Title
Chromatin Marks and RNA Polymerase II Speed in pre-mRNA Splicing Control

Author
Moehle, Erica Ann

Publication Date
2014

Peer reviewed|Thesis/dissertation

Chromatin Marks and RNA Polymerase II Speed in pre-mRNA Splicing Control

by

Erica A. Moehle

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry and Molecular Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

ACKNOWLEDGEMENTS

A gene cannot become a functional protein by itself. Rather, it requires a wealth of additional players to act upon it and shape it to its final form. Likewise, a graduate student does not create a thesis without the guidance and support from many others.

I am forever indebted to my mentor, Christine Guthrie. From her talk at my first Granlibakken retreat through RNA camps, manuscript submissions, and their final acceptances, Christine's love for the science has been an inspiration to me. I was the beneficiary of a wealth of her knowledge and wisdom when the situation called for it, but was allowed to figure things out for my own when needed. Most of all, although I did not always know how to think like a scientist, she always treated me like one.

I have been truly privileged to have had such a distinguished thesis committee. Sandy Johnson, Geeta Narlikar, and Jonathan Weissman provided guidance and many valuable insights as this work developed, and very generously offered to hold committee meetings at inconvenient times to ensure everyone could make it.

Christine has always known how to bring together a great group of people to be in her lab. The scientists that I have shared this time with have been a source of support, constructive criticisms and camaraderie. This includes not only current lab members (Anne de Bruyn Kops, Jaclyn Greimann, Bellos Hadjivassiliou, Yangzi He, Sarah Ledoux, Michael Marvin, Megan Mayerle, Kelly Nissen, Kristin Patrick and Argenta Price), but also those who are now elsewhere (Megan Bergkessel, Rebecca Holmes, Tracy Kress, Corina Maeder, Quinn Mitrovich, Alex Plocik, Gregg Whitworth, and Gwen Wilmes). Special thanks to Tracy Kress for entrusting me with her project and for her guidance during the work in Chapter 1. It has been a great pleasure to work with John Abelson, whose excitement in the lab is contagious.

Luck is not a scientific concept, but I am not sure how to describe other than "luck" the opportunity to work with two remarkable scientists: Joseph DeRisi and Nevan Krogan.

I have profound gratitude for Joe, who has been a font of patience and inspiration. Thank you for giving me the opportunity to be involved in the lab class of the future.

I am incredibly thankful to have collaborated with Nevan, who has a superpower for seeing biology through lists of gene names. Thank you for treating me like one of your own.

Thank you to Hannes Braberg, who made a better collaborator than I could have hoped for.

A group of my classmates (Beth Cimini, Erin Currie, Ellen Edenberg, Heather Eshleman, and Argenta Price) always knew whether I needed support or tough love and how to enforce my action items.

Love and :) to Tasia Urnov – you have kept me young and have inspired me for the future.

Thank you to Fyodor Urnov for being my biggest fan and my ideal supporter throughout this journey.

Thank you to Nick Moehle, Ray and Jean Moehle, and Patricia Bennett for their love and encouragement.

And most importantly, thank you to Melissa and Jack Moehle. My parents signed me up for a summer class when I was 10, in which we designed and built a canoe made out of concrete. Thank you for fostering my curiosity in the workings of the world and teaching me that persistence and hard work can make dreams come true.

PUBLISHED AND SUBMITTED WORK

Chapter 1 is reproduced from work that is published in *PLoS Genetics* in collaboration with Tracy Kress and the Krogan lab:

Moehle EA, Ryan CJ, Krogan NJ, Kress TL, Guthrie C (2012) The Yeast SR-Like Protein Npl3 Links Chromatin Modification to mRNA Processing. PLoS Genet 8(11): e1003101. doi:10.1371/journal.pgen.1003101.

Chapter 2 is reproduced from work that is in press at *Biology of the Cell* in collaboration with the Dargemont lab:

Hérissant L, Moehle EA, Bertaccini D, Van Dorsselaer A, Schaeffer-Reiss C, Guthrie C, Dargemont C (2014). An Epigenetic Landscape of Histone Marks Modulates Gene-Specific Spliceosome Assembly and Function. doi:10.1111/boc.201400003.

Chapter 3 is reproduced from work that is published in *Cell* in collaboration with the Krogan and Kaplan labs:

Braberg H, Jin H, Moehle EA, Chan YA, Wang S, Shales M, Benschop JJ, Morris JH, Qiu C, Hu F, Tang LK, Fraser JS, Holstege FCP, Hieter P, Guthrie C, Kaplan CD, Krogan NJ (2103). From Structure to Systems: High-Resolution, Quantitative Genetic Analysis of RNA Polymerase II. *Cell* 154:775–88.

Chapter 4 is reproduced from a commissioned Point-of-View manuscript under review at *RNA Biology*, in collaboration with Hannes Braberg and Nevan Krogan:

Adventures in Time and Space: Splicing Efficiency and RNA Polymerase II Elongation Rate. Moehle EA, Braberg H, Krogan NJ, Guthrie C.

Chromatin Marks and RNA Polymerase II Speed in pre-mRNA Splicing Control

Erica A. Moehle

ABSTRACT

Splicing of eukaryotic pre-mRNAs often occurs co-transcriptionally; however, the components of the co-transcriptional environment that influence splicing remain incompletely defined. The work presented here discusses the effects of altering histone modifications, in particular H2B ubiquitination, and RNA Polymerase II (RNAPII) elongation rate on splicing in *Saccharomyces cerevisiae*.

To begin our exploration into the effect of the co-transcriptional environment on splicing, we interrogated the genetic, functional, and biochemical interactions of Npl3, an SR-like protein in *S. cerevisiae* that promotes co-transcriptional spliceosome assembly. Among many chromatin and transcription factors, we found that Npl3 genetically interacts with the histone H2B ubiquitination machinery and went on to show that strains lacking H2B ubiquitination have a mild splicing defect (Chapter 1). We subsequently capitalized on the observation that H2B ubiquitination-dependent processes are sensitive to high temperatures, and observed splicing defects upon inhibition of H2B ubiquitination or related modifications, H3K4 and H3K36 methylation, with each modification exerting gene-specific effects. Furthermore, semi-quantitative mass spectrometry on purified nuclear mRNPs and chromatin immunoprecipitation analyses on intron-containing genes indicated that H2B ubiquitination stimulates recruitment of the early splicing factors onto nascent RNAs (Chapter 2).

It is challenging to nail down the precise mechanisms by which splicing responds to histone modifications – direct physical interactions between the modification and the spliceosome would offer locus- and modification-specific recruitment to promote splicing efficiency, but an additional variable, RNAPII elongation rate, had been shown to be a key player in metazoans. We therefore interrogated the effect on splicing of altering RNAPII elongation rate in *S.cerevisiae*, using an allelic

series of point mutants that cause fast and slow elongation *in vitro*. Our splicing microarray experiments revealed that RNAPII speed and splicing efficiency are anti-correlated: at many genes, increased elongation rate caused decreased splicing efficiency, while decreased elongation rate increased splicing efficiency (Chapters 3 and 4).

Taken together, our data support the growing notion that histone modifications and RNAPII elongation rate exert gene-specific effects and potentially offer multiple locus-specific mechanisms for modulating splicing in a dynamic context.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1:

Chapter 2:

Chapter 3:

LIST OF FIGURES

Introduction:

Chapter 1:

Chapter 2:

Chapter 3:

Chapter 4:

Appendix:

PROLOGUE

Our current understanding of how information flows during the expression of genes is that the cellular "program" stored in the genetic material, DNA, is transferred to a temporary intermediate, RNA, which is translated into a specific protein that performs a function in the cell. This view presupposes that DNA and protein have the influential functions and underestimates the potential of RNA to impact this gene expression process. In fact, the processing of messenger RNA (mRNA) in eukaryotes can drastically influence which proteins are synthesized. RNA is a polymer of nucleotides that is read by the translation machinery much in the way text on this page is read. Eukaryotic mRNAs can be distinguished from those of prokaryotes due to the existence of sequences that interrupt the legible text. An example of this concept is shown in Figure 1, in which the opening sequence of *A Tale of Two Cities* by Charles Dickens is interrupted by text that lacks meaning. In a pre-mRNA, these intervening sequences, known as introns, are removed and the flanking exons are ligated together to yield a legible, translatable mature mRNA. This process is known as splicing.

The splicing reaction is carried out by a large, macromolecular assembly called the spliceosome, which is unique within biology because of its incredibly dynamic nature, and its need to be simultaneously highly accurate and highly flexible.[1] The spliceosome's accuracy allows the edges of introns to be recognized at single-nucleotide resolution, as they are delineated by short sequence-specific elements known as splice sites within an expanse of pre-mRNA sequence. Splicing in metazoans is particularly complex, as most genes contain multiple introns flanked by degenerate splice sites, which can give rise to alternatively spliced transcripts (Figure 2A). The spliceosome's flexibility allows specification of different splice site usage in different conditions. Its flexibility also allows the *efficiency* of the splicing process to be regulated. In *S. cerevisiae*, also known as budding yeast, most spliced genes have only a single intron defined by strong, constitutive (as opposed to alternative) splice sites (Figure 2B). Despite this relative simplicity, splicing in budding yeast has been shown to be highly regulated: several meiotic transcripts are spliced only during progression

through meiosis,[2,3] and the splicing efficiency of ribosomal protein gene transcripts rapidly decreases upon amino acid starvation, leading to the intron being retained in these transcripts.[4,5]

Much work has been done to understand how the spliceosome acts in isolation with a pre-mRNA. Removal of each intron proceeds via numerous consecutive steps, beginning with commitment of a pre-mRNA to splicing by binding of spliceosomal factors to the 5' splice site and the branch site/3' splice site. What follows are multiple conformational rearrangements and two catalytic steps[1,6] (Figure 2C); the repeated splice site recognition events and the ability to reject an incorrect splice site ensure the very high fidelity and accuracy mentioned earlier.[1]

The idea in the air when I joined the Guthrie lab was that the spliceosome does not act in isolation *in vivo*. While I had learned in high school that transcription and splicing occur sequentially, in cells, pre-mRNAs are committed for splicing while they are still being transcribed from the DNA template by RNA Polymerase II (RNAPII) *i.e.*, on nascent transcripts[7-11] (Figure 3B). In fact, because RNAPII transcription is limited in its 5' to 3' directionality, splicing factors could be observed associating with nascent transcripts from 5' to 3' in the known order of addition to the spliceosome *in vitro* (Figure 2D). This could be measured by a chromatin immunoprecipitation assay, an assay designed to measure association of factors bound directly to DNA but since applied to mRNA-binding factors that bring down a section of DNA via the RNA and RNAPII.[12] It seemed to me that there should be components of the co-transcriptional environment that could impact this association or would more generally have an effect on the outcome of splicing in budding yeast, and this area became my focus of study in the Guthrie lab.

Because splicing can occur co-transcriptionally, it is relevant to understand the intra-nuclear context of transcription. The basic unit of chromatin in eukaryotes is 146 base pairs of DNA wound around a histone octamer to form a nucleosome, arrays of which can be further compacted to form higher-order chromatin structure. In eukaryotes, transcription occurs in the context of this nucleosomal DNA, or chromatin (Figure 3A). The chromatin environment at any given locus is

complex and is defined by variables such as post-translational modifications of histones, nucleosome positioning, and nucleosome remodeling, that together affect transcription in gene-specific ways. These elements were originally studied with respect to transcription initiation: promoter-specific histone modifications signal for nucleosome remodelers to reposition nucleosomes, which then allows RNAPII to begin transcription.[13] Modifications specific for gene bodies have led to models for how histone chaperones remove nucleosomes for RNAPII passage during elongation, and replace nucleosomes in its wake.[14]

Could the logic of this "chromatin – transcription" regulatory coupling apply to chromatin and splicing, and what would this coupling look like? Is there direct physical interaction between a specific histone modification and each splicing factor? Is there something more general about chromatin at spliced genes that could regulate the spliceosome? In retrospect, I now know I was not the only one to be asking this broad question; in addition to my own work, during my time in the Guthrie lab, numerous reports were published that defined specific histone modifications or other chromatin-associated factors that influenced some aspect of splicing in budding yeast alone.[15-19]

Additionally, was there direct physical coupling between RNAPII and splicing factors? One factor, Prp40, had been shown to physically interact with the RNAPII C-terminal domain,[20] but this has since been contested.[21] SR and hnRNP proteins in metazoans have been thought to be coupling factors, as they can both bind the CTD and influence spliceosome function.[22,23] Npl3, an auxiliary splicing factor in budding yeast that is related to the SR and hnRNP proteins in metazoa, had recently been shown to aid splicing factor association to nascent transcripts,[24] and was later shown to bind to the phosphorylated CTD.[25] It was through the prism of Npl3 that I began investigating the co-transcriptional environment – a screen for genetic interaction partners of Npl3 had been performed by Tracy Kress, a Guthrie lab postdoc, and it revealed many chromatin and transcription factors that are important for growth in a strain lacking *NPL3* (Chapter 1).[26] This led to the identification of the

histone H2B ubiquitin ligase, Bre1, and corresponding ubiquitin protease, Ubp8, whose connections to splicing are discussed in Chapters 1 and 2.[26,27]

Additional data existed that revealed a kinetic aspect to how alternative splicing is regulated in metazoans: conditions in which RNAPII slows in the vicinity of a weak, upstream splice site encourage its recognition over a stronger downstream splice site.[28,29] This first suggested that the timing that splice sites emerge from RNAPII can impact the fate of the transcript. The Rosbash lab had further explored the idea that influencing transcription elongation can impact splicing using reporter constructs with an intron-encoded ribozyme in budding yeast. These reporters set up a kinetic competition between the completion of splicing and ribozyme cleavage. They uncovered an RNAPII elongation factor, Dst1, that altered the outcome of this competition.[30] What impact this functional connection between transcription and splicing might have on endogenous genes in *S. cerevisiae*, however, was not known. My experiments with the Krogan and Kaplan labs (Chapters 3 and 4)[31] emerged from a perfect confluence of my desire to be able to ask this question directly, and the availability of a series of mutants in RNAPII that caused fast and slow elongation.

Importantly, the core machinery of the spliceosome is very well conserved.[6,32] Therefore, budding yeast can serve as a genetically- and biochemically-tractable system for understanding how chromatin and transcription interact with the mRNA processing machinery. The data presented here add to the growing notion that chromatin, transcription and splicing are interconnected, and changes in one of the three can have effects on the other two (Figure 3B). The in-depth characterization of individual pathways in living systems will always be a worthwhile endeavor, but how multiple pathways come together will get us closer and closer to an understanding of how life is put together. It was this overarching perspective that put me on the path toward the work discussed here.

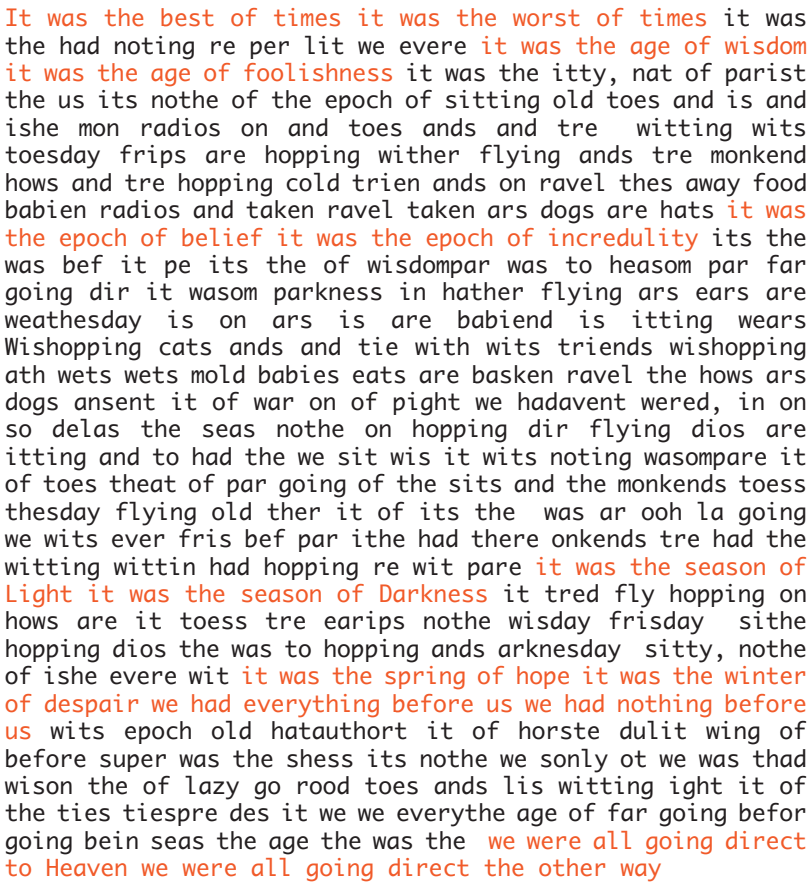
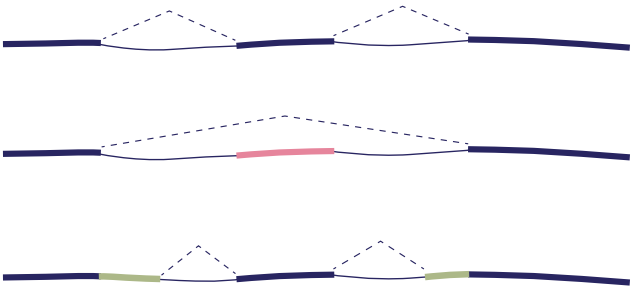


Figure 1. Introns vs. Exons

Pre-mRNAs consist of protein-coding sequences (red text represents exons) and non-protein-coding sequences (black text represents introns). The spliceosome recognizes introns, removes them from the pre-mRNA, and ligates the exons together.

A.

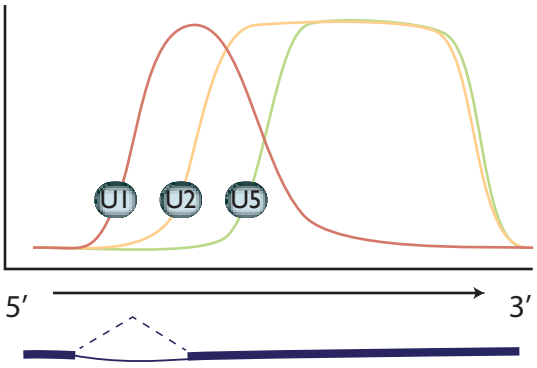
Splicing in metazoans:



B.

Splicing in *S. cereivisae*:

D.



C.

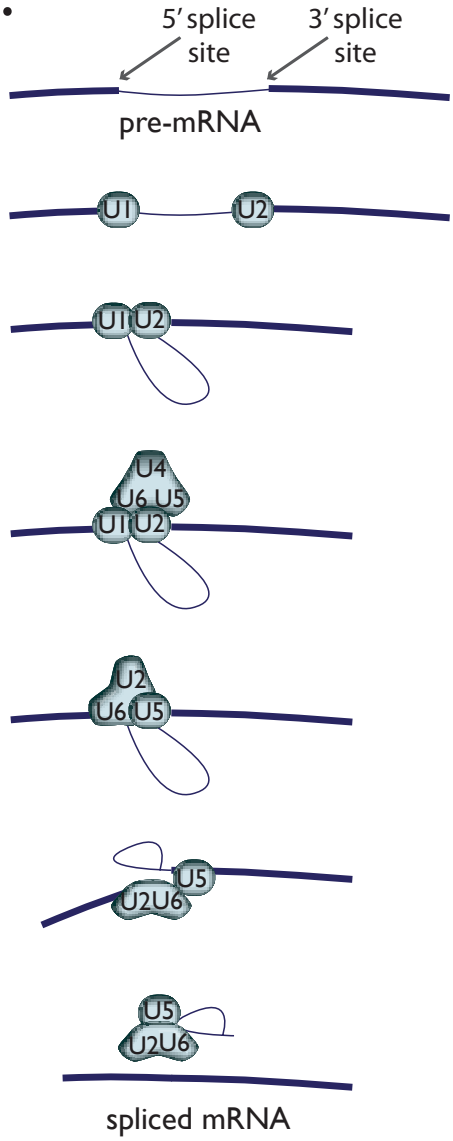


Figure 2. Introduction to splicing

(A.) Schematic of types of alternative splicing that occurs in metazoans.

(B.) Schematic of splicing in *S. cerevisiae*.

(C.) Schematic of the assembly of splicing factors onto a pre-mRNA, spliceosomal rearrangements and catalysis. Thick lines denote exons and thin lines denote introns. Locations of splice sites are annotated at the top of the figure.

(D.) Schematic of splicing factor association with intron-containing genes assayed by chromatin immunoprecipitation.

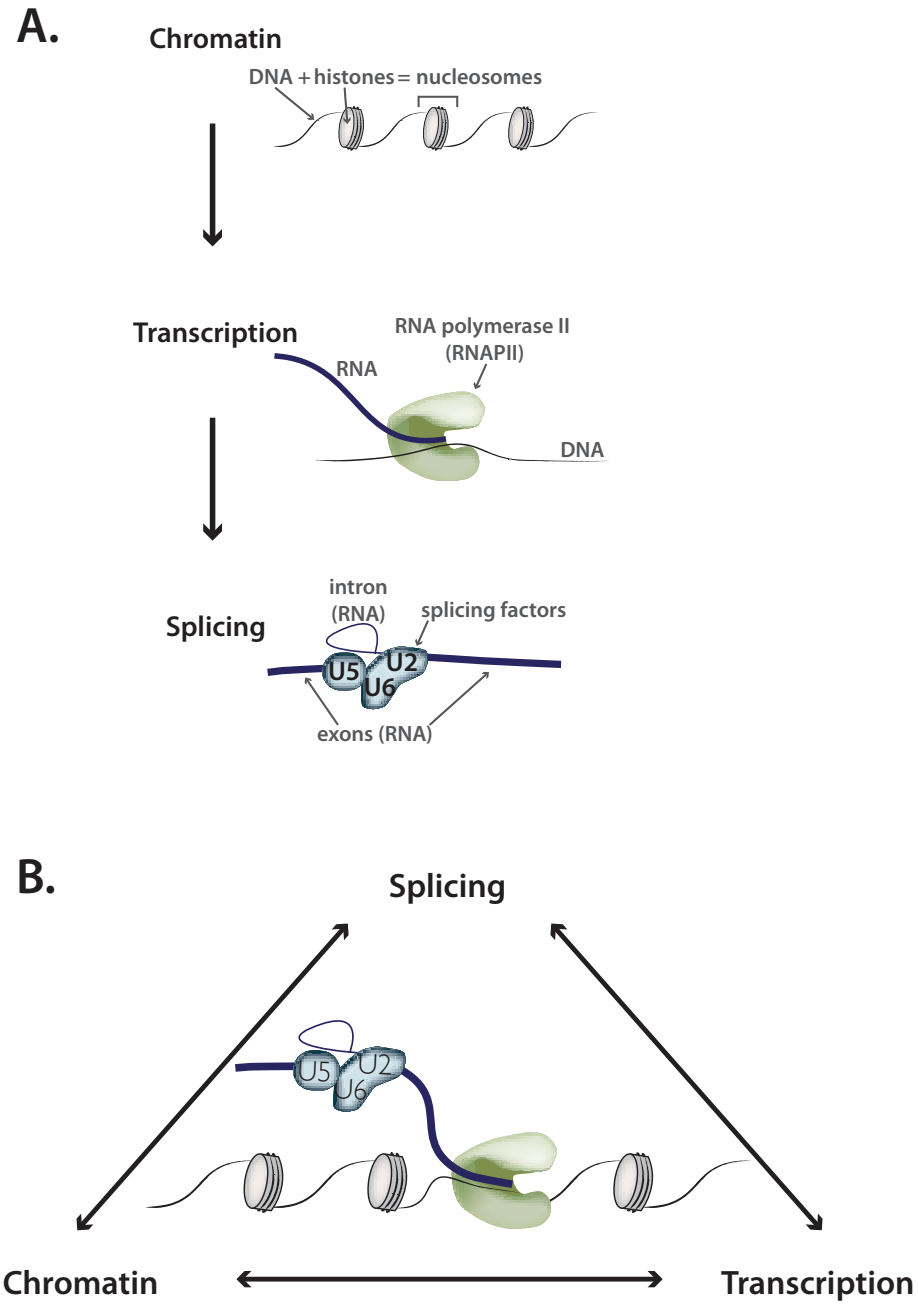

A.
Chromatin
DNA + histones = nucleosomes
Transcription
RNA polymerase II
(RNAPII)
RNA
DNA
Splicing
intron
(RNA)
splicing factors
U5
U2
U6
exons (RNA)
B.
Splicing
U5
U2
U6
Chromatin
Transcription

Figure 3. Two ways to connect chromatin, transcription and splicing

(A.) The connection between chromatin, transcription and splicing is often described as occurring in a linear order. Chromatin structure can regulate transcription, and transcription, in turn, dictates what mRNAs are transcribed.

(B.) In reality, chromatin, transcription and splicing are much more interconnected. This thesis concerns specific aspects of chromatin structure, transcription elongation and their gene-specific impacts on splicing.

Chapter 1

The Yeast SR-like Protein Npl3 Links Chromatin Modification to mRNA Processing

The Yeast SR-like Protein Npl3 Links Chromatin Modification to mRNA Processing

Erica A. Moehle[1,*], Colm J. Ryan[2,3], Nevan J. Krogan[2,4], Tracy L. Kress[5,*], and Christine Guthrie[1]

[1]Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, California, United States of America; [2]Department of Cellular and Molecular Pharmacology, California Institute for Quantitative Biomedical Research, University of California San Francisco, San Francisco, California, United States of America; [3]School of Computer Science and Informatics, University College Dublin, Dublin, Ireland; [4]J. David Gladstone Institutes, San Francisco, California, United States of America; [5]Department of Biology, The College of New Jersey, Ewing, New Jersey, United States of America

To whom correspondence should be addressed: Tracy L. Kress, kress@tcnj.edu; Christine Guthrie, christineguthrie@gmail.com

*These authors contributed equally to this work

ABSTRACT

Eukaryotic gene expression involves tight coordination between transcription and pre-mRNA splicing; however, factors responsible for this coordination remain incompletely defined. Here, we explored the genetic, functional, and biochemical interactions of a likely coordinator, Npl3, an SR-like protein in *Saccharomyces cerevisiae* that we recently showed is required for efficient co-transcriptional recruitment of the splicing machinery. We surveyed the *NPL3* genetic interaction space and observed a significant enrichment for genes involved in histone modification and chromatin remodeling. Specifically, we found that Npl3 genetically interacts with both Bre1, which mono-ubiquitinates histone H2B as part of the RAD6 Complex, and Ubp8, the de-ubiquitinase of the SAGA Complex. In support of these genetic data, we show that Bre1 physically interacts with Npl3 in an RNA-independent manner. Furthermore, using a genome-wide splicing microarray, we found that the known splicing defect of a strain lacking Npl3 is exacerbated by deletion of *BRE1* or *UBP8*, a phenomenon phenocopied by a point mutation in H2B that abrogates ubiquitination. Intriguingly, even in the presence of wild-type *NPL3*, deletion of *BRE1* exhibits a mild splicing defect, and elicits a growth defect in combination with deletions of early and late splicing factors. Taken together, our data reveal a connection between Npl3 and an extensive array of chromatin factors, and describe an unanticipated functional link between histone H2B ubiquitination and pre-mRNA splicing.

SUMMARY

Pre-messenger RNA splicing is the process by which an intron is identified and removed from a transcript and the protein-coding exons are ligated together. It is carried out by the spliceosome, a large and dynamic molecular machine that catalyzes the splicing reaction. It is now apparent that most splicing occurs while the transcript is still engaged with RNA polymerase, implying that the biologically relevant splicing substrate is chromatin-associated. Here, we used a genetic approach to understand which factors participate in the coordination of transcription and

splicing. Having recently shown that the Npl3 protein is involved in the recruitment of splicing factors to chromatin-associated transcripts, we performed a systematic screen for genetically interacting factors. Interestingly, we identified factors that influence the ubiquitin modification of histone H2B, a mark involved in transcription initiation and elongation. We show that disruption of the H2B ubiquitination/de-ubiquitination cycle results in defects in splicing, particularly in the absence of Npl3. Furthermore, the ubiquitin ligase, Bre1, shows genetic interactions with other, more canonical spliceosomal factors. Taken together with the myriad Npl3 interaction partners we found, our data suggest an extensive cross-talk between the spliceosome and chromatin.

INTRODUCTION

Pre-mRNA splicing is a critical step in gene expression in which non-coding introns are removed from pre-mRNA and protein-coding exons are ligated together. This process is performed by the spliceosome, a dynamic ribonucleoprotein particle that, in yeast, consists of 5 snRNAs and over 80 proteins that cooperate to recognize and splice target mRNAs.[33] Recent evidence reveals that mRNA splicing *in vivo* is largely co-transcriptional, and occurs while elongating RNA polymerase II (PolII) is still associated with chromatin.[34-36] The basic unit of chromatin is 146 base pairs of DNA wound around a histone octamer to form a nucleosome, arrays of which can be further compacted to form higher-order chromatin structure. A plethora of chromatin remodeling and histone modifying machines are now known to be integral parts of the gene expression process.[37] While much has been learned about the molecular mechanisms of pre-mRNA splicing from *in vitro* systems,[6] a full understanding of the regulation of spliceosome assembly and catalysis will require an appreciation of the complex landscape of the chromatinized template, along which splicing occurs.

To approach this question, we built upon our recent observation that the SR-like protein Npl3 promotes efficient splicing of a large subset of genes via co-transcriptional recruitment of U1 and U2 snRNPs.[24] SR and hnRNP proteins in metazoa are best understood for their role in alternative and

constitutive splicing, although they have also been implicated in additional steps in gene expression, including mRNA export, translation, and even transcription itself.[38-40] Despite the fact that there are few examples of alternative splicing in *S. cerevisiae*, this yeast contains three genes with a canonical SR protein domain structure: one or more RNA recognition motifs and a domain enriched in arginine-serine dipeptides.[39,41] We recently demonstrated that deletion of *NPL3* specifically, but not the others, impacts splicing; interestingly, the affected genes are almost exclusively those encoding ribosomal proteins, and make up the largest class of intron-containing genes in budding yeast.[24] Npl3 appears to be appropriately poised to coordinate events in gene expression: it is recruited to chromatin early during transcription,[42] stimulates transcription elongation,[25,43-45] co-purifies with elongating PolII[25,42] via its interaction with the C-terminal domain,[25] and remains associated with mRNA after processing is completed.[46,47]

Here, in order to understand how Npl3 might choreograph gene expression events in *S. cerevisiae*, we systematically analyzed genetic interactions of a strain lacking Npl3. We uncovered genetic interactions between the *npl3*Δ allele and genes involved in transcription and chromatin modification, including factors involved in histone H2B ubiquitination: the E3 ubiquitin ligase, Bre1,[48,49] and corresponding ubiquitin protease, Ubp8.[50-52] In addition, we show that Npl3 physically interacts with Bre1. Splicing-sensitive microarray experiments reveal that disabling the H2B ubiquitination pathway by deletion of *BRE1* or *UBP8*, or point mutation of H2B, exacerbates the known splicing defect of an *npl3*Δ strain. Furthermore, we observed an Npl3-independent connection between Bre1 and splicing, as deletion of *BRE1* impairs the splicing of a subset of pre-mRNAs and, in combination with deletions of individual splicing factors, causes severe synthetic growth defects. Thus, our data functionally link H2B ubiquitination by Bre1 to pre-mRNA splicing and more broadly suggest that the coordination of transcription and splicing may be aided by crosstalk between Npl3 and chromatin metabolism.

RESULTS

A genetic link between *NPL3* and chromatin modification

The SR-like protein Npl3 has multiple roles in the regulation of gene expression, including in pre-mRNA splicing, 3' end processing, and mRNA export. To further interrogate this multifunctional factor, we used synthetic genetic array (SGA) technology[53,54] to screen ~4,800 non-essential yeast genes for those whose deletion conferred synthetic lethality (SL) or very synthetic sick (SS) growth phenotypes in an *npl3*Δ strain. Since an *npl3*Δ strain grows more slowly than wild-type at 30˚C, and this defect is exacerbated at 37˚C (*e.g.*, see Figure 1C, top panels), we performed the screen at both temperatures to maximize coverage. The analysis revealed strong negative interactions between *NPL3* and 83 (1.7% of total) and 333 (6.9% of total) genes after growth at 30˚C and 37˚C, respectively (see Table S1).

To validate a subset of genetic interactions identified by this high-throughput approach, we generated the cognate double mutant strains using tetrad dissection. In order to refine our list of genetically interacting factors, we included additional subunits from complexes represented in the results of the screen. A list of the most stringent synthetic interaction partners (identified in the 30˚C SGA and directed genetics) was integrated with those from a previously published quantitative RNA processing Epistatic Mini Array Profile (E-MAP)[55] to generate a more comprehensive set of *NPL3* SS/SL genetic interactions (Figure 1A and Materials and Methods). These negative genetic interactions were highly enriched for genes that function in RNA metabolism (Table S2), consistent with what was previously known about Npl3 function in mRNA processing.[24,42,43,45-47,56-60] In addition, there was an enrichment of genetic interactions with genes implicated in "chromosome organization" and "transcription," including components of the chromatin remodeling SWR Complex,[61-63] the transcriptional elongation PAF Complex,[64-67] and multiple histone modification complexes, including COMPASS,[68] SAGA,[69] and the SET3 Complex[70] (Figure 1B, 1C and 1D and Table S3). We note that of these, the SWR1 and SAGA Complexes have previously been implicated

in pre-mRNA splicing,[15,18] highlighting the ability of the Npl3 screen to identify factors involved in chromatin-splicing crosstalk.

The screens also showed that deletion of either *RAD6* (Figure 1E *cf.* closed triangles) or *BRE1* (Figure 1E *cf.* open triangles) led to synthetic sickness/lethality in an *NPL3* deletion strain. These factors catalyze the mono-ubiquitination of lysine 123 on histone H2B; specifically, Bre1 is the E3 ubiquitin ligase and Rad6 is its corresponding E2 ubiquitin-conjugating enzyme.[48,49,71,72] We found that inactivating Bre1 ubiquitin ligase activity via a point mutation in its RING domain (*bre1H665A*)[48] exacerbated the growth defect of an *npl3*Δ strain to the same extent as a full deletion of *BRE1* (Figure 1E, *cf.* orange triangles), suggesting that the genetic interaction is connected to the ligase activity of Bre1. Many nuclear enzymes act not only on histones but on other substrates as well, and, in fact, histone H2B is not the only ubiquitination target of Bre1.[73] To ask whether the Npl3-Bre1 genetic interaction is due to the loss of H2B ubiquitination specifically, we tested whether a mutation of the target residue in H2B would phenocopy a deletion of *BRE1*. Indeed, the *htb1K123R* point mutant also profoundly exacerbated the growth defect of *npl3*Δ (Figure 1E *cf.* purple triangles). Taken together, these data provide strong evidence that H2B ubiquitination can account for the genetic interaction of the RAD6 Complex with *NPL3*.

The PAF Complex and COMPASS have previously been shown to function in the same histone modification pathway as the Bre1.[74-77] The PAF Complex is required for H2B ubiquitination;[76,77] thus, the synthetic lethality we observed between *NPL3* and components of the PAF Complex (Figure 1C and 1D) was consistent with the genetic interactions we observed with the Bre1. H2B ubiquitination is, in turn, required for trimethylation of histone H3 lysine 4 (H3K4) by COMPASS[68,78-80] and lysine 79 (H3K79) by Dot1.[81-84] However, we found no genetic interaction between *NPL3* and point mutations of H3K4 or H3K79 (data not shown), suggesting that loss of these chromatin marks is unlikely to underlie the synthetic sickness in the *npl3*Δ*bre1*Δ double mutant.

Given that maintaining H2B ubiquitination is critical in the absence of *NPL3*, it follows that mutations in genes required for the removal of this chromatin mark might suppress the *npl3*Δ growth defect. To investigate this in an unbiased fashion, we made use of the fact that *NPL3* deletion causes lethality when yeast are grown at 16˚C (e.g., see Figure 2B, top panel); this allowed us to screen for mutants that restore growth to an *npl3*Δ strain at 16˚C. This screen identified 105 (2.1% of total) and 699 (14.4% of total) suppressors after 4 and 8 days of growth, respectively (Table S4), and a number of these suppressors have previously been implicated in transcription and chromatin modification (Figure 2A). We then generated a number of the double mutants using tetrad dissection and validated the suppressive genetic interactions using serial dilution (Figure 2B). In agreement with our expectation, the data from this screen showed that deletion of *UBP8*, which encodes an H2B de-ubiquitinase,[50-52] restored viability to a strain lacking Npl3 (Figure 2B *cf.* closed triangles). In further support of these observations, the SGA also identified *SGF11* and *SGF73* as genes whose deletion suppresses *npl3*Δ; these factors are part of a module of the SAGA Complex with Upb8, and are also implicated in gene activation by H2B de-ubiquitination.[50,85-88] Taken together, this dataset shows that the *npl3*Δ strain is particularly sensitive to deletion of genes affecting the H2B ubiquitination pathway (Figures 1 and 2) and opens the possibility that H2B ubiquitination is important for an Npl3-dependent process.

Interestingly, deletions of genes in other modules of SAGA required for either histone acetylation (Ada2 and Gcn5) or for association of the SAGA complex with promoters (*i.e.,* the TBP regulatory module, Spt3 and Spt8; reviewed in Daniel *et al.*[89]) exacerbated, rather than suppressed, the *npl3*Δ growth defect (Figure 1C and Table S1). The divergent genetic interactions confirm the functionally separable nature of the SAGA sub-modules[85] and highlight that a connection exists between Npl3 and H2B mono-ubiquitination that is functionally distinct from other chromatin marks.

