consistently, double mutants shifted **ADH1** TSS distributions to similar extent as the **tfg2Δ146-180** single mutant (**Figure 7C**, **Figure 7 – Figure supplement 1**) as predicted for an increase in initiation efficiency buffering against effects of increase in scanning processivity.

Sub1, a conserved factor (yeast homolog of mammalian PC4) was previously found to facilitate Pol II transcription in a variety of ways, to be recruited to the PIC, and to alter accessibility of promoter single-stranded DNA, consistent with initiation functions. **sub1Δ** has extensive genetic interactions with initiation factors and itself causes TSSs to shift downstream, though its actual role in initiation is unknown. We previously found **sub1Δ** to confer a His* phenotype for the **imd2Δ::HIS3** initiation reporter and furthermore found that Pol II GOF alleles appeared epistatic to **sub1Δ**, leading to the proposal that **sub1Δ** effects in initiation were distinct from TFIIB or TFIIF alleles. Because we have observed similar epistatic interactions between **ssl2** and Pol II alleles, we considered that Sub1 might also be behaving as a scanning processivity factor. Therefore, we predicted the possibility of additive effects between two types of processivity alleles, namely **ssl2** upstream and downstream shifting alleles and **sub1Δ**, if they are acting independently. First, no strong genetic interactions (lethality) were observed between **ssl2** and **sub1Δ** alleles (**Figure 7A**, **Figure 7 – Figure supplement 2A**). Second, the majority of **sub1Δ** interactions with **ssl2** alleles appear to be additive when examining TSS distributions at **ADH1** as predicted for factors are acting on processivity independently. Third, and notably, we identified allele-specific interactions between **sub1Δ** and specific **ssl2** alleles within classes of **ssl2** allele that until these experiments have not been distinguishable. For example, most upstream shifting **ssl2** alleles were additive with **sub1Δ** for TSS distributions at **ADH1**, resulting in mutual suppression of TSS distribution shifts (**Figure 7D**). In contrast, **sub1Δ** was epistatic to **ssl2** Y750*, suggesting that a putative block to processivity due to C-terminal truncation of Ssl2 can be relieved by **sub1Δ**, and potentially may be due to altered Sub1 function in **ssl2** Y750*. Finally, allele specificity of **ssl2** F498L was revealed by these genetic experiments. This TSS upstream shifting **ssl2** allele was unexpectedly synthetic lethal with both **sua7-1** and **sub1Δ** suggesting heretofore undetected phenotypic differences from other alleles of the same class.

### Two networks controlling initiation by promoter scanning

Results of genetic interaction studies are consistent with two distinct networks controlling TSS selection...
by scanning (Figure 8). Additive-suppressive interactions were observed within networks while specific classes of mutants showed epistatic interactions between networks. One network impinges on Pol II catalysis and initiation efficiency, and genetic analyses suggest that the Pol II active site collaborates with activities of TFIIB and TFIIH in this process, consistent with experiments indicating effects of TFIIB and TFIIIF on Pol II catalytic activity e.g.\(^{21,75,76}\). The other, we propose, impinges on scanning processivity through TFIIH with the participation of Sub1. Our genetic interactions also uncover functional connections between TFIIB and TFIIIF and Ssl2 that are distinct from Pol II active site mutants. These results support predictions of altered PIC function for TFIIB and TFIIIF mutants beyond phosphodiester bond formation and will be interesting to test in biophysical experiments. Extensive epistasis observed between networks (Figure 8) supports predictions for how efficiency and processivity should interact during initiation by promoter scanning (Figure 9, see Discussion).

**ssl2 alleles shift positioning of PIC-components genome-wide**

We found previously that polar shifts of TSS distribution in Pol II catalytic activity mutants was accompanied with alteration in PIC localization as detected by ChIP-exo\(^{13}\). Shift in PIC components upon alteration to Pol II catalytic activity suggested that extent of scanning might be coupled to Pol II initiation, or that alteration to Pol II initiation kinetics affects observed distributions of GTFs. We performed these same ChIP-exo experiments on Sua7 and Ssl2 for two ssl2 alleles, N230D and N230I. (Figure 9 – Figure supplement 1A, B). Both shifted PIC localization genome-wide with the same polarity as they shift TSS distribution. We note that TAP-tagging Ssl2 confers slight phenotypes on its own and slight enhancement of ssl2 N230D and slight suppression of ssl2 N230I (Figure 9 – Figure supplement 2). However, each tagged mutant was compared to the tagged WT and the results are robust and distinct for each mutant. The extent of ChIP-exo shifts were as strong or stronger than Pol II mutant shifts although Pol II mutants have stronger effects on TSS distributions\(^{13}\). These results could be consistent with scanning by TFIIH on DNA uncoupled from the Pol II initiation decision, i.e. ssl2 mutants extend PIC scanning to downstream positions even though initiation has occurred. Such a result would be consistent with similar behavior of DNA compaction in optical tweezer analysis of initiation wherein dATP-supported reactions (presumptive TFIIH scanning-driven DNA translocation through Ssl2 use of dATP) and NTP-supported initiation reactions (TFIIH translocation and Pol II initiation allowed) have similar behavior\(^{17}\). The extent or mechanism of uncoupling between promoter scanning upon productive initiation are unknown and represent an open question in initiation mechanisms.

**DISCUSSION**

Our studies now reveal the impact of altered Ssl2 function on initiation by promoter scanning in S. cerevisiae. We find distinct classes of Ssl2 allele that alter initiation genome-wide with distinct behaviors, with many alleles being in highly conserved residues. Our genomic and genetic data support a model wherein Ssl2 function as a DNA translocase can be genetically modulated and this modulation is consistent with TFIIH having either increased or decreased processivity during promoter scanning. The positioning and genetic behaviors of our allele classes are consistent with one class behaving biochemically as a loss of function and therefore truncating the scanning process prematurely and narrowing TSS distributions genome-wide. This is exactly the predicted outcome for a translocase with decreased processivity. Conversely, our other class increases downstream TSS usage and alters PIC localization at promoters by extending it downstream. These behaviors are consistent with increased translocase processivity. Genetic interactions between ssl2 alleles, especially between ssl2 and GTF mutants and sub1Δ suggest further distinctions between allele classes or within allele classes, generating testable predictions for biochemical studies. Putative increased SSL2 activity alleles are also dominant or codominant genetically, consistent with being able to function on promoters in an increased capacity of some sort.

**Interpretation of Pol II and Ssl2 functions in the Shooting Gallery model for initiation by scanning**

We have previously described how Pol II determines the efficiency of a TSS in a “Shooting Gallery” model, where the rate at which a TSS (conceived of as a target) passes the active site, the rate of firing (catalytic activity), and the size of the target (innate sequence strength) together contribute to the
probability a target is hit (initiation happens) (Figure 9A)\textsuperscript{13,40}. Alteration of enzymatic activities supporting initiation, either the Pol II active site or TFIIH translocation will have predictable effects on individual TSS usage and the overall TSS distributions when initiation proceeds by scanning. In Pol II mutants with altered catalytic activity that is known to affect transcription efficiency, we observed polar changes to TSS distributions\textsuperscript{13}. Distributions will also necessarily be shaped by an additional factor: Pol II flux. Pol II flux describes the relative number of polymerases encountering a given start site, which has a higher value at upstream TSSs and a lower value at downstream TSSs, resulting in reduced apparent usage at downstream position distinct from their inherent efficiencies (Figure 9B).

Additionally, the potential upstream and downstream constraints for defining the scanning window will have effects on TSS distributions. Studies suggest that very upstream TSSs close to the presumed location of PIC assembly show reduced transcription initiation\textsuperscript{27}. The physical basis for defining the upstream boundary of the scanning window has not yet been determined. An obvious constraint is the minimum space required for PIC assembly. Moreover, we hypothesize that downstream constraints for defining the scanning window could be TFIIH’s processivity, the +1 nucleosome, or both. Previous single molecule studies suggested that TFIIH drives downstream scanning distances similar in length to the distribution of TSSs at yeast promoters\textsuperscript{17}. We propose that TSS distribution of a promoter is established by the cooperation of Pol II’s catalytic activity and TFIIH’s processivity for reaching and activating TSSs at promoter sites.

When Pol II has increased catalytic activity, e.g. in Pol II catalytic activity GOF alleles, upstream TSSs will increase in efficiency (Figure 9C, D, Pol II GOF). In this allele class, usage of downstream TSSs also will decrease due to reduction in Pol II flux reaching downstream sites due to prior initiation. Conversely, when Pol II has decreased catalytic activity, TSSs at upstream sites will be less efficiently used, more slowly reducing Pol II flux (Figure 9C, D, Pol II LOF). Inability to initiate earlier in scanning will result in increased TSS usage at downstream sites and a flattening and spreading of the TSS distribution (as demonstrated by an efficiency curve with decreased slope). We hypothesize that alleles with increased processivity (processivity GOF allele) will expand the scanning window by allowing the PIC to scan further downstream while attempting initiation, increasing the probability that downstream TSSs are reached during any individual scanning event (Figure 9E, processivity GOF). As a consequence, processivity GOF alleles increase the potential for scanning downstream but only if Pol II flux (Pol II molecules still scanning) persists to reach those sites. In contrast, a processivity LOF allele would limit the Pol II machinery’s access to downstream TSSs sites by reducing the scanning window (Figure 9E, processivity LOF). Consequently, there would be an upstream shift in TSS distribution compared to WT, without the activation of additional upstream TSSs.

Prior biochemical and more recent structural analyses indicate that TFIIH is a fascinating complex with numerous contacts suggested or predicted to modulate or control TFIIH enzymatic subunits’ activities\textsuperscript{77}. For example, Ssl2/XPB must be activated during transcription initiation to allow promoter opening. Recent structures suggest that interactions with both Mediator and TFIIID may position parts of TFIIH for different functions in initiation\textsuperscript{78-80}. Genetic and biochemical studies also suggest that TFIIH may itself impose a block to initiation that is then relieved by TFIIH activity through Ssl2/XPB\textsuperscript{81,82}. Both TFIIH ATPase subunits, Rad3/XPD and Ssl2/XPB, must also be regulated for TFIIH’s function in NER with Rad3/XPD held inactive during transcription and released for NER\textsuperscript{84,77}. XPB mutations in patients that result in Xeroderma Pigmentosum are straightforwardly interpreted as conferring NER defects, however transcriptional phenotypes may also be present depending on mutation\textsuperscript{51-52}. Mutations in XBP that cause Trichothiodystrophy (TTD) localize to conserved residues in the XBP N-terminus where we have identified a number of mutations. TTD mutations in XBP have been interpreted as reducing the amount of TFIIH in the cell through potential destabilization, while one appears to impact folding and activity of XBP\textsuperscript{84}. Our identification of putative loss and gain of function mutations in this domain in S. cerevisiae underscore the idea that observed conservation in this region may control key inputs to Ssl2/XPB activity. The substitutions we have identified are largely in residues conserved from yeast to humans (Figure 2 – Figure supplement 3), and we suggest these residues detect potential paths for allosteric regulation of Ssl2/XPB. Only a subset of our alleles confer UV sensitivity, suggestive of NER defects\textsuperscript{83}. These are C-terminal and this suggests that our alleles uniquely alter Ssl2 modulation in transcription or
that transcriptional functions of Ssl2 are sensitized to defects that do not appreciably lead to UV sensitivity. DNA translocases are the engines for chromatin remodeling and much regulation of chromatin remodelers relates to coupling of ATP hydrolysis to translocation potential\textsuperscript{84, 85}. Mutations in remodelers that increase or decrease coupling have strong effects on remodeling. We posit that it is likely that a number of our alleles will act through altered coupling of ATPase activity and translocation, with the end result being increase or decrease in scanning processivity. Biochemical studies will reveal specific aspects of TFIIH activity that are altered by these substitutions.

Our putative gain-of-function alleles are concentrated in the TFB2C-like N-terminal domain and DRD of Ssl2. The TFB2C-domain has been implicated as a target of Tfb3/Mat1 in restricting Ssl2/XBP activity in holoTFIIH\textsuperscript{62, 77, 86}. However, upon assembly into the PIC, Tfb3/Mat1 releases the N-terminus of Ssl2\textsuperscript{78, 87}. This Ssl2 domain is also targeted by Tfb2/p52\textsuperscript{61}, 63, 64, 87, which has long been implicated in modulating Ssl2/XBP activity in addition to assembling it into TFIIH. We have identified one His\textsuperscript{+} allele at the Ssl2-Tfb2 interface, though most are in internal interfaces at the putative nexus of Ssl2 NTD, DRD, and HD1/ATPase Lobe 1 (Figure 2).

Key open questions relate to how the PIC communicates to Ssl2/XBP to engage and open promoter DNA, what the mechanistic basis for any imposition of initiation block by Ssl2/XBP is, what the basis of its subsequent relief is, and how translocation is terminated upon or subsequent to productive initiation, i.e. are the processes coupled in anyway. Aibara, Schilbach et al have recently imaged the human PIC captured in two states, suggestive of pre- and post-translocation intermediates of TFIIH\textsuperscript{88}. These structures show loss of contact between TFIIH through the MAT1 RING domain and the Pol II stalk/TFIIE in the proposed post-translocated state. This raises an attractive model for uncoupling TFIIH translocase from Pol II after a single translocation step, functionally limiting initiation to the small window of exposed TSSs within reach of the Pol II active site. These structures were from a minimal PIC lacking Mediator and TFIID and therefore it remains to be determined if this translocase is in fact uncoupled, as other potential TFIIH/PIC contacts remain. Other events during initiation may also propagate changes to the PIC, such as lengthening of the nascent RNA to potentially clash with TFIIIB and potential subsequent reorganization of the PIC by this event, or due to Pol II CTD phosphorylation.

In many organisms, nucleosome-depleted regions promote initiation bidirectionally and these regions are flanked by positioned nucleosomes\textsuperscript{89-92}. Nucleosomes would potentially act as competitors for double stranded DNA being translocated by TFIIH. Transcription activity drives histone dynamics at promoters, consistent with TFIIH translocation proposed to function akin to a chromatin remodeler\textsuperscript{93}. How the +1 nucleosome might feed back on scanning or regulate TFIIH translocation is an open question. In the absence of the remodeler RSC’s function in yeast, nucleosomes move upstream into normally nucleosome depleted regions and inhibit or narrow TSS usage at a number of promoters\textsuperscript{94}. These results are consistent with nucleosomes competing with TFIIH for promoter DNA. However, these results are conditional on RSC depletion under normal conditions obviates the ability for +1 nucleosomes at active promoters to provide a block to scanning or initiation on the edges of their positions.

How Pol II specifies multiple TSSs at individual promoters across eukaryotes has long been an open question. The observation of downstream-located TSS sites relative to where DNA melting occurs in S. cerevisiae led to the original proposal of a scanning mechanism for TSS identification\textsuperscript{9}. In contrast to how Pol II finds TSSs in S. cerevisiae for all promoters, in other eukaryotes including other fungi, a scanning process is not required for promoters with defined architecture specified by a TATA element\textsuperscript{95, 96}. These promoters use highly focused TSSs immediately and precisely downstream of the DNA melting sites ~30 basepairs downstream of the TATA element +1 position. For example, Lu et al have pinpointed the split between scanning from TATA-promoters and non-scanning species within the Saccharomycetes\textsuperscript{95}. However, most eukaryotic promoters are TATA-less and use multiple TSSs\textsuperscript{89, 97-107}. Therefore, whether or not scanning is also a mechanism in higher eukaryotic promoters, or minimally for a subset of eukaryotic promoters or within a specified window, is still an unanswered question as there have been no formal tests of this mechanism. It has been suggested, however, that each individual TSS is recognized as an individual promoter due to sequence signatures apparent in
comparison of thousands of TSSs in humans\textsuperscript{108}. Very recent results of cryo-EM studies on human PICs, especially in structures visualizing TFIID, indicate that promoter classes may assemble PIC components in distinct fashion within a single organism\textsuperscript{78}, yet these assembly pathways similarly position an upstream TSS proximal to the Pol II active site, consistent with proposals that human promoters could contain information for assembling PICs individually\textsuperscript{108}. Recent results suggest that there may be plasticity in TSS selection from individual PICs upon mutation of Inr sites in mouse, potentially supporting a type of scanning in mammals\textsuperscript{109}. That diverse initiation mechanisms are supported by highly conserved factors suggests that we may yet to find additional unexpected plasticity in initiation across evolution in eukaryotes.

METHODS

Yeast strains

Yeast strains are derived from a GAL\textsuperscript{+} derivative of S288C (FY2)\textsuperscript{110}. Yeast strains used in this study are listed in Supplementary File 1.

Plasmids and bacterial strains

Bacterial strains and plasmids used in this study are listed in Supplementary File 2.

Yeast media

Yeast media used in this study were made as previously described\textsuperscript{39, 41, 47, 111}. Briefly, YP medium is made of yeast extract (1% w/v; BD) and peptone (2% w/v; BD). Solid YP medium contains bacto agar (2% w/v; BD), adenine (0.15 mM, Sigma-Aldrich) and tryptophan (0.4 mM, Sigma-Aldrich). YPD medium uses YP medium components supplemented with dextrose (2% w/v, VWR). YPRaf medium uses YP medium components supplemented with raffinose (2% w/v, Amresco) and antimycin A (1 mg/ml; Sigma-Aldrich). YPRafGal medium uses YP medium components and supplemented with raffinose (2% w/v), galactose (1% w/v; Amresco) and antimycin A (1 mg/ml; Sigma-Aldrich). Minimal media (SC-\textsuperscript{-}) was made with a slightly modified “Hopkins mix” (0.2% most amino acids w/v), and supplemented with Yeast Nitrogen Base containing ammonium sulfate (without amino acids, BD), bacto agar (2% w/v; BD) and dextrose (2% w/v, VWR). The original “Hopkins mix” and the slight modification were as previously described\textsuperscript{47, 111}. All amino acids are from Sigma-Aldrich. SC-Leu+5FOA is minimal medium of SC-Leu supplemented with 5-fluoroorotic acid (5-FOA), (1 mg/ml, Gold Biotechnology). SC-Leu+MPA media is minimal media of SC-Leu supplemented with mycophenolic acid (20 \textmu g/ml, Sigma-Aldrich, from a 10 mg/ml stock in ethanol). SC-His+3AT is minimal medium of SC-His supplemented with 3-aminotriazole (3-AT, 0.5mM, Sigma-Aldrich).