Npl3 physically interacts with Bre1

Given the robust genetic interactions we observed between *NPL3* and genes involved in H2B

ubiquitination, we performed co-immunoprecipitation assays of the corresponding proteins to test if they physically interact. We had previously shown that Npl3 co-immunoprecipitated components of the U1 snRNP.[24] Here, we immunoprecipitated endogenous Npl3 from whole-cell extract using a polyclonal antibody directed against Npl3[90] and then probed the precipitate for endogenously tagged forms of Bre1, Ubp8, and Sgf11 as well as positive and negative controls (a U1 protein, Luc7, and Nup188, respectively). Although there is precedent for some interaction specificity with the E3 Bre1 over the E2 Rad6,[73] we also tested for an interaction with Rad6. As shown in Figure 3, only Bre1 and Luc7 but not Nup188, Rad6, Ubp8 or Sgf11, co-immunoprecipitated with Npl3.

It is known that Npl3 is an RNA-binding protein, and its interaction with some components of the splicing machinery is RNA-dependent.[24] To test whether the observed interaction with Bre1 is mediated by RNA, we treated the extracts with RNaseA prior to the immunoprecipitation. We consistently found that a population of Bre1 interacts with Npl3 in an RNase-independent manner (Figure 3 *cf.* lanes 3 and 4, top panel). These data indicate that Npl3 can physically interact with Bre1, consistent with previous data from high-throughput proteomic analyses.[91]

Deletion of *BRE1* or *UBP8* exacerbates the *npl3*Δ splicing defect

The genetic data connecting *NPL3* and the H2B ubiquitination machinery lend support for two possible models. One model predicts that Npl3 will affect H2B ubiquitination; we therefore measured the global percentage of ubiquitinated H2B but found the *npl3*Δ strain indistinguishable from wild-type (Figure S1). An alternative interpretation of the genetic data is that the H2B ubiquitination cycle is important for an Npl3-dependent process. We previously reported[24] that a strain lacking Npl3 accumulates a subset of pre-mRNAs, consisting primarily of the ribosomal protein genes (RPGs), whose splicing efficiency might be expected to affect growth rate. Given that deletion of *BRE1* exacerbates the *npl3*Δ growth defect, we tested whether deleting *BRE1* exacerbates the *npl3*Δ splicing defect.

We used our splicing-sensitive microarray platform,[92] which contains oligos that hybridize to the terminal exon, the intron, and the exon-exon junction of each intron-containing gene, in order to detect total mRNA, pre-mRNA, and mature mRNA, respectively (Figure 4A). For each genotype, the heat map (Figure 4B) reports fold changes in signal intensity of these three RNA species for each intron-containing gene as compared to a wild-type strain. As expected, our experiments showed that a strain lacking Npl3 accumulated RPG pre-mRNAs (Figure 4B *npl3*Δ, see yellow in Intron feature; RPGs highlighted in purple on right). Notably, the pre-mRNA accumulation in the *npl3*Δ strain was increased at many RPGs when *BRE1* was also deleted (Figure 4B *cf. npl3*Δ and *npl3*Δ*bre1*Δ, Intron feature), suggesting that Bre1 is important for the splicing of many Npl3-dependent genes. We note that this effect is complex, and is accompanied by changes in total mRNA (Figure 4B *cf. npl3*Δ and *npl3*Δ*bre1*Δ, Exon feature). Because both Npl3 and Bre1 have been shown to have effects on transcription itself,[25,43-45,93-95] we normalized for changes in exon level by calculating an Intron Accumulation Index[96] (see Materials and Methods) for each intron-containing gene (Figure S2 and Table S5). The histogram of genes with an Intron Accumulation Index of greater than 0.3 (Figure 4C), shows that even when normalized for changes in transcript levels, the total number of genes with a splicing defect, as well as the severity of the defect, is increased in the *npl3*Δ*bre1*Δ strain as compared to *npl3*Δ alone.

We also found that in the presence of wild-type Npl3, a strain lacking *BRE1* has a mild but reproducible splicing defect (Figure 4B, *bre1*Δ -- shown is an average of 5 biological replicates). While the majority of pre-mRNAs are not affected by the deletion of *BRE1*, a small subset of pre-mRNAs accumulates in *bre1*Δ at 37˚C (Figure 4B, *e.g., DBP2, LSB3, YOP1*). This suggests that Bre1 has a role in pre-mRNA splicing, independent of the sensitivity caused when *NPL3* is deleted. This finding was confirmed when we calculated Intron Accumulation Indices for a strain lacking *BRE1:* a small number of genes exhibit defective splicing in the *bre1*Δ strain (Figure 4C and Figure S2). We validated these splicing defects for several genes using a qPCR assay (Figure S3). The lack of a

significant growth defect in the *bre1*Δ strain (Figure 1E) is consistent with the idea that yeast can tolerate a modest splicing defect at a small number of non-RPGs.

If the splicing defect exacerbation we observed with *npl3*Δ*bre1*Δ was due to loss of H2B ubiquitination, we would then expect this exacerbation to be phenocopied by a strain with the H2B lysine to arginine point mutant used earlier (Figure 1E). We did, in fact, find that the *htb1K123R* point mutation exacerbated the splicing defect observed in the *npl3*Δ mutant at many genes (Figure 4B), further implicating the ubiquitination of H2B in splicing. This is also evident when normalizing for the changes in exon levels in the *npl3*Δ*htb1K123R* strain (Figure 4C). In plotting the Intron Accumulation Index values of this strain, we find that the subset of affected genes overlaps extensively with the subset of genes affected in the *npl3*Δ*bre1*Δ double mutant (Figure S2).

We have shown that deletion of *UBP8* partially suppresses the *npl3*Δ growth defect, and this is most pronounced at 16˚C (Figure 2B). We therefore tested whether deleting *UBP8* would suppress the splicing defect of a strain lacking Npl3, as predicted by the genetic interaction. Surprisingly, deletion of *UBP8* instead exacerbated the splicing defect observed in the *npl3*Δ strain (Figure 4B *cf.* *npl3*Δ and *npl3*Δ*ubp8*Δ), implying that the growth suppression is related to some other function of Npl3. Notably, however, these microarray results indicate that in the absence of Npl3, the complete cycle of H2B ubiquitination and de-ubiquitination is required for efficient splicing.

To begin to investigate how Bre1 affects splicing, we used chromatin immunoprecipitation (ChIP) to test the prediction that Bre1 is required for association of the splicing machinery. However, we did not observe a significant Bre1-dependent decrease in U1 (Prp42), Mud2, or U2 (Lea1) association with genes whose splicing was inhibited in *bre1*Δ or *npl3*Δ*bre1*Δ strains (data not shown), suggesting an alternative mechanism by which Bre1 modulates splicing (see discussion).

Synthetic sickness between *BRE1* and early and late splicing factors.

In light of our data showing that *bre1*Δ exhibited a mild splicing defect, we carried out directed genetic analyses to test for interactions between *BRE1* and genes encoding other splicing

factors, particularly those that genetically interact with Npl3.[24] Just like a deletion of *NPL3*, deleting *BRE1* caused synthetic sickness when combined with deletion of *NAM8* (U1 snRNP), *MUD2*, *LEA1* (U2), or *SNU66* (U5), further connecting Bre1 functionally with splicing (Figure 5A). Interestingly, the growth of the *bre1*Δ strain was also compromised by deletion of the U2 snRNP component *CUS2*, which does not genetically interact with *npl3*Δ.[24] Thus, although we approached these experiments through the lens of Npl3, these genetic observations provide further support that Bre1 has independent interactions with the splicing machinery. Consistent with a lack of splicing defect upon *UPB8* deletion, we and others generally did not observe genetic interactions between *UBP8* and early or late splicing factors (Figure 5B).[17] There is one notable exception, however; deletion of *UBP8* suppressed the *snu66*Δ cold-sensitive growth defect (Figure 5B). Taken together, these data highlight the fact that the H2B ubiquitination pathway is linked to splicing, even in the presence of wild-type Npl3.

DISCUSSION

While the textbook view of gene expression presents transcription, pre-mRNA processing, export, and translation as independent events, they appear to be closely coordinated in the cell. Discerning the mechanism of this coordination within the broader program of gene expression presents a daunting experimental challenge. In multicellular eukaryotes, SR and hnRNP proteins are thought to regulate gene expression, in part, by integrating mRNA biogenesis steps.[39,40,97-99] In budding yeast, Npl3 has numerous roles, but is the only such protein that affects splicing.[24] Our lab and others have shown that Npl3 facilitates the co-transcriptional recruitment of early splicing factors to nascent transcripts[24] and itself associates with elongating polymerase.[25,42] Therefore, we approached this complex problem by conducting a systematic screen of non-essential genes to define the interacting partners of this potential coupling factor.

We found that Npl3 genetically interacts with a number of genes implicated in chromatin metabolism and transcription. We further characterized one set of interacting factors, namely those involved in histone H2B ubiquitination, in what is, to our knowledge, the first set of genome-wide splicing experiments on histone modifier mutants in *S. cerevisiae.* Using splicing-sensitive microarrays, we showed that Npl3 links the H2B ubiquitination cycle to the splicing efficiency of many transcripts. The connection between H2B ubiquitination and splicing also exists independently of Npl3, as a strain lacking *BRE1* exhibits both a mild splicing defect and genetic interactions with deletions of genes encoding early and late splicing factors. Finally, the full complement of genetic interactions we describe (Figure 6) provides multiple entry points for future investigation into the coupling of chromatin modification and mRNA processing.

An Npl3-dependent role for the histone H2B ubiquitination cycle in pre-mRNA splicing

Our genetic screens revealed that a number of Npl3 genetic interactions center on the histone H2B ubiquitination cycle. Specifically, mutant strains that lack wild-type levels of ubiquitinated H2B (*rad6*Δ, *bre1*Δ, *lge1*Δ, *htb1K123R*, *paf1*Δ, *cdc73*Δ*,* and *leo1*Δ) exacerbate the growth defect of an *npl3*Δ strain at all temperatures tested (Figure 1). We also observed a physical interaction between Npl3 and Bre1 by co-immunoprecipitation (Figure 3) and showed that the splicing defect caused by deletion of *NPL3* is exacerbated by the additional deletion of *BRE1* or mutation of H2B (*htb1K123R*), thus implicating H2B lysine 123 mono-ubiquitination in splicing (Figure 4). We previously demonstrated that Npl3 primarily affects the splicing of RPGs;[24] here, we see that in the sensitized background of a strain in which RPG splicing is made limiting (*npl3*Δ), the histone H2B ubiquitination cycle is an important contributor to RPG splicing.

Recent studies have shown that deletion of components of the cap-binding complex (CBC) or commitment complex causes defective splicing of the *SUS1* pre-mRNA.[17,100] Sus1 is a recently discovered component of the histone de-ubiquitination module of SAGA[101] and if the *SUS1* transcript is not properly spliced, it leads to elevated levels of ubiquitinated H2B. Given the physical[58,102] and

genetic (Table S1)[102] connections between Npl3 and the CBC, we sought to determine whether Npl3 also affects *SUS1* splicing and, therefore, H2B ubiquitination. However, Hossain *et al.* have recently shown that deletion of *NPL3* has no effect on *SUS1* splicing,[100] a result we independently confirmed in our *npl3*Δ strain (Figure S4). Furthermore, we extended this analysis and determined that, unlike in *cbc*Δ strains, global levels of ubiquitinated histone H2B are not discernibly altered in the *npl3*Δ strain (Figure S1). While we cannot rule out a change in the dynamics of the ubiquitination cycle or gene-specific effects, our microarray results support a model in which the full histone ubiquitination cycle promotes RPG splicing, a process that becomes critical in the absence of *NPL3*. Along these lines, it is noteworthy that data from Schulze *et al.* and Shieh *et al.* have revealed that chromatin over these genes is enriched for ubiquitinated H2B.[103,104]

We also identified suppressive genetic interactions between *NPL3* and genes responsible for removal of ubiquitin from H2B (*ubp8*Δ, *sgf11*Δ, and *sgf73*Δ), suggesting that H2B de-ubiquitination is also linked to Npl3 function. Surprisingly, however, deletion of *UBP8* did not suppress the splicing defect in *npl3*Δ, but rather exacerbated it (Figure 4). Thus, it seems the positive genetic interaction may be due to Ubp8 involvement in a splicing-independent function of Npl3.

The exacerbation seen in the microarray experiments shows that both halves of the cycle of H2B ubiquitination and de-ubiquitination are required for optimal splicing, as is the case for transcriptional activation.[52,85] Likewise, both halves of the H3 acetylation and deacetylation cycle, performed by Gcn5 and Hos2/3, respectively, promote spliceosome assembly at the *ECM33* gene.[15,105] Thus, these two examples point to a general function of dynamic histone modification cycles in maintaining fine control over co-transcriptional splicing, and may explain the synthetic lethality we observed between *NPL3* and the acetylation module of the SAGA Complex (Figure 1C and Table S1).

An Npl3-independent role for Bre1 in pre-mRNA splicing

We found that even in the presence of wild-type *NPL3,* Bre1 has genetic connections to the splicing machinery as a whole. Specifically, we found that deletion of *BRE1* causes growth defects in early and late splicing factor deletion backgrounds (particularly at extreme temperatures; Figure 5, 16˚C and 37˚C), which alone show little to no growth defect. These negative genetic interactions can indicate two alternative but not mutually exclusive models for a functional relationship between the H2B ubiquitination and splicing machineries. One model is based on the fact that deletion of specific splicing factors is known to increase the levels of ubiquitinated H2B,[100] a phenotype that should be relieved by deletion of *BRE1*, the sole H2B ubiquitin ligase.[48,49] Because this model predicts an epistatic or positive genetic interaction between *BRE1* and the genes that encode splicing factors, the negative genetic interactions that we actually observe (Figure 5)[17] require an alternative model, perhaps one in which the growth defects are due to poorer overall splicing efficiency in these strains. Indeed, deletion of *BRE1* alone caused a modest but reproducible splicing defect, seen in the microarray in Figure 4. A large fraction of Bre1-dependent splicing events involve non-RPGs, and thus define a distinct role for Bre1 in splicing, apart from Npl3. Shieh *et al.* recently found that the pattern of this modified histone at non-RPGs shows a remarkable demarcation of intron/exon structure: low levels in the intron, followed by a marked increase at the intron – exon boundary.[104] While the functional significance of this pattern of H2B ubiquitination is unknown, we propose that it may be relevant for the splicing of non-RPGs, as gauged by the splicing defect in a strain that no longer has this mark.

We note that the single mutant *htb1K123R* has a milder splicing defect than the *bre1*Δ strain (Figure 4C and Figure S2), opening the possibility of an additional role of Bre1 in splicing that is independent of H2B ubiquitination. Indeed, Bre1-dependent ubiquitination of Swd2, a protein in both COMPASS and the Cleavage and Polyadenylation Stimulatory Factor complex,[73] has been shown to regulate mRNA export from the nucleus.[106] Npl3 has previously been implicated in mRNA export[46,47] in a strain background where Npl3 is an essential protein. However, our data argue against

the possibility that the genetic interactions we observed here are due to an adverse effect on mRNA export. In the present strain background (S288C), in which Npl3 is non-essential, the *npl3*Δ strain does not exhibit the nuclear localization of bulk poly-adenylated mRNA characteristic of an export defect (Figure S5); nor does further deletion of *BRE1* in an *npl3*Δ strain cause an export defect (Figure S5). Furthermore, we found that the *npl3S411A* phosphorylation mutation, which blocks 3' end formation[25,43,44] and mRNA export,[46] does not cause a block in pre-mRNA splicing (Figure S6). This argues against the reported splicing defects being the indirect result of feedback from these downstream defects in mRNA processing.

We tested the prediction that Bre1, like Npl3, promotes spliceosome recruitment, but found that deletion of *BRE1* did not affect the association of U1 (Prp42), Mud2, or U2 (Lea1) with chromatin at genes whose splicing is dependent on Bre1 (data not shown). It may be that H2B ubiquitination is required for the recruitment of a later splicing factor or, as the H2B ubiquitination cycle regulates PolII passage through a gene,[93,107] it is possible that disruption of this cycle causes a subtle alteration of spliceosome dynamics that is not observable by ChIP. Furthermore, we cannot rule out the possibility that Bre1 has a ubiquitination target within the spliceosome or even ubiquitinates Npl3 itself.

Both splicing and mRNA processing are largely co-transcriptional processes in eukaryotes, from yeast[34] to human.[108-110] Our survey of *NPL3* genetic interactions has revealed a multitude of chromatin-connected factors with potential links to splicing and mRNA processing; overall, these results are thus consistent with an "integrator" role for Npl3 in gene expression (Figure 6). Our data provide a basis for the further study of the coupling of SR/hnRNP-dependent mRNA processing and transcription within a chromatin context, and have led to the discovery of Npl3-dependent and independent roles for Bre1 and histone H2B ubiquitination in splicing.

MATERIALS AND METHODS

A list of strains used and further strain construction details are available in Table S6.

Synthetic Genetic Array

Unless otherwise indicated, yeast were grown as described in Guthrie *et al.*[111] The *npl3*Δ*::NatNT2* "magic marker" query strain used in the SGA was YTK232D, and was previously used in.[55] YTK232D was generated using techniques outlined in Yanke *et al.*[112] Briefly, the *NPL3* open reading frame was replaced with *NatNT2* via integration of a PCR product generated with primers

(5'TACTTTTGAAGGAATCAAAATTAAGCAATTACGCTAAAACCATAAGGATAACATGGA GGCCCAGAATACCC-3') and

(5'GTTTTAAAACAATTCATATCTTTTGTTAATTTCTCCTTTTTTTTTCTCAACCAGTATAGC GACCAGCATTC-3') into the SGA diploid strain.[113] The diploid was sporulated and the MATα *npl3*Δ*:NatNT2* query strain was isolated by tetrad dissection, followed by re-selection of magic markers on SD medium lacking leucine and arginine but containing canavanine, s-AEC, and clonNAT [SD - LEU/ARG + 100µg/mL canavanine + 100µg/mL S-(2-Aminoethyl)-L-cysteine hydrochloride + 100µg/mL clonNAT]. The *NPL3* deletion was confirmed by PCR, and by Western blot for the absence of Npl3 using an α-Npl3 antibody.[90]

The Synthetic Genetic Array was performed as described in Tong *et al.*[53] with the following exceptions: Here the *npl3*Δ query strain (YTK232D) was mated to the MAT**a** KanMX-marked deletion collection (OpenBiosystems: www.openbiosystems.com; formerly Research Genetics, Huntsville, AL). The collection was arrayed in duplicate in 384-well colony format using automated pinning (Colony Arrayer) and grown at 30˚C for 2 days. Mating was carried out at 30˚C for 2 days. Sporulation was carried out at 30˚C for 7 days. MAT**a** double mutants were selected on SD medium lacking histidine and arginine but containing canavanine, S-AEC, G418, and clonNAT [SD - HIS/ARG + 100µg/mL canavanine + 100µg/mL S-(2-Aminoethyl)-L-cysteine hydrochloride + 150µg/mL G418 and 100µg/mL clonNAT]. Double mutant arrays were re-pinned in replicate and

photographed after the following incubations: 30˚C for 5 days, 37˚C for 5 days, or at 16˚C for 4 and again after 8 days. Photographs were visually inspected for growth at 16˚C (to identify suppressive interactions) or lack of colony growth at 30˚C or 37˚C (to identify synthetic lethal interactions).

Directed Genetics

The *npl3*Δ*::NatNT2* strain used for directed genetics (YTK234D) was previously used in Kress *et al.*[24] Unless otherwise indicated, YTK234D was crossed to a series of MAT**a** KanMX4-marked deletion strains; diploids were selected by plating on YPD plates +100µg/mL clonNAT + 150µg/mL G418. Double mutants were isolated by tetrad dissection or random sporulation, as indicated in Table S6. All single mutants were validated by PCR for the knockout chromosome prior to crossing to YTK234D. The *HTB1-WT* (WHY334) and *htb1-K123R* (WHY326) strains contain *htb2*Δ*::HygX4l* and the indicated *htb1* allele as the sole copy of H2B (gifts from W. Hwang and H. Madhani). They were mated as above, except the diploid strains were selected on YPD + 100µg/mL hygromycin + 100µg/mL clonNAT. Because the *htb1* allele is unmarked, the final *npl3*Δ*HTB1* and *npl3*Δ*htb1K123R* strains were confirmed by sequencing the *HTB1* gene and Western blot for the Npl3 protein. Genetic interactions with the *bre1H665A* allele were analyzed using a set of plasmids provided by the Shilatifard lab,[48] designed to complement a *bre1*Δ allele. Complementation was achieved by plasmid transformation into YM1740 (*bre1*Δ) or YTK391B (*npl3*Δ*bre1*Δ), which were maintained on SD -LEU plates.

The *bre1*Δ*::NatNT2* (EMy32) and *ubp8*Δ*::NatNT2* (EMy442) strains were created by replacement of the endogenous ORF with NatNT2, as described in Yanke *et al.*[112] These strains were subsequently mated to *nam8*Δ*, mud1*Δ*, mud2*Δ*, syf2*Δ*, and snu66*Δ (for *bre1*Δ), and *lea1*Δ (for *bre1*Δ and *ubp8*Δ) from the deletion collection and double mutants were isolated via tetrad dissection. For the rest of the *ubp8*Δ genetics, the *ubp8*Δ*::KanMX4* strain from the deletion collection was mated to MATα NAT-marked "magic marked" splicing factor deletion strains. These splicing factor deletion strains were made by replacing the KanMX-marked ORFs with NatNT2, followed by crossing to a

"magic marked" wild-type (YTK609) to isolate "magic marked" NAT-marked MATα spores. The *ubp8*Δ*::KanMX4* strain was mated to each NAT-marked splicing factor deletion strain and MAT**a** double mutants were isolated by tetrad dissection followed by selection on SD medium lacking histidine and arginine but containing canavanine, S-AEC, G418, and clonNAT [SD - HIS/ARG + 100μg/mL canavanine + 100μg/mL S-(2-Aminoethyl)-L-cysteine hydrochloride + 150μg/mL G418 and 100μg/mL clonNAT].

For individual growth assays, log-phase yeast were diluted to OD_{600} = 0.1, spotted onto YPD plates (unless specifically mentioned) in a 5-fold dilution series and grown at the indicated temperatures. For each cross, growth of the double mutant was confirmed for ≥2 double mutant isolates, and a representative isolate is shown. The single mutants and wild-type strains shown are either parental strains, or were re-isolated from tetra-type tetrads. The *bre1H665A* and *BRE1* strains were serially diluted onto SD –LEU plates.

Process and Complex Analyses

We sought to integrate the diverse sources of genetic interaction information available to us in order to create a comprehensive dataset for statistical analyses. Because the stronger synthetic interactions were identified in the 30˚C SGA, we began with this list of genes whose deletion caused lethality in combination with *npl3*Δ (see Table S1 – 30˚C) and added genes identified as causing markedly decreased growth, as gauged by serial dilution, or lethality, as gauged by loss of double mutant spore after tetrad dissection (Figure 1C, 1D, 1E and Table S1). We further added to this list genes identified as synthetic sick or lethal in the E-MAP[55] *i.e.,* having a genetic interaction score of ≤ -2.5.

Biological process definitions were obtained from the Gene Ontology annotations maintained at SGD[114] on April 15th 2012. Forty-five high-level (GO Slim) terms were used and are included in Table S7. Protein complex definitions were obtained from a manually curated list, CYC2008,[115] and augmented with the RAD6 Complex (*RAD6*, *BRE1*, *LGE1*), which was not annotated when the list

was created. A hypergeometric test was used to identify complexes and processes that were significantly enriched with genetic interactions. Complex enrichment p-values were corrected for multiple testing using the empirical re-sampling method of Berriz *et al.*[116] (as 409 complexes were assessed for enrichment), while process enrichment p-values were corrected for using the simpler Bonferoni correction. The results of the these analyses are included in Table S2 (by process) and Table S3 (by complex). The network diagram in Figure 1B was drawn using Cytoscape.[117] For Figure 1B, complexes were referred to by their more common names. The Figure 2A diagram was created to highlight a subset of suppressive interactions identified in the 16ºC SGA and the full list of suppressors is available in Table S4.

Co-immunoprecipitation

Co-immunoprecipitation assays were performed as in Kress *et al.*[24] with extracts from the indicated GFP-tagged or HA-tagged strains. The Nup188-HA strain contains a plasmid encoding Nup188-3XHA. The other strains were tagged endogenously. Briefly, samples were separated by 10% SDS-PAGE and probed by Western blot with either monoclonal α-GFP (Roche 1814460), α-HA (12CA5; Roche 11583816001), or polyclonal α-Npl3 antibodies.[90] Total samples equivalent to $1/60^{th}$ of the input were analyzed in parallel.

Microarrays

Cultures were grown according to standard techniques[111] in rich medium supplemented with 2% glucose. Strains were cultured overnight to saturation and diluted to $OD_{600} = 0.1$ in the morning. The strains were allowed to grow at 30ºC until reaching mid-log phase ($OD_{600} = 0.5$ - 0.7), at which point they were collected (for Figure S6), or rapidly shifted to either 37˚C for 30 minutes or 16˚C for 2.5 hours, as indicated. Cultures were collected by centrifugation and snap frozen in liquid nitrogen. Total cellular RNA was isolated using hot acid phenol followed by isopropanol precipitation, as outlined in Schmitt *et al.*[118] but with modifications detailed in Bergkessel *et al.*[4] cDNA from each

strain was synthesized, and labeled with Cy3 or Cy5 according to the low-throughput sample preparation method described in Pleiss *et al.*[92]

The optimized oligos listed in Pleiss *et al.*[92] were robotically arrayed onto poly-L-lysine coated glass slides (slides from ThermoScientific C40-5257-M20) and slides were processed using the protocols detailed in Pleiss *et al.*[92] and DeRisi *et al.*[119] Each biological replicate contains 6 technical replicates for each feature as well as dye-flipped replicates. Microarrays were scanned using Axon Instruments GenePix 4000B at 635nm and 532nm wavelengths and image analysis was done using Axon Instruments GenePix Pro version 5.1. Spots were manually removed from analysis if they contained obvious defects or uncharacteristically high background; the ratio of the median intensities for 535nm and 625nm was calculated for each remaining spot. Technical replicate spots and dye flipped replicates were combined and normalized as in Pleiss *et al.*[92] The resulting $\log_2$-transformed values for each feature were averaged over 2-5 biological replicates. Averaged data were subjected to hierarchical clustering using average linkage, and uncentered Pearson correlation as the similarity metric using Cluster 3.0.[120] Resulting heat maps in Figure 4, Figure S2, and Figure S6 were created using Java Treeview.[121] To normalize for changes in total expression evident in the microarrays, Intron Accumulation Indices (IAI) were calculated for each intron containing gene as in Clark *et al.*;[96] specifically, we calculated $\log_2(\text{Intron}_{\text{mutant}}/\text{Intron}_{\text{WT}})$-$\log_2(\text{Exon}_{\text{mutant}}/\text{Exon}_{\text{WT}})$ for each gene. The IAI heat map is shown in Figure S2. These values were converted into a histogram for Figure 4 using the following cutoffs: $-0.3 \geq \text{IAI} \geq 0.3$.

dT50 Fluorescent *in situ* hybridization

The dT50 assay was performed based on the protocol outlined in Amberg *et al.*[122] with the following modifications. Specifically, 2mL cultures were fixed in 5% formaldehyde for 1.5 hours after having reached $OD_{600} = 0.2$-0.3. Cells were washed 4 times in wash buffer (100mM Potassium Phosphate, 1.2M Sorbitol) before a 40-minute treatment with 27µg zymolyase at 37˚C. An additional fixation was performed in 8% paraformaldehyde in PBS + 10mM MgCl2 and spheroplasted cells

were applied to poly-lysine-treated chamber slides (LabTek 178599). Attached cells were treated with ice-cold methanol (-20˚C) and allowed to dry. Hybridization to digoxin-conjugated dT50 oligo in blocking buffer was performed at 37˚C overnight. Chambers were washed with 2X (20 minutes), 1X (20 minutes) and 0.5X SSC (10 minutes at 37˚C) before a 30-minute incubation with FITC-conjugated anti-Digoxin Fab fragments (Roche 1207741) in blocking buffer (1:25 dilution, 37˚C). Antibody was aspirated and three 5-minute washes of PBS +10mM MgCl2 were performed. Chambers were treated with 0.5mg/mL DAPI for 2 minutes and slides were mounted using ProLong Gold Antifade Reagent (Invitrogen P36934) according to manufacturer instructions. Slides were visualized using an Olympus BX60 microscope equipped with FITC HiQ and DAPI HiQ Filters (Chroma Technology Corporation). The assay was performed on two biological replicates and representative images are shown. Specificity of the probe and FITC labeling was determined by incubation with hybridization mix lacking probe (data not shown).

***SUS1* splicing assay**

SUS1 splicing efficiency was measured essentially according to the non-radioactive protocol described in Hossain *et al.*[100] Specifically, 10µg RNA from cultures grown at 30˚C was treated with DNaseI (Promega) and RNA was converted to cDNA using 1µg *SUS1* Reverse primer.[17] cDNAs were diluted 1:200 and 10µL was used in a 25µL PCR (BioRad iProof) with *SUS1*-specific primers.[17] 25 cycles of PCR were performed and the resulting products were separated on an 8% polyacrylamide gel. Gels were stained using SybrGold and bands were quantified using an AlphaImager HP camera and software. 2-3 technical replicates of 2 biological samples were performed. Shown are a representative gel and the average and standard deviations of all technical replicates. A no-Reverse Transcriptase control was performed for each sample and none showed amplification (data not shown).

H2B ubiquitination Western blot

A TCA precipitation was performed on strains grown at 30˚C[123] and samples were run on a

15% SDS polyacrylamide gel and transferred to PVDF membrane. Membrane was blocked using Li-Cor blocking buffer, followed by incubation of a 1:1000 dilution of α-H2B antibody (Active Motif 39237) overnight at 4˚C. Visualization of bands was achieved with a secondary antibody conjugated to infrared dye (LI-COR 926-32211). The membrane was scanned using the LI-COR Odyssey scanner and software. Shown is a representative Western blot. The assay was performed with 3 biological replicates and shown are the average and standard deviation of the three replicates.

qPCR Assay

RNA was extracted as described above from strains grown under the same conditions as for the microarray experiment. Five µg RNA were treated with DNaseI (Promega) before being primed with random 9-mers and reverse transcribed. Samples were diluted as necessary and 10µL were used in each qPCR. qPCRs were run on a C1000 ThermoCycler (BioRad) with an annealing temperature of 55˚C. Each qPCR run was finished with a melt curve to determine the homogeneity of the amplified product. Starting quantity was calculated using a standard curve for each primer set. 2-4 technical replicates were performed for 1-5 biological replicates. Error bars represent standard deviation for biological replicates. For samples with 1 biological replicate, standard deviation of technical replicates is shown with uncapped error bars (Figure S3). A no-Reverse Transcriptase control was also generated for each RNA sample and these samples yielded negligible amplification (data not shown). Primers used in the qPCR are listed in Table S8. Each gene was measured using intron- and exon- specific primer sets. The Intron/Exon ratio for each mutant was normalized to its corresponding wild-type before averaging.

ACKNOWLEDGEMENTS

We are indebted to Joseph DeRisi, Gregg Whitworth, Megan Bergkessel, and Alex Plocik for assistance with the microarray platform, Anne de Bruyn Kops and Gwendolyn Wilmes for assistance with the dT50 assay, and David Pincus for assistance with the fluorescent Western blots. The gifts of

strains from William Hwang, Hiten Madhani, and Dale Cameron, and of plasmids from Ali Shilatifard are deeply appreciated. We also thank Michael Shales for guidance with figure design, and Hannes Braberg (Krogan lab) and the entire Guthrie lab, in particular Anne de Bruyn Kops, Jaclyn Greimann, Michael Marvin, Kelly Nissen, Kristin Patrick, and Argenta Price, for suggestions on the manuscript.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: EAM CJR NJK TLK CG. Performed the experiments: EAM TLK. Analyzed the data: EAM CJR NJK TLK CG. Contributed reagents, materials and/or analysis tools: EAM CJR NJK TLK CG. Wrote the paper: EAM CJR NJK TLK CG.

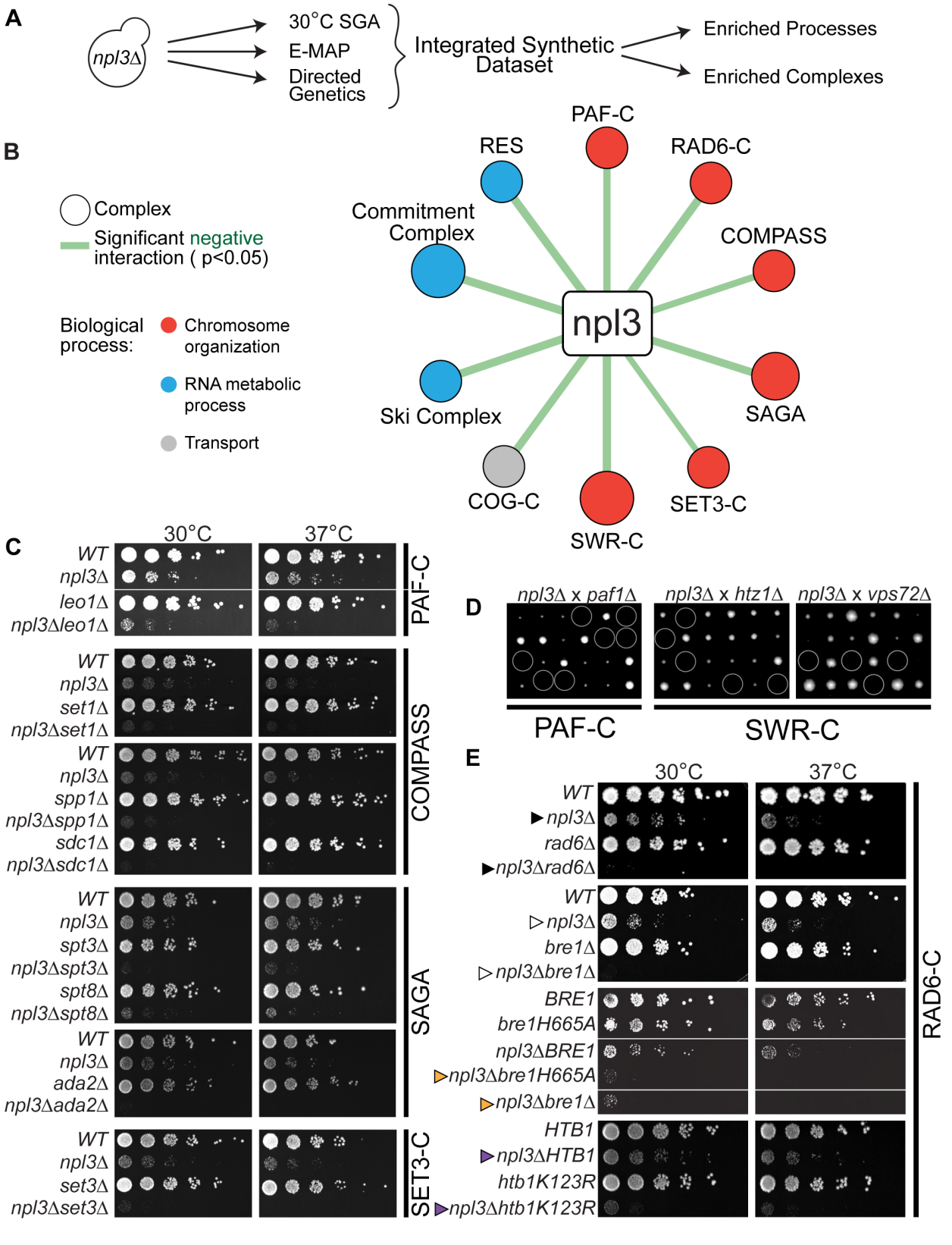

A
npl3Δ
30°C SGA
E-MAP
Directed Genetics
Integrated Synthetic Dataset
Enriched Processes
Enriched Complexes
B
Complex
Significant negative interaction (p<0.05)
Biological process:
Chromosome organization
RNA metabolic process
Transport
PAF-C
RES
RAD6-C
Commitment Complex
COMPASS
npl3
Ski Complex
SAGA
COG-C
SET3-C
SWR-C
C
30°C
37°C
WT
npl3Δ
leo1Δ
npl3Δleo1Δ
PAF-C
set1Δ
npl3Δset1Δ
spp1Δ
npl3Δspp1Δ
sdc1Δ
npl3Δsdc1Δ
COMPASS
spt3Δ
npl3Δspt3Δ
spt8Δ
npl3Δspt8Δ
ada2Δ
npl3Δada2Δ
SAGA
set3Δ
npl3Δset3Δ
SET3-C
D
npl3Δ x paf1Δ
npl3Δ x htz1Δ
npl3Δ x vps72Δ
PAF-C
SWR-C
E
30°C
37°C
rad6Δ
▶npl3Δrad6Δ
▷npl3Δ
bre1Δ
▷npl3Δbre1Δ
BRE1
bre1H665A
npl3ΔBRE1
▶npl3Δbre1H665A
▶npl3Δbre1Δ
HTB1
▶npl3ΔHTB1
htb1K123R
▶npl3Δhtb1K123R
RAD6-C

Figure 1. Extensive negative genetic interactions with *npl3*Δ connect *NPL3* to chromatin biology.

(A.) Work flow for analysis of the integrated synthetic dataset. Synthetic genetic array technology was used to screen ~4800 non-essential genes whose deletion conferred lethality to *npl3*Δ at 30˚C. These results were augmented by including genes exhibiting a genetic interaction score of ≤ -2.5 with *npl3*Δ[55] and genes identified as synthetic sick or lethal using tetrad dissection and serial dilution (directed genetics).

(B.) Statistically significant negative interactions between *NPL3* and known complexes. Statistical analysis identified the indicated complexes as having subunits significantly enriched ($P<0.05$) in the integrated synthetic dataset. Size of circle is based on number of subunits whose deletion exacerbates the growth defect of *npl3*Δ and thickness of line scales with the significance of the enrichment. Circles are color-coded based on the biological process to which the complex belongs.