Plate phenotyping and growth heatmaps

Yeast phenotyping assays were performed by spotting 10-fold serial dilutions of saturated YPD-liquid yeast cultures on various solid media, as previously described\textsuperscript{47}. Yeast cells on various media were cultured at 30 °C except for temperature sensitivity phenotypes, which were at 16 °C (YPD 16) and at 37 °C (YPD 37). Yeast growth on specific media was recorded by taking pictures every 24 hours after an initial 16h of growth, from day 2 (40h) to day 7 for all media except for YPRaf/Gal (pictures to day 9). Growth phenotypes on specific media were scored on days when WT yeast reached mature colony sizes, as follows: YPD on day 2 (40h after spotting); SC-Leu, SC-His and SC-Trp on day 3 (64h); YPRaf on day 4 (88h); SC-Lys and SC-Leu+MPA on day 5 (112h); and YPRaf/Gal on day 7 (160h). To illustrate the strength and distribution of mutants on the two-dimensional structure of SsI2 (Figure 2), growth phenotypes are converted to a numerical score, using the scale 1-5 to indicate the level of growth, where 1 indicates no growth and 5 indicates full growth. The level of growth is positively correlated with the strength of phenotypes for SC-His, SC-Lys and YPRaf/Gal medium, so the “growth score” is directly used as “phenotyping score” for making a heat map. For other media, the level of growth is negatively correlated with the strength of growth, thus growth score 1-5 is inversely converted to the phenotyping strength score 5-1, with 5 growth score converted into phenotyping score 1 to indicate no phenotypes, with 4 growth score converted into phenotyping score 2 to show a weak phenotype, and so on. The heatmap uses light to dark color showing weak (phenotyping score 2) to strong phenotypes (phenotyping score 5), no phenotype (phenotyping score 1) is not shown.
**Primer extension**

To detect putative usage of TSSs in yeast, a primer extension (PE) assay was performed as previously described\(^\text{112}\), modifying original protocol in \(^{112}\). Briefly, 30μg of total RNA isolated from yeast cells was used for each PE reaction. A primer complementary to ADH1 mRNA was end-labeled with gamma-P\(^32\)ATP and T4 PNK and annealed to total RNA. Reverse-transcription was then performed by adding M-MLV reverse-transcriptase (Fermentas/ThermoFisher) and RNase Inhibitor (Fermentas/ThermoFisher). RNase A was added to remove RNA after reverse-transcription. Products were detected by running an 8% acrylamide gel (19:1 acrylamide:bisacrylamide) (Bio-Rad), 1X TBE, and 7M urea. PE gels were visualized by phosphorimaging (GE Healthcare or Bio-Rad) and quantified by ImageQuant 5.1 (GE) or Image Lab software (Bio-Rad).

**ssl2 mutant screening**

ssl2 mutants were created by PCR-based random mutagenesis coupled with a gap repair. Briefly, mutation of SSL2 (ssl2\(^+\)) was accomplished by standard PCR reactions using Taq polymerase (New England Biolabs). ssl2\(^+\) PCR products were then transformed into yeast along with a linearized pRS315 SSL2 LEU2 plasmid with most of the wild-type SSL2 sequence removed by restriction digest. Leu\(^+\) transformants were selected. Homologous sequences on each end of the ssl2\(^+\) PCR products and the gapped SSL2 vector allowed homologous recombination, resulting in a library of gap-repaired plasmids containing potential ssl2\(^+\) alleles. Since SSL2 is essential, these yeast cells are pre-transformed with a pRSI316 SSL2 URA3 plasmid to support growth, while the genomic SSL2 was deleted to allow plasmid SSL2 alleles to exhibit phenotypes. After gap repair, cells retaining pRSI316 SSL2 URA3 plasmids were killed by replica-plating transformants to medium containing 5FOA. Yeast cells were then plated on YPD media for growth and replica-plated to a variety of media to screen for mutants that have transcription-related or conditional phenotypes. Plasmids from yeast mutants were recovered and transformed into E. coli for amplification, followed by sequencing to identify mutations. Mutant yeast candidates were additionally mated with yeast cells that contain a WT SSL2 URA3 plasmid to create diploid strains and perform phenotyping again to determine dominance/recessivity of ssl2 mutations. Plasmid shuffling on diploid strains was performed by adding 5FOA on the medium so that presumably only presumptive ssl2\(^+\) was kept. This was followed by an additional phenotyping to determine if the mutant phenotype is plasmid linked or not. All mutants described here were verified by retransformation into a clean genetic background.

**Transcription start site sequencing (TSS-seq)**

Yeast cell cultures were grown in triplicates and cells were harvested at mid-log phase at a density of 1X10\(^7\) cells/mL, as determined by cell counting. For S. cerevisiae TSS-seq, cells collected from 50 mL of S. cerevisiae culture and 5 mL of S. pombe culture were mixed and total RNA was extracted as described\(^{113}\). We performed cDNA library construction for TSS-seq essentially as described by Vvedenskaya et al\(^{67}\), steps are described as follows. 100 μg of the isolated total RNA was treated with 30 U of DNase I (QIAGEN) and purified using RNeasy Mini Kit (QIAGEN). A Ribo-Zero Gold rRNA Removal Kit (Illumina) was used to deplete rRNAs from 5 μg of DNase-treated RNAs. The RNA-depleted RNA was purified by ethanol precipitation and resuspended in 10 μL of nuclease-free water. To remove RNA transcripts carrying a 5\(^\prime\) monophosphate moiety (5\(^\prime\)-P), 2 μg of RNA-depleted RNA were treated with 1 U Terminator 5\(^\prime\)-Phosphate-Dependent Exonuclease (Epipcentric) in the 1x Buffer A in the presence of 40 U RNaseOUT in a 50 μL reaction at 30 °C for 1 h. Samples were extracted with acid phenol-chloroform pH 4.5 (ThermoFisher Scientific), and RNA was recovered by ethanol precipitation and resuspended in 30 μL of nuclease-free water. Next, to remove 5\(^\prime\)-terminal phosphates, RNA was treated with 1.5 U CIP (NEB) in 1x NEBuffer 3 in the presence of 40 U RNaseOUT in a 50 μL reaction at 37 °C for 30 min. Samples were extracted with acid phenol-chloroform and RNA was recovered by ethanol precipitation and resuspended in 30 μL of nuclease-free water. To convert 5\(^\prime\)-capped RNA transcripts to 5\(^\prime\)-monophosphate RNAs ligatable to 5\(^\prime\) adaptor, CIP-treated RNAs were mixed with 12.5 U CapClip (Cellscript) and 40 U RNaseOUT in 1X CapClip reaction buffer in a 40 μL reaction and incubated at 37 °C for 1 h. RNAs were extracted with acid phenol-chloroform, recovered by ethanol precipitation and resuspended in 10 μL of nuclease-free water. To ligate the 5\(^\prime\) adapter, the
CapClip-treated RNA products were combined with 1 μM 5' adapter oligonucleotide s1086 (5'-
GUUCAGAGUUCAGAGCGAGAUCNNNNN3'), 1X T4 RNA ligase buffer, 40 U RNaseOUT, 1 mM ATP, 10% PEG 8000 and 10 U T4 RNA ligase 1 in a 30 μL reaction. The mixtures were incubated at 16°C for 16 h and the reactions were stopped by adding 30 μL of 2X RNA loading dye. The mixtures were separated by electrophoresis on 10% 7 M urea slab gels in 1X TBE buffer and incubated with SYBR Gold nucleic acid gel stain. RNA products migrating above the 5' adapter oligo were recovered from the gel as described34, purified by ethanol precipitation and resuspended in 10 μL of nuclease-free water. To generate first strand cDNA, 5'-adaptor-ligated products were mixed with 0.3 μL of 100 μM s1082 oligonucleotide (5'-GCCTTGGCACCAGAATTCCANNNNNNNNN3', N=A/T/G/C) containing a randomized 9-nt sequence at the 3' end, incubated at 65°C for 5 min, and cooled to 4°C. A solution containing 4 μL of 5X First-Strand buffer, 1 μL (40 U) RNaseOUT, 1 μL of 10 mM dNTP mix, 1 μL of 100 mM DTT, 1 μL (200 U) of SuperScript III Reverse Transcriptase and 1.7 μL of nuclease-free water was added to the mixture. Reactions were incubated at 25°C for 5 min, 55°C for 60 min, 70°C for 15 min, and cooled to 25°C. 10 U RNase H was added, the mixtures were incubated 20 min at 37°C and 20 μL of 2X DNA loading solution (PippinPrep Reagent Kit, Sage Science) were added. Nucleic acids were separated by electrophoresis on 2% agarose gel (PippinPrep Reagent Kit, external Marker B) to collect species of ~90 to ~550 nt. cDNA was recovered by ethanol precipitation and resuspended in 20 μL of nuclease-free water. To amplify cDNA, 9 μL of gel-isolated cDNA was added to the mixture containing 1X Phusion HF reaction buffer, 0.2 mM dNTPs, 0.25 μM Illumina RP1 primer (5'-.
AATGATACGCGACCGAGATCTACAGTGTTCATGATCTACAGTCCGAGATCC-3'), and 0.02 U/μL Phusion HF polymerase in 30 μL reaction. PCR was performed with an initial denaturation step of 10 s at 98°C, amplification for 12 cycles (denaturation for 5 s at 98°C, annealing for 15 s at 62°C and extension for 15 s at 72°C), and a final extension for 5 min at 72°C. Amplified cDNAs were isolated by electrophoresis on 2% agarose gel (PippinPrep Reagent Kit, external Marker B) and products of ~180 to ~550 nt were collected. cDNA was recovered by ethanol precipitation and resuspended in 13 μL of nuclease-free water. Barcoded libraries were pooled and sequenced on an Illumina NextSeq platform in high output mode using custom primer s1115 (5'-.
CTACACGTTCAGAGTTCACTACAGTCCGAGATCC-3').