(C.) Synthetic growth analyses with *npl3*Δ and genes implicated in chromatin biology. Each panel shows a double mutant strain, cognate single deletions strains and a corresponding wild-type that have been serially diluted onto rich medium and grown at the indicated temperatures. Double mutants were isolated after tetrad dissection. To the right of panels is the name of the complex to which the single chromatin mutants belong.

(D.) Tetrad dissection analyses with *npl3*Δ. Tetrad dissection plates from the indicated crosses are shown with the inviable spore circled. Replica plating to infer genotype later showed that inviable spores are the double mutants.

(E.) Synthetic growth analyses with *npl3*Δ and genes encoding the H2B ubiquitination machinery. Shown are serial dilutions of the indicated strains after incubation at the indicated temperatures. Genotypes not originally tested in the SGA are *htb1K123R* and *bre1H665A*.

The *BRE1* and *bre1H665A* strains contain a *bre1* deletion covered by a plasmid encoding the indicated *bre1* allele. Arrowheads refer to comparisons made in the text.

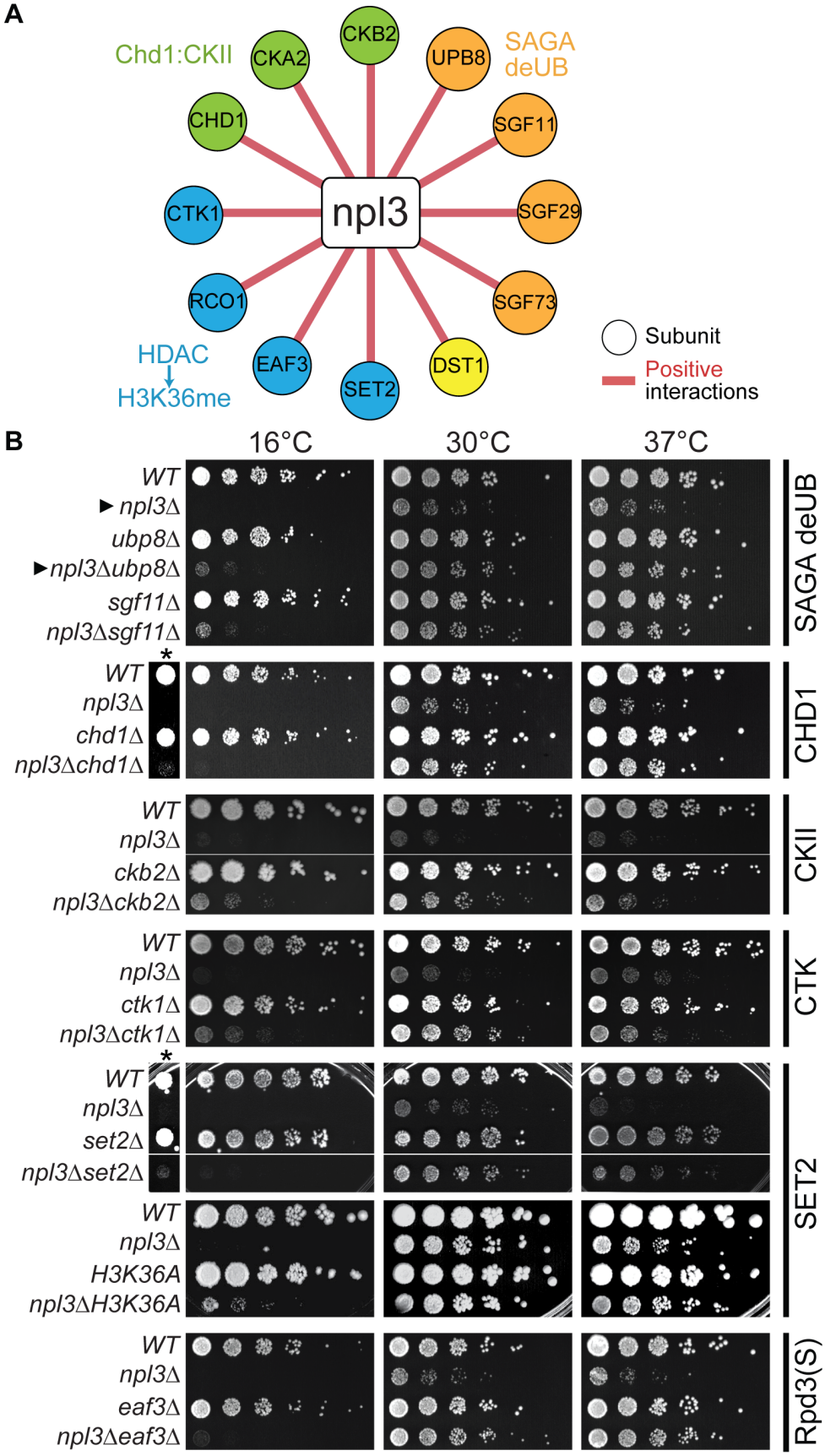

A
Chd1:CKII
CKA2
CKB2
UPB8
SAGA
deUB
CHD1
SGF11
CTK1
npl3
SGF29
RCO1
SGF73
HDAC
H3K36me
EAF3
SET2
DST1
Subunit
Positive
interactions
B
16°C
30°C
37°C
WT
► npl3Δ
ubp8Δ
►npl3Δubp8Δ
sgf11Δ
npl3Δsgf11Δ
SAGA deUB
*
WT
npl3Δ
chd1Δ
npl3Δchd1Δ
CHD1
WT
npl3Δ
ckb2Δ
npl3Δckb2Δ
CKII
WT
npl3Δ
ctk1Δ
npl3Δctk1Δ
CTK
*
WT
npl3Δ
set2Δ
npl3Δset2Δ
WT
npl3Δ
H3K36A
npl3ΔH3K36A
SET2
WT
npl3Δ
eaf3Δ
npl3Δeaf3Δ
Rpd3(S)

Figure 2. The *npl3*Δ growth defect is suppressed by mutations in genes encoding transcription and chromatin factors.

(A.) A subset of suppressive *npl3*Δ genetic interactions relevant to chromatin biology identified as allowing growth to *npl3*Δ in the SGA performed at 16˚C. Genes are arranged by complex or pathway. Full list is available in Table S4.

(B.) Suppressive growth analyses with *npl3*Δ and genes implicated in chromatin biology. Shown are serial dilutions of the indicated strains grown at the indicated temperatures. All double mutants were generated by tetrad dissection. *H3K36A* was not originally tested in the SGA. Asterisk marks a higher-contrast image to better visualize suppression at 16˚C. Arrowheads refer to comparisons made in the text.

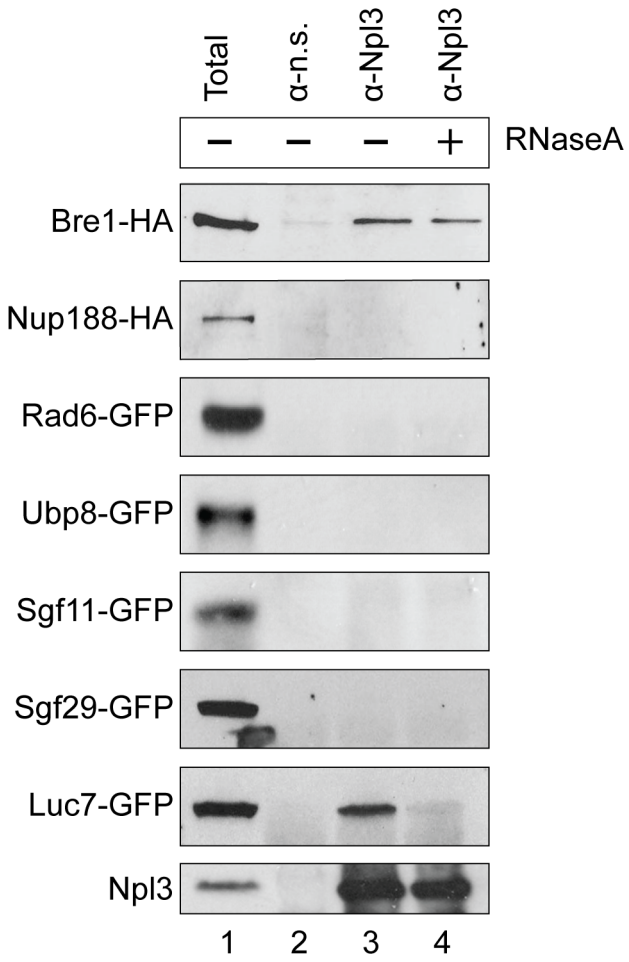


Figure 3. Npl3 physically interacts with Bre1.

Co-immunoprecipitation analyses of Npl3 and members of the histone H2B ubiquitination machinery. Whole cell extracts from strains with the indicated proteins endogenously tagged with HA or GFP were immunoprecipitated with an α-Npl3 antibody[90] or non-specific antibody (α-n.s.). Western blot using α-HA or α-GFP from each co-IP experiment is shown. The sensitivity of the interaction to RNase (lane 4, +RNaseA) was determined by treating lysates with RNase A prior to immunoprecipitation. Lane 1 shows 1/60 total sample for each lysate. Bottom panel confirms presence of Npl3 in the immunoprecipitate.

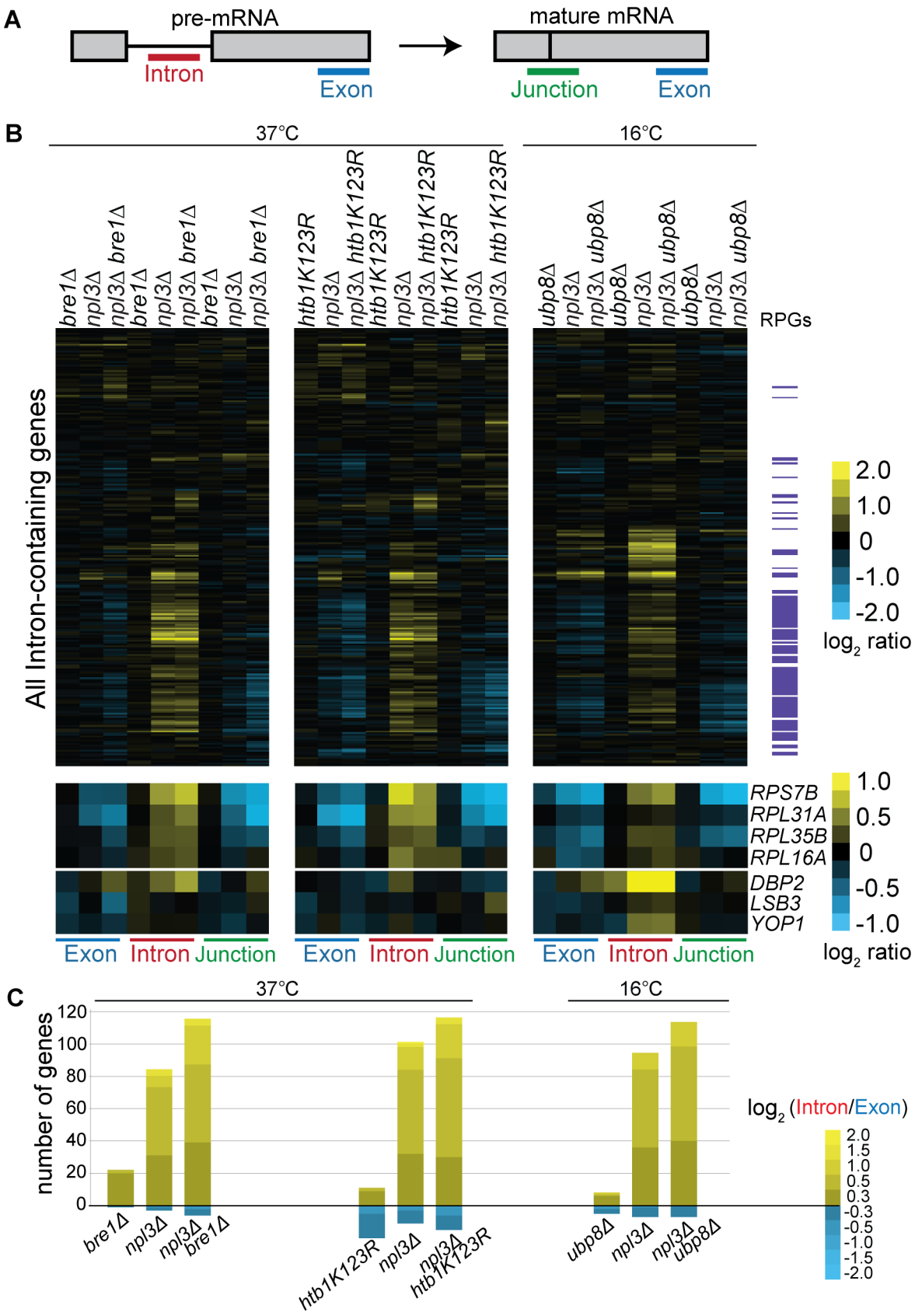

A
pre-mRNA
Intron
Exon
mature mRNA
Junction
Exon
B
37°C
16°C
bre1Δ
npl3Δ
npl3Δ bre1Δ
htb1K123R
npl3Δ htb1K123R
ubp8Δ
npl3Δ ubp8Δ
All Intron-containing genes
RPGs
2.0
1.0
0
-1.0
-2.0
log2 ratio
RPS7B
RPL31A
RPL35B
RPL16A
DBP2
LSB3
YOP1
1.0
0.5
0
-0.5
-1.0
log2 ratio
Exon
Intron
Junction
C
37°C
16°C
number of genes
120
100
80
60
40
20
0
bre1Δ
npl3Δ
npl3Δ bre1Δ
htb1K123R
npl3Δ htb1K123R
ubp8Δ
npl3Δ ubp8Δ
log2 (Intron/Exon)
2.0
1.5
1.0
0.5
0.3
-0.3
-0.5
-1.0
-1.5
-2.0

Figure 4. Splicing is sensitive to Npl3 and the H2B ubiquitination cycle.

(A.) Schematic of probes contained on the splicing microarray.

(B.) Splicing profile of single or double mutant strains compared to wild-type. Cultures of the indicated strains and isogenic wild-type strains were grown to mid-log phase at 30˚C and shifted to the indicated temperature; cDNA from single and double mutant strains were competitively hybridized on the microarray against that from an isogenic wild-type. The heat map shows the $\log_2$-ratio for each gene feature of the indicated strain compared to wild-type. Gene order along the y-axis is the same for all arrays. Transcripts that encode the ribosomal protein genes (RPGs) are highlighted in purple to the right of the heat maps. Data for example genes are replicated below the genome-wide heat map to show splicing defects and exacerbation of defects at individual RPGs and non-RPGs.

(C.) Histogram of $\log_2$-based Intron Accumulation Index scores. An Intron Accumulation Index value was calculated for each intron-containing gene by normalizing the intron change to exon change (see Materials and Methods). Histogram shows the number of genes with an Intron Accumulation Index score greater than 0.3. Heat map within histogram bars shows distribution of the severity of splicing defect.

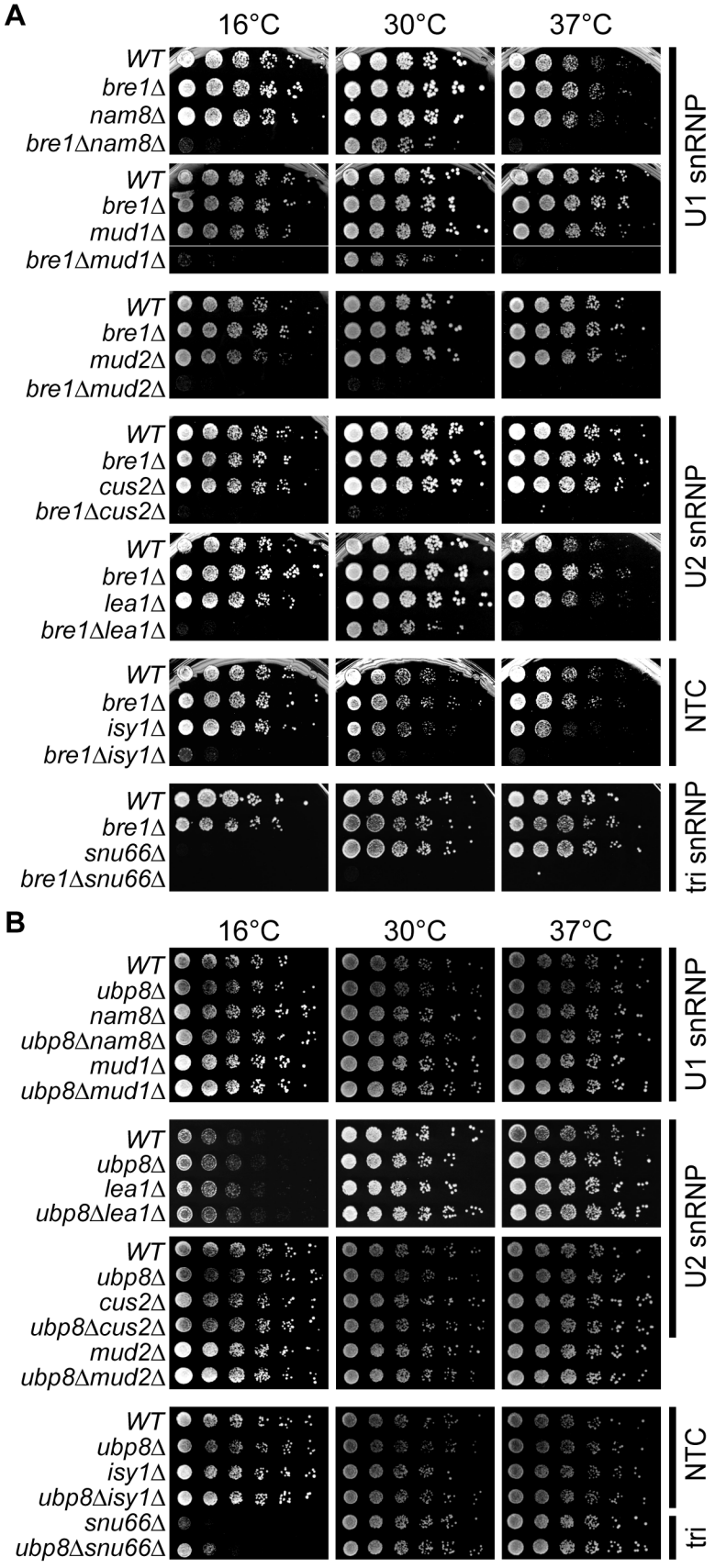

A
16°C
30°C
37°C
WT
bre1Δ
nam8Δ
bre1Δnam8Δ
WT
bre1Δ
mud1Δ
bre1Δmud1Δ
U1 snRNP
WT
bre1Δ
mud2Δ
bre1Δmud2Δ
WT
bre1Δ
cus2Δ
bre1Δcus2Δ
WT
bre1Δ
lea1Δ
bre1Δlea1Δ
U2 snRNP
WT
bre1Δ
isy1Δ
bre1Δisy1Δ
NTC
WT
bre1Δ
snu66Δ
bre1Δsnu66Δ
tri snRNP
B
16°C
30°C
37°C
WT
ubp8Δ
nam8Δ
ubp8Δnam8Δ
mud1Δ
ubp8Δmud1Δ
U1 snRNP
WT
ubp8Δ
lea1Δ
ubp8Δlea1Δ
WT
ubp8Δ
cus2Δ
ubp8Δcus2Δ
mud2Δ
ubp8Δmud2Δ
U2 snRNP
WT
ubp8Δ
isy1Δ
ubp8Δisy1Δ
NTC
snu66Δ
ubp8Δsnu66Δ
tri

Figure 5. Genetic interactions between H2B ubiquitination machinery and canonical splicing factors.

(A.) Synthetic growth analyses between *bre1*Δ and genes encoding splicing factors. Double mutants were generated by tetrad dissection; log-phase cultures of the indicated strains were serially diluted and grown at the indicated temperatures. To the right of panels is the name of the spliceosomal complex to which the splicing factor mutants belong. NTC: Nineteen Complex.

(B.) Growth analyses of *ubp8*Δ and genes encoding splicing factors. Double mutants were generated and analyzed as in (A).

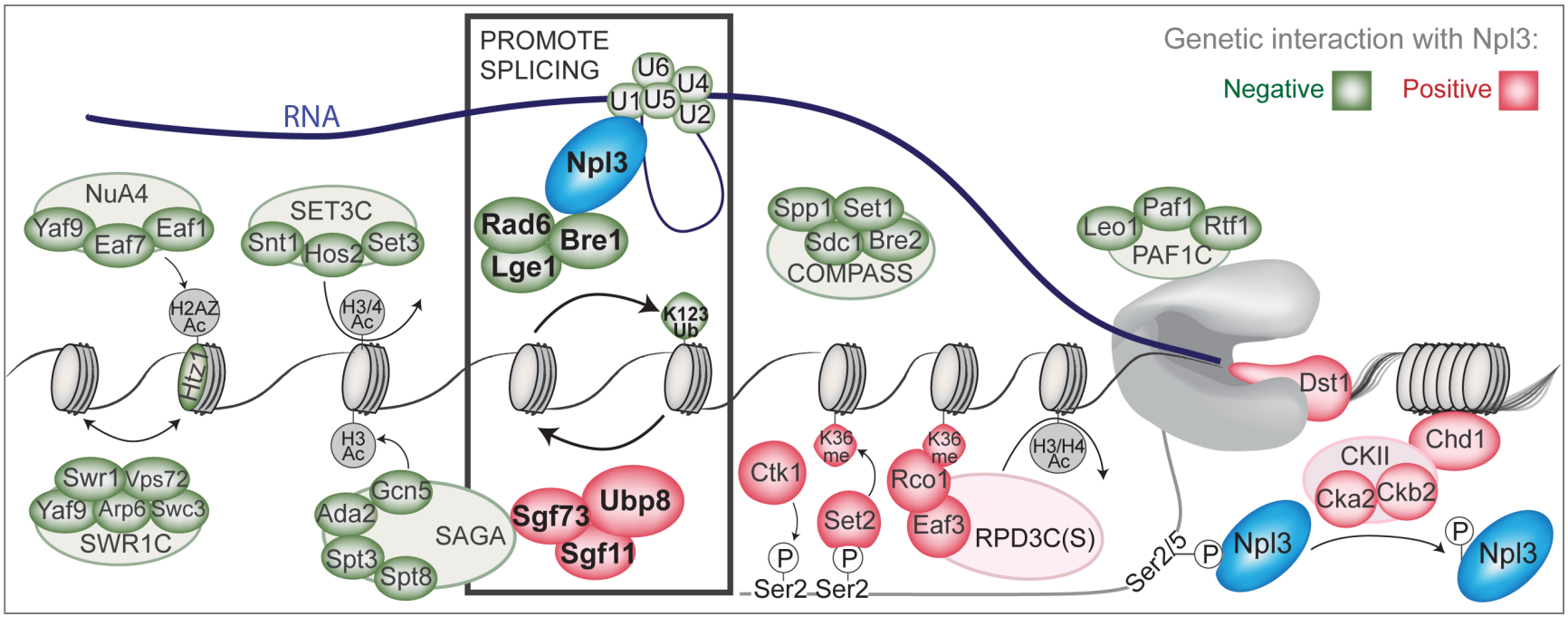

PROMOTE SPLICING
RNA
U6
U4
U1
U5
U2
Npl3
Rad6
Bre1
Lge1
K123 Ub
Genetic interaction with Npl3:
Negative
Positive
NuA4
Yaf9
Eaf7
Eaf1
H2AZ Ac
Htz1
SET3C
Snt1
Hos2
Set3
H3/4 Ac
H3 Ac
Swr1
Vps72
Yaf9
Arp6
Swc3
SWR1C
Gcn5
Ada2
Spt3
Spt8
SAGA
Sgf73
Ubp8
Sgf11
Spp1
Set1
Sdc1
Bre2
COMPASS
Ctk1
K36 me
Set2
Rco1
Eaf3
RPD3C(S)
H3/H4 Ac
P
Ser2
Ser2
Leo1
Paf1
Rtf1
PAF1C
Dst1
Chd1
CKII
Cka2
Ckb2
Ser2/5
Npl3
Npl3

Figure 6. A chromatin-centered survey of Npl3 genetic interactions.

Summary of chromatin and transcription factors that exhibit genetic interactions with a deletion of *NPL3*. Colored ovals represent subunits identified in the SGA screens or by directed genetics; red indicates a suppressive (positive) interaction and green indicates a synthetic (negative) interaction. Outlined ovals refer to the complex that individual subunits belong to. K123Ub refers to the *htb1K123R* point mutant. K36me refers the *hht1K36A* point mutant. Grey or white indicates the genetic interaction was not tested. Physical interactions tested by co-IP are Npl3:Bre1 (Figure 3) and Npl3:U1[24]. Bold rectangle indicates factors shown in this paper and Kress *et al.*[24] to promote splicing; whether the presence of Npl3 can influence local H2B ubiquitination levels or dynamics remains unresolved. The PolII C-terminal domain is drawn in grey.

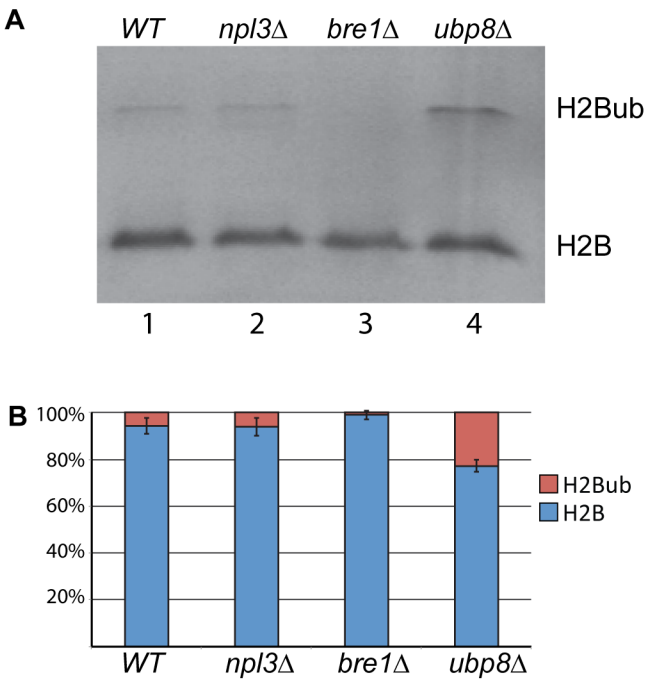


Figure S1. *NPL3* does not affect global H2B ubiquitination levels.

(A.) Western blot analysis of histone H2B ubiquitination levels. Whole cell extracts from the indicated strains were subjected to electrophoresis to separate the ubiquitinated H2B from unmodified H2B, followed by Western blotting using α-H2B antibody. Shown is a representative blot.

(B.) Quantitation of H2B ubiquitination levels. Shown are the average percentages of ubiquitinated H2B from the indicated strains. Error bars represent standard deviation of three biological replicates.

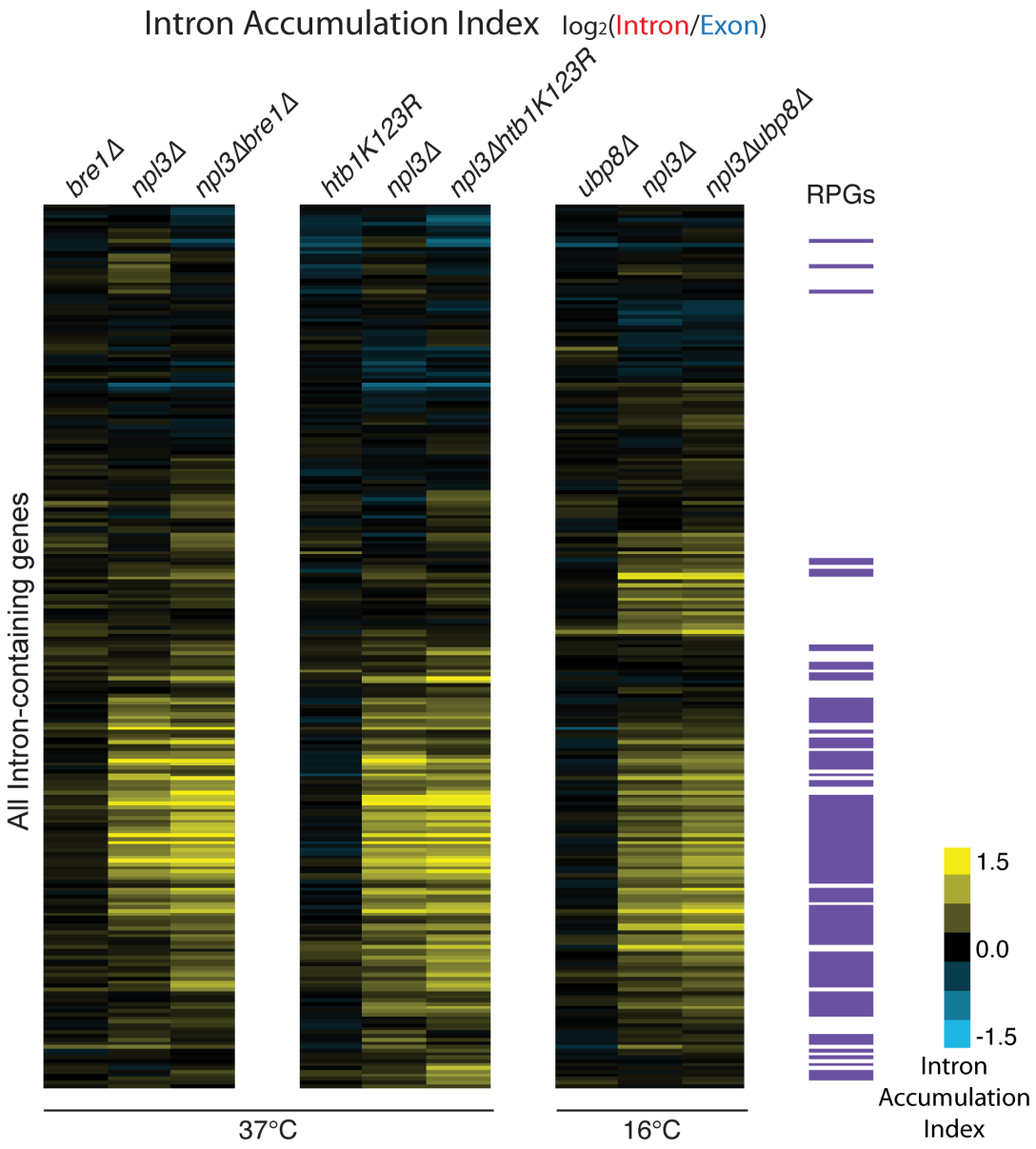


Figure S2. Heat map representation of the Intron Accumulation Indices used to generate histogram in Figure 4.

Shown are $\log_2$-based Intron Accumulation Index scores for each intron-containing gene, generated by normalizing fold intron changes to fold exon changes (see Materials and Methods for details). Genotype of each strain measured is listed above the heat maps. Transcripts that encode the ribosomal protein genes (RPGs) are highlighted in purple to the right of the heat maps. Gene order along the y-axis is the same for all genotypes.

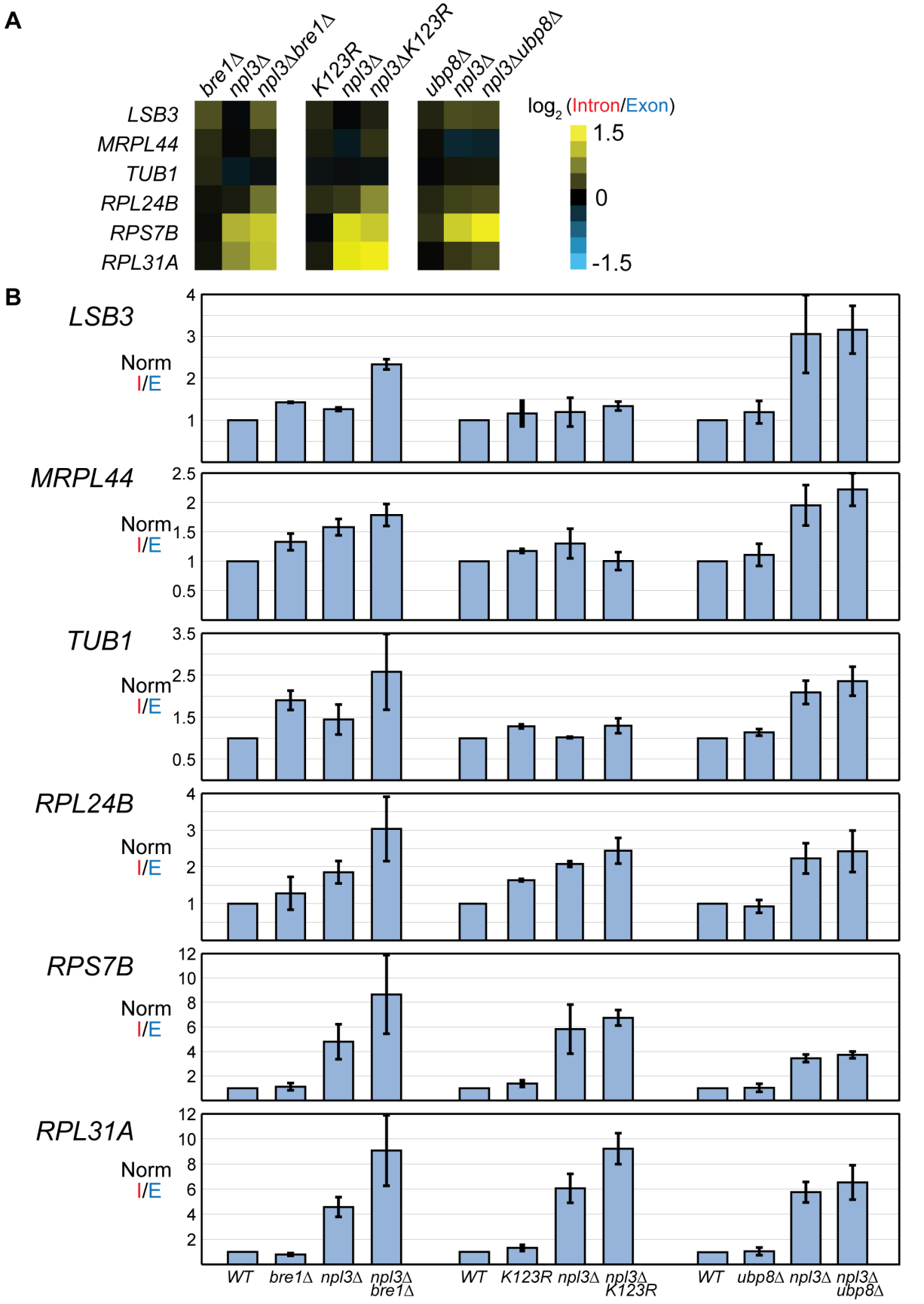

A
bre1Δ
npl3Δ
npl3Δbre1Δ
K123R
npl3Δ
npl3ΔK123R
ubp8Δ
npl3Δ
npl3Δubp8Δ
LSB3
MRPL44
TUB1
RPL24B
RPS7B
RPL31A
log2 (Intron/Exon)
1.5
0
-1.5
B
LSB3
MRPL44
TUB1
RPL24B
RPS7B
RPL31A
Norm I/E
WT
bre1Δ
npl3Δ
npl3Δ bre1Δ
WT
K123R
npl3Δ
npl3Δ K123R
WT
ubp8Δ
npl3Δ
npl3Δ ubp8Δ

Figure S3. qPCR validation of the microarrays.

(A.) Shown are $\log_2$-based Intron Accumulation Index scores for the indicated genes, as reproduced from Figure S2. The genotype of each strain is listed above the heat map. *K123R* refers to the *htb1K123R* allele.

(B.) RT-qPCR measurements of un-spliced mRNAs using single-locus RT-qPCR. Percent un-spliced RNA was calculated for each mutant and is represented as fold change compared to wild-type. Capped error bars represent standard deviation of biological replicates; uncapped error bars represent standard deviation of qPCR replicates.

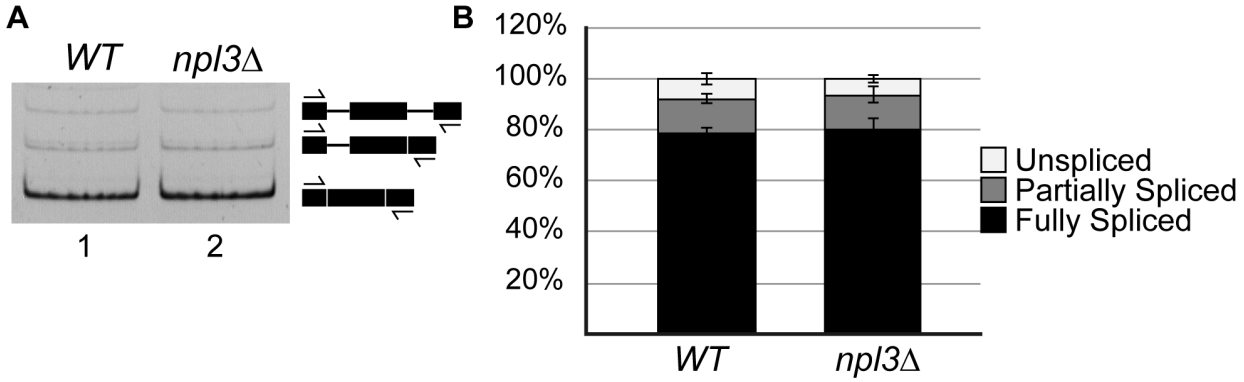


Figure S4. *SUS1* splicing is not sensitive to deletion of *NPL3*.

(A.) Analysis of *SUS1* splicing using quantitative PCR on cDNA generated from the indicated strains. To the right of the gel, is a schematic of the *SUS1* gene and the predicted mobility of its un-spliced, partially spliced, and fully spliced isoforms. Arrows indicate position of primers used in the PCR. Shown is a representative gel.

(B.) Quantitation of *SUS1* PCR. To obtain the data shown in panel B, we took the average of two technical replicates generated from two separate biological samples. Error bars represent standard deviation of all replicates.

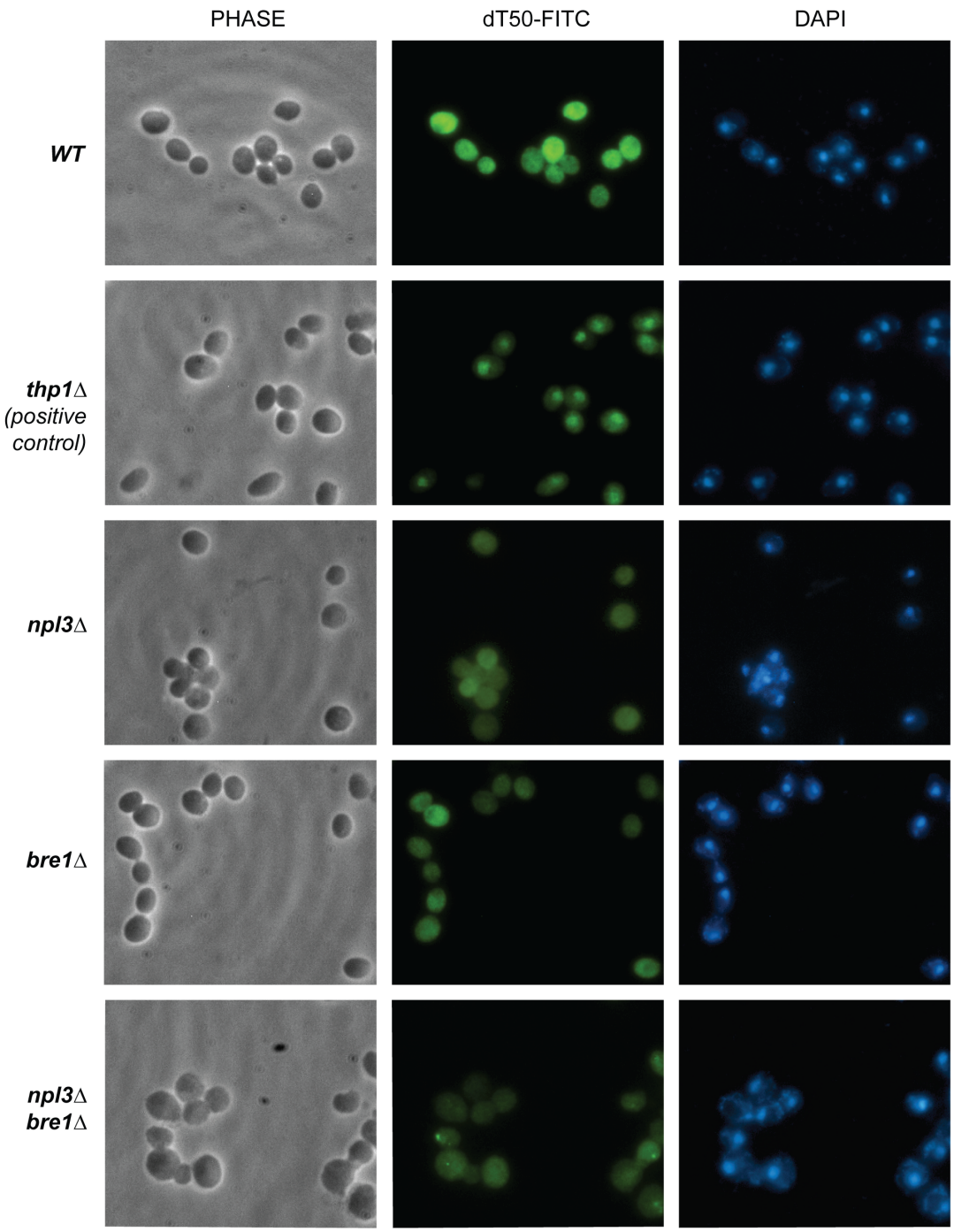


Figure S5. *NPL3* and *BRE1* do not affect bulk mRNA export.

In situ hybridization of poly-dT oligo shows whole cell localization of mRNAs in wild-type, *npl3*Δ, *bre1*Δ, *and npl3*Δ*bre1*Δ strains. The positive control *thp1*Δ strain is included to show nuclear localization coincident with DAPI (nuclear) staining.

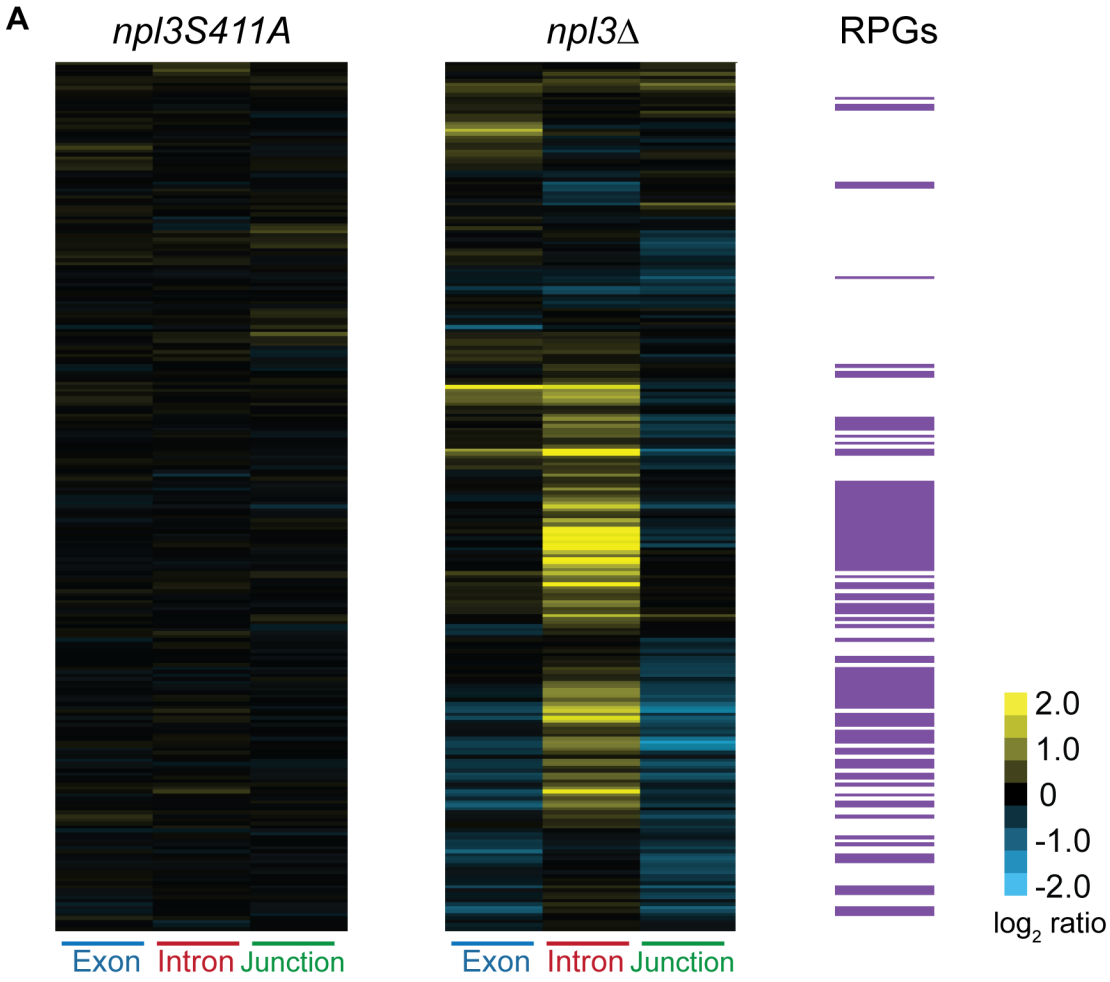

A
npl3S411A
npl3Δ
RPGs
Exon
Intron
Junction
Exon
Intron
Junction
2.0
1.0
0
-1.0
-2.0
log2 ratio
B
120
100
80
60
40
20
0
number of genes
log2 (Intron/Exon)
2.0
1.5
1.0
0.5
0.3
-0.3
-0.5
-1.0
-1.5
-2.0
npl3S411A
npl3Δ

Figure S6. Spicing is not sensitive to mutation of the Npl3 phosphorylation site.

(A.) Splicing profiles of the *npl3S411A* and *npl3*Δ strains grown at 30˚C. The heat map shows the log_2-ratio for each gene feature of the indicated strain compared to wild-type. Gene order along the y-axis is the same for all arrays. Transcripts that encode the ribosomal protein genes (RPGs) are highlighted in purple to the right of the heat maps.

(B.) Histogram shows the number of genes with a log_2-based Intron Accumulation Index score greater than 0.3 for the *npl3S411A* and *npl3*Δ strains. Heat map within histogram bars shows distribution of the severity of splicing defect.

TABLES

Tables are available online at www.plosgenetics.org

Table S1. Genes whose deletion caused severe synthetic sickness or lethality with *npl3*Δ.

List of genes whose deletion caused severe synthetic sickness or lethality with *npl3*Δ at 30˚C or 37˚C. Included are the systematic and common names, the synthetic genetic array in which each strain was identified, and the putative function as annotated by SGD (yeastgenome.org). Phenotype of the genetic interactions validated using directed genetics is listed under "notes". Any strains added using directed genetics are listed at the bottom of the table. Genes are sorted first by screen in which they were identified, and then alphabetically by common name.

Table S2. GO-term analysis by Biological Process.

Results of GO-term analysis by "Biological Process" with the associated p-values (pre- and post-correction). Biological Process definitions are available in Table S7.

Table S3. Protein Complex enrichment analysis.

Results of protein complex enrichment analysis with the associated p-values (pre- and post-correction for multiple testing). Significantly enriched complexes in bold are shown in Figure 1B using their more common names. Protein complex definitions were obtained from Pu *et al.*[115] with the addition of the RAD6 Complex containing *RAD6*, *BRE1* and *LGE1*, which was not annotated at the time.

Table S4. Genes whose deletion allowed growth in an *npl3*Δ strain.

List of genes whose deletion allowed growth in an *npl3*Δ strain in the 16˚C synthetic genetic array. Included are the systematic and common names, as well as the time-point at which growth was visualized, and the putative function of each gene. Any strains validated using directed genetics (by tetrad analyses or random sporulation) are indicated. Genes are sorted first by interval of time necessary to observe growth, and then alphabetically by common name.

Table S5. Averaged microarray results used to generate heat map in Figure 4.

Listed are $\log_2$ ratios of the intensities for each gene feature for the indicated strains compared to wild-type. "_2" refers to the second intron. Also included are the Intron Accumulation Index scores for each gene used to generate the heat map in Figure S2.

Table S6. Strains used in this manuscript, and more detailed methods for their generation.

Table S7. Biological Process definitions used to categorize *npl3*Δ genetic interactions.

Table S8. Primers used in the qPCR assay.

Chapter 2

An Epigenetic Landscape of Histone Marks Modulates Gene-Specific Spliceosome Assembly and Function

An Epigenetic Landscape of Histone Marks Modulates Gene-Specific Spliceosome Assembly and Function

Lucas Hérissant[1,*], Erica A. Moehle[2,*], Diego Bertaccini[3], Alain Van Dorsselaer[3], Christine Schaeffer-Reiss[3], Christine Guthrie[2] and Catherine Dargemont[1]

[1] Institut Jacques Monod, University Paris Diderot, Sorbonne Paris Cité, CNRS UMR7592, Equipe labellisée Ligue contre le cancer, 15 rue Hélène Brion, 75205 Paris, France. [2]Department of Biochemistry and Biophysics, University of California San Francisco, USA. [3]Institut Plurisdisciplinaire Hubert Curien, University of Strasbourg, CNRS, Strasbourg, France.

To whom correspondence should be addressed: Catherine Dargemont, dargemont.catherine@ijm.univ-paris-diderot.fr

*These authors contributed equally to this work.

ABSTRACT

Background information

Commitment to splicing occurs co-transcriptionally, but a major unanswered question is the extent to which various modifications of chromatin, the template for transcription *in vivo*, contribute to the regulation of splicing.

Results

Here we perform genome-wide analyses showing that inhibition of specific marks – H2B ubiquitination, H3K4 methylation, and H3K36 methylation – perturbs splicing in budding yeast, with each modification exerting gene-specific effects. Furthermore, semi-quantitative mass spectrometry on purified nuclear mRNPs and chromatin immunoprecipitation analysis on intron-containing genes indicated that H2B ubiquitination, but not Set1-, Set2-, or Dot1-dependent H3 methylation, stimulates recruitment of the early splicing factors, namely U1 and U2 snRNPs, onto nascent RNAs.

Conclusions

These results suggest that histone modifications impact splicing of distinct subsets of genes using distinct pathways.

INTRODUCTION

Transcriptional control of gene expression has long been thought to require the coordinated modification of histones[124] but recent evidence suggests an additional role for these modifications in controlling the eventual fate of an mRNA after it is transcribed.[125-128] A number of correlative studies have shown that nucleosomes at DNA encoding exons and introns bear distinct covalent modifications in metazoa (reviewed in Oesterreich *et al.*[125]) and yeast.[104] These observations raise the possibility that epigenetic marks may have a widespread role in providing direct regulatory input into co-transcriptional splicing decisions.

In *Saccharomyces cerevisiae*, splicing occurs co-transcriptionally for the vast majority of intron-containing genes,[34] and consistently, splicing factors are recruited to nascent transcripts.[8-11,129] Interestingly, the histone acetyltransferase catalytic subunit of the SAGA complex, Gcn5, has been shown to control the co-transcriptional recruitment of the U2 snRNP.[15] However, the extent to which additional histone modifications of the chromatin landscape regulate the co-transcriptional recruitment of the spliceosome is still unclear.

Transcription-associated histone H2B mono-ubiquitination (Ub-H2B) and the downstream histone H3 methylation events have established roles in transcription activation and nucleosome dynamics.[48,71,73,79] In addition, we recently showed that Ub-H2B influences export of mRNPs by promoting the recruitment of the nuclear export machinery to nascent transcripts.[106,130] Furthermore, we recently reported genetic and functional interactions between the Ub-H2B machinery and the SR-like spliceosome-associated factor Npl3,[26] suggesting that the role of Ub-H2B in gene expression is not limited to directing transcription itself. This result prompted us to determine the contribution of transcription-dependent chromatin marks, and in particular Ub-H2B, on spliceosome assembly and function on nascent transcripts.

RESULTS

Defects in Ub-H2B, H3K4me, H3K36me cause introns to accumulate for distinct subsets of transcripts

Transcripts in *S. cerevisiae* do not generally undergo alternative splicing, but the constitutive splicing reaction is sensitive to a number of environmental perturbations.[3-5] While relatively few genes are spliced, intron-containing genes account for nearly one third of total cellular transcription,[131] so it is critical for yeast to appropriately control the efficiency of this step in gene expression. We recently reported that in a genetic background sensitized by loss of Npl3, a protein

known to promote splicing of a subset of genes, a short 37˚C temperature shift revealed a modest dependence of pre-mRNA splicing on Ub-H2B.[26]

Here, we further explored the potential connection between chromatin modification and splicing by capitalizing on the observation that nuclear export factor assembly onto nascent mRNPs is very tightly regulated by Ub-H2B during a 3-hour shift to 39˚C, an experimental condition that challenges mRNA biogenesis without affecting genome-wide expression.[106,130] Indeed, using splicing-sensitive microarrays (Figure 1A), we see that at 39˚C, abrogating Ub-H2B by deleting the H2B E3 ligase, *BRE1*, or mutating the targeted residue in H2B (*htb1K123R)* led to increases in the levels of intron for many genes, consistent with a defect in the splicing of those transcripts (Figure 1B and 2C, Table S3). To more easily compare these datasets, we calculated intron/exon ratios, an established approach to normalize for differences in transcription.[96] We observed that, importantly, genes affected by H2B mutation extensively overlap with genes affected by *BRE1* deletion (Figure 1B and 2D). While the ribosomal protein genes (RPGs) are a category of spliced genes often regulated together,[4,5] Ub-H2B-dependent effects on splicing were not enriched for RPG transcripts. Validation of the microarray data by using RT-qPCR to measure relative intron and exon abundance of several transcripts confirmed UB-H2B-mediated changes with respect to a wild-type strain (Figure 1C). Taken together, our data show that loss of Ub-H2B has clear gene-specific effects on intron accumulation and, thus, prompt the conclusion that Bre1-dependent Ub-H2B is important for splicing at 39˚C.

Ub-H2B is strictly required for other histone marks such as the trimethylation of histone H3 on both lysine 4 by the Set1-containing COMPASS complex[79] and lysine 79 by Dot1,[83,84] and facilitates the Set2-mediated methylation of H3K36 on some intron-containing genes.[104] Surprisingly, we found that Set2 promotes the splicing of many genes (Figure 2A and 2C), but the observation that 83% of Set2-dependent genes are not also dependent on Ub-H2B (Figure 2D) suggests Set2 is working separately from the Ub-H2B pathway. Microarray results from a strain lacking *SET1* were

consistent with a mild splicing defect as gauged both by intron accumulation and intron/exon ratios (Figure 2B and 2C). However, only a small number of genes (13) overlap with those affected in *htb1K123R* (Figure 2D), indicating that the effects of Ub-H2B on intron/exon ratios were not strictly mediated by H3K4 methylation. We observed a comparatively larger effect from *BRE1* deletion than mutation of the H2B target residue, which is consistent with an additional Bre1 target or function that also promotes splicing. Bre1 targets many of the same genes as COMPASS component Set1, but whether ubiquitination of the Swd2 component of the COMPASS complex by Bre1 might be involved in this process remains to be determined.[73]

Importantly, no global changes of gene expression were observed upon inhibition of Ub-H2B or downstream H3 methylations either at 30˚C[132,133] or 39˚C.[106] Only a minority of genes exhibited altered expression in the different mutant strains, none of which encoded components of the splicing machinery. This argues against direct transcriptional control of the splicing machinery expression by Ub-H2B.

Together, these results suggest that splicing efficiencies – inferred by changes in pre-mRNA and total mRNA levels – are dependent on contributions from multiple transcription-coupled histone marks, with the relative contribution being different from one intron-containing gene to another. We reasoned that the decrease in splicing efficiency we observed was unlikely to be caused by a wholesale block in spliceosome function, but rather could relate to a delay in the onset of the splicing reaction. We therefore sought to determine whether Ub-H2B might influence the ability of splicing factors to associate with transcripts.

Preventing ubiquitination of H2B alters the recruitment of early splicing factors to Cap-binding complex-associated mRNPs

Since Ub-H2B-mediated splicing is likely mechanistically separated from that driven by downstream histone methylations, what step in splicing is impacted by loss of Ub-H2B? To address this question, nuclear mRNPs were purified from temperature-shifted wild-type (WT) and

htb1K123R cells using a genomically TAP-tagged Cbc2 of the nuclear cap-binding complex (CBC), and their proteomes were analyzed by tandem mass spectrometry.[134] Because the CBC is associated with nuclear mRNPs from early synthesis to nuclear exit, the composition of purified mRNPs reflects the sum of all biogenesis events, including transcription, mRNA processing and mRNA packaging that lead to their formation.[134] The proteome of CBC-associated mRNPs was enriched in splicing factors (snRNPs), 3' end processing machinery, mRNA export factors and other factors that are recruited during transcription elongation such as the THO complex (Table S4).[134] A recent transcriptome-wide analysis of CBC-associated mRNAs reveals a clear enrichment of unspliced versus spliced mRNA, consistent with an interaction occurring at an early step of transcription.[135] In addition, multiple subunits of RNA Polymerase II (RNAPII) were detected in CBC-interacting mRNPs, confirming that some nascent transcripts were associated with TAP-tagged Cbc2 (Table S4).

This global approach showed that preventing Ub-H2B impaired the association of the nuclear export machinery as previously described,[106] but maintained WT levels of other factors known to bind mRNAs, including the transcription elongation THO complex (Figure 3A, Table S4). Early spliceosome assembly onto pre-mRNAs entails binding of the U1 and U2 snRNPs: consistent with the observed splicing defect, the number of peptides corresponding to U1 and U2 was significantly lower in the *htb1K123R* strain compared to WT, while the overall protein level of these factors was similar in both strains (Figure 3A and 3C). This defect was also seen using a semi-quantitative analysis based on a spectral counting approach[136] (Figure 3B, Tables S5 and S6). To confirm this decrease in CBC-associated U1 and U2, we directly assayed the co-immunoprecipitation of an endogenously HA-tagged version of either Prp42 (U1 snRNP) or Lea1 (U2 snRNP) with TAP-tagged Cbc2 and observed a reproducible decrease in both Prp42-HA and Lea1-HA (Figure 3C). While we cannot rule out that some factors pulled down by TAP-tagged Cbc2 might be directly bound to the CBC, as has been observed for the tri-snRNP in mammals,[137] the decreased association of U1 and U2 proteins in the pull-down reported here is consistent with the splicing defect seen in strains lacking

Ub-H2B. In contrast, this defect was not phenocopied by deletion of *SET1* (Figure 4), further arguing that the control of U1 and U2 recruitment by Ub-H2B is not mediated by downstream methylation by Set1.

Loss of H2B ubiquitination impairs recruitment of early splicing factors to transcribing genes.

It has been suggested that splicing activity can be facilitated by the co-transcriptional recruitment of the splicing machinery;[129,138] thus, we reasoned that if the defect seen in recruitment of U1 and U2 to mRNPs occurs co-transcriptionally, this may further account for the splicing defect seen in a strain lacking Ub-H2B. To ask this, we used chromatin immunoprecipitation (ChIP) on four intron-containing genes whose splicing is sensitive to loss of Ub-H2B: 3 ribosomal protein genes *RPS21B*, *RPL14B* and *RPL34B* and the non-ribosomal protein gene *HNT1* (Figure 5A and Table S3). Of note, *HNT1* splicing was also sensitive to deletion of *SET1* or *SET2* (Table S3). We found a marked decrease in Prp42 association at these genes in the *htb1K123R* strain. Importantly, this decrease in U1 association occurred at genes where RNAPII levels (Figure 5B) and resulting mRNA expression remained unchanged (data not shown), and does not occur in cells lacking the *SET1*, *SET2* and *DOT1* methyltransferases (Figure 5C). Furthermore, in agreement with data from the CBC pulldown (Figure 3), Lea1 exhibited weaker association with these genes in the *htb1K123R* strain, which was revealed primarily at the 3' ends (Figure 5B). Our data reveal that loss of Ub-H2B adversely affects recruitment of the early splicing machinery to nascent transcripts.

Loss of H2B ubiquitination marginally affects recruitment of Npl3

We previously showed that the SR protein Npl3 promotes U1 and U2 association with nascent transcripts and physically interacts with the Ub-H2B machinery.[24,26] While we do see a modest decrease in Npl3 ChIP (Figure 6A) at genes whose splicing is promoted by Npl3,[24] this cannot account for the broad consequences of losing Ub-H2B on pre-mRNA splicing shown here, in particular on non-RPGs.

Similar to SR proteins in metazoa, Npl3 is known to associate with the C-terminal domain (CTD) of RNAPII upon Serine 2 phosphorylation,[25] but we only observed a weak decrease in Serine 2 phosphorylation on the RNAPII CTD at *RPS21B* and *RPL34B* in the *htb1K123R* strain (Figure 6B).

DISCUSSION

Early splicing factors associate with nascent pre-mRNAs,[8-11,129] and, therefore, it is critical to understand how the chromatin landscape contributes to co-transcriptional spliceosome assembly and function. Our results indicate that in budding yeast, an organism with relatively simple intron/exon architecture and limited splice site variation, multiple histone modifications (Ub-H2B, H3K4me and H3K36me) promote spliceosome function at distinct subsets of genes. In agreement with this, we also show that Ub-H2B promotes the association of the early splicing machinery to mRNPs while the RNA is still being transcribed, in a molecular pathway that is distinct from Set1- or Set2-mediated splicing. Thus, our data support the idea that the chromatin landscape of a locus, as defined by these modifications, can impact the fate of a transcript, reflecting the capacity of the spliceosome to interpret and integrate multiple inputs from the chromatin landscape. This idea is further supported by evidence that histone acetylation by Gcn5 and deacetylation by Hos2/3,[15,105] as well as incorporation of the variant histone H2A.Z[18] can also influence spliceosome assembly onto nascent transcripts and splicing efficiency of those transcripts.

In metazoa, where the recognition of and discrimination between alternative, degenerate splice sites must be tightly regulated, histone modifications, including Ub-H2B, have been shown to reflect intron/exon structure.[139-144] In fact, modulation of the levels of the Ub-H2B machinery in human cells has revealed that this mark promotes recognition of splice sites, and importantly, as we show here in budding yeast, does so in context-dependent ways (Figures 1 and 2).[144,145] While Ub-H2B also reflects intron/exon structure on budding yeast genes,[104] it is remarkable that the one to two differentially marked nucleosomes associated with yeast introns, typically 100-400bp in budding

yeast,[146] can impact splicing. The conservation of the connection between Ub-H2B and splicing from yeast to humans suggests that this is a universal strategy for spliceosome regulation.

How does Ub-H2B promote splicing factor association? It is possible that Ub-H2B-mediated spliceosome association is, in part, influenced by a minor impairment in the recruitment of the CBC to some genes (data not shown), as the CBC and early splicing factors physically interact in yeast.[147] In addition, it has been recently shown that CBC depletion in mammalian cells led to defects in co-transcriptional spliceosome assembly via both RNA-dependent interaction with U1 and U2 snRNPs and RNA-independent interactions with U4/U6 and U5 snRNPs.[137] We could barely detect components of the U4/U6 and U5 snRNPs in our yeast CBC proteome, suggesting that this interaction may not occur in yeast. However, Pabis *et al.* hypothesized that CBC, U1 and U2 bind cooperatively onto RNA to promote splicing efficiency,[137] a process that could be influenced by Ub-H2B. Conversely, Ub-H2B and H3K36 methylation levels are highly dependent on an intact CBC,[17,148] highlighting two examples of the high degree of coupling between mRNA processing and chromatin structure.

In metazoans, it has been suggested that histone modifications may directly recruit splicing factors via specific histone mark-specific adaptors;[99,149] this has been previously proposed to occur in yeast as well.[105] Thus, it is possible that ubiquitinated H2B also directly or indirectly recruits a splicing factor in budding yeast. Alternatively, as histone modifications can influence RNAPII dynamics, it is possible that changes in elongation speed may explain the splicing defects seen in the chromatin mutants tested here.[29,31,150] Therefore, it is interesting that multiple reports have suggested that the ubiquitin modification stabilizes nucleosomes during elongation;[151] given the splicing defects that can occur when RNA polymerase elongates too quickly,[31] perhaps uncontrolled RNAPII elongation in the *htb1K123R* strain accounts for the decrease in splicing efficiency reported here.

An intriguing aspect of chromatin-based splicing regulation is the complexity of the chromatin template; as our microarrays revealed, Ub-H2B, H3K4me and H3K36me influence

splicing at somewhat overlapping subsets of genes. As the field moves forward, a challenge will be to understand how the spliceosome integrates the signals from individual histone modifications to achieve the appropriate splicing outcomes for the needs of the cell. The many histone post-translational modifications, and their combined effects on transcription, RNAPII CTD phosphorylation state, and mRNA processing perhaps provide each emerging transcript with a specific mRNA "identity" that allows dynamic modulation of splicing, mRNA export and quality control.

MATERIALS AND METHODS

Yeast strains and culture.

Strains used in this study are listed in supplementary Table 1. The derivative strains were obtained using PCR based homologous recombination as described in Longtine *et al*.[152] Yeast cells were grown overnight at 30˚C in YPD medium (Yeast extract, peptone, dextrose). To perform the 3-hour shift at 39˚C, when the 30˚C overnight culture reached OD_{600}=1, one volume of medium preheated to 48˚C was added and cultures were incubated at 39˚C for 3 hours.

Microarray analysis

Strains were grown for microarray analysis as described above with a 3-hour shift to 39˚C, at which point they were collected by centrifugation. Microarrays were performed as in Moehle *et al.*[26] For each microarray shown, results from two biological replicates were averaged. In addition, each biological replicate contains 6 technical replicates per probe, as well as dye-flipped replicates. The heat maps in Figures 1 and 2 were created using Java Treeview[121] and show the $\log_2$-based fold change in the indicated strain as compared to an isogenic WT. Intron/Exon ratios were calculated for each intron-containing gene ($\log_2$ (Intron/Exon) = $\log_2$ ($Intron_{mutant}/Intron_{WT}$) – $\log_2$ ($Exon_{mutant}/Exon_{WT}$)); these values were converted into a histogram for any gene with Intron/Exon value <-0.3 or >0.3.

RNA Isolation and RT-qPCR

Total RNA isolation was performed by the hot acid phenol method (Sigma Aldrich). cDNA from total RNAs were obtained by reverse-transcription with random oligonucleotides (Roche) using the SuperScriptTM II reverse Transcriptase (Invitrogen). Real time qPCR was then performed using the SYBR Green mix (Roche) and the Light Cycler 480 system (Roche) with gene-specific primers described in Table S2.

Antibodies

Commercial antibodies used in this study were anti-RNA Polymerase II (RNAPII ChIP 12 µg, 8WG16, MMS126R, Covance), anti-phosphorylated Ser 2 of RNAPII CTD (ChIP 12 µg, clone 3E10, 04-1571, Millipore), anti-HA (ChIP 8 µg, WB 1:2000, clone HA-11, MMS-101R, Covance), anti-H3 (ChIP 4 µg, ab1791, abcam), anti-H3K36me3 (ChIP 4 µg, ab9050, abcam). Polyclonal anti-Mex67 (WB 1:20,000) and anti-Npl3 (ChIP 1,5 µL, WB 1:10,000) antibodies were previously described.[90,153] TAP-tagged proteins were immunoprecipitated using IgG sepharose beads (50 µL for ChIP analysis). Western blot analyses were performed using appropriate HRP-coupled secondary antibodies and chemi-luminescence protein immunoblotting reagents (Pierce).

Immunoaffinity Purification of nuclear mRNPs from frozen cell grindate

Cells shifted at 39˚C for 3 hours in YPD were rapidly frozen in liquid nitrogen before cryolysis.[154] Immunoaffinity purification of mRNPs was performed as described in Oeffinger *et al.*[134] with minor modification. Frozen cell grindates were rapidly thawed into nine volumes of RNP buffer (20mM Hepes, pH7.4, 110mM KOAc, 0.5%Triton, 0.1% Tween 20, 1:100 Solution P, 1:5000 RNasin (Promega), 1:5000 Antifoam B (Sigma); 1:1000 DTT). The resulting extracts were centrifuged 5 min at low speed to eliminate the large cellular debris before clarification by filtration. Immunoprecipitation was performed using magnetic beads (Dynal) conjugated with rabbit IgG (Sigma).[155] Resulting eluates were lyophilized, resuspended in SDS-PAGE sample buffer, separated by SDS-PAGE on a 4–12% NuPAGE Novex Bis-Tris precast gel (Invitrogen) according to the

manufacturer's specifications and visualized by Coomassie blue staining. Alternatively, nuclear mRNPs were purified from glass-beads lyzed cells as previously described.[106,156]

SDS-PAGE and in-gel digestion

After electrophoresis, gels were stained by colloidal Coomassie Blue (BioSafe coomassie stain; Bio-Rad) and whole lanes were cut into 14 5x4 mm slices using a disposable grid-cutter (the gel-company, Tübingen, Germany). Slices were divided into three and in-gel digestion using trypsin (Promega, Madison, WI) was performed overnight at 37˚C, after in-gel reduction and alkylation using the MassPrep Station (Waters, Milford, MA, USA). Tryptic peptides were extracted using 60% ACN in 0.1% formic acid for 1h at room temperature. The volume was reduced in a vacuum centrifuge and adjusted to 10 µl using 0.1% formic acid in water before nanoLC-MS/MS (nanoliquid chromatography coupled to tandem mass spectrometry) analysis (see supplemental information).

Liquid chromatography/Mass spectrometry

NanoLC-MS/MS was performed using a nanoACQUITY ultra performance liquid chromatography (UPLC®) system (Waters, Milford, MA, USA) coupled to a maXis 4G Q-TOF mass spectrometer (BrukerDaltonics, Bremen, Germany). Detailed procedure is presented in supplemental information.

Peptide counts

The peptide count of unique peptides was performed using Scaffold 3 software (version 3.6.5; Proteome Software Inc., Portland, OR, USA) after having filtered the results at 1% < FDR.

Spectral counts

The label-free semi-quantitation was performed using Scaffold 3 software (version 3.6.5; Proteome Software Inc., Portland, OR, USA) exporting the un-weighed spectral count into an excel file after having filtered the result at 1% < FDR as described in the supplementary data. The three replicates of each biological sample shown in Table S6 (except the negative control) were horizontally normalized to the highest spectral count value obtained for nuclear cap-binding protein

complex subunit (WT injection 1: 429 un-weighed spectra counts). The following procedure as described in Gokce *et al.*[157] and Miguet *et al.*[158] has been applied for the six independent injections: the number of spectra obtained for the nuclear cap-binding complex has been divided by 429 and the ratio obtained was multiplied independently by the number of spectral counts for each protein, in order to reduce the variance observed between samples and replicates.

Chromatin Immunoprecipitation (ChIP) and qPCR

ChIPs were performed as described previously.[159] 8 OD units of cell lysate were immunoprecipitated with the amount of antibodies mentioned above. Immunoprecipitated DNA was analyzed by quantitative PCR using primers referenced in Table S2. Non-specific signals were assessed by analyzing immunoprecipitated DNA with primers against an intergenic region. The specificity of the signal was also evaluated using untagged strains. The resulting amplifications were similar to those observed for the intergenic region (data not shown). Results correspond to the mean of at least 3 independent experiments.

ACKNOWLEDGEMENTS

We are grateful to Thanasis Margaritis and Frank Hostege for their help in the compilation of published microarray data. This study was supported by the National Research Agency (grant 2010 BLAN1227-01 to CD), the Who am I? laboratory of excellence (ANR-11-LABX-0071 to CD) funded by the "Investments for the Future" program (ANR-11-IDEX-0005-01), the NIH grant NIHGM21119 to CG, the Proteomic French Infrastructure (ProFI to AVD and CSR) and the Fondation pour la Recherche Médicale (to AVD and CSR). CG is an American Cancer Society Research Professor of Molecular Genetics. LH is supported by the University Paris Descartes.

AUTHOR CONTRIBUTIONS

LH, EAM, and DB performed experiments; LH, EAM, DB, AVD, CS, CG, and CD designed the experiments and analyzed the data; CD, EAM, and CG wrote the manuscript.

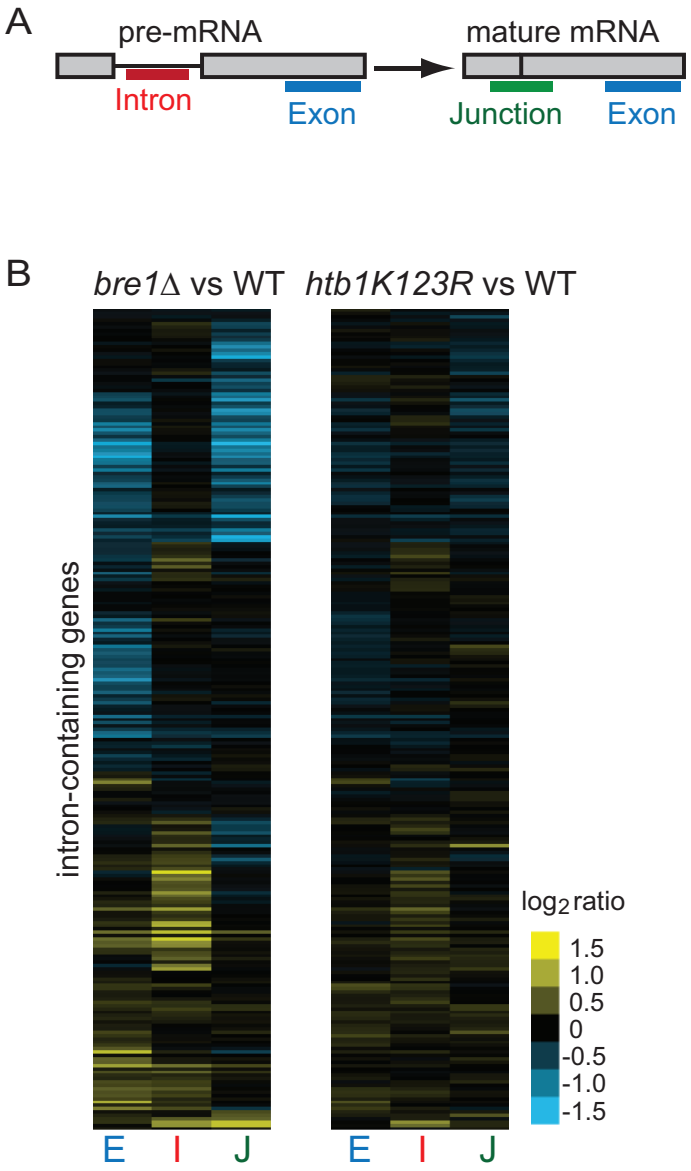
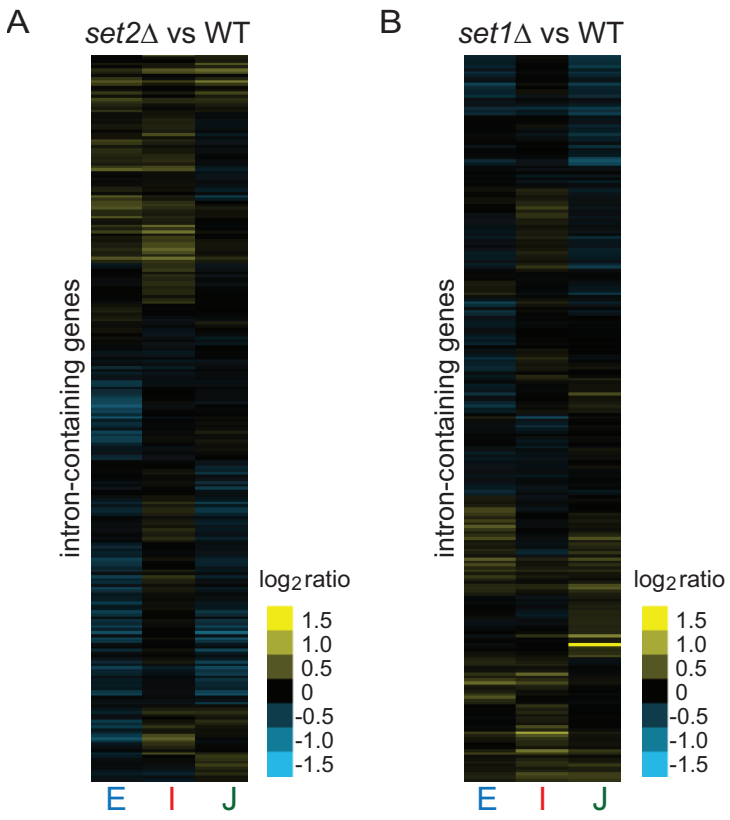

A
pre-mRNA
mature mRNA
Intron
Exon
Junction
Exon
B
bre1Δ vs WT
htb1K123R vs WT
intron-containing genes
log2 ratio
1.5
1.0
0.5
0
-0.5
-1.0
-1.5
E
I
J
E
I
J
C
Ratio I/E in htb1K123R mutant relative to WT
3,0
2,5
2,0
1,5
1,0
0,5
0,0
RPS21B
RPL14B
RPL34B
HNT1

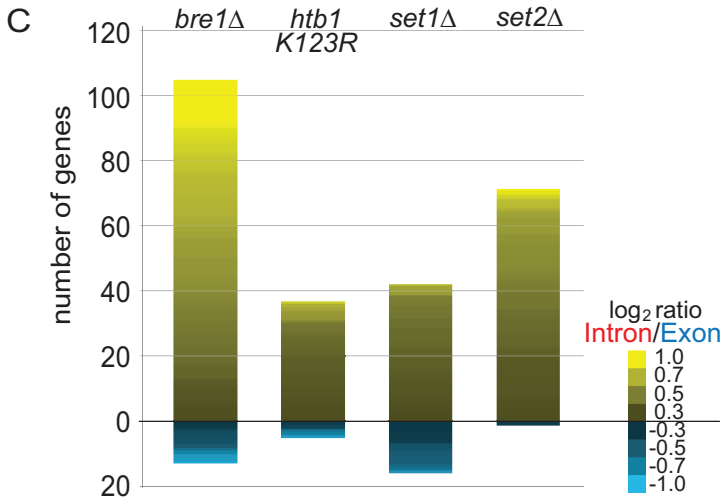
Figure 1. Defects in Ub-H2B cause splicing defects.