TSS-seq data processing

Quality control on TSS sequencing library FASTQ files was performed to remove reads with low quality using fastq_quality_filter in the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) package with parameters 'fastq_quality_filter -v -q 20 -p 75'. Cutadapt114 was then used to remove the 6 nucleotide 5' linker with parameter of 'cutadapt -u 6'. The resulting reads were trimmed from 3' end to 35 nucleotides long with parameter of 'cutadapt -l 35 --minimum-length=35'. Trimmed reads were mapped to the S. cerevisiae R64-1-1 (SacCer3) genome using Bowtie115 with allowance of no more than two mismatches with suppression of non-uniquely mapped reads 'bowtie -p3 -v2 -m1 -q -sam --un', reported in sam files. Uniquely mapped reads were then extracted from sam files using SAMTools116 and output in bam format 'samtools view -F 4 -S -b'. Bam files were then sorted and converted into bed files by SAMTools 'samtools sort -o', and BEDTools117 'bedtools bamtobed -cigar'. Customized commands were then used on bed files to identify the genomic coordinate of the 5' end of each uniquely mapped read 'awk 'BEGIN{FS=OFS="\t"} $6=="+" {$3==$2+1} $6=="-" {\$2<=$3-1} (print)"'. BEDTools was then used to determine pileup (TSS coverage) across the genome with parameters of 'bedtools genomecov -g R64.new.genome -i -bg -strand - and 'bedtools genomecov -g R64.new.genome -i -bg -strand +', resulting in stranded bedGraph files. FASTQ files of individual library were directly processed or contacted by strains/mutants to generate bedGraph files for correlation analysis. For each of 5979 selected yeast promoters, TSS usage was examined within 401-nt wide window, spanning 250 nt upstream and 150 nt downstream of the previously annotated median TSS. Using customized bash and R scripts, TSS coverage from the bedGraph files of a library or a mutant were assigned into the defined windows to generate a 401×5979 TSS count table, with each row representing one of the 5979 promoters in the same order as the promoter annotation file, each column represents a promoter.
position, and the number in each cell representing 5’ ends mapping to that position. These count tables were stored in csv files. Using customized R script and the count table of a library or a mutant, an **expression-spread-median file** containing promoter expression, median TSS position of the promoter, TSS spread of the promoter was generated. The median TSS position was defined as the actual TSS containing the 50th percentile of the promoter window. The spread of TSS, which measures the width of the middle 80% of TSS distribution, was calculated by subtracting positions of 10th percentile and 90th percentile of TSS counts in 401-nt promoter window and adding 1. The positions of 10th percentile and 90th percentile of TSS counts in each promoter window were also stored in this expression-spread-median file. The streamlined codes to generate bedGraph files, the prompter annotation file and the customized scripts to generate count tables and expression-spread-median files can be found at the Github repository https://github.com/Kaplan-Lab-Pitt/Ssl2_scanning.

**TSS correlation**

TSS coverage data in library-based bedgraph files were used to examine the correlation between TSS libraries by pairwise comparison. A custom R script was used to filter bedgraph files to examine genome positions with greater than two counts in each library. Log2 transformed TSS counts at the same genomic location in two examined libraries were plotted for all TSS sites to create a heat scatter plot using the LSD R package117 and the Pearson correlation coefficient was calculated. The correlation coefficients deriving from all pairwise comparisons were plotted by a web-based heatmap tool Morpheus (https://software.broadinstitute.org/morpheus/) and clustered by Euclidian distance. Replicates with correlation coefficient greater than 0.85 and the shortest Euclidian distance to each other in the clustering analysis among all the analyzed libraries were recognized as having good sequencing reproducibility and used for downstream analysis.

**TSS count table and heatmaps**

For each of 5979 selected promoters, TSS usage was examined within a 401-nt wide window, spanning 250-nt upstream and 150-nt downstream of the previously annotated median TSS15. Using BEDTools and customized R scripts, TSS coverage from the bedgraph files of a library or mutant were assigned into the defined windows to generate a 401×5979 TSS count table, with each row representing one of the 5979 promoters, each column represents a promoter position, and the number in each cell representing 5’ ends mapping to that position. The count table was filtered to keep data from n=4392 promoters with ≥100 sequence reads on average per WT library and used for downstream analyses. The filtered count table was row-normalized to get the relative TSS usage at each promoter position. TSS distribution differences were determined by subtracting normalized WT data (concatenated from both RPB1 and SSL2 WT libraries) from normalized mutant data and visualized using heatmaps (Morpheus).

**TSS metrics**

Data analyses for TSS distributions were based on customized R scripts and results were plotted in Graphpad Prism 8 (https://www.graphpad.com/scientific-software/prism/) unless otherwise indicated. The median TSS position was defined as the actual TSS containing the 50th percentile of the TSS distribution and was determined for each promoter. “TSS shift” represents the difference in nucleotide of the median TSS position for each promoter between two libraries or mutants. The distributions of TSS shifts for n=4392 promoters with ≥100 sequence reads on average per WT library or selected promoter classes in each library or mutant were illustrated by both heatmap (Morpheus) and boxplots.

**TSS spread**

The spread of TSS, which measures the width of the middle 80% of TSS distribution, was calculated by subtracting positions of 10th percentile and 90th percentile of TSS reads in 401-nt promoter window and adding 1. TSS spread of selected promoters are shown in boxplot and compared between libraries by performing one-way ANOVA. The differences of TSS spreads between WT and the mutant for selected promoter classes were presented in heatmap (Morpheus) and/or boxplot.
ChIP-exo data processing

ChIP-exo data processing was performed as described by Rossi et al. in Nature Communications, 2018 and Qiu et al. in Genome Biology, 2020. Briefly, ChIP-exo libraries were sequenced on a NextSeq 500 in paired-end mode to generate 40 (read1) × 36 bp (read2) reads. Reads passing Q30 quality threshold were then aligned to the sacCer3 genome using the BWA-MEM alignment algorithm (v0.7.9a) with default parameters. After alignment, PCR duplicates were removed using Picard and SAMtools assuming unique combinations of read1 and read2 were PCR duplicates. Using ScriptManager v0.12 (https://github.com/CEGRcode/scriptmanager), BAM files of a library were assigned into two 401×5979 matrices and saved in CDT files, which stores counts of 5' position of protein binding on top and bottom strands, respectively. The same as in TSS-seq data analysis, each row of 401×5979 matrix representing one of 5979 promoters and each column representing a position in the 401-nt promoter window. These 401×5979 matrices were also saved in csv format. Matrices from the same mutant were combined into a single matrix by adding counts in library matrices at the same dimension and saved in csv files. Similar to TSS-seq analysis, the customized R script and the matrix of a library or a mutant were used to generate an expression-spread-median file containing promoter expression, median 5' position of protein binding, the binding site of the spread of the promoter and saved in txt files.

Availability of data and materials

Genomics datasets generated in the current study are available in the NCBI BioProject and SRA, under the accession numbers of PRJNA681384 and SRP295731, respectively. The processed genomic data files are available in GEO, under the accession number of GSE182792. The streamlined commands to generate TSS-seq bedGraph files, count tables, tables of expression, spread and median TSS can be found at https://github.com/Kaplan-Lab-Pitt/Ssl2_scanning. ChIP-exo and data analysis was performed as described by Rossi et al. and Qiu et al. Source data files are listed in Supplementary File 3.

Acknowledgements

The authors thank Chenxi Qiu for R scripts used in statistical analysis of Kruskal-Wallis test with Dunn’s correction and Mann-Whitney U test. We are deeply grateful to Scott Kuerten and Fred Hyde (Illumina) for advice and give of legacy Ribozero reagents for yeast rRNA removal.

Additional information

BFP has a financial interest in Peconic, LLC, which utilizes the ChIP-exo technology implemented in this study and could potentially benefit from the outcomes of this research. All other authors declare that they have no competing interests. All other authors declare that they have no competing interests.

Funding

<table>
<thead>
<tr>
<th>Funder</th>
<th>Grant reference number</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Institutes of Health</td>
<td>NIH R01GM120450</td>
<td>Craig D. Kaplan</td>
</tr>
<tr>
<td>National Institutes of Health</td>
<td>NIH R35GM118059</td>
<td>Bryce E. Nickels</td>
</tr>
<tr>
<td>National Institutes of Health</td>
<td>NIH R01GM059055</td>
<td>B. Frank Pugh</td>
</tr>
</tbody>
</table>
References


925 24(12):1139-1145. PMCID or (doi): PMC5741190.
927 16. Fishburn J, Tomko E, Galburt E, Hahn S: Double-stranded DNA translocase activity of
928 transcription factor TFIH and the mechanism of RNA polymerase II open complex
931 17. Fazal FM, Meng CA, Murakami K, Kornberg RD, Block SM: Real-time observation of the
933 (doi): (10.1038/nature14882) http://www.ncbi.nlm.nih.gov/pubmed/26331540
934 18. Struhl K: Promoters, activator proteins, and the mechanism of transcriptional initiation in
936 19. Goel S, Krishnamurthy S, Hampsey M: Mechanism of start site selection by RNA
937 polymerase II: interplay between TFIIB and Ssl2/XPB helicase subunit of TFIH. J Biol
940 20. Yang C, Ponticelli AS: Evidence that RNA polymerase II and not TFIIB is responsible for
941 the difference in transcription initiation patterns between Saccharomyces cerevisiae and
945 cerevisiae TFIIF during transcription start site utilization. Mol Cell Biol 2008, 28(11):3757-
947 22. Pal M, Ponticelli AS, Luse DS: The role of the transcription bubble and TFIIB in promoter
950 23. Majovski RC, Khapersky DA, Ghazy MA, Ponticelli AS: A functional role for the switch 2
951 region of yeast RNA polymerase II in transcription start site utilization and abortive
954 24. Freire-Picos MA, Krishnamurthy S, Sun ZW, Hampsey M: Evidence that the Tfg1/Tfg2 dimer
955 interface of TFIIF lies near the active center of the RNA polymerase II initiation complex.
959 yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with
962 26. Chen BS, Hampsey M: Functional interaction between TFIIB and the Rpb2 subunit of RNA
963 polymerase II: implications for the mechanism of transcription initiation. Mol Cell Biol
964 2004, 24(9):3983-3991. PMCID or (doi): 387735.
966 27. Faitar SL, Brodie SA, Ponticelli AS: Promoter-specific shifts in transcription initiation
967 conferred by yeast TFIIB mutations are determined by the sequence in the immediate
970 28. Pappas DL, Jr., Hampsey M: Functional interaction between Ssu72 and the Rpb2 subunit
973 29. Wu WH, Pinto I, Chen BS, Hampsey M: Mutational analysis of yeast TFIIB. A functional
974 relationship between Ssu72 and Sub1/Tsp1 defined by allele-specific interactions with