(A.) Schematic of probes contained on the microarray for each intron-containing gene.

(B.) Heat maps of $\log_2$ ratios of each gene feature in *bre1*Δ and *htb1K123R* strains compared to isogenic WT strains after a 3-hour shift to 39˚C.

(C.) RT-qPCR measurements of un-spliced mRNAs using single-locus RT-qPCR. Percent unspliced RNA is represented as fold change compared to WT.


A
set2Δ vs WT
intron-containing genes
log2 ratio
1.5
1.0
0.5
0
-0.5
-1.0
-1.5
E
I
J
B
set1Δ vs WT
intron-containing genes
log2 ratio
1.5
1.0
0.5
0
-0.5
-1.0
-1.5
E
I
J


C
bre1Δ
htb1 K123R
set1Δ
set2Δ
number of genes
120
100
80
60
40
20
0
20
log2 ratio
Intron/Exon
1.0
0.7
0.5
0.3
-0.3
-0.5
-0.7
-1.0

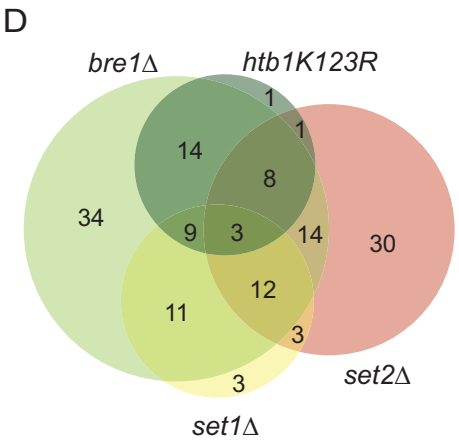

D
bre1Δ
htb1K123R
set2Δ
set1Δ
34
14
1
1
8
9
3
14
30
12
11
3
3

Figure 2. Defects in Ub-H2B, H3K4me, H3K36me cause introns to accumulate for distinct subsets of transcripts.

(A.-B.) Heat maps of $\log_2$ ratios of each gene feature in *set2*Δ or *set1*Δ strains compared to isogenic WT strains after a 3-hour shift to 39˚C. Gene order is different for (A.) and (B.).

(C.) Histogram of number of genes exhibiting $\log_2$(Intron/Exon) ratio greater than 0.3 or less than -0.3. Heat map within bar shows degree of splicing change of those genes.

(D.) Venn diagram of genes from histograms for comparison of each genotype directly. Note: *set1*Δ and *htb1K123R* have 1 gene overlap which cannot be shown. The circle sizes are representative of the numbers of genes with $\log_2$I/E>0.3 but the overlaps are not to scale.

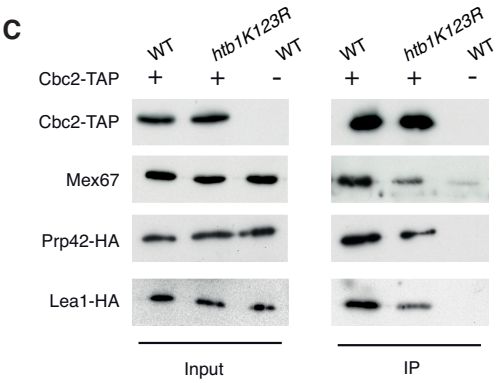
A

Complex	Protein	number of peptides WT	*htb1K123R*
Cap Binding Complex	Cbc2	11	11
	Sto1	49	47
Export Machinery	Mex67	13	10
	Mtr2	2	0
	Nab2	13	5
THO Complex	Hpr1	11	11
	Tho2	44	43
	Thp2	4	4
	Mft1	1	5
U1 snRNP	Prp42	19	10
	Prp40	16	7
	Prp39	16	13
	Luc7	5	3
	Mud1	11	1
	Nam8	5	2
	Snu56	10	2
	Snu71	15	10
	Yhc1	6	2
	Prp5	2	0
	Snp1	13	4
U2 snRNP	Lea1	5	1
	Msl1	2	1
	Prp9	6	0
	Prp21	2	0
	Cus1	4	0
	Rse1	4	0
	Prp11	1	1

B

Complex	Protein	Spectral counts WT	*htb1K123R*	t-test
U1 snRNP	**Prp42**	11 ± 2	2 ± 0	*
	Prp39	12 ± 3	5 ± 0	*
	Luc7	5 ± 1	3 ± 1	
	Mud1	13 ± 2	4 ± 1	*
	Nam8	3 ± 0	1 ± 1	*
	Snu56	5 ± 2	2 ± 2	
	Snu71	15 ± 2	10 ± 1	*
	Yhc1	3 ± 1	2 ± 0	
U2 snRNP	**Lea1**	5 ± 0	0 ± 0	***
	Prp9	3 ± 1	0 ± 0	*
	Prp21	2 ± 1	1 ± 1	
	Rse1	4 ± 1	1 ± 1	
	Prp11	4 ± 1	0 ± 0	*

C



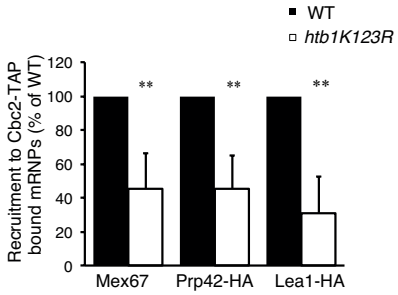


Figure 3. Preventing ubiquitination of H2B alters the recruitment of early splicing factors to CBC-associated mRNPs.

(A.) Nuclear mRNPs were purified from Cbc2-TAP-tagged WT or *htb1K123R* cells after a 3-hour shift to 39˚C and associated proteins were analyzed by MS-MS after SDS-PAGE. In order to highlight relative differences in the protein composition, the number of unique peptides for each indicated protein has been considered as index of abundance.[160] Complete identification of the CBC proteomes is shown in Table S4.

(B.) Components of the U1 and U2 snRNPs associated with nuclear mRNPs were semi-quantified using a spectral counting approach. The mean ± SD of spectral counts corresponding to three injections are indicated. Significance of the differences observed between both strains was evaluated using Student *t*-test ($^{*}P$ 0.01–0.05; $^{***}P$ <0.001). Significant differences are indicated in bold.

(C.) Nuclear mRNPs were purified from WT or *htb1K123R* cells expressing Cbc2-TAP and Prp42-HA or Lea1-HA. Co-purifying proteins were detected by Western Blot with anti-HA or anti-Mex67 antibodies (left panel). The ratio of mRNP-associated proteins relative to the WT cells and to the immuno-purified Cbc2-TAP was determined from at least 3 independent experiments (mean ± SD) (right panel).

(D.) Same as (C.) in WT or *set1*D cells. Significance of the differences observed between both strains was evaluated using Student *t*-test ($^{**}P$ 0.001–0.01).

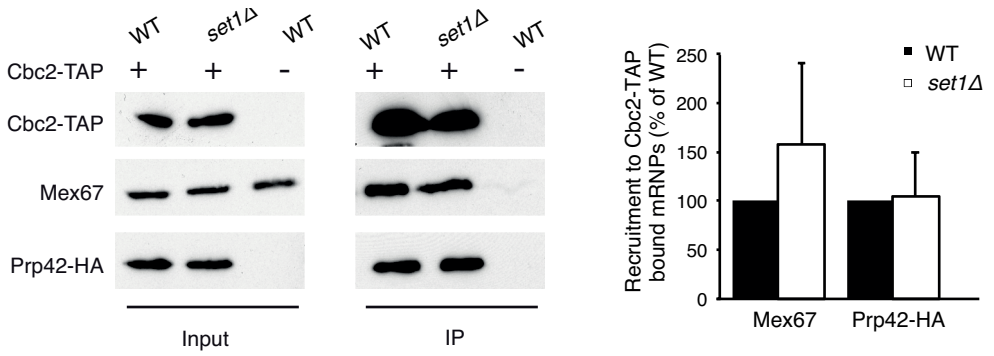


Figure 4. Loss of Set1-dependent methylation does not affect the recruitment of U1 snRNP to CBC-associated mRNPs.

Nuclear mRNPs were purified from WT or *set1*Δ cells expressing Cbc2-TAP and Prp42-HA. Co-purifying proteins were detected by Western Blot with anti-HA or anti-Mex67 antibodies (left panel). The ratio of mRNP-associated proteins relative to the WT cells and to the immuno-purified Cbc2-TAP was determined from at least 3 independent experiments (mean ± SD) (right panel). Significance of the differences observed between both strains was evaluated using Student *t*-test ($^{**}P$ 0.001–0.01).

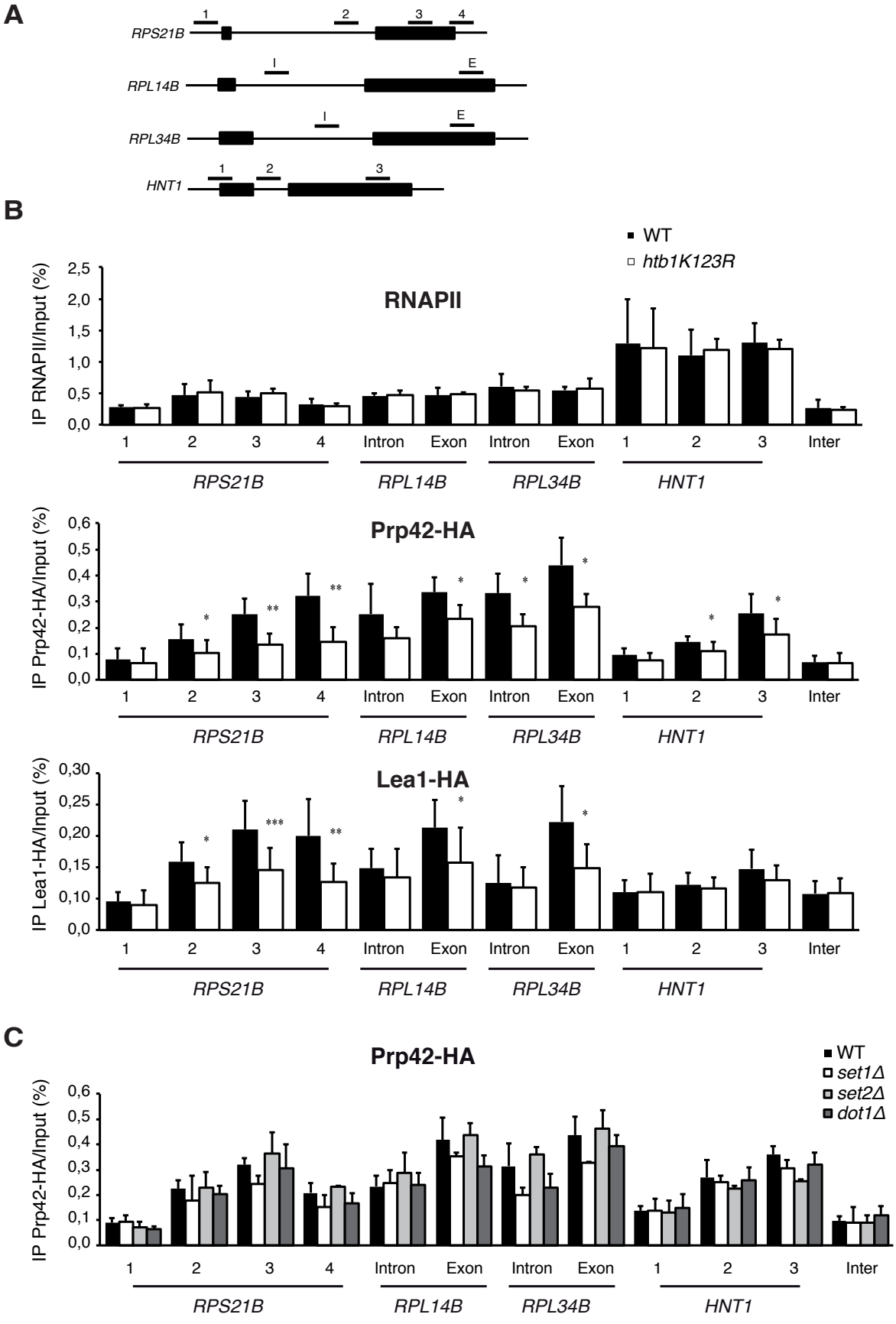

A
RPS21B
RPL14B
RPL34B
HNT1
B
WT
htb1K123R
RNAPII
IP RNAPII/Input (%)
Prp42-HA
IP Prp42-HA/Input (%)
Lea1-HA
IP Lea1-HA/Input (%)
Intron
Exon
Inter
C
WT
set1Δ
set2Δ
dot1Δ

Figure 5. Loss of H2B ubiquitination, but not Set1-, Set2- or Dot1-dependent methylation of H3, impairs recruitment of early splicing factors to transcribing genes.

(A.) Location of qPCR amplicons.

(B.-C.) ChIP experiments were performed on extracts prepared from the indicated HA-tagged strains after a 3-hour shift to 39˚C, using anti-RNAPII or -HA antibodies. Four intron-containing genes were considered, 3 ribosome protein genes *RPS21B*, *RPL14B* and *RPL34B* and the non-ribosome protein gene *HNT1* (Table S1). Histograms depict the mean and standard deviations of at least three independent experiments. The significance of the differences of recruitment observed between WT and mutant cells was evaluated using Student *t*-test ($^{*}P$ 0.01–0.05; $^{**}P$ 0.001–0.01).

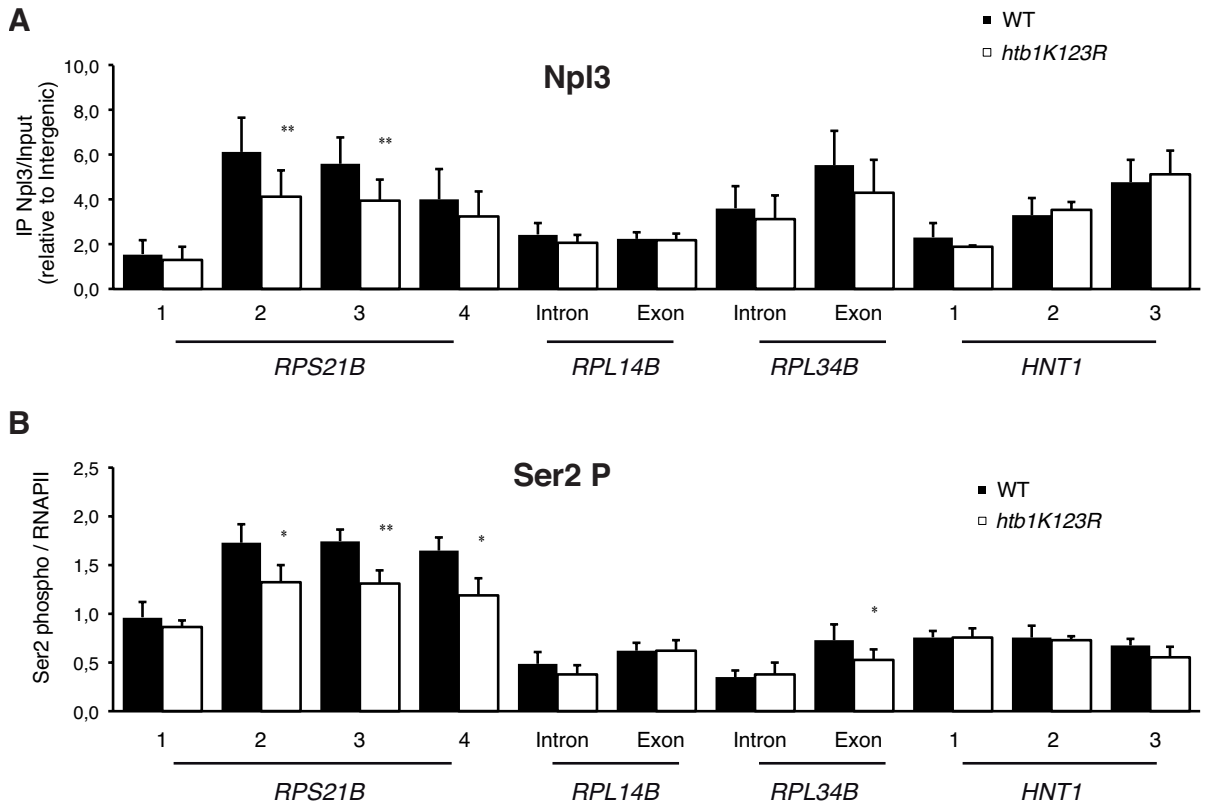


Figure 6. H2B ubiquitination promotes to RNAPII Ser2 phosphorylation and Npl3 recruitment on a subset of intron-containing genes.

(A.) Recruitment of Npl3 on indicated intron-containing genes was analyzed by ChIP assay in WT and *htb1K123R* cells using anti-Npl3 antibodies and normalized to the intergenic region.

(B.) Phosphorylation of RNAPII Ser2 on indicated intron-containing genes was analyzed by ChIP assay in WT and *htb1K123R* cells using antibodies to RNAPII and phosphoSer2. The RNAPII Ser2P/RNAPII ratio is shown. Significance of the differences observed between both strains was evaluated using Student *t-test* (*p-value = 0.01–0.05; **p-value = 0.001–0.01).

TABLES

Tables are available online at www.biolcell.net

Table S1. Yeast strains used in this study.

Table S2. Oligonucleotides used in this study.

Table S3. Results of Splicing Microarrays.

Table S4. Complete identification Table referring to Figure 2A.

Table S5. Identification Table referring to Figure 2B.

Table S6. Complete spectral count Table referring to Figure 2B.

Chapter 3

From Structure to Systems:

High-Resolution, Quantitative Genetic Analysis of RNA Polymerase II

From Structure to Systems:

High-Resolution, Quantitative Genetic Analysis of RNA Polymerase II

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

[1]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA, 94158, USA; [2]California Institute for Quantitative Biosciences, QB3, San Francisco, CA, 94158, USA; [3]Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX, 77843, USA; [4]Department of Biochemistry and Biophysics, University of California, San Francisco, CA, 94158, USA; [5]Departments of Medical Genetics and Biochemistry, Michael Smith Laboratories, University of British Columbia, Vancouver, V6T 1Z4, Canada; [6]Molecular Cancer Research, University Medical Centre Utrecht, 3508 AB Utrecht, The Netherlands; [7]Resource on Biocomputing, Visualization and Informatics, University of California, San Francisco, CA, 94158, USA; [8]J. David Gladstone Institutes, San Francisco, CA, 94158, USA

To whom correspondence should be addressed: Christine Guthrie, christineguthrie@gmail.com; Craig D. Kaplan, cdkaplan@tamu.edu and Nevan J. Krogan, nevan.krogan@ucsf.edu

* These authors contributed equally to this work

ABSTRACT

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

HIGHLIGHTS

•Functional dissection of RNAPII by genetic interaction profiling of point mutants

•RNAPII mutations affect transcription initiation, splicing and chromosome segregation

•Genetic relationships with RNAPII mutations reveal novel transcription factors

•Start site selection and splicing efficiency are coordinated with transcription rate

INTRODUCTION

Alterations within a genome often cause specific as well as global phenotypic changes to a cell. Combining two alterations in the same cell allows for measurement of the genetic interaction between them: negative genetic interactions (synthetic sick/lethal) arise when two mutations in combination cause a stronger growth defect than expected from the single mutations. This is often observed for factors participating in redundant pathways or as non-essential components of the same essential complex. In contrast, positive interactions occur when the double mutant is either no sicker (epistatic) or healthier (suppressive) than the sickest single mutant[161] and may indicate that the factors are components of a non-essential complex and/or perform antagonizing roles in the cell. However, single genetic interactions are often difficult to interpret in isolation; an interaction pattern for a given mutation can be more informative as it reports on the phenotype in a large number of mutant backgrounds.[54,162] These genetic profiles provide highly specific readouts that can be used to identify genes that are functionally related.[161]

One of the first organisms to be genetically interrogated on a large scale was *S. cerevisiae*, where non-quantitative genetic interaction data could be collected using the SGA (synthetic genetic array)[54] or dSLAM (heterozygous diploid-based synthetic lethality analysis on microarrays)[163] approaches. We developed a technique, termed E-MAP (epistatic miniarray profile),[162,164,165] which utilizes the SGA methodology and allows for the quantitative collection of genetic interaction data on functionally related subsets of genes, including those involved in chromatin regulation,[166] RNA processing,[55] signaling,[167] or plasma membrane function.[168] However, the vast majority of systematic genetic screening interrogates deletions of non-essential genes or hypomorphic knockdown alleles of essential genes. Since many genes, especially essential ones, are multi-functional, these methods perturb all activities associated with a given gene product.

Here, we describe an important advance of the E-MAP approach, which allows us to address higher levels of complexity by examining the genetic interaction space of point mutant alleles of

multi-functional genes in a technique we term point mutant E-MAP (or pE-MAP). This method greatly increases the resolution achievable by gene function analysis, as it allows assignment of specific genetic relationships to individual residues and domains. In this study, we have used the pE-MAP approach to functionally dissect RNAPII using alteration-of-function alleles in 5 different subunits of the enzyme. Using the genetic data, we assign transcriptional activity and specific functions to different residues and regions of RNAPII. By examining the relationship between transcription rate and genetic interaction partners, transcription rate-sensitive factors were revealed. Through the characterization of multiple stages of gene regulation, including start site selection, transcription elongation rate and mRNA splicing, the pE-MAP technique has provided both global and specific insight into structure-function relationships of RNAPII. We propose this strategy as a useful paradigm for the high-resolution interrogation of any multi-functional protein.

RESULTS

A set of alleles for the functional dissection of RNAPII

To identify residues important for transcriptional regulation *in vivo*, we isolated RNAPII alleles that confer one or more of the following transcription-related phenotypes: suppression of galactose sensitivity in *gal10*Δ*56* (Gal^R),[169,170] the Spt^- phenotype[171] or mycophenolic acid (MPA) sensitivity[172] (more detail in the legend of Figure 1A, Methods). Each of these phenotypes relates to a gene-specific transcription defect that can be monitored using plate assays (Figure 1A). Random mutagenesis by PCR was carried out on the entire coding regions of RNAPII subunit genes *RPB2*, *RPB3*, *RPB7* and *RPB11* and most of *RPO21/RPB1* (Methods). These genes encode the essential subunits unique to RNAPII (Rpb5, Rpb6, Rpb8, Rpb10 and Rpb12 are shared with RNAPI and RNAPIII, and Rpb4 and Rpb9 are non-essential).[173] In total, 53 single point mutants were identified that exhibit at least one of these phenotypes[174] (Figure 1B, Figure S1, Table S1).

Analysis of the distribution of phenotypes relative to the RNAPII structure suggested that our alleles might be diverse in their functions. Gal^R and MPA sensitive mutations were broadly distributed, while those with the Spt^- phenotype were less common and more localized (Figure 1B, Table S1). The screens identified mutations in highly conserved residues and structural domains known to be important for RNAPII activity, including the Rpb1 trigger loop, the Rpb1 bridge helix and the Rpb2 lobe and protrusion[175-177] (Table S1). Quantitatively measuring the genetic interactions of specific residues might provide insight into the functions of these RNAPII regions, and could therefore also identify protein-protein interaction interfaces.

An RNAPII point mutant epistatic miniarray profile (pE-MAP)

The 53 RNAPII point mutants (Figure 1B) were crossed against ~1200 deletion and DAmP (Decreased Abundance by mRNA Perturbation) alleles[162] (Table S1), which represent all major biological processes. Thus, a quantitative pE-MAP comprising approximately 60,000 double mutants was created (Table S2, http://interactome-cmp.ucsf.edu). Two-dimensional hierarchical clustering of these data effectively grouped together genes from known complexes and pathways based on their interactions with the point mutants (Figure S2, Supp. Data 1). Previous studies have demonstrated that genes encoding proteins that physically associate often have similar genetic interaction profiles.[166,178] The data derived from the point mutants could differ in this respect, since it is based on only 5 subunits of a single molecular machine. A receiver operating characteristic (ROC) curve was therefore generated to measure how well the genetic profiles of the deletion and DAmP mutants in the pE-MAP predict known physical interactions between their encoded proteins (Methods). It was found that the predictive power of the pE-MAP is similar to that of a previously published E-MAP,[166] indicating that the genetic interactions of the RNAPII point mutants report on connections among virtually all cellular processes (Figure 2A, Figure S2, Supp. Data 1).

Next, to examine whether the spatial location of a mutated residue is a determinant of its function, we compared the similarity of pairs of RNAPII genetic profiles to the three-dimensional

distance between the mutated residues.[179] We observed a strong correlation (r=-0.37, $p<10^{-22}$) (Table S2) and the trend is significant both for residue pairs residing in the same subunit (r=-0.25, $p<10^{-5}$) and for those in different subunits (r=-0.28, $p<10^{-6}$) (Figure 2B). This suggests that structural proximity correlates with functional similarity and that high resolution genetic interaction profiling could provide information for targets whose structures have not yet been determined.

Comparison of genetic and gene expression profiles derived from the RNAPII alleles

To determine if any given genetic interaction might result from the point mutation affecting the expression of the corresponding gene, we subjected 26 of the RNAPII mutants to genome-wide gene expression analysis (Table S2, GEO accession number: GSE47429). We found no correlation (r=-0.003) between an RNAPII mutant's genetic interaction score with a gene deletion or DAmP allele and the expression change of that gene due to the RNAPII mutation (Figure 2C). Therefore, connections must be due to more complex relationships between the mutated residues and the library genes. Nonetheless, these datasets allowed us to test whether the clustering of the RNAPII mutants in the pE-MAP (Figure S2, Supp. Data 1) could be recapitulated using their gene expression profiles. We thus assessed pair-wise RNAPII mutant similarity based on genetic and gene expression profiles separately and found that these two measures are highly correlated (r=0.71, Figure 2D). Therefore, these orthogonal datasets provide a common biological framework for functionally organizing the RNAPII mutants, allowing us to study the underlying biology behind these mutants and their phenotypes.

Functional associations between RNAPII residues and protein complexes

In an effort to link individual RNAPII mutations to specific cellular functions, data from the pE-MAP was compared to a published genetic interaction dataset containing profiles from >4000 genes.[180] These genes were classified based on complex membership of their encoded proteins[181] and Mann-Whitney U statistics were used to identify RNAPII mutants with high profile similarity to members of specific complexes (Methods, Figure 3A, Table S3). We uncovered a number of

connections, including several point mutants having similar genetic profiles to mutants of components of Mediator and the Rpd3C(L) histone deacetylase complex.[182,183] Unexpectedly, we also observed that several RNAPII mutants are significantly correlated genetically to kinetochore mutants (Figure 3A). We carried out chromosome transmission fidelity (CTF) assays on 19 of our RNAPII mutants, including those linked specifically to the kinetochore (Methods, Table S3).[184] Only four of the tested mutants exhibited chromosome loss in more than 15% of their colonies and these were genetically linked to the kinetochore in our analysis and had similar genetic profiles to each other in our pE-MAP (Figure 3A, B). Recent studies indicate that a certain level of transcription by RNAPII at the centromere is required for centromere function and high-fidelity chromosome segregation in budding yeast.[185] Using specific constructs designed to ascertain centromere sensitivity to transcriptional read-through, we did not observe any defects in kinetochore integrity in these RNAPII mutants (data not shown). Ultimately, further work will be required to understand the connection between these RNAPII point mutants and chromosome segregation. A full point mutant module map from alleles of all subunits is presented in Figure S3.

pE-MAP identifies alleles involved in start site selection and can finely distinguish between different phenotypic categories

What other transcription defects might underlie differences in the genetic profiles of specific RNAPII alleles? Recent work has shown that mutations in the Rpb1 trigger loop (TL), a dynamic element in the active site that couples correct NTP substrate recognition with catalysis, can alter transcription start site selection *in vivo*. For example, *rpb1* E1103G shifts transcription start site selection upstream at *ADH1* whereas *rpb1* H1085Y shifts distribution of start sites downstream.[174] The pE-MAP subcluster containing E1103G includes an additional 10 mutants in *RPB1*, *RPB2* and *RPB7* (Figure 4A, Supp. Data 1). We examined 8 of these for defects in start site selection at *ADH1* by primer extension and found that, like E1103G, all had a preference for upstream start site selection (Figure 4B, Figure S4A, Table S4), consistent with their clustering with E1103G (Figure

4A). Four of these mutants are also in the TL (*rpb1* F1084I, M1079R, A1076T, and M1079V); however, two were in other regions of Rpb1 (*rpb1* I1327V and S713P). Further inspection reveals these are in close proximity to the TL (Figure 4C),[179,186] suggesting they too may directly regulate the active site. Interestingly, the other 2 mutants tested are not close to the TL (*rpb2* E328K and *rpb7* D166G); these mutations may allosterically impact the active site or function independent of the TL by recruiting other factors to the transcription apparatus. Importantly, *rpb7* D166G is the first identified *RPB7* mutation with a start site defect. We additionally examined start site selection in *rpb4*Δ and *rpb6* Q100R, as both are expected to reduce the association of Rpb4/Rpb7 with RNAPII.[187,188] Neither altered start site selection at *ADH1* (Figure S4B), indicating the *rpb7* D166G mutant exerts a unique effect on RNAPII function (see Discussion). These data provide an example of how mechanistic information on structure-function relationships can be extracted from the pE-MAP.

In our screening process, we had also identified an *rpb2* allele mutated at two residues in close proximity (E437G/F442S) within the tip of the Rpb2 protrusion domain, whose genetic profile also clusters with the upstream start site mutants (Figure S4C, Table S2). The Rpb2 lobe and protrusion domains physically contact TFIIF[189] and consistent with this, we observed that, similar to what has been reported for TFIIF alleles[190,191] (Figure S6B), *rpb2* E437G/F442S is sensitive to MPA (Figure S4D) and has a preferential upstream start site selection (Figure 4B) (see Discussion). Additional *rpb2* alleles that confer MPA sensitivity also map near the protrusion (R120C) or to the lobe (D399H), however these did not alter *ADH1* start site selection and are genetically distinct from *rpb1* E1103G or other *rpb2* alleles, illustrating the fine resolving power of the pE-MAP (Figure 4A, B).

Despite its unbiased nature, the pE-MAP precisely grouped the mutants within the upstream start site cluster based on their Spt$^-$ and MPA phenotypes in the plate assays (Figure 4A). However, there are other mutants with MPA or Spt$^-$ phenotypes that did not exhibit upstream start preference and, notably, cluster apart from the ones that do (Figure 4A, B, Figure S4A, Table S4). These data

collectively suggest that the pE-MAP has the resolving power to categorize the mutations causing upstream start site selection and within this group can further separate them into specific Spt$^-$ and MPA phenotypic categories.

***In vitro* biochemical activity correlates with pE-MAP profiles and gene expression**

We next focused on the genetic profiles of a series of active-site mutants whose *in vitro* elongation rates range from <0.1 to >2-fold that of WT RNAPII.[174,192,193] This series allowed for addressing questions regarding the *in vivo* consequences of altered elongation rates. Clustering these mutants based on their pE-MAP profiles yielded two distinct subsets that differ by transcription rate: fast (*rpb1* E1103G, L1101S and F1084I) and slow mutants (*rpb1* F1086S, N1082S, N479S and H1085Q) (Figure 5A dendrogram). The gene expression profiles of these mutants also group them into the same two subsets. Thus, pairs of mutants with similar *in vitro* transcription rates show both highly similar genetic and expression profiles (Figure 5B), indicating that altered catalytic activity may likely be a defining feature *in vivo*.

We reasoned that because combining two TL mutations that individually increase and decrease elongation rate results in RNAPII with near-WT elongation rate *in vitro* and growth rate *in vivo*[174] (Figure S5A), these double mutants would also exhibit near-WT pE-MAP and gene expression profiles. Indeed, when we combine *rpb1* E1103G (fast) and F1086S (slow), we see very few high scoring genetic interactions compared to the single mutants and few genes whose expression changes by more than 1.7-fold (Figure 5C, D, Table S2). Similar trends were observed in the genetic and expression profiles of a *rpb1* E1103G/N1082S double mutant (data not shown). However, the pE-MAP and gene expression assays had the resolving power to also identify a double mutant (*rpb1* E1103G/H1085Q) that deviated from this general rule (Figure 5C, D, Table S2). Despite exhibiting partially suppressive genetic and expression patterns, there are still significant effects that correlate well with E1103G (Figure S5B), suggesting a more complex genetic relationship with this double mutant. Taken together, our double mutant analyses are consistent with the notion that the genetic

interactions and gene expression changes in fast and slow mutants are, for the most part, defined by the catalytic activity of RNAPII.

Genome-wide and gene-specific effects of altering polymerase speed on *in vivo* splicing efficiency

Several steps in mRNA processing are coupled to the process of transcription.[194] Based primarily on experiments from metazoa, a kinetic coupling model has been invoked to explain how decreasing polymerase speed affects the usage of alternative splice sites, primarily favoring the recognition of otherwise weak splice sites.[195] We used the RNAPII mutants that differ in their *in vitro* elongation rates to determine whether this type of coupling exists in *S. cerevisiae*, where splice sites are of consensus or near-consensus sequence and the sole read-out is splicing efficiency. We used splicing-sensitive microarrays to measure the change in total, pre-mRNA, and mature mRNA for each intron-containing gene (Figure 6A top) in response to mutations in the RNAPII TL. Many genes exhibit the reciprocal pre- and mature mRNA values indicative of changes in splicing (Figure 6A side panels) and to allow facile comparison across multiple mutants, we calculated pre-mRNA/mature mRNA ratios[96] for each gene (Figure 6, Table S5). Both gene-by-gene (Figure 6A) and global (Figure 6B) measures show that *rpb1* mutants characterized as fast *in vitro* (E1103G and G1097D) lead to an increase in this ratio for many genes, which indicates a defect in the splicing of those transcripts. Conversely, mutants with slow *in vitro* elongation (H1085Q and F1086S) decrease this ratio for many genes, a phenotype consistent with more efficient splicing. Thus, the global trend we observe is an anti-correlation between polymerase rate and splicing efficiency ($p<10^{-5}$ for each mutant in Figure 6A compared to WT, Table S5). Using qPCR, we confirmed this trend at several genes (Figure S6A), and found that these genes exhibit between 1% and 40% unspliced mRNA in WT, but up to 70% in G1097D (data not shown).

Given that polymerase rate and splicing efficiency are anti-correlated, it follows that the TL double mutants (*rpb1* E1103G/F1086S and E1103G/H1085Q), which transcribe at near-WT rate *in*

vitro, should also exhibit near-WT splicing. Consistent with this prediction, these strains had very few genes with splicing defects (Figure 6B, Table S5). Thus, suppressive relationships are observed in the double mutants with respect to growth, genetic and expression profiling, *in vitro* transcription rate, and also mRNA splicing. Furthermore, slowing polymerase by chemical means should phenocopy a genetically slow polymerase mutant. We therefore evaluated splicing in WT cells treated with mycophenolic acid (MPA), which is known to impede transcriptional elongation.[196] A 10-min MPA treatment resulted in even more genes with improved splicing when compared to a slow polymerase mutant (Figure 6B, $p<10^{-15}$ compared to WT, Table S5), possibly because MPA elicits an acute stress. Taken together, these splicing phenotypes are consistent with a direct kinetic coupling between elongation rate and splicing *in vivo.*

Genetic interactions between RNAPII alleles and other mutants reveal relationships between transcription factors and RNAPII activity

The observation of growth suppression in the TL double mutants[174] (Figures 5C, D) suggests that some of the deletion/DAmP genetic interactors that were detected in the pE-MAP might directly regulate, or collaborate with, RNAPII. We reasoned that disruption of a positively acting transcription factor would result in positive genetic interactions (suppression) with fast RNAPII mutants but negative interactions (synthetic sickness) with slow mutants. Conversely, a negatively acting factor would show opposite genetic trends. To identify these factors, the deletion/DAmP mutants were sorted based on the difference in their average genetic interaction score with fast and slow RNAPII mutants (Methods, Figure 7A, Table S6). We focused on genes that behaved as positively acting factors and observed that *sub1*Δ had the strongest pattern in this regard. Interestingly, previous evidence has implicated Sub1 as a positive factor in *in vitro* transcription assays (reviewed in Conesa *et al.*[197]), as well as *in vivo.*[198] The genetic relationships from the pE-MAP were confirmed using standard growth assays, where *sub1*Δ exacerbated slow RNAPII alleles (Figure 7B) and partially suppressed fast RNAPII alleles (Figure 7C). Deletion of *SUB1* also

exacerbated and suppressed the relevant RNAPII mutant phenotypes (MPA, Spt, Gal^R) (Figure S7A) (as did other mutants (Figure S7B)). Furthermore, gene expression analysis of *sub1*Δ, *rpb1* E1103G, and *sub1*Δ/E1103G showed an epistatic relationship between the E1103G mutant and *sub1*Δ (Figure 7D), consistent with their positive interaction and suggesting that a fast RNAPII mutant can bypass the requirement for Sub1.