1031 43. Qiu H, Park E, Prakash L, Prakash S: The Saccharomyces cerevisiae DNA repair gene
1034 44. Lee BS, Lichtenstein CP, Faiola B, Rincikel LA, Wysock W, Curcio MJ, Garfinkel DJ:
1035 Posttranscriptional inhibition of Ty1 retrotransposition by nucleotide excision
1038 45. Kuehner JN, Brow DA: Regulation of a eukaryotic gene by GTP-dependent start site
1040 46. Jenks MH, O'Rourke TW, Reines D: Properties of an intergenic terminator and start site
1043 47. Kaplan CD, Jin H, Zhang IL, Belyanin A: Dissection of Pol II trigger loop function and Pol II
1046 48. Eichner J, Chen HT, Warfield L, Hahn S: Position of the general transcription factor TFIIIF
1049 49. Kaplan CD, Larsson KM, Kornberg RD: The RNA polymerase II trigger loop functions in
1055 structure to systems: high-resolution, quantitative genetic analysis of RNA polymerase II.
1058 XPB DNA helicase gene (ERCC3): xeroderma pigmentosum without and with Cockayne
1062 52. Cleaver JE, Thompson LH, Richardson AS, States JC: A summary of mutations in the UV-sensitive disorders: xeroderma pigmentosum, Cockayne syndrome, and
associated with mutations in the DNA repair and transcription gene ERCC3. *Am J Hum
1080 56. Swedler KS, Hanawalt PC: The COOH terminus of suppressor of stem loop (SSL2/RAD25)
in yeast is essential for overall genomic excision repair and transcription-coupled repair.


https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6078178/

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3902107

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5580651


https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6315372

https://www.ncbi.nlm.nih.gov/pmc/articles/33219055


Kitamura H, Kitano H, Kollias G, Krishnan SP, Kruger A, Kummerfeld SK, Kurochkin IV, Lareau  
N, Nakano N, Nakauchi H, Ng P, Nilsson R, Nishiguchi S, Nishikawa S, Nori F, Ohara O,  
J, Reid JF, Ring BZ, Ringwald M, Rost B, Ruan Y, Salzberg SL, Sandelin A, Schneider C,  
Schonbach C, Sekiguchi K, Semple CA, Seno S, Sessa L, Sheng Y, Shibata Y, Shimada H,  
Grindom SM, Teasdale RD, Liu ET, Brusic V, Quackenbush J, Wahlestedt C, Mattick JS,  
Y, Consortium F, Group RGER, Genome Science G: The transcriptional landscape of the  
mammalian genome. Science 2005, 309(5740):1559-1563. PMCID or (doi):  
106. Kadonaga JT: Perspectives on the RNA polymerase II core promoter. Wiley Interdiscip Rev  
107. Haberle V, Stark A: Eukaryotic core promoters and the functional basis of transcription  
108. Luse DS, Parida M, Spector BM, Nilson KA, Price DH: A unified view of the sequence and  
functional organization of the human RNA polymerase II promoter. Nucleic Acids Res  
2020, 48(14):7767-7785. PMCID or (doi): PMC7641323.  
109. Chou S-P, Alexander AK, Rice EJ, Choate LA, Cohen PE, Danko CG: Genetic Dissection of  
or (doi): (10.1101/2021.05.23.445279) http://biorxiv.org/content/early/2021/05/30/2021.05.23.445279.abstract  
111. Amberg DC, Burke DJ, Strathern JN: Methods in Yeast Genetics: A Cold Spring Harbor  
Spring Harbor Press; 2005  
112. Ranish JA, Hahn S: The yeast general transcription factor TFIIA is composed of two  
or (doi): http://www.ncbi.nlm.nih.gov/pubmed/1918049  
from Saccharomyces cerevisiae. Nucleic acids research 1990, 18(10):3091-3092. PMCID or  
114. Martin M: Cutadapt removes adapter sequences from high-throughput sequencing reads.  
EMBNetjournal; Vol 17, No 1: Next Generation Sequencing Data AnalysisDO -  
1014806/ej/171200 2011. PMCID or (doi):  
115. Langmead B, Trapnell C, Pop M, Salzberg SL: Ultrafast and memory-efficient alignment of  
short DNA sequences to the human genome. Genome Biol 2009, 10(3):R25. PMCID or (doi):  
Genome Project Data Processing S: The Sequence Alignment/Map format and SAMtools.  
26

Li H: Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv 2013, 1303.3997. PMCID or (doi):

Supplementary Files

Supplementary File 1. Yeast strains used in this study.

Supplementary File 2. Bacterial strains and plasmids used in this study.

Supplementary File 3. List of sources data files and their association with figures in this study.

Source data files

Figure 1-source data 1. Figure 1E Primer Extension Gel (annotated).

Figure 1-source data 2. Figure 1G Primer Extension Gel (annotated).

Figure 1-source data 3. Figure 1E,G Graph data.

Figure 1-source data 4. Figure 1E Primer Extension Gel (raw).

Figure 1-source data 5. Figure 1G Primer Extension Gel (raw).

Figure 1-Figure supplement 1 source data 1. Figure 1-Figure supplement 1A Graph data.

Figure 2-source data 1. Figure 2F,H Graph data.

Figure 2-Figure supplement 2-source data 1. Figure 2-Figure supplement 2A Primer Extension Gel (annotated).

Figure 2-Figure supplement 2-source data 1. Figure 2-Figure supplement 2A Primer Extension Gel (annotated).

Figure 2-Figure supplement 2-source data 3. Figure 2-Figure supplement 2B Graph data.

Figure 2-Figure supplement 2-source data 4. Figure 2-Figure supplement 2A Primer Extension Gel (raw).

Figure 2-Figure supplement 2-source data 5. Figure 2-Figure supplement 2A Primer Extension Gel (raw).

Figure 3-source data 1. Figure 3C Heat map data.

Figure 3-Figure supplement 1-source data 1. Figure 3-Figure supplement 1B Heat map data.

Figure 4-source data 1. Figure 4B ssl2 N230D Heat map data.

Figure 4-source data 2. Figure 4B ssl2 N230I Heat map data.

Figure 4-source data 3. Figure 4DE Taf1-Enriched Heat map and graph data.

Figure 4-source data 4. Figure 4DE Taf1-Depleted Heat map and graph data.

Figure 4-Figure supplement 1-source data 1. Figure 4-Figure supplement 1 rpb1 E1103G Heat map data.
Figure 4-Figure supplement 1-source data 2. Figure 4-Figure supplement 1 rpb1 H1085Y Heat map data.

Figure 4-Figure supplement 1-source data 3. Figure 4-Figure supplement 1 ssl2 L225P Heat map data.

Figure 4-Figure supplement 1-source data 4. Figure 4-Figure supplement 1 ssl2 D522V Heat map data.

Figure 4-Figure supplement 1-source data 5. Figure 4-Figure supplement 1 ssl2 R636C Heat map data.

Figure 4-Figure supplement 1-source data 6. Figure 4-Figure supplement 1 ssl2 Y750* Heat map data.

Figure 4-Figure supplement 2-source data 1. Figure 4-Figure supplement 2A Heat map data.

Figure 5-source data 1. Figure 5B Taf1-Enriched Graph data.

Figure 5-source data 2. Figure 5B Taf1-Depleted Graph data.

Figure 5-source data 3. Figure 5CD Taf1-Enriched Heat map and graph data.

Figure 5-source data 4. Figure 5CD Taf1-Depleted Heat map and graph data.

Figure 5-Figure supplement 1-source data 1. Figure 5-Figure supplement 1A Heat map data.

Figure 5-Figure supplement 2-source data 1. Figure 5-Figure supplement 2A Graph data.

Figure 5-Figure supplement 2-source data 2. Figure 5-Figure supplement 2B Data table.

Figure 5-Figure supplement 2-source data 3. Figure 5-Figure supplement 2B Graph data.

Figure 5-Figure supplement 2-source data 4. Figure 5-Figure supplement 2C Data table.

Figure 5-Figure supplement 2-source data 5. Figure 5-Figure supplement 2C Graph data.

Figure 5-Figure supplement 2-source data 6. Figure 5-Figure supplement 2D Graph data.

Figure 5-Figure supplement 2-source data 7. Figure 5-Figure supplement 2E Data table.

Figure 5-Figure supplement 2-source data 8. Figure 5-Figure supplement 2E Graph data.

Figure 5-Figure supplement 2-source data 9. Figure 5-Figure supplement 2F Data table.

Figure 5-Figure supplement 2-source data 10. Figure 5-Figure supplement 2F Graph data.

Figure 6-source data 1. Figure 6G Heat map data.

Figure 6-source data 2. Figure 6H Primer extension gel (annotated).

Figure 6-source data 3. Figure 6H Primer extension gel (annotated).

Figure 6-source data 4. Figure 6I rpb1 singles Heat map data.
Figure 6-source data 5. Figure 6I N230D single Heat map data.

Figure 6-source data 6. Figure 6I N230D double Heat map data.

Figure 6-source data 7. Figure 6J Primer extension gel (annotated).

Figure 6-source data 8. Figure 6J Primer extension gel (annotated).

Figure 6-source data 9. Figure 6K rpb1 singles Heat map data.

Figure 6-source data 10. Figure 6K N230I single Heat map data.

Figure 6-source data 11. Figure 6K N230I double Heat map data.

Figure 6-source data 12. Figure 6H Primer extension gel (raw).

Figure 6-source data 13. Figure 6H Primer extension gel (raw).

Figure 6-source data 14. Figure 6J Primer extension gel (raw).

Figure 6-source data 15. Figure 6J Primer extension gel (raw).

Figure 6-Figure supplement 3-source data 1. Figure 6-Figure supplement 3A Graph data.

Figure 6-Figure supplement 3-source data 2. Figure 6-Figure supplement 3B Graph data.

Figure 7-source data 1. Figure 7A Heat map data.

Figure 7-source data 2. Figure 7B Heat map data.

Figure 7-source data 3. Figure 7C Heat map data.

Figure 7-source data 4. Figure 7D Heat map data.

Figure 7-Figure supplement 1-source data 1. Figure 7-Figure supplement 1C Primer extension gel (annotated).

Figure 7-Figure supplement 1-source data 2. Figure 7-Figure supplement 1D Graph data.

Figure 7-Figure supplement 1-source data 3. Figure 7-Figure supplement 1G Primer extension gel (annotated).

Figure 7-Figure supplement 1-source data 4. Figure 7-Figure supplement 1H Graph data.

Figure 7-Figure supplement 1-source data 5. Figure 7-Figure supplement 1C Primer extension gel (raw).

Figure 7-Figure supplement 1-source data 6. Figure 7-Figure supplement 1G Primer extension gel (raw).

Figure 7-Figure supplement 2-source data 1. Figure 7-Figure supplement 2B Primer extension gel (annotated).
Figure 7-Figure supplement 2-source data 2. Figure 7-Figure supplement 2B Primer extension gel (annotated).

Figure 7-Figure supplement 2-source data 3. Figure 7-Figure supplement 2B Primer extension gel (annotated).

Figure 7-Figure supplement 2-source data 4. Figure 7-Figure supplement 2B Primer extension gel (annotated).

Figure 7-Figure supplement 2-source data 6. Figure 7-Figure supplement 2B Primer extension gel (raw).

Figure 7-Figure supplement 2-source data 7. Figure 7-Figure supplement 2B Primer extension gel (raw).

Figure 7-Figure supplement 2-source data 8. Figure 7-Figure supplement 2B Primer extension gel (raw).

Figure 7-Figure supplement 2-source data 9. Figure 7-Figure supplement 2B Primer extension gel (raw).

Figure 7-Figure supplement 2C Graph data.

Figure 9-Figure supplement 1-source data 1. Figure 9-Figure supplement 1A Graph data.

Figure 9-Figure supplement 1C Graph data.
Figure 1: Genetic screening identifies novel ssl2 alleles with transcriptional defects. (A) Schematics illustrating transcriptional phenotypes of four transcriptional phenotypes utilized in this study (IMD2, imd2Δ::HIS3, gal10Δ56, lys2-128Δ). IMD2, In GTP replete conditions, IMD2 transcription initiates at upstream TSSs utilizing GTP as the first nucleotide. These are non-functional due to the presence of a terminator prior to the IMD2 coding sequence. Upon GTP starvation induced by the drug MPA, initiation shifts to a downstream ATP-initiated TSS, enabling functional IMD2 expression.
confering resistance to MPA. Inability to shift to the downstream IMD2 TSS in the presence of MPA leads to MPA sensitivity (MPA$^S$), commonly found in mutants that shift TSSs upstream. imd2Δ::HIS3, The IMD2 ORF is replaced by the HIS3 ORF creating a transcriptional fusion between the IMD2 promoter and HIS3. If the downstream IMD2 TSS is used constitutively, HIS3 mRNA production to supports growth on medium lacking histidine (SC-His). Inappropriate use of the downstream IMD2 start site by downstream TSS shifting mutants in the absence of MPA can be detected by a His+ phenotype in cells with the imd2Δ::HIS3 reporter. gal10Δ56, Deletion of GAL10 polyadenylation signal at p(A) site interferes with GAL10 3'-end formation and GAL7 initiation, resulting in galactose toxicity$^{77}$. Transcriptional mutants that increase GAL10 3'-end formation or GAL7 initiation allow suppression of galactose toxicity and display galactose resistance (GAL$^R$). lys2-128δ, Insertion of a Ty transposon δ element into LYS2 causes premature termination of Pol II initiating at LYS2, resulting in lysine auxotrophy. Certain mutants allow expression of LYS2 from a cryptic site within the Ty δ element and allow yeast growth on medium lacking lysine (SC-Lys), conferring the “Suppressor of Ty” (Spt) phenotype. (B) Schematic illustrating the plasmid shuffle assay to examine ssl2 mutant phenotypes (see Materials and Methods). (C) Transcription-related growth phenotypes of five classical ssl2 alleles. All spot dilutions shown throughout manuscript are representative of at least two independent transformants (D) Schematic illustrating the TSS region of S. cerevisiae ADH1. The ADH1 promoter contains two commonly used TSSs (red letters), which are 37nt (-37) and 27nt (-27) upstream from the translational start codon. For quantification of changes to ADH1 TSS distribution by primer extension, ADH1 TSS usage is separated into six bins. (E) Left panel, primer extension-detected TSS usage of the WT and five existing ssl2 mutants at ADH1. Right panel, quantitative analysis of five classical ssl2 alleles at ADH1. Average of ≥ 3 biological replicates +/- standard deviation are shown. (F) Transcription-related growth phenotypes of ssl2 alleles homologous to human disease alleles of XPB. (G) Primer extension detection of TSS usage at ADH1 for alleles shown in (F). Average of ≥ 3 biological replicates +/- standard deviation are shown.
Figure 2: ssl2 alleles have distinct behavior from Pol II and other GTFs alleles for TSS usage.