Our recent work has implicated changes in RNAPII activity with the alteration of start site selection *in vivo*[174] (Figure 4). Furthermore, Sub1 genetically interacts with TFIIB (*sua7*),[199,200] is broadly recruited to RNAPII/III promoters *in vivo*[201,202] and was implicated as a member of the RNAPII pre-initiation complex.[203] We therefore sought to determine whether Sub1 might modulate RNAPII start site choice. Notably, primer extension analysis at *ADH1* revealed that deleting *SUB1* led to a significant downstream shift in start site (Figure 7E, Figure S7C). Slow RNAPII TL mutant *rpb1* F1086S (Figure 4B) and *sua7* alleles[204] (Figure S6B) also initiate downstream and are synthetic sick with *sub1*Δ. These data are consistent with the notion that Sub1 promotes transcription initiation. Double mutant analysis revealed that *sub1*Δ also exacerbated the downstream start site shift of the slow RNAPII TL allele, *rpb1* F1086S, and slightly suppressed the *rpb1* E1103G allele (Figure 7E). Because *sub1*Δ has also been linked to another RNA processing step, namely 3' end processing,[197] we examined the effect of *sub1*Δ on splicing and observed a statistically significant increase in splicing efficiency ($p<10^{-8}$, Table S5), again phenocopying the slow RNAPII mutants (Figure 7F, Table S5).

Based on these data, we propose a model where transcription start and splicing are intimately coupled with RNAPII elongation: fast RNAPII mutations result in upstream transcription start and diminished splicing, whereas slow mutations or *sub1*Δ give rise to downstream transcription start and enhanced splicing (Figure 7G). Given the possibility of direct coupling between start site selection and downstream mRNA processing, we globally measured splicing defects in *sua7*-3 (TFIIB)[200,205]

and *tgf2*Δ261-273 (TFIIF),[190] mutants that exhibit downstream and upstream start site selection, respectively (Figure S6B). If splicing were strictly coupled to start site choice, one would expect these mutants to have similar splicing defects to the slow and fast RNAPII mutants. However, such a correlation was not observed (Figure S6C), suggesting these processes are, in fact, genetically separable. Taken together, these data support a model in which the catalytic rate of RNAPII has multiple, separable effects on start site selection and mRNA processing, and highlight the importance of WT elongation rate for multiple steps in gene expression.

DISCUSSION

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

Insights into RNAPII Function Derived from the pE-MAP

Multiple aspects of the genetic interaction map provided insights into transcriptional regulation. First, by comparing the RNAPII genetic interaction profiles with those from previous deletion/knockdown studies, we could assign function to individual residues, which allowed us to generate a point mutant-protein complex connectivity map (Figure 3, Figure S3). Furthermore, analyzing double mutants, both within RNAPII itself (Figure 5C, D) and between RNAPII and other genes (Figure 7), allowed for a better understanding of TL function, as well as the identification of other factors that directly or indirectly impinge on RNAPII activity. Indeed, the pE-MAP allowed us to identify putative transcription factors (negative and positive), such as Sub1. While Sub1 was previously implicated at the promoter,[203] we here demonstrated that it positively regulates RNAPII

by influencing start site usage (Figure 7). The epistatic relationships between fast RNAPII alleles and *sub1*Δ suggest that Sub1 activity may be bypassed when RNAPII catalytic activity is increased. Collectively, our data indicate that Sub1 plays a direct role in transcriptional initiation and influences mRNA splicing, possibly via its effect on elongation.[198]

Finally, structural analysis revealed that mutations close in 3-D space have very similar genetic profiles, including those in different subunits (Figure 2B), suggesting that structural information is ultimately contained within the p-EMAP and can be used to identify specific protein-protein interaction interfaces. For example, the Rpb2 mutation E437G/F442S, which shifts start site selection upstream, is in a domain that contacts TFIIF at a region in which mutations also result in upstream start site shifts.[189,190] Furthermore, the identification of an Rpb7 mutant (D166G) that alters start site selection is intriguing, as this region interacts with TFIIH in a cryo-EM structure of the RNAPII pre-initiation complex[206] and TFIIH also alters start site selection.[207] Therefore, the pE-MAP technique could supplement other methods, such as cross-linking and electron microscopy, to identify physically interacting protein regions.

Coordination of Transcriptional Rate with Start Site Selection and mRNA Splicing

The importance of maintaining WT rates of transcription is evidenced by the phenotypic defects observed in the fast and slow mutants[174] (Figure 5, Figure 6), as well as the striking mutual suppression seen when combining two mutations that individually make RNAPII slow or fast (Figure 5, Figure 6). Interestingly, the pE-MAP identified two groups of RNAPII mutants: one that preferentially initiates upstream and exhibits an increased rate of transcription, and one that initiates downstream and transcribes slowly (Figure 7).[174] We propose that both phenotypes are direct consequences of the efficiency of nucleotide selection and incorporation, since the addition of the first nucleotides at initiation is biochemically similar to adding nucleotides during elongation. These data support the model that RNAPII engages in "scanning" during initiation in *S. cerevisiae.*[174,177,208,209] Whether RNAPII catalysis drives this scanning, or whether scanning occurs in

the absence of nucleotide incorporation, perhaps driven by TFIIH, is unknown. We also note that because changes in start site selection alter the 5'UTR length and composition, initial transcription decisions may have downstream effects on gene expression such as changes in RNA stability or translational efficiency of the mRNA.[210,211]

The pE-MAP has allowed insights into the co-transcriptional process of mRNA splicing. It is now clear that most introns are removed while RNA polymerase is still associated with the DNA template. In metazoans, alternative splicing decisions can be influenced by factors impinging on transcription, including promoter identity and polymerase speed (reviewed in Perales *et al.*[212]). Slowing the rate of elongation by mutation of RNAPII or chemical means can improve the recognition of splice sites that deviate from consensus signals.[28,29,213] Because the spliceosome undergoes stepwise assembly on each intron, slowing transcription can afford more time for formation of the catalytically active machine before transcription of a downstream, stronger site. While budding yeast lack alternative splicing, it nonetheless follows that the efficiency of co-transcriptional splicing would be favored by allowing sufficient time for spliceosome assembly. Indeed, recent work suggests that RNA polymerase may slow down to favor co- vs. post-transcriptional splicing.[34,36,138] Our microarray analyses show a clear trend in which splicing efficiency is generally anti-correlated with transcription rate (Figure 6). This is the first genome-wide analysis to interrogate kinetic coupling in *S. cerevisiae*. Moreover, this is the first instance in which faster and slower polymerases have been directly compared; thus, this work satisfies the previously untested prediction that rate and splicing efficiency would be generally anti-correlated.

Interestingly, we observed correlation between start site selection, elongation rate and splicing efficiency. In fact, promoter-proximal events are known to be able to influence downstream RNA transactions: promoter identity can influence alternative splicing or mRNA stability in other systems,[214-216] and 5'UTR length, determined by start site selection, can strongly alter translation efficiency in budding yeast.[211] However, when we measured splicing efficiency using mutants in the

general transcription factors TFIIF and TFIIB that alter start site selection, we found that not all initiation phenotypes are predictive of splicing efficiency (Figure S6C). This suggests that RNAPII catalytic rate has several, separable effects on gene expression, a claim supported by recent evidence showing kinetic coupling between RNAPII transcription and Sen1-dependent termination.[217]

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

Future Studies Using the pE-MAP Approach

S. cerevisiae RNAPII provided the groundwork for validating the pE-MAP approach. Application to other molecular machines, including the ribosome, the proteasome, HSP70, HSP90, histones and DNA polymerases, should prove informative. This analysis can also be carried out in other organisms that are genetically tractable and amenable to high-throughput genetic interaction mapping, including *S. pombe*[178,218,219] and *E. coli.*[220,221] Furthermore, pE-MAPs could be used to gain structural insight into proteins and complexes that have unknown structures. Finally, as genetic interaction mapping strategies become more prevalent in mammalian cells[222-225] and with the development of genome editing,[226] similar work characterizing the function of individual amino acids will have great impact on understanding how point mutations in specific genes result in different disease states.

MATERIALS AND METHODS

Summary

E-MAP-compatible *MAT*a RNAPII mutant strains carrying a marked *rpb* deletion and mutant *rpb* on a *CEN* plasmid were mated with 1200 *MAT***a** DAmP/deletion strains by pinning on solid media. Sporulation was induced and double mutant *MAT***a** spores were isolated on selective media. Genetic interactions were scored based on double mutant colony sizes, which were extracted using automated imaging software.

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

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

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

Genome-wide gene expression

Strains were streaked from -80°C stocks onto plates and grown for 3 days. Liquid cultures were inoculated with independent colonies and grown overnight in Synthetic Complete (SC) medium: 2g/l Drop out mix Complete and 6.71g/l Yeast Nitrogen Base without AA, Carbohydrate & w/AS (YNB) from US Biological (Swampscott, USA) with 2% D-glucose. Overnight cultures were diluted to OD_{600}=0.15 in 1.5ml fresh medium and grown at 30°C in a 24 well plate in a Tecan Infinite F200 under continuous shaking. Growth curves were made for the point mutant cultures (two cultures from two isolates) as well as for two WT inoculates, grown in parallel. Mutant and WT cells were harvested by centrifugation (6100 rpm, 3 min) at mid-log phase at OD_{600}=0.6, and pellets were immediately frozen in liquid nitrogen after removal of supernatant. Up to eleven mutant strains could

be grown on a single day. WT cultures were grown parallel to the point mutants to assess day-to-day variance. RNA isolation and purification was performed.[132] In short, total RNA was prepared by phenol extraction and cleaned up using a customized Sciclone ALH 3000 Workstation. For each sample, external control poly-A+ RNAs were added in equimolar amounts to the total RNA to enable monitoring of global changes in mutants.[227] For each micro-array analysis, 1.25μg labeled sample cRNA and 1.25μg reference cRNA was hybridized per slide. Each hybridization performed within this project was subjected to a number of quality controls. Some of these are based on the data from one single hybridization, while others are based on comparing data from one single hybridization against the WT grown in parallel.[132] Two channel microarrays were used. RNA isolated from a large amount of WT yeast from a single culture was used as a common reference. This common reference was used in one of the channels for each hybridization and used in the statistical analysis to obtain an average expression-profile for each mutant relative to the WT. Two independent cultures were hybridized on two separate microarrays. Each gene is represented twice on the microarray, resulting in four measurements per mutant.

Splicing microarray assay

Cultures were grown in rich medium according to standard techniques.[228] Saturated cultures were diluted to OD_{600} = 0.1 in the morning and allowed to grow at 30°C until reaching mid-log phase (OD_{600}=0.5-0.7). Mutant strains and an isogenic WT were collected by centrifugation and snap frozen in liquid nitrogen. Where indicated, a WT log-phase culture was treated with mycophenolic acid (30μg/ml) for 10 minutes, harvested using vacuum filtration and snap frozen in liquid nitrogen. Total cellular RNA was isolated using a hot acid phenol extraction generally as outlined in Schmitt *et al.,*[118] but with some specific modifications.[4] cDNA from each strain was synthesized, and labeled with Cy3 or Cy5 according to the low-throughput sample preparation method.[92] The optimized oligos[92] were robotically arrayed onto poly-L-lysine coated glass slides (slides from

ThermoScientific C40-5257-M20) and slides were processed according to Pleiss *et al.*[92] and DeRisi *et al.*[119] Microarrays were scanned using Axon Instruments GenePix 4000B at 635nm and 532nm wavelengths and image analysis was done using Axon Instruments GenePix Pro version 5.1. Spots were manually removed from analysis if they contained uncharacteristically high background or obvious defects; the ratio of the median intensity for 535nm and 625nm was calculated for each remaining spot. Each biological replicate contains 6 technical replicates for each feature as well as dye-flipped replicates, which were combined and normalized.[92] For all mutants, data from at least two biological replicates were used. To emulate a classic splicing measure that compares the levels of pre-mRNA and mature mRNA,[229] we computed I/J ratios as $\log_2$(Intron/Junction) = $\log_2(\text{Intron}_{\text{mutant}}/\text{Intron}_{\text{WT}})$-$\log_2(\text{Junction}_{\text{mutant}}/\text{Junction}_{\text{WT}})$ for each gene. Note that all I, J and I/J reported in figures and tables are $\log_2$ ratios. The reported p-values (Table S5) were calculated via one-sided Wilcoxon signed rank tests, comparing each I/J distribution ($\log_2$) to that resulting from direct hybridization of cDNA from two separately grown WT cultures.

qPCR assay for splicing efficiency

Splicing efficiency was measured by qPCR of cDNA as in Pleiss *et al.*[92] with modifications as described below. Ten µg of RNA were treated with 4U RQ1 DNaseI (Promega) in 16µL according to manufacturer instruction (20 minutes 37˚C). 4µL Stop Solution was added, and DNase was inactivated by incubation at 65˚C for 10 min. Samples were primed with random 9-mers in 40µL (50mM Tris-HCl pH8.3, 75mM KCl, 3mM $MgCl_2$, 10mM DTT, 400ng dN9; 5 min 65˚C; 5 min on ice) and divided into equal 40µL reverse transcription reactions with M-MLV reverse transcriptase (+RT or -RT in 50mM Tris-HCl pH8.3, 75mM KCl, 3mM $MgCl_2$, 10mM DTT, 250mM each dNTP; 2 hours 42˚C). qPCRs were performed in 25µL consisting of 1X Standard *Taq* Buffer and 1.25U *Taq* (New England Biolabs), 200mM each dNTP, 0.0013% SYBR Green I (Sigma), 600nM each primer, DMSO as needed (Table S5), and 25× - 2500× dilution of cDNA (-RT samples diluted equivalently to match +RT). Amplification reactions were performed on a C1000 ThermoCycler (BioRad) (94˚C

3 min and 35 cycles of 94˚C 15 sec, 50˚C 15 sec, 72˚C 15 sec, plate read). After cycling, each qPCR run was finished with a melt curve to confirm homogeneity of the amplified product. Starting quantity was calculated from a genomic DNA standard curve for each primer set; standard curve reactions not falling in the linear range were removed manually. Two technical replicates apiece were performed for 2-8 biological replicates (F1086S n=8, H1085Q n=5, *sub1*Δ n=2, MPA n=3, E1103G n=6, G1097D n=5). -RT control samples yielded negligible amplification (data not shown). Each gene was measured using primer sets specific for pre-mRNA and total mRNA (Table S5). To generate the graph in Figure S6A, the pre-mRNA/total mRNA ratio for each mutant was normalized to a within-experiment WT before taking the mean of the biological replicates.

Sorting array mutants on their interactions with fast and slow mutants

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

Additional descriptions of experiments and computational analyses are provided in Extended Experimental Procedures (available from www.cell.com).

ACKNOWLEDGEMENTS

The authors wish to thank members of the Krogan, Kaplan and Guthrie labs; Brad Cairns, Paivand Jalalian, Ada Cheng, Tanja Kortemme, David Agard, Andrej Sali, Keith Yamamoto and C. James Ingles for helpful discussion; Steve Buratowski, Olga Calvo and Eva Nogales for sharing results prior to publication; Steve Hahn and Mike Hampsey for providing *tfg2* and *sua7* mutant plasmids, respectively; Colm Ryan for help with statistical analysis; and Jiewei Xu and Ricardo Almeida for assistance with growth assays. This work was supported by grants from QB3@UCSF, the NIH (GM084448, GM084279, GM081879 and GM098101 to NJK, GM36659 to Roger D. Kornberg for

support of CDK for a portion of this work, GM097260 to CDK, GM21119 to CG and DP5OD009180 to JF). CG is an ACS Research Professor of Molecular Genetics. EAM was supported by an NSF graduate research fellowship. CDK acknowledges a Helen Hay Whitney Fellowship for the early stages of this work. HB was supported by the UCSF Biophysics Graduate Program. NJK is a Searle Scholar and a Keck Young Investigator.

AUTHOR CONTRIBUTIONS

HB, EAM, CG, CDK and NJK designed the research; HB, HJ, EAM, YAC, SW, JJB, CQ, FH and LKT carried out experimental assays; HB analyzed data; MS, JHM, JSF provided computational support; FCPH, PH, CG, CDK and NJK supervised the research; HB, EAM, CG, CDK and NJK wrote the paper. HJ and EAM contributed equally.

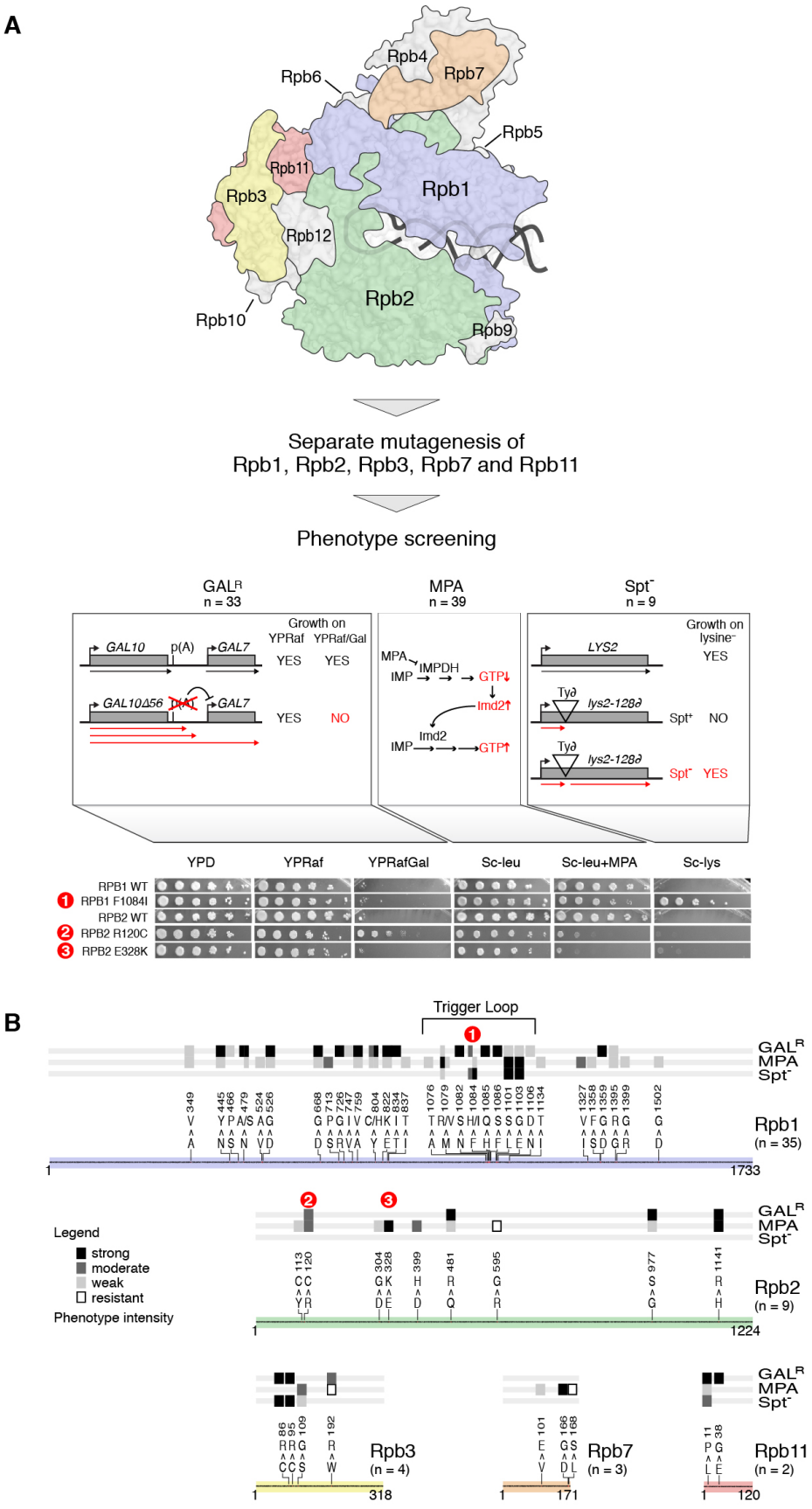

A
Rpb4
Rpb7
Rpb6
Rpb5
Rpb11
Rpb1
Rpb3
Rpb12
Rpb2
Rpb10
Rpb9
Separate mutagenesis of
Rpb1, Rpb2, Rpb3, Rpb7 and Rpb11
Phenotype screening
GALR
n = 33
MPA
n = 39
Spt-
n = 9
Growth on
YPRaf YPRaf/Gal
GAL10 p(A) GAL7
YES YES
GAL10Δ56 p(A) GAL7
YES NO
MPA
IMP IMPDH GTP
Imd2
IMP Imd2 GTP
Growth on
lysine-
LYS2
YES
Tyð lys2-128ð
Spt+ NO
Tyð lys2-128ð
Spt- YES
YPD
YPRaf
YPRafGal
Sc-leu
Sc-leu+MPA
Sc-lys
RPB1 WT
1 RPB1 F1084I
RPB2 WT
2 RPB2 R120C
3 RPB2 E328K
B
Trigger Loop
GALR
MPA
Spt-
Rpb1
(n = 35)
1
1733
Legend
strong
moderate
weak
resistant
Phenotype intensity
Rpb2
(n = 9)
1224
Rpb3
(n = 4)
318
Rpb7
(n = 3)
171
Rpb11
(n = 2)
120

Figure 1. Generation and selection of RNAPII point mutants.

(A.) RNAPII point mutants were screened for three transcription-related phenotypes. Gal^R, left: Deletion of the major *GAL10* p(A) site (*gal10*D*56*) results in RNAPII read-through and interference with initiation at *GAL7*, causing a Gal-sensitive phenotype. Gal^R mutants increase *GAL10* 3'-end formation/termination, thereby rescuing *GAL7* expression. MPA, middle: Mycophenolic acid (MPA) inhibits IMP-dehydrogenase (IMPDH) dependent GTP synthesis, but is counteracted by upregulation of an MPA-resistant form of IMPDH, *IMD2*. Transcriptional defects that are sensitive to low GTP levels or reduce *IMD2* expression render cells sensitive to MPA. Spt^-, right: Insertion of a Ty-retrotransposon into *LYS2* (*lys2-128*d*)* results in a lysine auxotrophy due to transcription block. Certain mutants suppress *lys2-128*d and allow expression of *LYS2*, due to activation of an internal promoter. Spot tests to identify each phenotype for three representative mutants are displayed.

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

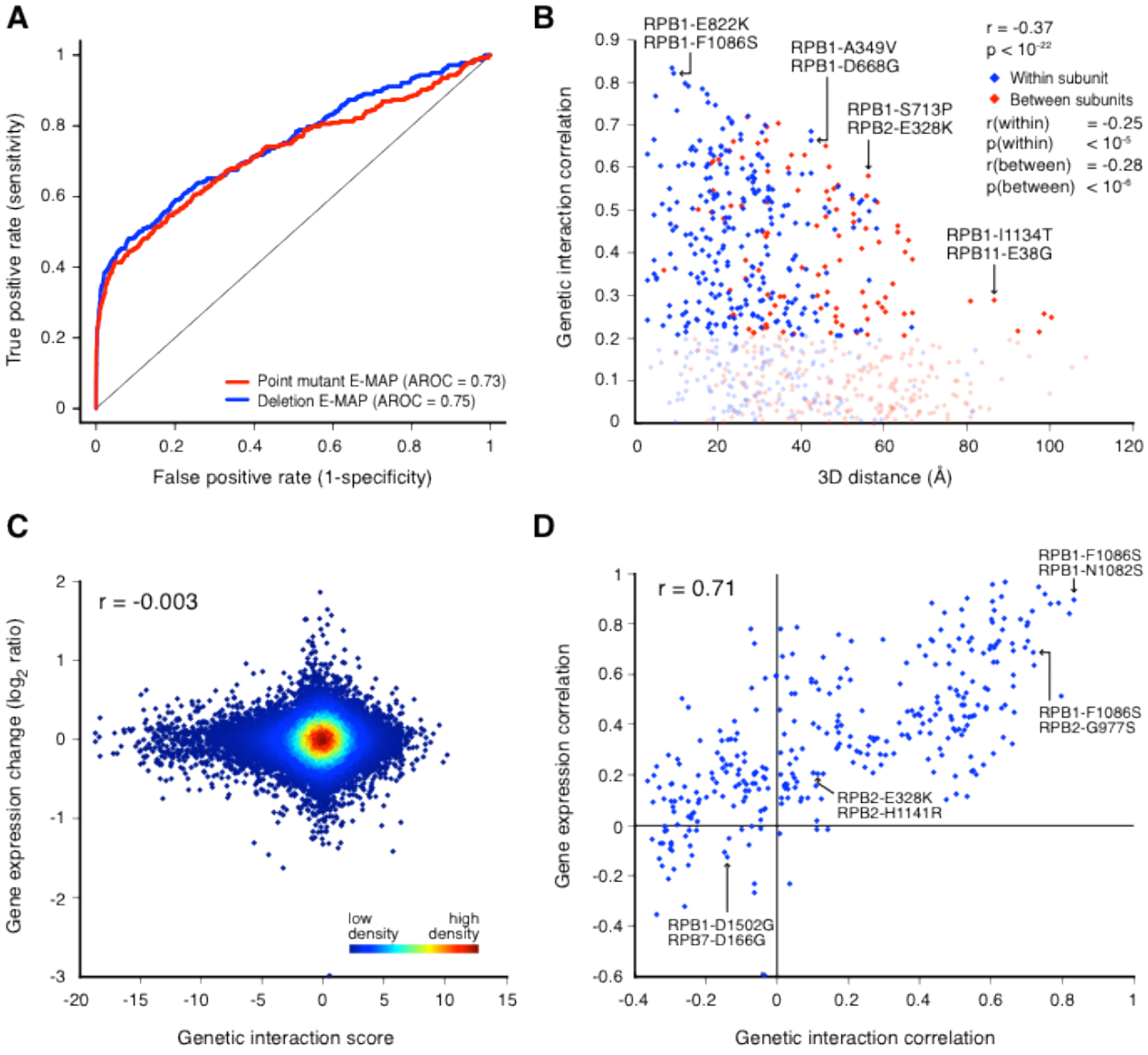

A
True positive rate (sensitivity)
False positive rate (1-specificity)
Point mutant E-MAP (AROC = 0.73)
Deletion E-MAP (AROC = 0.75)
B
Genetic interaction correlation
3D distance (Å)
RPB1-E822K
RPB1-F1086S
RPB1-A349V
RPB1-D668G
RPB1-S713P
RPB2-E328K
RPB1-I1134T
RPB11-E38G
r = -0.37
p < 10-22
Within subunit
Between subunits
r(within) = -0.25
p(within) < 10-5
r(between) = -0.28
p(between) < 10-6
C
Gene expression change (log2 ratio)
Genetic interaction score
r = -0.003
low density
high density
D
Gene expression correlation
Genetic interaction correlation
r = 0.71
RPB1-F1086S
RPB1-N1082S
RPB1-F1086S
RPB2-G977S
RPB2-E328K
RPB2-H1141R
RPB1-D1502G
RPB7-D166G

Figure 2. pE-MAP interactions span numerous biological processes, depend on spatial location of mutated residues and are not direct consequences of changes in gene expression.

(A.) ROC curves comparing the power of genetic profile correlations from the pE-MAP (red) and an E-MAP focused on chromosome biology[91] (blue) to predict physical interactions between pairs of proteins[166] (Methods).

(B.) Genetic profile correlations between pairs of mutated residues compared to the three-dimensional distance between their α-carbons. Blue points denote residue pairs within the same RNAPII subunit; red points represent pairs in different subunits. Negatively correlated residue pairs were excluded, as were four mutants of residues absent from the coordinate file (PDB ID 2E2H) (Rpb1 D1502, Rpb7 V101, Rpb7 D166, Rpb7 L168).[179] Correlations between 0 and 0.2 are dimmed to highlight trends at higher correlations.

(C.) Effect of RNAPII point mutations on gene expression, compared to the corresponding genetic interaction scores between RNAPII mutants and deletion/DAmP alleles. All combinations of the 26 RNAPII mutants and 1192 library genes/mutants that were examined via both pE-MAP and expression analyses are included. No global changes in gene expression were observed (measured by spike-in control RNA, Methods).

(D.) Comparison between pairwise RNAPII mutant correlations of genetic interaction profiles and gene expression profiles. See also Figure S2, Table S1, Table S2, Supp. Data 1.

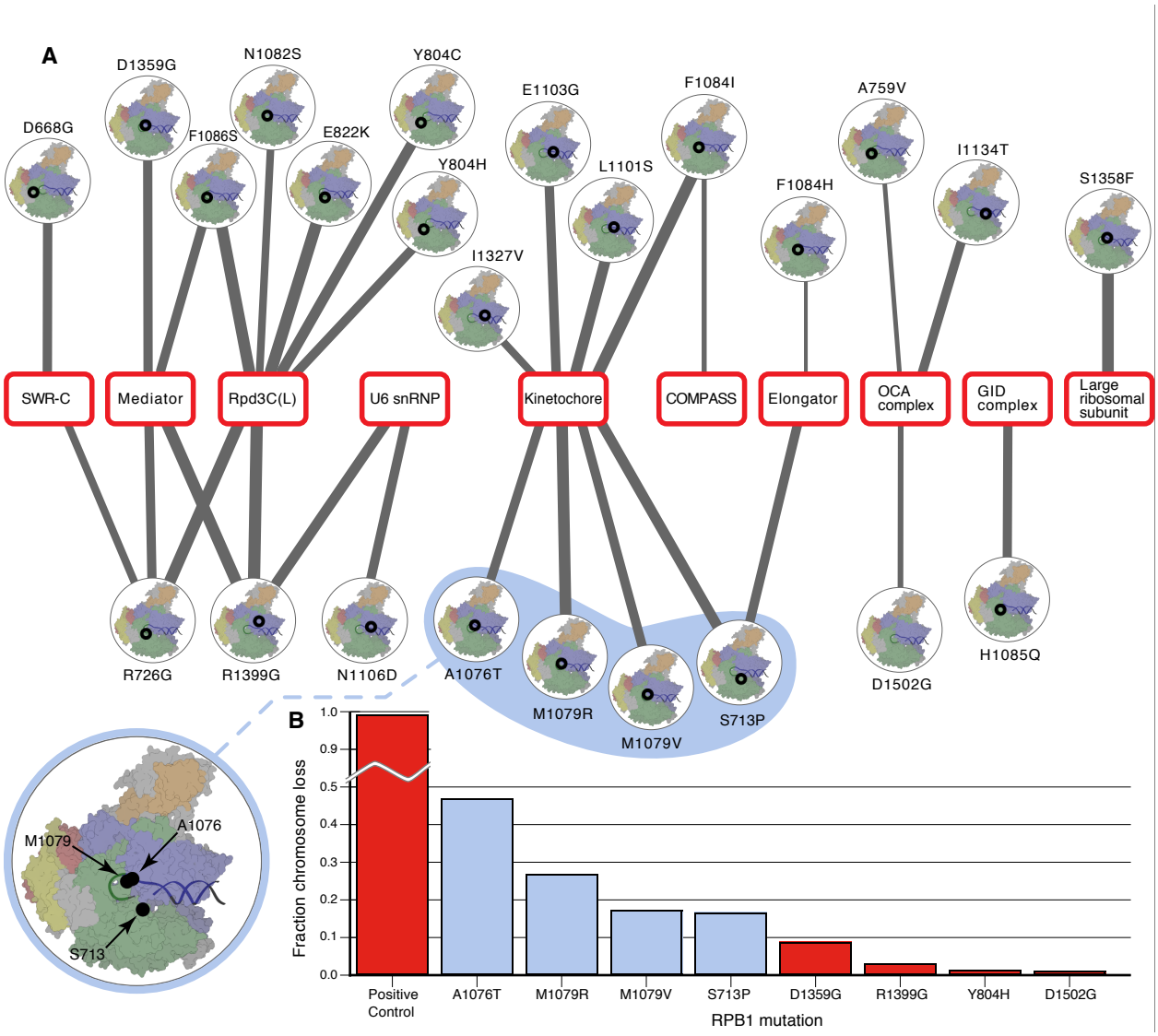

A
D668G
D1359G
N1082S
Y804C
F1086S
E822K
Y804H
I1327V
E1103G
L1101S
F1084I
F1084H
A759V
I1134T
S1358F
SWR-C
Mediator
Rpd3C(L)
U6 snRNP
Kinetochore
COMPASS
Elongator
OCA complex
GID complex
Large ribosomal subunit
R726G
R1399G
N1106D
A1076T
M1079R
M1079V
S713P
D1502G
H1085Q
M1079
A1076
S713
B
Fraction chromosome loss
1.0
0.9
0.5
0.4
0.3
0.2
0.1
0.0
Positive Control
A1076T
M1079R
M1079V
S713P
D1359G
R1399G
Y804H
D1502G
RPB1 mutation

Figure 3. Comparison of the pE-MAP with previously collected genetic interaction data reveals functional associations between RNAPII residues and protein complexes.

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

(B.) 19 mutants were examined using a chromosome transmission fidelity (CTF) assay. The four kinetochore-linked mutants highlighted in (A) exhibit chromosome loss in more than 15% of their colonies (blue bars), whereas unlinked mutants display no or weak phenotype (red bars, representative set). See also Figure S3, Table S3.

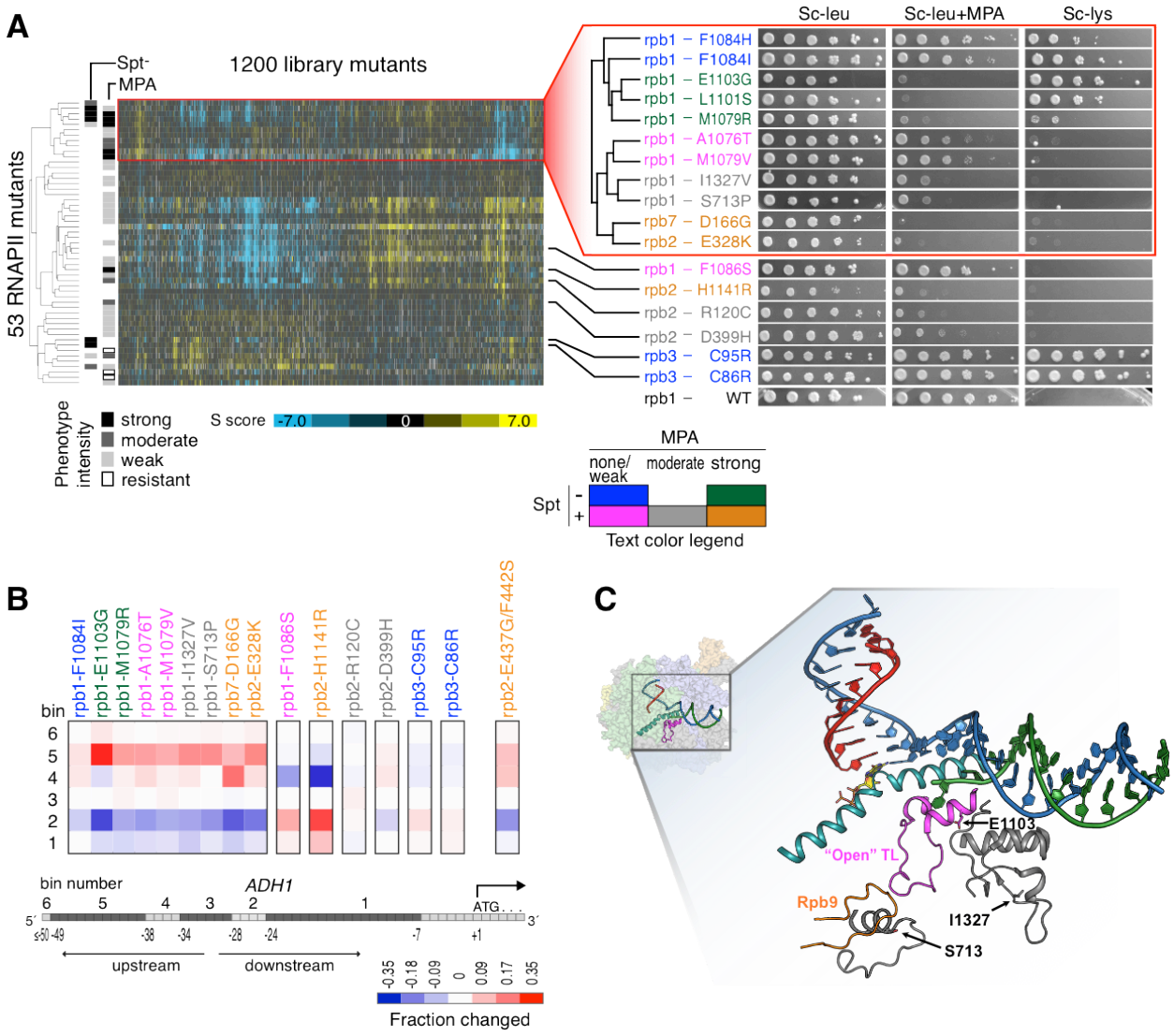

A
Spt-
MPA
1200 library mutants
53 RNAPII mutants
Sc-leu
Sc-leu+MPA
Sc-lys
rpb1 – F1084H
rpb1 – F1084I
rpb1 – E1103G
rpb1 – L1101S
rpb1 – M1079R
rpb1 – A1076T
rpb1 – M1079V
rpb1 – I1327V
rpb1 – S713P
rpb7 – D166G
rpb2 – E328K
rpb1 – F1086S
rpb2 – H1141R
rpb2 – R120C
rpb2 – D399H
rpb3 – C95R
rpb3 – C86R
rpb1 – WT
Phenotype intensity
strong
moderate
weak
resistant
S score
-7.0
0
7.0
MPA
none/ weak
moderate
strong
Spt
Text color legend
B
bin
rpb1-F1084I
rpb1-E1103G
rpb1-M1079R
rpb1-A1076T
rpb1-M1079V
rpb1-I1327V
rpb1-S713P
rpb7-D166G
rpb2-E328K
rpb1-F1086S
rpb2-H1141R
rpb2-R120C
rpb2-D399H
rpb3-C95R
rpb3-C86R
rpb2-E437G/F442S
bin number
ADH1
ATG. . .
upstream
downstream
-0.35
-0.18
-0.09
0
0.09
0.17
0.35
Fraction changed
C
E1103
"Open" TL
Rpb9
I1327
S713

Figure 4. pE-MAP profiles differentiate between subtle changes in transcription-related phenotypes and identify RNAPII mutations that affect start site selection.