(A) Identified ssl2 substitutions and their phenotypes relative to Ssl2 domains. Structure colored by identified Ssl2 domains as in B4. Light grey areas in the schematic have not yet been observed in any Ssl2 or homolog structures. (B) Position of Ssl2 relative to TFIIH and Pol II in the yeast PIC (PDB 7O4I). (C) Identified substituted residues illustrated as spheres on cartoon of the Ssl2 structure from (B). Residue numbers are color coded by Ssl2 domain color from (A). (D) Rotation of (C) illustrating the mutants clustered at NTD/DRD/Helicase Domain (HD) 1 interface. (E) Spot assay showing example rpb1 mutant transcription phenotypes. (F) Primer extension and quantification showing rpb1 mutant effects on ADH1 TSS distribution. Color coding of rpb1 allele class in this graph is used throughout the figures. Green bars represent upstream shifting rpb1 alleles when used to annotate figures. Blue bars represent downstream shifting rpb1 alleles. Average of ≥ 3 biological replicates +/- standard deviation are shown. (G) Spot assay showing example ssl2 mutant transcription phenotypes. (H) Primer extension and quantification showing example ssl2 mutant effects on ADH1 TSS distribution. Color coding of ssl2 allele class in this graph is used throughout the figures. Orange bars represent downstream shifting ssl2 alleles when used to annotate figures. Purple bars represent upstream shifting ssl2 alleles. Average of ≥ 3 biological replicates +/- standard deviation are shown.
Figure 3: Transcription start site sequencing (TSS-seq) identifies global ss/l2 initiation effects. (A) TSS-seq library construction as in67. See Materials and Methods for details. (B) Scatter plot showing the correlation of log2 transformed reads at individual genome positions for all positions ≥3 reads for each position for example replicate pairs for SSL2 WT, ssl2 N230D, or ssl2 N230I, see Figure supplements for other libraries and description of biological replicates performed for all genotypes. (C) Hierarchical cluster of Pearson r correlation coefficients for libraries of combined biological replicates.
**Figure 4**: *ssl2* alleles shift transcription start site distribution genome-wide. (A) Schematic indicating TSS distribution within promoter window defined by median of the WT TSS distribution. (B) Heatmaps of TSS read differences between WT and *ssl2* mutants within defined promoter windows. Promoter classes defined by enrichment or depletion of Taf1 were rank-ordered according to total reads in WT. The promoter-normalized read differences between mutant and WT are shown by a color scheme ranging from cyan (negative)-white-orange (positive). (C) Schematic illustrating median TSS upstream or downstream shift in mutant relative to WT. (D) Heatmap and hierarchical clustering of median TSS shifts for *ssl2* or *rpb1* mutants for Taf1-enriched or -depleted promoter classes (promoters as in (B)). The shift of median TSS is indicated by cyan (upstream) or orange (downstream). (E) TSS shifts in *ssl2* mutants are less strong than in Pol II mutants. Median TSS shifts across promoters, regardless of promoter class are statistically distinguished from zero at p<0.0001 for all mutants (Wilcoxon Signed Rank Test).
Figure 5: Distinct alterations to TSS distribution in ss12 mutants. (A) Schematic illustrating TSS "spread" reflecting the distance encompassing 80% of TSSs in a promoter window and the measurement of mutant changes in TSS spread. (B) TSS spreads in ss12 and Pol II mutants at Taf1-enriched or -depleted promoters. All mutants show a statistical difference in medians from WT at p<0.05 (Kruskal-Wallis test with Dunn's multiple comparisons test). Asterisks indicate differences in means from WT (p<0.05 *, p<0.01 **, p<0.005 ***, p<0.001 ****). (C) Heatmap showing TSS spread...
changes for \textit{ssl2} or \textit{rpb1} mutants by promoter class (hierarchical clustering by mutant on the \textit{x-axis} and promoter on the \textit{y-axis}). (D) \textit{ssl2} upstream and downstream shifting mutants narrow and widen TSS distributions at promoters as measured in (A). The median TSS spread changes of all the \textit{ssl2} and Pol II mutants are statistically distinguished from zero at \(p<0.0001\) (Wilcoxon Signed Rank Test).
Figure 6: Genetic interactions between ssl2 and Pol II initiation alleles suggest distinct functions of each in initiation by scanning. (A) Growth phenotypes of rpb1, ssl2 N230D, ssl2 N230I single or double mutants. rpb1 mutants represent known catalytically hyperactive alleles or genetically similar (G1097D, E1103G, L1101S, F1084I) and four with reduced catalytic activity (F1086S, H1085Q, N1082S, H1085Y). Strains are arranged according to measured Pol II elongation rate in vitro (slowest at top). (B) ssl2 MPA sensitive alleles are epistatic to Pol II LOF alleles’ His⁺ phenotypes (double mutants retain MPA⁻ of ssl2 single mutant while His⁺ phenotypes of rpb1 mutants are suppressed. Conversely, Pol II TSS upstream shifting alleles appear epistatic/non-additive with ssl2 MPA⁻ alleles and do not show synthetic growth phenotypes. (C) Pol II upstream TSS shifting alleles appear epistatic to ssl2 N230I phenotypes (MPA⁻ retained and His⁺ suppressed in double mutants). There are only minor synthetic defects between ssl2 N230I and Pol II downstream TSS shifting mutants suggesting lack of synergistic defect and either mild additivity or epistasis. (Double mutant of N230I and H108Y is nearly dead and was not tested here or in E.) (D,E) Schematic (D) indicating how qualitative growth data of mutants encoded (E) for visualization in heat maps. (F) Phenotyping heat map legend. (G) Qualitative Heatmaps for ssl2 and rpb1 genetic interactions. Growth Phenotypes are detected using reporters described in Figure 1. (H) Primer extension of ssl2 N230D and rpb1 mutants at ADH1. ssl2 N230D appears to truncate distribution of TSSs on downstream side and is epistatic to downstream shifting rpb1 alleles (blue bar) while upstream shifting rpb1 alleles (green bar) are non-additive or epistatic to ssl2 N230D. Numbered regions indicate TSS positions that were binned for quantification in (I). Representative primer extension of ≥ 3 independent biological replicates is shown. (I) Quantification of (H) with heat map showing relative differences in TSS distribution binned by position (bins are numbered and shown in (H). Mean changes of ≥ 3 independent biological replicates are shown in the heat map. (J) Primer extension of ssl2 N230I and rpb1 mutants at ADH1. ssl2 N230I appears to enhance usage of downstream TSSs and is additive with downstream shifting rpb1 alleles (blue bar) while upstream shifting rpb1 alleles (green bar) are epistatic to ssl2 N230I. Numbered regions indicate TSS positions that are binned for quantification in (K). Representative primer extension of ≥ 3 independent biological replicates is shown. (K) Quantification of (J) with heat map showing relative differences in TSS distribution binned by position (bins are numbered and shown in (J). Mean changes of ≥ 3 independent biological replicates are shown in the heat map.
Figure 7: Complex genetic interactions between GTF initiation alleles suggest multiple distinct activities in initiation by scanning. (A) Genetic interactions between ssl2, TFIIB, TFIIF, and sub1 mutants shown as a heat map indicating phenotypic strength of single and double mutants. Scaling as in 6F. (B-D) Heatmaps showing quantified \textit{ADH1} primer extension data for ssl2, sua7-1, tfg2, sub1Δ single and double mutants. Primer extension as in Figure 1E etc. Mean changes of ≥ 3 independent biological replicates are shown in the heat maps.
Figure 8: Two major functional networks controlling initiation by scanning. In our genetic experiments, additive/suppressive effects are mainly observed between alleles predicted to function by alteration to initiation efficiency (\(rpb1\) and tested alleles of TFIIB/TFIIF). Multiple lines between classes indicates allele specific interactions between a factor and individuals of an allele class e.g. \(sub1\Delta\) and \(ssl2\) LOF alleles. In contrast to interactions within the “efficiency” network, widespread epistasis was observed between \(ssl2\) and other factors as predicted for interactions between processivity and efficiency alleles. \(sub1\Delta\) generally shows additivity/suppression with \(ssl2\) alleles, consistent with it functioning as a scanning processivity factor. Unique lethal interactions between \(ssl2\) LOF upstream shifter F498L and \(sua7-1\) and \(sub1\Delta\) indicate distinct behavior within the \(ssl2\) LOF class.
Figure 9: Model for interaction between initiation efficiency and scanning processivity. (A) The “Shooting Gallery” model. The Pol II active site controls initiation efficiency, i.e. “the rate of firing”. TFIIH controls the rate of the extent of scanning, i.e. “the speed of target passage and number of targets reached.” (B) Reduction in relative TSS usage as scanning Pol II initiates. As Pol II (WT) scans from upstream to downstream, successful initiation at upstream positions will reduce the amount of Pol II continuing to scan downstream. Increasing initiation efficiency at each position as is predicted for increased Pol II catalytic activity will result in a more rapid decrease in observed initiation from upstream to downstream. Conversely, reducing initiation efficiency at each position will flatten observed TSS distribution because more Pol II will reach downstream positions. (C) TSS distributions during promoter scanning in the “Shooting Gallery” model. The TSS distribution (black arrows) of a promoter window can be affected by Pol II catalytic activity, PIC scanning rate and processivity, TSS strength, Pol II flux, and additional observed (upstream limitation on initiation too close to PIC assembly) or potential (downstream limitation through chromatin structure) constraints. (D) Effects of Pol II catalytic activity on TSS distributions. Increased Pol II catalytic activity increases the efficiency of upstream TSSs that are encountered by Pol II and decreases the usage of downstream TSSs due to quickly reduced Pol II flux (changes indicated as green arrows). Decreased Pol II catalytic activity decreases TSS efficiency of upstream TSSs encountered by Pol II and increases apparent TSS usage at downstream sites due to failed upstream initiation, resulting in a downstream shifted TSS distribution within a window determined by PIC scanning potential (changes shown as blue arrows). (E) Effects of altered scanning processivity on TSS distributions. Increased processivity alleles are hypothesized to increase the probability of Pol II scanning further downstream if Pol II flux remains, thus expanding the scanning window and allowing Pol II usage of downstream TSSs if Pol II flux is not limiting (orange TSS). In contrast decreased processivity will limit Pol II scanning downstream, truncating the distribution of observed TSSs (purple TSS).
Figure 1 – Figure supplement 1. Growth phenotypes of human disease related and RED motif ssl2 mutants. (A) Example of TSS usage quantification at ADH1. Left panel, primer extension detected TSS usage at ADH1 in WT and rpb1 E1103G, an upstream shifting TSS mutant. Middle panel, distribution of WT and mutant TSS signal within six ADH1 bins. Each bin is quantified as the percentage of the total signal in the lane and displayed in a bar graph. Average of ≥ 3 biological replicates +/- standard deviation are shown. Right panel, quantitative showing the changes of TSS usage in rpb1 E1103G compared to WT. (B) Viability of ssl2 alleles with mutations homologous to human disease but inviability of RED motif substitutions (allele marked as “RED” is a triple substitution of R519A/E520A/D521A).
Figure 2 – Figure supplement 1. Growth phenotypes of ssl2 mutants. Growth phenotypes of novel ssl2 mutants illustrated by spotting of serial dilution (10-fold) of strains on different media. Phenotypes and media are as described in Figure 1. Serial dilution phenotypes shown are representative of at least two independent transformants (biological replicates).
Figure 2 – Figure supplement 2. TSS usage of representative ssl2 single mutants. (A) TSS usage for selected ssl2 alleles at ADH1 detected by primer extension. The purple bar indicates mutants showing MPA sensitivity (upstream shifting) and the orange bar indicates mutants showing the His⁺ phenotype (downstream shifting). Primer extensions shown are representative of ≥ 3 independent biological replicates. (B) Quantification of TSS usage for selected ssl2 alleles at ADH1 in (A). Error bars indicate average +/- SD of at least three independent biological replicates.
Figure 2 – Figure supplement 3. Alignment of Ssl2 homologs illustrating position of substitutions and the conservation of affected residues. Protein sequences of Ssl2 and homologs from a selection of eukaryotes were aligned and displayed in Jalview. Locations of ssl2 mutations identified in this research were red-boxed on the yeast sequence (“RAD25_YEAST”). Locations of ssl2 mutations were marked with red or green stars if they were annotated in the yeast PIC structure.
* Technical replicates
Figure 3 – Figure supplement 1. Correlation of read counts between TSS-seq replicates at TSSs across the genome. (A) Scatter plot showing the correlation of sequencing reads across all genome positions for all positions ≥3 reads between replicates in ssl2 and Pol II mutants. (B) Matrix heatmap and the hierarchical clustering of Pearson correlation coefficients between TSS-seq libraries from independent biological replicates labeled by substitution and replicate number (H1085Y and E1103G are rpb1 alleles, all others are ssl2). SSL2WT_2,3,4 are technical replicates as are SSL2WT_5,6,7. All other libraries represent biological replicates. Color bars indicate allele class as described in Figure 1.

Figure 4 – Figure supplement 1. Analysis of TSS shift in ssl2 and Pol II mutants. Difference heatmaps of TSS distributions between WT and other ssl2 mutants as described in Figure 4B. The number of points in the figure is lower than the resolution so some averaging will be present. Note that the color scale has been set with sensitive thresholds to have effects be more visually obvious.
Figure 4 – Figure supplement 2. Analysis of TSS shift in ssl2 and Pol II mutants. (A) Heatmap and hierarchical clustering of median TSS shift for ssl2 or rpb1 mutant in individual replicates. As described in Figure 4D except performed for individual replicate TSS-seq libraries. SSL2WT_2,3,4 are technical replicates as are SSL2WT_5,6,7. All other libraries represent biological replicates. (B) PCA analysis for median TSS shift of ssl2 and Pol II mutants in individual replicate TSS-seq libraries. TSS-seq libraries are separated into five major classes (ssl2 upstream and downstream mutants, WT, Pol II upstream and downstream mutants) by the first dimension consistent with direction of shift, and the second dimension consistent with the different genes.
Figure 5 – Figure supplement 1. Analysis of TSS spread in ssl2 and Pol II mutants. (A) Heatmap and hierarchical clustering of TSS spread change for ssl2 or rpb1 mutant in individual replicate. As described in Figure 5C. SSL2WT_2,3,4 are technical replicates as are SSL2WT_5,6,7. All other libraries represent biological replicates. (B) PCA analysis for change of TSS spread in ssl2 and Pol II individual TSS-seq libraries. TSS-seq libraries are separated into five major classes (ssl2 upstream and downstream mutants, WT, Pol II upstream and downstream mutants).
Figure 5 – Figure supplement 2. Number of promoters affected by *ssl2* or Pol II mutants. (A) Bar plot showing number of promoters that shift median TSS position either upstream or downstream in each mutant for any degree of shift. (B) Bar plot indicating number of promoters that shift median TSS position significantly upstream or downstream in each mutant, as determined by Kruskal-Wallis test with Dunn’s correction for multiple comparisons (p≤0.05). The test was performed individually on each promoter and all mutants were taken into account. (C) Bar plot showing number of promoters that shift median TSS position significantly upstream or downstream in each mutant as in (B) but determined by Mann-Whitney U test (p≤0.05). (D-F) As tested in A-C but examined for change of TSS spread instead of median TSS shift. (D) Bar plot showing number of promoters that have narrower or wider TSS spread in each mutant. TSS spread was determined as described in Figure 5A. (E) Bar plot indicating number of promoters that have significantly increased or decreased TSS spread in each mutant as determined by Kruskal-Wallis test with Dunn’s correction for multiple comparisons (p≤0.05). The test was also performed individually on each promoter and all the mutants were taken into account. (F) Bar plot showing number of promoters that have significantly increased or decreased TSS spread in each mutant as in E but determined by Mann-Whitney U test (p≤0.05).