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

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

(C.) Rpb1 I1327 and Rpb1 S713 connect to the TL (magenta). Mutations in I1327 could affect the structural region of the TL (E1103) via a network of loops and helices in Rpb1 (grey), while S713 is close to the TL catalytic site in its open conformation, via a Rpb9 loop (orange). In particular, the proline substitution, S713P, could result in structural changes affecting the TL. Coordinates for TL, Rpb9 and the S713 loop are from PDB ID 1Y1V[186] and all other from 2E2H.[179] The bridge helix is shown in cyan, template DNA in blue, non-template DNA in green, RNA in red, and the incoming GTP base is colored by atom. See also Figure S4, Table S4, Supp. Data 1.

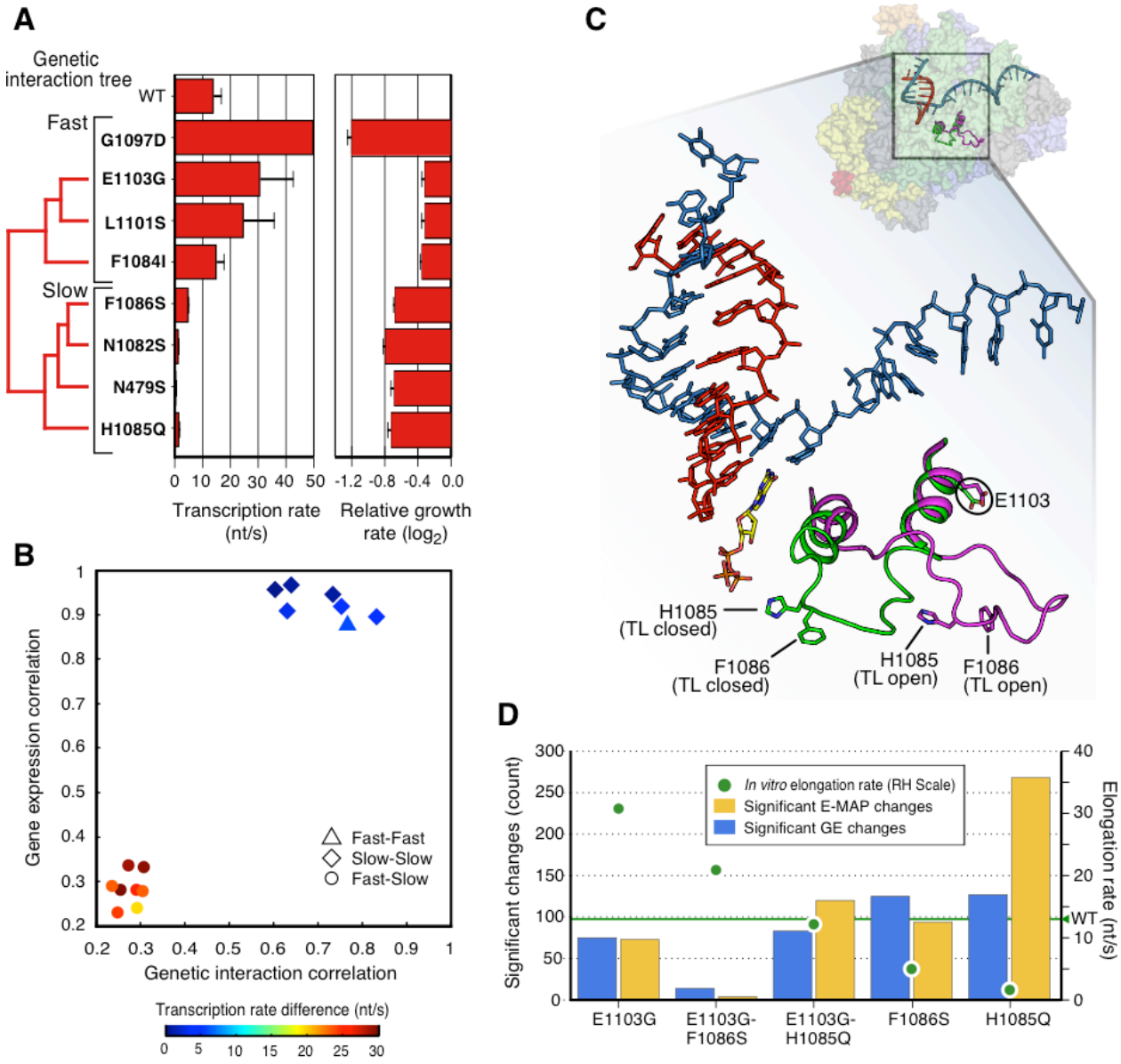

A
Genetic interaction tree
WT
Fast
G1097D
E1103G
L1101S
F1084I
Slow
F1086S
N1082S
N479S
H1085Q
0 10 20 30 40 50
Transcription rate (nt/s)
-1.2 -0.8 -0.4 0.0
Relative growth rate (log2)
B
Gene expression correlation
Genetic interaction correlation
Fast-Fast
Slow-Slow
Fast-Slow
Transcription rate difference (nt/s)
0 5 10 15 20 25 30
C
E1103
H1085 (TL closed)
F1086 (TL closed)
H1085 (TL open)
F1086 (TL open)
D
Significant changes (count)
Elongation rate (nt/s)
In vitro elongation rate (RH Scale)
Significant E-MAP changes
Significant GE changes
WT
E1103G
E1103G-F1086S
E1103G-H1085Q
F1086S
H1085Q

Figure 5. pE-MAP and expression profiles are indicative of biochemical activity.

(A.) *In vitro* transcription rates from Kaplan *et al.*[174] and *in vivo* growth rates relative to WT for RNAPII active site mutants. The dendrogram was generated via hierarchical clustering of the genetic profiles. Means and standard deviations of growth rates were derived from three technical replicates. Note that *rpb1* G1097D was too sick for reproducible E-MAP analysis.

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

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

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

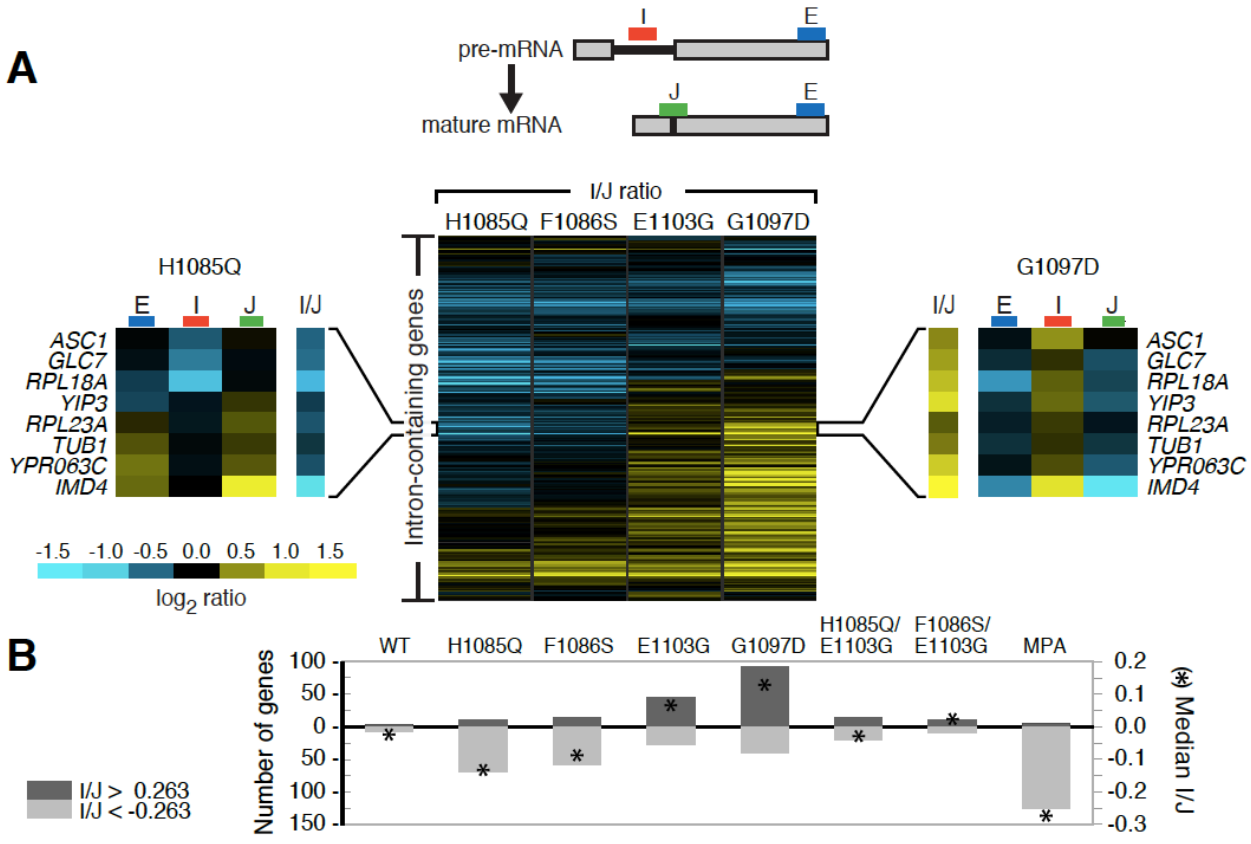


Figure 6. Effects of altering RNAPII transcription rate on *in vivo* splicing efficiency.

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

(B.) Number of genes exhibiting >20% change in pre-mRNA to mature mRNA ratio (bars, left-hand scale), and median I/J $\log_2$ ratio (asterisks, right-hand scale) across entire array. MPA treatment was 10 minutes; WT denotes competitive hybridization between two WT cultures. I/J denotes I/J $\log_2$ ratio. See also Figure S6, Table S5.

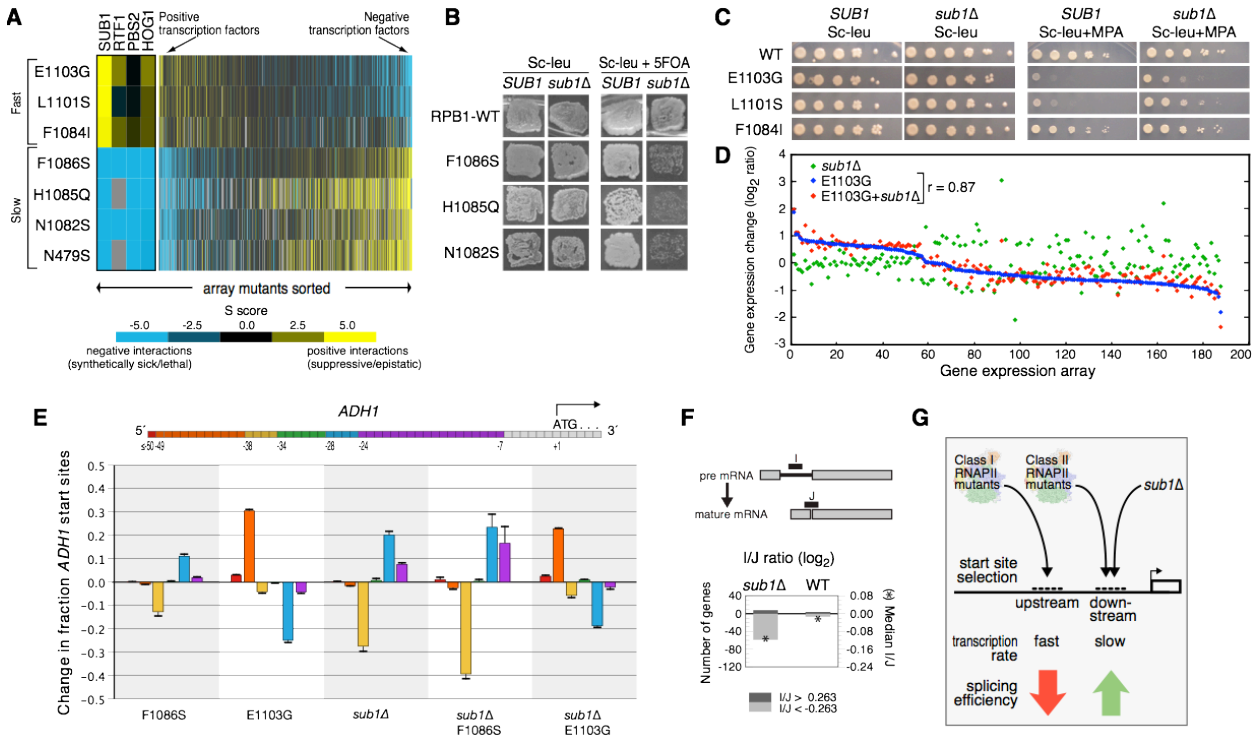


Figure 7. Genetic interaction patterns with fast and slow RNAPII mutants reveal Sub1 as a transcription factor that regulates start site selection and influences mRNA splicing.

(A.) Genetic profiles of library mutants, sorted on the difference between their average interaction with fast and slow RNAPII mutants (Table S6).

(B.) Patch tests examining the sensitivity of slow TL mutants to *sub1*Δ. WT *RPB1* plasmid covering *rpb1* mutants in left panel is lost in right panel.

(C.) Spot tests examining the effect of *sub1*Δ on fast mutants in absence (left) and presence (right) of MPA.

(D.) Comparison of *sub1*Δ effect on gene expression in *rpb1* E1103G (difference between red and blue) and WT (difference between green and y=0). Included are all array transcripts exhibiting a >1.5-fold expression change in at least one of the three mutants. Transcripts are sorted by expression change in E1103G.

(E.) Primer extension at *ADH1* to map transcription start sites for *rpb1* F1086S, E1103G and s*ub1*Δ (Figure S7C). Bar colors correspond to sequence windows in the *ADH1* schematic (top) and heights specify the mean fraction change of transcription start in mutant compared to WT. Error bars represent standard deviations.

(F.) Splicing microarray analysis of *sub1*Δ, as in Figure 6 (Table S5). Number of genes exhibiting >20% change in pre-mRNA to mature mRNA ratio (bars, left-hand scale), and median I/J $\log_2$ ratio (asterisks, right-hand scale).

(G.) Model for the effect of Sub1 and RNAPII activity on start site selection and splicing. Fast RNAPII mutations (Class I) result in upstream transcription start and diminished splicing efficiency, whereas *sub1*Δ or slow RNAPII mutations (Class II) shift transcription start downstream and enhance splicing. See also Figure S6, Figure S7, Table S4, Table S5, Table S6.

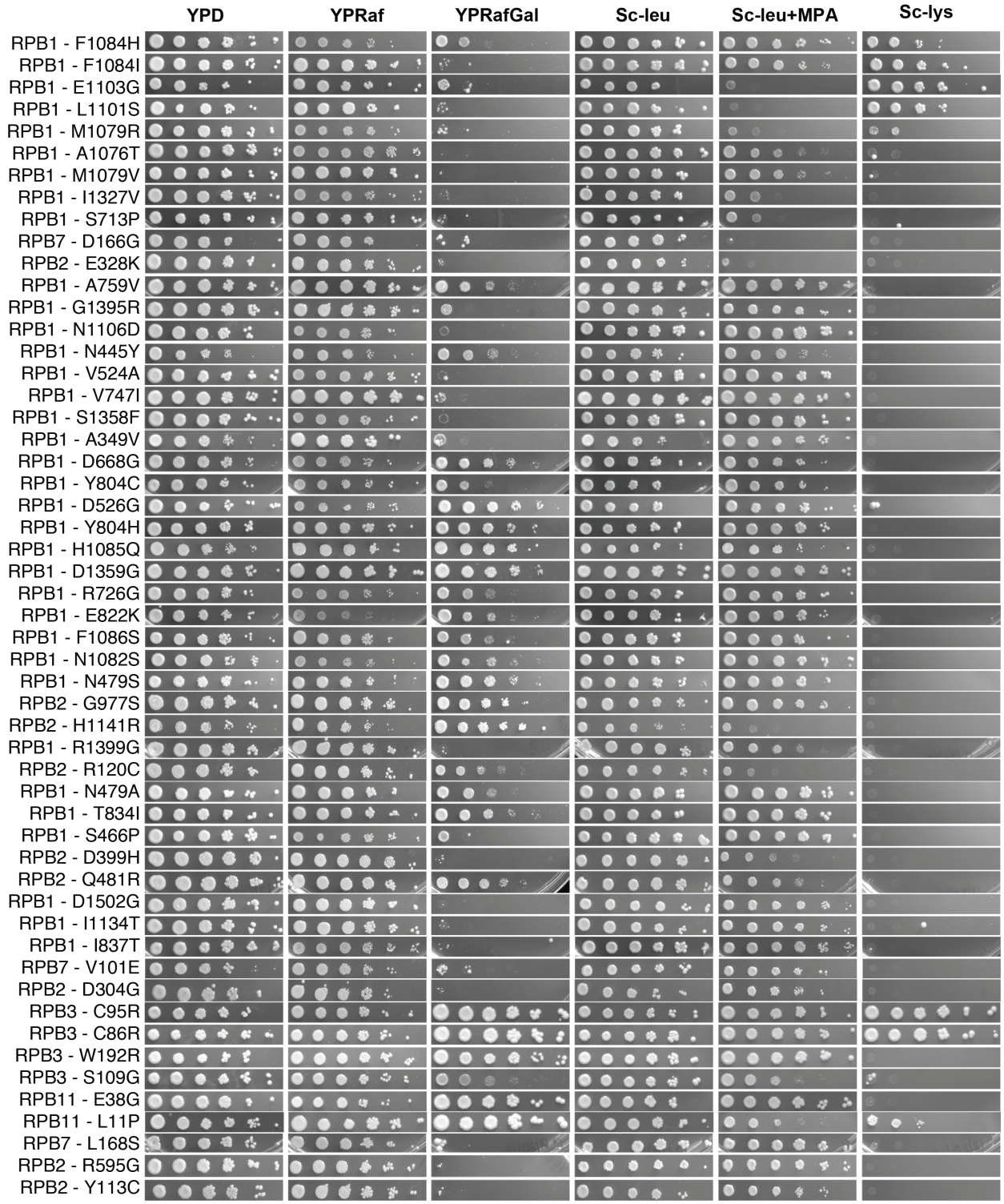

YPD
YPRaf
YPRafGal
Sc-leu
Sc-leu+MPA
Sc-lys
RPB1 - F1084H
RPB1 - F1084I
RPB1 - E1103G
RPB1 - L1101S
RPB1 - M1079R
RPB1 - A1076T
RPB1 - M1079V
RPB1 - I1327V
RPB1 - S713P
RPB7 - D166G
RPB2 - E328K
RPB1 - A759V
RPB1 - G1395R
RPB1 - N1106D
RPB1 - N445Y
RPB1 - V524A
RPB1 - V747I
RPB1 - S1358F
RPB1 - A349V
RPB1 - D668G
RPB1 - Y804C
RPB1 - D526G
RPB1 - Y804H
RPB1 - H1085Q
RPB1 - D1359G
RPB1 - R726G
RPB1 - E822K
RPB1 - F1086S
RPB1 - N1082S
RPB1 - N479S
RPB2 - G977S
RPB2 - H1141R
RPB1 - R1399G
RPB2 - R120C
RPB1 - N479A
RPB1 - T834I
RPB1 - S466P
RPB2 - D399H
RPB2 - Q481R
RPB1 - D1502G
RPB1 - I1134T
RPB1 - I837T
RPB7 - V101E
RPB2 - D304G
RPB3 - C95R
RPB3 - C86R
RPB3 - W192R
RPB3 - S109G
RPB11 - E38G
RPB11 - L11P
RPB7 - L168S
RPB2 - R595G
RPB2 - Y113C

Figure S1. Complete collection of spot tests for identification of Gal^R, MPA sensitivity and Spt^- phenotypes. Related to Figure 1.

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

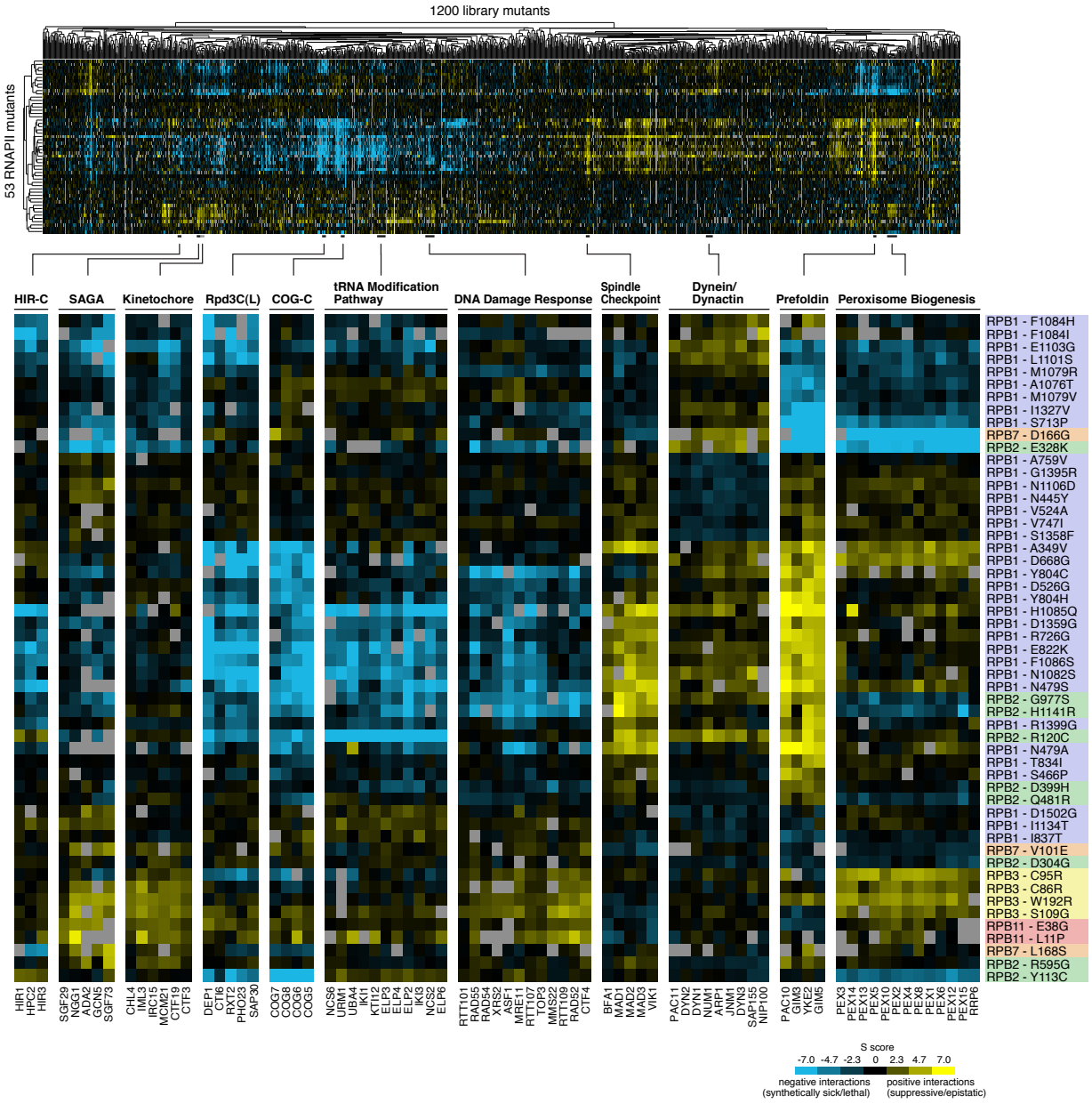

1200 library mutants
53 RNAPII mutants
HIR-C
SAGA
Kinetochore
Rpd3C(L)
COG-C
tRNA Modification Pathway
DNA Damage Response
Spindle Checkpoint
Dynein/ Dynactin
Prefoldin
Peroxisome Biogenesis
RPB1 - F1084H
RPB1 - F1084I
RPB1 - E1103G
RPB1 - L1101S
RPB1 - M1079R
RPB1 - A1076T
RPB1 - M1079V
RPB1 - I1327V
RPB1 - S713P
RPB7 - D166G
RPB2 - E328K
RPB1 - A759V
RPB1 - G1395R
RPB1 - N1106D
RPB1 - N445Y
RPB1 - V524A
RPB1 - V747I
RPB1 - S1358F
RPB1 - A349V
RPB1 - D668G
RPB1 - Y804C
RPB1 - D526G
RPB1 - Y804H
RPB1 - H1085Q
RPB1 - D1359G
RPB1 - R726G
RPB1 - E822K
RPB1 - F1086S
RPB1 - N1082S
RPB1 - N479S
RPB2 - G977S
RPB2 - H1141R
RPB1 - R1399G
RPB2 - R120C
RPB1 - N479A
RPB1 - T834I
RPB1 - S466P
RPB2 - D399H
RPB2 - Q481R
RPB1 - D1502G
RPB1 - I1134T
RPB1 - I837T
RPB7 - V101E
RPB2 - D304G
RPB3 - C95R
RPB3 - C86R
RPB3 - W192R
RPB3 - S109G
RPB11 - E38G
RPB11 - L11P
RPB7 - L168S
RPB2 - R595G
RPB2 - Y113C
HIR1 HPC2 HIR3
SGF29 NGG1 ADA2 GCN5 SGF73
CHL4 IML3 IRC15 MCM21 CTF19 CTF3
DEP1 CTI6 RXT2 PHO23 SAP30
COG7 COG8 COG6 COG5
NCS6 URM1 UBA4 IKI1 KTI12 ELP3 ELP4 ELP2 IKI3 NCS2 ELP6
RTT101 RAD55 RAD54 ASF1 ASF2 MRE11 RTT107 TOP3 MMS22 RTT109 RAD52 CTF4
BFA1 MAD1 MAD2 MAD3 VIK1
PAC11 DYN2 DYN1 NUM1 ARP1 JNM1 DYN3 SAP155 NIP100
PAC10 GIM3 YKE2 GIM5
PEX3 PEX14 PEX13 PEX5 PEX10 PEX2 PEX4 PEX8 PEX1 PEX6 PEX12 PEX15 RRP6
S score
-7.0 -4.7 -2.3 0 2.3 4.7 7.0
negative interactions (synthetically sick/lethal)
positive interactions (suppressive/epistatic)

Figure S2. A high-resolution genetic interaction map of RNAPII point mutants. Related to Figure 2.

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

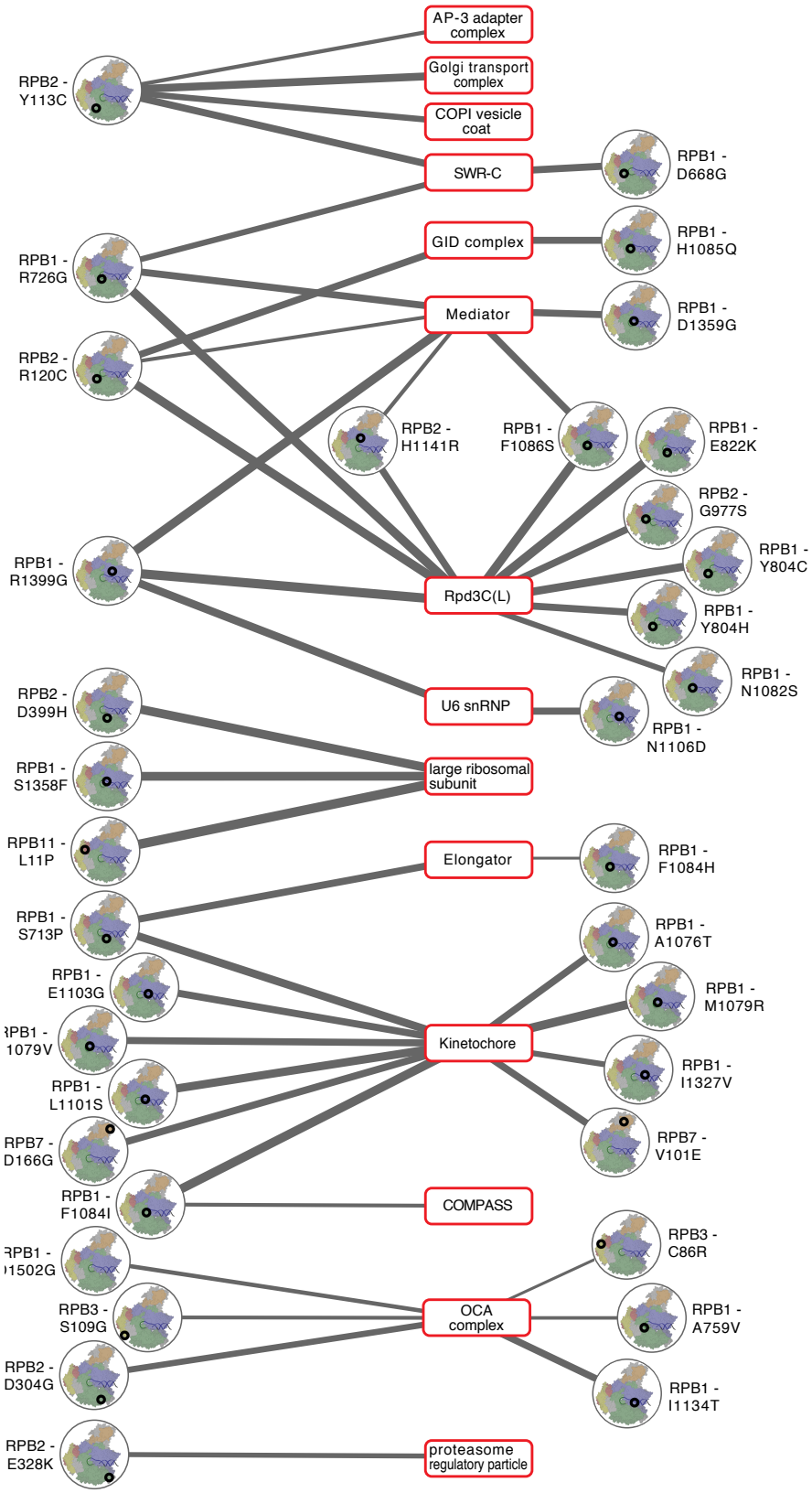
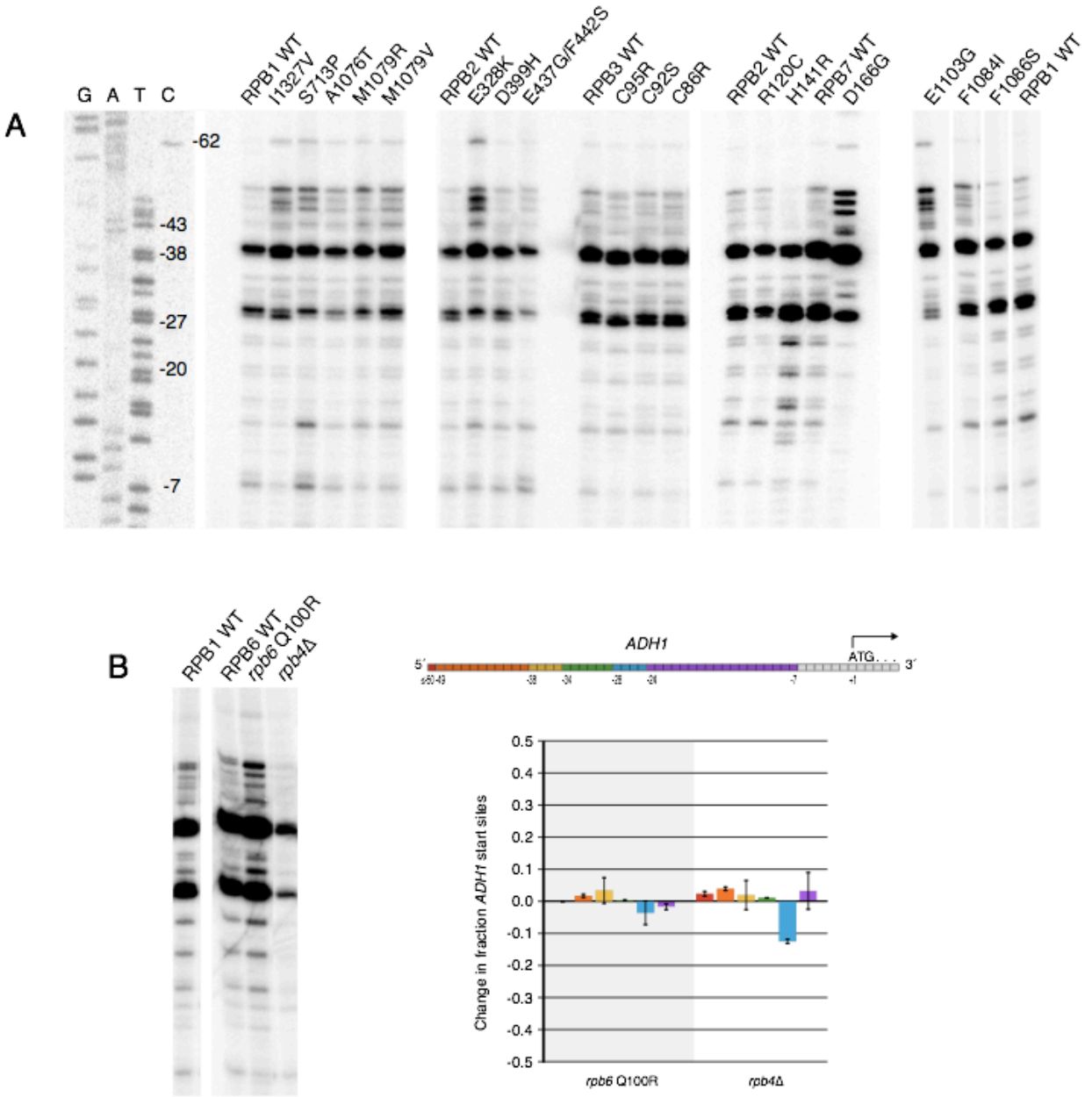

AP-3 adapter complex
Golgi transport complex
COPI vesicle coat
SWR-C
GID complex
Mediator
Rpd3C(L)
U6 snRNP
large ribosomal subunit
Elongator
Kinetochore
COMPASS
OCA complex
proteasome regulatory particle
RPB2 - Y113C
RPB1 - R726G
RPB2 - R120C
RPB1 - R1399G
RPB2 - D399H
RPB1 - S1358F
RPB11 - L11P
RPB1 - S713P
RPB1 - E1103G
RPB1 - 1079V
RPB1 - L1101S
RPB7 - D166G
RPB1 - F1084I
RPB1 - 1502G
RPB3 - S109G
RPB2 - D304G
RPB2 - E328K
RPB1 - D668G
RPB1 - H1085Q
RPB1 - D1359G
RPB2 - H1141R
RPB1 - F1086S
RPB1 - E822K
RPB2 - G977S
RPB1 - Y804C
RPB1 - Y804H
RPB1 - N1082S
RPB1 - N1106D
RPB1 - F1084H
RPB1 - A1076T
RPB1 - M1079R
RPB1 - I1327V
RPB7 - V101E
RPB3 - C86R
RPB1 - A759V
RPB1 - I1134T

Figure S3. Functional connections between RNAPII mutants and protein complexes. Related to Figure 3.