Figure 6 – Figure supplement 1. Design of *ssl2* genetic interaction tests with efficiency alleles. (A) Schematic showing of a promoter window with TSS. (B) Increased processivity is hypothesized to increase TSS usage downstream. (C) Decreased processivity is hypothesized to reduce TSS usage downstream. (D) Increased efficiency is expected to activate TSS usage upstream. (E) Decreased efficiency is expected to increase apparent TSS usage downstream. (F) A processivity GOF allele and efficiency GOF double mutant is expected to show efficiency GOF single mutant allele’s effects on activating upstream TSSs (epistasis). (G) A processivity GOF and efficiency LOF double mutant is expected to allow or enhance efficiency LOF single allele’s effect on increasing apparent TSS usage downstream (potential additivity). (H) A processivity LOF and efficiency GOF double mutant is expected to show efficiency GOF single allele’s effect on activation TSS usage upstream (epistasis). (I) A
processivity LOF and efficiency LOF double mutant is expected to show processivity LOF single allele’s effect on reducing downstream TSS usage (epistasis).

Figure 6 – Figure supplement 2. The scoring method used to quantify yeast growth phenotypes and make phenotypic heatmaps.
Figure 6 – Figure supplement 3: Pol II efficiency alleles are able to increase TSS efficiency within the processivity defined scanning window. (A) Quantification of TSS usage at ADH1 in rpb1 single mutant, ssl2 N230D single mutant and the double mutants. TSS usage in single or double mutants is compared to WT TSS usage. Bars are average +/- standard deviation of ≥ 3 independent biological replicates. (B) Quantification of TSS usage at ADH1 in rpb1 single mutant, ssl2 N230I single mutant and the double mutants. TSS usage in double mutants is compared to WT TSS usage. Bars are average +/- standard deviation of ≥ 3 independent biological replicates.
Figure 7 – Figure supplement 1: TFIIB and TFIIF alleles show strong and distinct genetic interaction behavior with ssl2 alleles. (A) Lethality is broadly observed between ssl2 downstream shifting alleles and the sua7-1 downstream shifting allele. Patch assay showing that six downstream shifting alleles I170T, V226D, N230I, E340G, L225P and R636C, show synthetic lethality phenotypes when combined with sua7-1 allele (ssl2 I170T and sua7-1 double mutant is very sick on 5FOA plate but cannot be propagated to single colonies). Four ssl2 upstream shifting alleles N230D, V473D, D522V, Y750* show normal growth phenotypes when sua7-1 is incorporated. One ssl2 upstream shifting allele F498L showed unique (lethality) genetic interactions compared to other ssl2 upstream shifting alleles (normal growth). This allele additionally shows lethality with both TFIIF and sub1a alleles as discussed in the corresponding sections. (B) ssl2 upstream shifting alleles show epistasis with the sua7-1 downstream shifting allele for MPA^3 phenotypes. Spot assay shows that four upstream shifting ssl2 alleles' strong MPA^3 phenotypes (N230D, V473D, D522V and Y750*) are almost completely unaffected.
when sua7-1 allele, which shows a His\(^\ast\) phenotype by itself, is incorporated. Primer extensions shown are representative of ≥ 3 independent biological replicates. (C) Double mutants of ssl2 upstream shifting alleles and sua7-1 downstream shifting allele shift TSS distribution downstream. Primer extension results show that ssl2 alleles' effect on shifting TSS distribution upstream is almost completely reversed to downstream shifting after incorporation of sua7-1 allele, with only slight additive effect seen in double mutants' TSS usage. Primer extensions shown are representative of ≥ 3 independent biological replicates. (D) Quantification of primer extension detected TSS usage at ADH1 in ssl2, sua7-1 single or double mutants. Bars are average +/- standard deviation of ≥ 3 independent biological replicates. (E) Lethality is observed between ssl2 upstream shifting alleles and the tfg2Δ146-180 upstream shifting allele. Patch assay illustrates that upstream shifting alleles N230D, V473D, D522V and F498L show synthetic lethality phenotypes when combined with tfg2Δ146-180 allele (ssl2 Y750\(^\ast\) and tfg2Δ146-180 double mutant shows moderate level of growth on 5FOA plate, however, shows strong sickness when propagated to single colonies). Three ssl2 downstream shifting alleles L225P, N230I and R636C show normal growth phenotypes when tfg2Δ146-180 is incorporated. (F) ssl2 downstream shifting alleles show both epistasis and additive effects with tfg2Δ146-180 downstream shifting allele. Spot assay shows that the double mutants of ssl2 downstream alleles and tfg2Δ146-180 allele show MPA\(^\ast\) phenotypes; however, double mutants' MPA\(^\ast\) phenotypes are not as strong as observed in single tfg2Δ146-180 allele. (G) Double mutants of ssl2 downstream shifting allele and tfg2Δ146-180 upstream shifting allele shift TSS distribution upstream. Primer extension results show that ssl2 alleles' effect on shifting TSS distribution upstream is almost completely reversed to upstream shifting after incorporation of tfg2Δ146-180 allele. Primer extensions shown are representative of ≥ 3 independent biological replicates. (H) Quantification of primer extension detected TSS usage at ADH1 in ssl2, tfg2Δ146-180 single or double mutants. Bars are average +/- standard deviation of ≥ 3 independent biological replicates.
Figure 7 – Figure supplement 2: Multiple genetic interactions between ssl2 and sub1Δ alleles.

(A) Three types of genetic interactions are observed between ssl2 alleles and sub1Δ for the IMD2 promoter. First, three ssl2 upstream shifting alleles (N230D, V473D or D522V) are epistatic to sub1Δ allele for genetic phenotypes. Spot assay shows that the double mutants of sub1Δ (His*) and ssl2 upstream shifting alleles (MPA5) show MPA5 phenotypes like ssl2 upstream single alleles. Second, sub1Δ is epistatic to ssl2 allele Y750*. Double mutant of sub1Δ (His*) and Y750* (MPA5) shows a His* phenotype. Third, sub1Δ is epistatic to ssl2 downstream shifting alleles (L225P, N230I, R636C). Double mutants of sub1Δ (His*) and ssl2 downstream shifting alleles (His*) show His* phenotypes that are not stronger than in sub1Δ alone. (B) Genetic interactions between ssl2 and sub1Δ alleles at ADH1 promoter. First, additive effects are observed between ssl2 upstream shifting alleles and sub1Δ. Double mutants of ssl2 upstream shifting alleles and sub1Δ show a compromised TSS distribution. Second, additive effects are also observed between ssl2 downstream shifting alleles and sub1Δ. Double mutants show stronger downstream shift than either single alone. Third, epistasis is observed between ssl2 Y750* and sub1Δ. In the double mutant, ssl2 Y750*'s effect on shifting TSS distribution upstream is completely reversed to downstream shifting by sub1Δ. Primer extensions shown are representative of ≥ 3 independent biological replicates. (C) Quantification of primer extension detected TSS usage at ADH1 in ssl2, sub1Δ single or double mutants. Bars are average +/- standard deviation of ≥ 3 independent biological replicates.
Figure 9 – Figure supplement 1: ssl2 alleles shift PIC-component positioning genome-wide as predicted for mutants altering scanning processivity. (A) Positions of PIC components Sua7 and Ssl2 determined by ChIP-exo in WT and ssl2 mutant strains. Aggregated genome-wide ChIP-exo signals on top (TOP) or bottom (BOT) strand at ± 150 nt TSS centered promoter windows. Enrichment of ChIP-exo signals from the top or bottom strand of Sua7 or Ssl2 sites are shown for two biological
replicates per genotype. Curves are LOWESS smoothing of aggregated ChIP-exo reads from the top 50% promoters in WT. (B) ssl2 alleles shift positioning of PIC components Sua7 and Ssl2 at most promoters. Showing is the shift of median Sua7 or Ssl2 position determined by ChIP-exo as illustrated in (A). Statistical analysis shows that the shift of the average median positions of Sua7 and Ssl2 in two replicates are significantly different from zero (Wilcoxon Signed Rank test) at p<0.0001 (****).

**Figure 9 – Figure supplement 2. Effects of TAP-tagging SSL2.** A TAP-tag was integrated at the 3’ end of SSL2 and CRISPR-Cas9 was used to introduced mutations at codon 230 along with an adjacent silent PAM site mutation to prevent Cas9 recutting. It was discovered that the TAP-tag modulated SSL2 phenotypes, behaving like a slight loss of function, causing mild MPA sensitivity, enhancing MPA sensitivity of ssl2 N230D (top row, SC+MPA plate), suppressing the Spt phenotype of N230I (SC-Lys plates), but having no effect on the His* phenotype of N230I in the background of the imd2∆::HIS3 initiation reporter (bottom row, SC-His, SC-His 1 mM 3-AT). 3-aminotriazole (3-AT) is a competitive inhibitor of the HIS3 gene product. All serial dilution phenotyping results are representative of at least two independently constructed strains (biological replicates).