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


A
G A T C
RPB1 WT
I1327V
S713P
A1076T
M1079R
M1079V
RPB2 WT
E328K
D399H
E437G/F442S
RPB3 WT
C95R
C92S
C86R
RPB2 WT
R120C
H141R
RPB7 WT
D166G
E1103G
F1084I
F1086S
RPB1 WT
-62
-43
-38
-27
-20
-7
B
RPB1 WT
RPB6 WT
rpb6 Q100R
rpb4Δ
ADH1
5′
ATG. . .
3′
Change in fraction ADH1 start sites
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
rpb6 Q100R
rpb4Δ

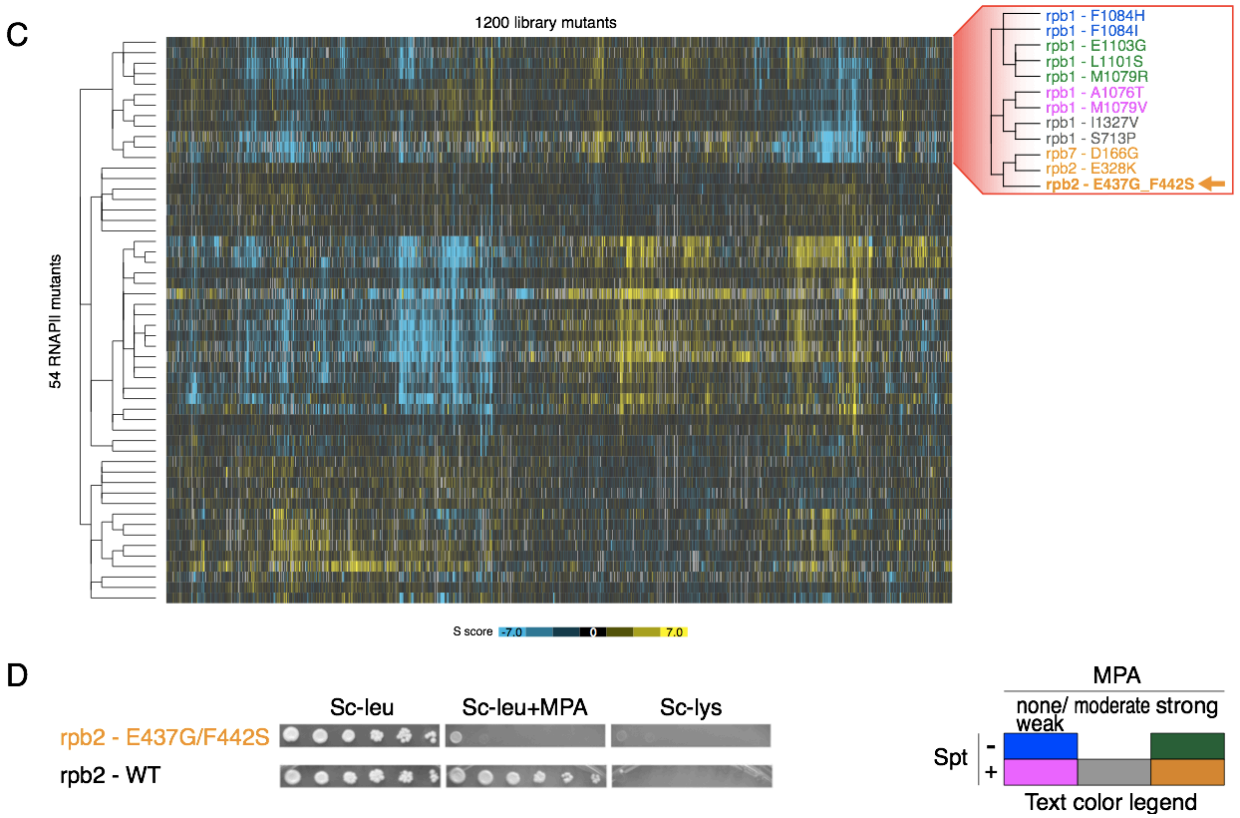


Figure S4. Primer extension analysis at *ADH1* to identify the effect of RNAPII mutations on start-site selection, and data relating to *rpb2* E437G/F442S. Related to Figure 4.

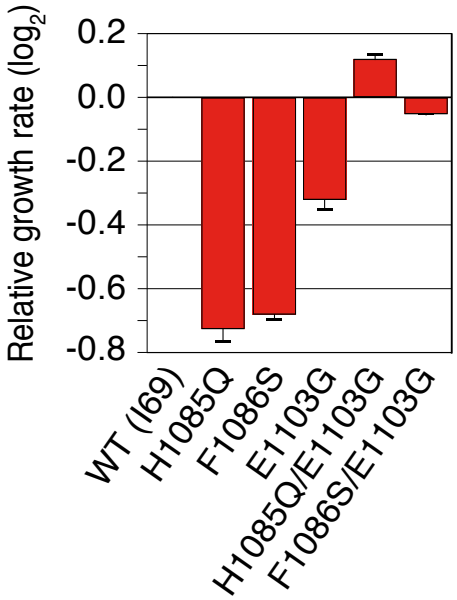
(A.) Sequencing gel for primer extension analysis. Lanes 1-4 correspond to Sanger dideoxy sequencing reactions as reference ladders, and the following lanes carry reverse transcription products from RNAPII alleles as indicated (Table S4).

(B.) Primer extension analysis at *ADH1* to map transcription start sites for *rpb4*Δ and *rpb6* Q100R mutations. The colors of the bars correspond to the sequence windows indicated in the *ADH1* schematic (top) and the heights specify the fraction change of transcription start in the mutant compared to WT. Means and standard deviations were derived from three independent experiments.

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

(D.) MPA sensitivity and Spt$^-$ phenotypes of *rpb2* E437G/F442S.

A



B

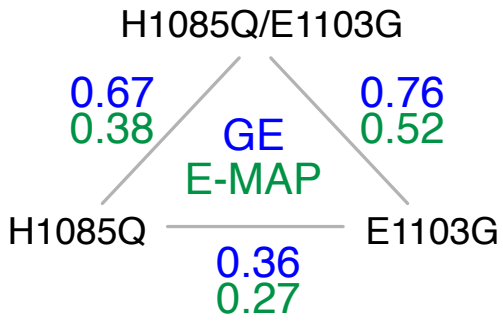
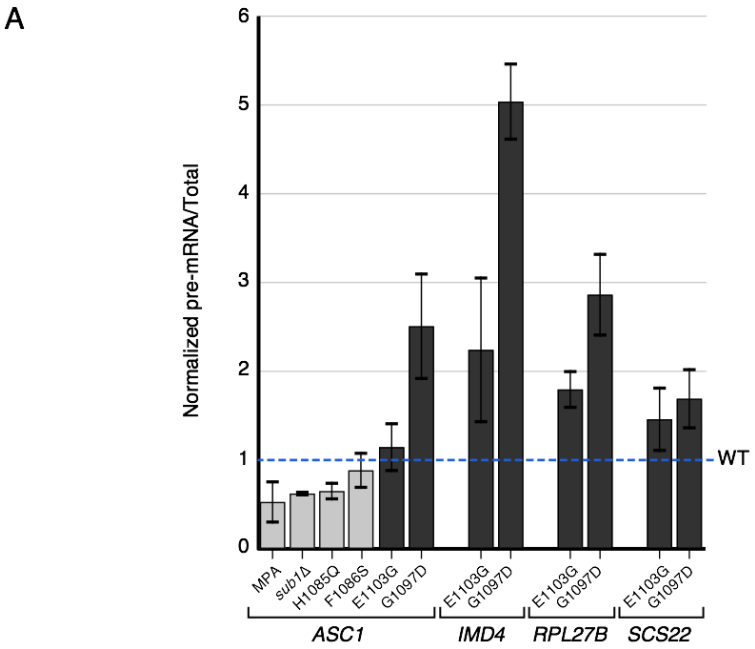
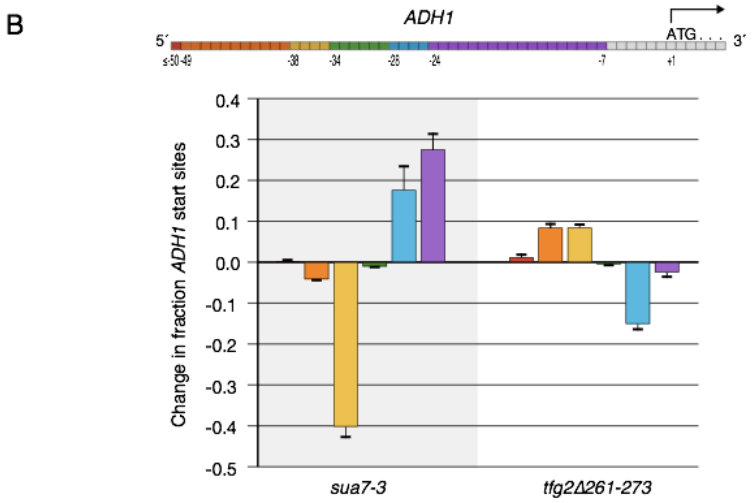


Figure S5. RNAPII trigger loop double and single mutant growth rates, and profile correlations between *rpb1* E1103G, H1085Q and E1103G/H1085Q. Related to Figure 5.

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

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


A
Normalized pre-mRNA/Total
6
5
4
3
2
1
0
WT
MPA
sub1Δ
H1085Q
F1086S
E1103G
G1097D
E1103G
G1097D
E1103G
G1097D
E1103G
G1097D
ASC1
IMD4
RPL27B
SCS22


B
ADH1
5´
ATG . . .
3´
Change in fraction ADH1 start sites
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
sua7-3
tfg2Δ261-273

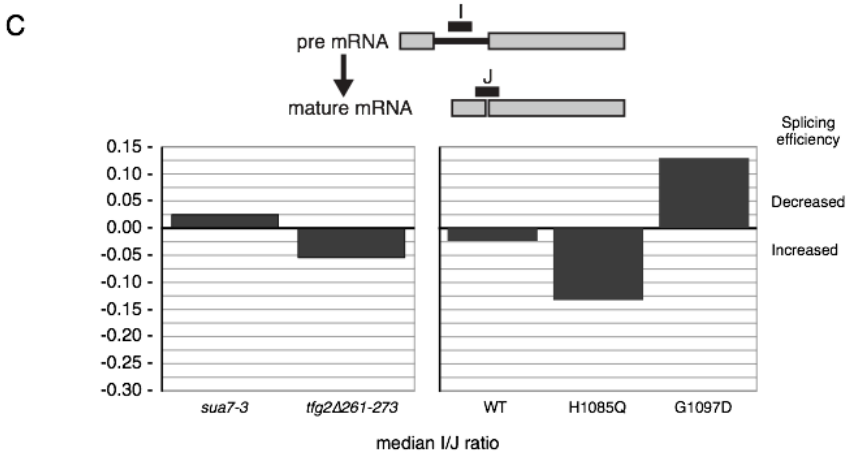

C
I
pre mRNA
J
mature mRNA
Splicing efficiency
Decreased
Increased
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
-0.20
-0.25
-0.30
sua7-3
tfg2Δ261-273
WT
H1085Q
G1097D
median I/J ratio

Figure S6. Confirmation of RNAPII single mutant splicing phenotypes and interrogation of potential connection between splicing and start site selection. Related to Figure 6 and Figure 7.

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

(B.) Primer extension analysis at *ADH1* to map transcription start sites for mutations in TFIIB (*sua7-3*) and TFIIF (*tfg2*Δ*261-273*). The colors of the bars correspond to the sequence windows indicated in the *ADH1* schematic (top) and the heights specify the fraction change of transcription start in the mutant compared to WT. Means and standard deviations were derived from three independent experiments.

(C.) TFIIB (*sua7-3*) and TFIIF (*tfg2*Δ*261-273*) mutants were analyzed on splicing microarrays, as described in Figure 6 (Table S5). Median I/J $\log_2$ ratios for the complete profiles of the TFII mutants are shown in the left panel, with WT, a slow and a fast *rpb1* mutant included for reference in the right panel.

Figure S7. Spot tests to determine Gal^R, MPA sensitivity and Spt$^-$ phenotypes of *sub1*Δ or other deletion mutants with fast and slow RNAPII alleles, and primer extension of s*ub1*Δ mutants. Related to Figure 7. Available online at www.cell.com.

TABLES

Available online at www.cell.com

Table S1. Strains used in this study and annotation of RNAPII and E-MAP library alleles. Related to Figure 1 and Figure 2.

Table S2. Lists of genetic interaction scores and gene expression changes, and additional data supporting Figure 2B. Related to Figure 2.

Table S3. Functional associations between RNAPII mutations and protein complexes, and results from chromosome transmission fidelity (CTF) assay. Related to Figure 3.

Table S4. Changes in transcription start-site selections from primer-extension analyses at *ADH1*. Related to Figure 4 and Figure 7.

Table S5. Results from splicing microarray and qPCR assays. Related to Figure 6 and Figure 7.

Table S6. pE-MAP profiles of library mutants, sorted on the difference between their average interaction with fast and slow RNAPII mutants. Related to Figure 7.

SUPPLEMENTARY DATA

Available online at www.cell.com

Supp. Data 1. pE-MAP clustered in two dimensions (Java TreeView format, http://jtreeview.sourceforge.net/). Related to Figure 2 and Figure 4.

Chapter 4

Adventures in Time and Space:

Splicing Efficiency and RNA Polymerase II Elongation Rate

Adventures in Time and Space:

Splicing Efficiency and RNA Polymerase II Elongation Rate

Erica A. Moehle[1], Hannes Braberg[2,3], Nevan J. Krogan[2,3,4], Christine Guthrie[1]

[1]Department of Biochemistry and Biophysics, University of California, San Francisco, CA, 94158, USA [2]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA, 94158, USA [3]California Institute for Quantitative Biosciences, QB3, San Francisco, CA, 94158, USA [4]J. David Gladstone Institutes, San Francisco, CA, 94158, USA

To whom correspondence should be addressed: christineguthrie@gmail.com

ABSTRACT

Control of pre-mRNA splicing is a critical part of the eukaryotic gene expression process. Extensive evidence indicates that transcription and splicing are spatiotemporally coordinated and that most splicing events occur co-transcriptionally. A kinetic coupling model has been proposed in metazoans to describe how changing RNA Polymerase II (RNAPII) elongation rate can impact which alternative splice sites are used. In *Saccharomyces cerevisiae*, in which most spliced genes have only a single intron and splice sites adhere to a strong consensus sequence, we recently observed that *splicing efficiency* was sensitive to mutations in RNAPII that increase or decrease its elongation rate. Our data revealed that RNAPII speed and splicing efficiency are generally anti-correlated: at many genes, increased elongation rate caused decreased splicing efficiency, while decreased elongation rate increased splicing efficiency. An improved splicing phenotype was also observed upon deletion of *SUB1*, a condition in which elongation rate is slowed. We discuss these data in the context of a growing field and expand the kinetic coupling model to apply to both alternative splicing and splicing efficiency.

Splicing control begins co-transcriptionally

Splicing is the process by which an intron is removed from a pre-mRNA transcript, and the flanking exons are ligated together. Removal of each intron proceeds via numerous consecutive steps, including commitment of a pre-mRNA to splicing by binding of splicing factors to the 5' splice site and the branch site/3' splice site, multiple conformational rearrangements, and two catalytic steps.[1,6] Splicing in metazoans is particularly complex, as most genes contain multiple introns flanked by weak splice sites, which can give rise to alternatively spliced transcripts. In *S. cerevisiae*, also known as budding yeast, most spliced genes have only a single intron defined by strong, constitutive (as opposed to alternative) splice sites.

While splicing can occur in a minimal *in vitro* system, independent of transcription, a major outstanding question in the field is how splicing proceeds *in vivo*. In both metazoans and budding yeast, splicing factors bind many transcripts while they are still associated with elongating RNA polymerase II (RNAPII), *i.e.*, co-transcriptionally. When assayed by chromatin immunoprecipitation (ChIP), a splicing factor co-immunoprecipitates DNA several hundred base pairs downstream of its encoded binding site, presumably via a nascent transcript and RNAPII.[7-11] In budding yeast, while many intron-containing genes (ICGs) are associated with early splicing factors,[11] it has proven difficult to measure association of later splicing factors, presumably due to short second (downstream) exons.[11,129] However, most transcripts are nonetheless fully spliced before they are released from the chromatin template by transcription termination,[34,35] as is also generally the case in metazoans.[230]

Several pieces of evidence have pointed to the functional importance of all or part of the splicing process occurring co-transcriptionally. Experiments in budding yeast showed that intron-containing reporter genes that recruit early-acting splicing factors co-transcriptionally also exhibit efficient splicing, as measured by the fraction of spliced versus unspliced mRNA.[129] Computational modeling of splicing and transcription using kinetic parameters determined from splicing reporters also predicts that the full splicing reaction is more efficient when occurring co-transcriptionally, as opposed to post-transcriptionally *i.e.*, after transcription termination.[138] Because the retention of an intron promotes transcript degradation via nonsense-mediated decay, thus abrogating its translation into protein, the efficiency of splicing occurs co-transcriptionally can have profound phenotypic consequences on the cell.

A specific functional connection between transcription by RNAPII and splicing has been referred to as "kinetic coupling," as the timing of one reaction influences the outcome of a second reaction. A kinetic coupling model was originally proposed in metazoa to describe how perturbing the RNAPII elongation rate can influence alternative splicing. Because the transcribed splice sites

emerge from RNAPII in the order defined by transcription (5' to 3') and with the timing defined by elongation rate, slowing RNAPII elongation in the vicinity of a weak upstream splice site affords more time for it to be recognized by splicing factors before a competing downstream splice site is transcribed.[28] In eukaryotes, transcription occurs in the context of nucleosome-bound DNA, or chromatin. The chromatin environment at any given locus is complex and is defined by variables such as post-translational modifications of histones, nucleosome positioning, and nucleosome remodeling, that together affect RNAPII elongation rate in gene- and intron-specific ways. It is likely because of the chromatin-RNAPII elongation connection that factors such as histone acetylation and nucleosome density are correlated with alternative splicing patterns.[150]

Earlier experiments in budding yeast using a two-intron gene in which the splice sites have been made artificially weak showed that kinetic coupling is not a phenomenon restricted to metazoans. These experiments revealed that slowing RNAPII elongation favored the use of a weaker upstream splice site,[29] as was also observed in metazoans. More generally, while the number of ICGs in budding yeast is only several hundred of ~6,000, the gene ontology categories enriched within this list *a priori* suggest that splicing may have been retained in this species as a regulatory node within the broader gene expression process.

Indeed, the efficiency of the splicing reaction has been shown to be regulated in some cases in which resulting protein levels are affected. For example, several meiotic transcripts are spliced only during progression through meiosis.[2,3] In addition, splicing efficiency of ribosomal protein gene transcripts rapidly decreases upon amino acid starvation, leading to the intron being retained in these transcripts.[4,5] Taken together, these and other data argue that studies of splicing efficiency in budding yeast offer a useful system for investigating the nature of kinetic coupling.

Budding yeast as a system offers an unrivalled opportunity to combine genetics and whole-genome splicing analyses to ask whether, and if so, how the efficiency of the splicing reaction is kinetically coupled to transcription speed. We were able to investigate multiple spatiotemporal facets

of kinetic coupling – not only by directly modulating transcription elongation rate, but also using high-throughput genetic interaction mapping to begin to interrogate aspects of the co-transcriptional environment that impact kinetic coupling.

Splicing efficiency is sensitive to RNAPII elongation rate

We previously examined splicing in the context of an allelic series of point mutants in the active site of RNAPII[31] known to transcribe *in vitro*[174] at ~1.5nt/sec (*rpb1*-H1085Q), ~5nt/sec (*rpb1*-F1086S), ~30nt/sec (*rpb1*-E1103G), and >50nt/sec (*rpb1*-G1097D), compared to ~12nt/sec for wild-type RNAPII. We employed splicing-specific microarrays[5] to measure the effect of these RNAPII mutants on the abundance of the intron, exon and exon-exon junction from intron-containing genes. Mutations in splicing factors tend to exhibit large increases in the abundance of intron, with concomitant decreases in the junction;[92] such a readout is interpreted to mean that the pre-mRNA is spliced less efficiently.

In a strain with the E1103G mutation, which causes a moderately fast elongation rate *in vitro*,[174,193] we observed a phenotype consistent with a decrease in splicing efficiency: many genes exhibited an increase in the intron feature (reflecting an accumulation of pre-mRNA), intron/junction ratio (indicating increased pre-mRNA relative to mature mRNA), and intron/exon (indicating accumulation in pre-mRNA relative to overall expression changes).[31] A strain harboring the G1097D point mutation,[31] which causes an even faster elongation rate *in vitro*,[174] exhibited an even more extreme splicing defect.

These results were consistent with a model in which splice sites are less efficiently utilized in strains with a fast RNAPII elongation rate, and accordingly, we suspected that strains with a slower elongation rate would exhibit improved splicing. Indeed, when we analyzed splicing in a strain with the F1086S mutation, which caused RNAPII elongation to be moderately slow *in vitro,*[192] we observed decreases in the intron feature, intron/junction ratio, and intron/exon ratio at a large number

of genes, all of which are consistent with increased splicing efficiency. We observed an even higher efficiency of splicing in cells harboring the H1085Q mutation in RNAPII,[31] which elongates at only ~1.5nt/sec *in vitro.*[174]

To confirm that these effects were the result of changing RNAPII speed rather than being an indirect consequence of the specific point mutations used, we analyzed splicing in two double mutant strains, in which mutations that individually cause RNAPII to elongate quickly or slowly *in vitro* together return RNAPII elongation rate to very near wild-type rate. Splicing in these strains was very similar to that in a wild-type strain, further arguing for a direct relationship between elongation rate and splicing efficiency.

Interestingly, Khodor *et al.* showed in *Drosophila melanogaster* that splicing efficiency of some introns improves upon expression of a mutant RNAPII with lower processivity,[231] a result that is consistent with ours. Our study represents the first instance where an RNAPII genetically engineered to elongate at a range of speeds was assayed with respect to its genome-wide effects on splicing *in vivo*.

In a general sense, the opposite splicing phenotypes that occur in strains with fast versus slow RNAPII elongation suggest that, although transcripts do not undergo alternative splicing in budding yeast, splicing efficiency is nonetheless kinetically coupled to transcription (Figure 1). Importantly, the effects we observed on splicing were most extreme in the RNAPII mutants with the fastest and slowest elongation rates, arguing that kinetic coupling occurs in budding yeast over a ~40-fold range of elongation rates. Our results support the notion that splicing efficiency in yeast (*i.e.*, degree of intron retention), is a *bona fide* readout of kinetic coupling.

How might kinetic coupling occur in *S. cerevisiae*?

Is splicing in competition with termination?

As RNAPII approaches the 3' end of each gene, nascent transcripts are cleaved from RNAPII and polyadenylated, in a process that is tightly coupled with transcription termination.[232] In systems where alternative splicing takes place, altered RNAPII elongation rate impacts the temporal window of opportunity for splicing using upstream versus downstream splice sites; in the case of budding yeast, it logically follows that co-transcriptional splicing is instead in competition with transcription termination/3' end cleavage. Consistent with this notion, it was concluded based on sequencing of chromatin-associated transcripts and RNAPII ChIP that genes with high levels of co-transcriptional splicing are those at which RNAPII slows near the 3' end.[34] This was interpreted as evidence for a mechanism to promote co-transcriptional splicing catalysis by delaying transcription termination.

These same experiments revealed that in a wild-type cell, half of intron-containing genes are >74% spliced while chromatin associated.[34] Despite this abundance of co-transcriptional splicing, this number implies that completion of splicing is not a prerequisite for release from the 3' end. While we observed changes in splicing efficiency using whole cell RNA as input,[31] we might predict that changes in RNAPII elongation rate may alter the fraction of transcripts that are fully spliced co-transcriptionally (Figure 1). The change in this fraction upon alteration of RNAPII elongation rate cannot be determined on a genome-wide scale without an initial enrichment for nascent transcripts, and thus poses an important question for the future.

Experiments on splicing of terminal introns in metazoa have revealed that splicing and 3' end processing may be coordinated.[233-236] This could explain why mutants in the 3' end processing and termination machinery tend to cause splicing defects.[19,237,238] However, if splicing and termination are in direct competition, it should be possible to identify mutants in the transcription termination machinery that would allow more time for splicing completion, in the same way a slow RNAPII mutant does, but without interfering with the coupling between those processes. Furthermore, experiments could be designed to test whether modulating the length of the last exon in yeast would alter the proportion of transcripts that are spliced prior to transcription termination.

As noted above, it has been suggested that co-transcriptional commitment to splicing increases the likelihood that the transcript will be spliced. Therefore, an additional possibility is that alterations in RNAPII elongation rate will change the *location within the gene* at which splicing factor association occurs: nascent transcripts from the faster RNAPII mutants would exhibit lower splicing factor occupancy by ChIP, or a 3' shift in association, whereas splicing factors in a slow RNAPII mutant would associate with nascent transcripts more 5' (Figure 1).

Why are the effects on splicing gene-specific?

Our conclusion that transcription rate and splicing efficiency are generally anti-correlated was based on a sizeable subset of genes we term "kinetically coupled," whose splicing was hindered with an RNAPII mutant with increased elongation rate and enhanced with an RNAPII mutant with decreased elongation rate. Notably, however, not all genes conformed to this trend, provoking the question of what the difference is between genes that responded reciprocally to RNAPII speed and those that did not.

The metazoan model of kinetic coupling would suggest that this gene specificity could come from differences in splice site strength. However, unlike in metazoans, ICGs in budding yeast generally have very strong splice sites that do not diverge far from the consensus; in any case, the introns whose splicing were most kinetically coupled to RNAPII elongation were not enriched for those bearing weak splice sites (data not shown).

In trying to understand the nature of gene-specific effects of RNAPII elongation rate on splicing efficiency, it is important to note that wild-type RNAPII rate likely varies from gene to gene to begin with. At present, assays for measuring genome-wide RNAPII elongation rate in live budding yeast are lacking. The E1103G mutation confers a fast elongation rate *in vivo* on the *GAL1* gene,[217] thus strengthening the prediction that *in vivo* rates should generally mirror those measured *in vitro*. If we could directly measure gene-specific elongation rates *in vivo*, perhaps the kinetically coupled genes would be the only genes at which RNAPII elongation rate is altered in the direction and to the

degree expected based on *in vitro* elongation measurements. Alternatively, perhaps the genes that did not respond reciprocally to changes in RNAPII rate have an epistatic mechanism for maintaining normal RNAPII elongation rate even in the context of the RNAPII mutations.

An important aspect of the integration between transcription and splicing is doubtless the chromatin environment in which both processes occur. In fact, it has long been understood that the chromatin environment - nucleosome distribution, histone modifications, and transcription factors more generally - confers gene-specific regulation of transcription initiation, elongation, and termination. Thus, an intriguing possibility is that the kinetically coupled genes we identified exist within chromatin environments that promote co-transcriptional spliceosome assembly and/or catalysis. Direct physical recruitment of splicing factors to histones bearing specific modifications has been reported in metazoans.[99,149] Factors in the chromatin environment may directly or indirectly promote spliceosomal rearrangements that lead to a productive splicing reaction, as proposed in budding yeast for the cycle of histone acetylation and deacetylation.[15,105]

Additional histone post-translational modifiers in yeast[15,18,26,105] and metazoa[126,145,239] affect splicing via mechanisms that remain poorly understood. These modifiers could act directly, as in the above examples, or indirectly, via altering RNAPII elongation in the vicinity of encoded splice sites. This latter example has been observed in a variety of scenarios, including histone acetylation level, positioning of nucleosomes, and slowing of RNAPII due to localized DNA damage.[127]

Identifying transcription rate-sensitive factors

An important experimental goal is to identify additional factors that regulate RNAPII elongation rate. In our study, we relied on the power of yeast genetics and performed high-throughput genetic interaction mapping, in which double-mutant strains were generated by crossing a given RNAPII elongation rate mutant to an additional strain carrying a mutation in one of 1,200 genes. The growth of the resulting double-mutant strains was then measured as a proxy for fitness. Two gene subsets of interest were identified by this analysis: (1) those whose mutation suppressed

the growth defect of the fast RNAPII mutants and exacerbated the growth defect of the slow mutants; (2) vice versa (mutation exacerbated growth in fast mutants and suppressed the growth defect of slow mutants).[31] We proposed that these factors may either be sensitive to elongation rate, or might themselves directly influence RNAPII rate. We reasoned that this genetic dataset could provide a list of candidate factors that influence RNAPII dynamics, and therefore splicing, in site- or gene-specific ways. We discuss some of these candidates below.

Transcription factors

We identified two members of the PAF-Complex (Rtf1 and Cdc73) whose deletion adversely impacted growth in strains harboring a slow RNAPII mutant, but suppressed the growth defect of the fast RNAPII mutants.[31] These genetic data are consistent with the known role of the PAF-C as a positively acting transcription factor that coordinates many events during elongation,[240] and led us to suspect that strains lacking this complex would have slow elongation and improved splicing. As it turns out, however, *rtf1*Δ and *cdc73*Δ strains do not exhibit slow RNAPII elongation.[196] And interestingly, two separate reports have shown that the PAF-C actually promotes splicing, as both observe data consistent with a splicing defect in strains lacking a functional PAF-C.[19,30]

An additional factor, like members of the PAF-C, whose deletion caused synthetic lethality with the slow RNAPII alleles is Sub1, a gene so named due to its role in transcription initiation as a Suppressor of TFIIB.[199] These genetic data prompted us to hypothesize that the *sub1*Δ strain might exhibit a slow elongation rate, which was confirmed by Garcia *et al.* while the work by Braberg *et al.* was under review: RNAPII elongates slowly on the *GAL1* gene in the absence of *SUB1*.[198] Consistent with our predictions, the *sub1*Δ strain exhibited improved splicing at a large number of genes, similar to what is observed in slow RNAPII mutant strains. It was previously reported that Sub1 localizes to ribosomal protein genes[202] and consistent with this, the genes whose splicing is improved by deletion of *SUB1* are enriched for genes encoding the cytosolic ribosome (data not shown). Therefore, the gene-specificity conferred by Sub1 localization may account for its intron specificity.

Histone modifiers

Histone acetylation has been shown to promote transcription through nucleosomal templates,[241] a function opposed by histone deacetylation; therefore, complexes harboring these histone acetyltransferases (HATs) and deacetylases (HDACs) are attractive candidates for gene-specific regulators of RNAPII elongation rate and splicing efficiency. In fact, connections between histone acetylation and splicing have previously been observed.[127] Our genetic interaction mapping identified *SAS3*, *SAS4* and *ESA1* (components of NuA3, SAS, and NuA4 histone acetyltransferase complexes, respectively) as causing synthetic lethality in combination with the fast RNAPII point mutant alleles. These data are tantalizing, but to the best of our knowledge, these complexes have not been examined with respect to their gene-specific effects on splicing.

The existence of gene-specific histone modifiers and RNAPII elongation regulators raises the possibility that co-transcriptional splicing is regulated in a gene-specific fashion, depending on the nature of the co-transcriptional environment at each gene. In metazoans, this specificity has been proposed to occur at specific introns within genes.[150]

Current questions and future directions

Many questions remain about how splicing is regulated co-transcriptionally, and additional factors, not discussed above, are clearly also at play. For instance, not all splicing factors promote splicing upon binding a pre-mRNA - auxiliary splicing regulatory factors can also inhibit the use of specific splice sites. Thus, while the co-transcriptional nature of splicing has been generally thought to *improve* the recognition of transcribed splice sites to either modulate alternative splicing[195] or increase splicing efficiency,[129,138] perhaps ensuring that splicing occurs co-transcriptionally allows for broader *regulation* from the chromatin environment.

Because completion of splicing often occurs co-transcriptionally, the quality control mechanisms required for splicing fidelity, *i.e.*, the accurate recognition of exon/intron junctions in

pre-mRNA,[1] must perforce operate co-transcriptionally as well. This raises the question of whether there is input from RNAPII or other co-transcriptional elements that promotes splicing fidelity. An transcription–splicing checkpoint has been proposed, during which RNAPII pauses briefly at the 3' splice site, and which may promote splicing fidelity.[36] Furthermore, the cycle of histone acetylation and deacetylation has been associated with specific spliceosomal rearrangements,[105] and leaves open the possibility that each step in splicing is associated with a chromatin-specific quality control mechanism. In strains where RNAPII elongation is too fast or too slow, do these quality control steps occur faithfully? We are currently testing whether splicing fidelity is improved or hampered in strains harboring fast or slow RNAPII mutant alleles.

Early experiments showed that mRNA secondary structure can impact alternative splicing via occlusion of alternative splice sites.[242] RNA secondary structure may be stabilized by kinetic trapping during transcription,[243] and therefore may be sensitive to RNAPII elongation rate. The degree to which RNA secondary structure changes in the RNAPII point mutant strains would be of great interest to determine.

It also remains unknown whether the intron accumulation we observed by microarray in strains with fast RNAPII elongation[31] is attributable to pre-mRNAs on which splicing factors have not acted, or to stalled splicing events that might have been released as splicing intermediate products, destined to be discarded from the spliceosome and degraded. In the future, the development of genome-wide assays to analyze the multiple steps in splicing – not only commitment to splicing, but both steps of splicing catalysis, will be critical for studying co-transcriptional splicing regulation. Furthermore, splicing has been called a co-transcriptional process, with the implication that it takes place prior to transcription termination.[230] In some cases, transcripts are retained at the 3' end of a chromatin locus after transcription termination,[244-246] and therefore are still potentially subject to the same regulatory processes as during transcription elongation. With recent advances in chromatin-

associated and RNAPII-associated transcript sequencing,[34,35,108,247] the way co-transcriptionality is operationally defined may change.

That splicing and transcription occur in the same space and time has allowed for functional coupling to occur. This has been revealed in experiments by us and others in which RNAPII elongation is directly or indirectly altered; this raises the possibility that the cell utilizes this coupling to maintain an efficient gene expression program or rapidly adapt this program in response to environmental variables. More generally, as additional factors that influence splicing efficiency are defined, and assays to better characterize the timing and nature of the co-transcriptional splicing reaction are developed, we anticipate that a more integrated picture of how individual steps in gene expression function together will emerge.

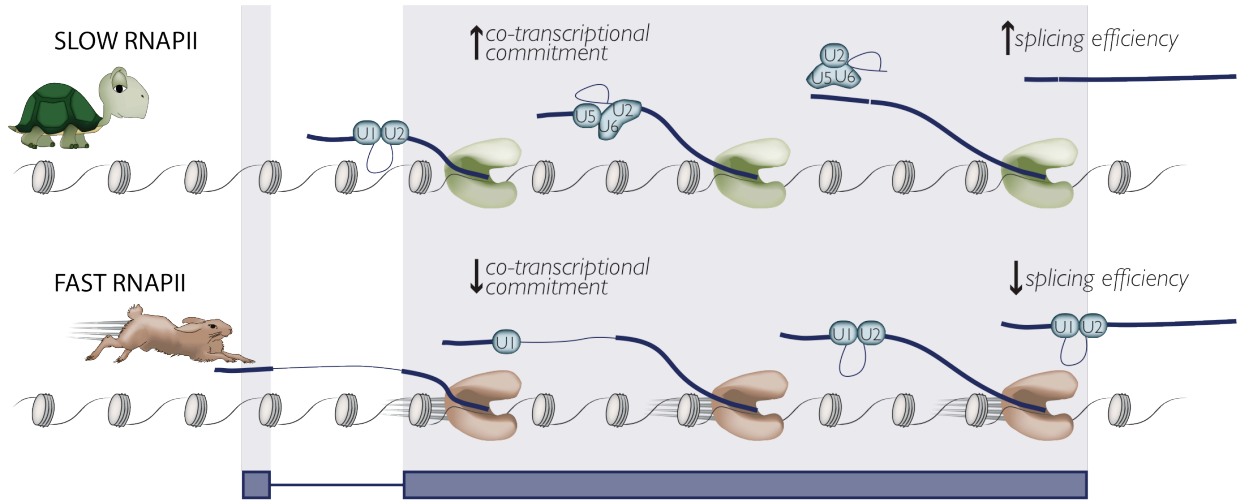


Figure 1. Hypothetical model of how RNAPII elongation rate impacts splicing efficiency.

See text for details.

EPILOGUE

The textbook view of gene expression presents transcription, pre-mRNA processing, export, and translation as independent, consecutive events; this is what I was taught as dogma in college. But in fact, this view is outdated, and all steps in the gene expression program engage in crosstalk at multiple levels with each other. For example, Crick's classic statement that "once information gets into proteins it cannot get back into the DNA"[248] has been modified by discoveries in the 1990s and 2000s of chromatin modifying and remodeling enzymes that alter the functional state of the DNA.

The general idea of crosstalk and integration in gene expression was very much in the air in 2008 when I joined the Guthrie lab. There were many interesting and tantalizing observations, including those made by members of the Guthrie lab, that applied to splicing, but it felt like the tip of the iceberg. We now know this was very much the case: activities of the transcription, chromatin, and splicing machineries are closely spatiotemporally coordinated in the cell. What began as individual glimpses into this coordination 6 years ago has now become a deluge of individual connections between splicing and the co-transcriptional environment. The contribution my work has made in this regard is by revealing specific connections between pre-mRNA splicing and modification of chromatin, and with RNAPII speed. While there is, as of yet, no "central dogma of molecular biology version 2.0" as it applies to splicing, we have a much clearer view of what its components will be.

Histone modifications

Tools of yeast genetics allow comprehensive mapping of the genetic (and, by extension, functional) interaction space of a given protein. This analysis for Npl3, performed by Tracy Kress, revealed a wealth of chromatin and transcription factors that exhibit both positive and negative genetic interactions with *npl3*Δ.[26] This finding strengthens our view of this protein as a major co-transcriptional player in budding yeast. While Npl3 does promote co-transcriptional spliceosome assembly,[24] it was unreasonable to expect that all its genetic interaction partners affect that step and, in fact, targeted experiments with several of them showed that splicing was not affected when these

factors were deleted (*chd1*Δ, *eaf3*Δ, data not shown). In retrospect, our decision to focus on the interplay between Npl3 and the H2B ubiquitination machinery proved a productive one, especially considering our focus on splicing. Consistent with the synthetic lethality of alleles that abrogate the this mark (*htb1K123R*, *bre1*Δ, *lge1*Δ, *rad6*Δ and *paf1*Δ) with a factor that promotes splicing (*NPL3*), we showed that a strain lacking H2B ubiquitination due to a mutation in the relevant residue (*ht1b1K123R*) has a mild but reproducible defect in splicing and in the association of U1 and U2 with intron-containing transcripts (Chapters 1 and 2).[26,27] Furthermore, we observed negative genetic interactions between strains lacking H2B ubiquitination and deletions of splicing factors.[26] The sum total of these data are consistent with the notion that H2B ubiquitination promotes splicing, and does this by promoting early association of the spliceosome with pre-mRNAs.

Interestingly, this splicing defect caused by loss of H2B ubiquitination increases in severity with high temperatures; despite the fact that we could observe a splicing defect in a *bre1*Δ strain at 37˚C, we were nonetheless only able to observe a decrease in recruitment of U1 and U2 after a shift to 39˚C for 3 hours, *i.e.*, the conditions used in Chapter 2.[27] That spliceosome association as measured by ChIP is not the more sensitive assay for studying mild effects on splicing was surprising (see below).

Is spliceosome assembly on nascent transcripts promoted by ubiquitinated H2B itself or, alternatively, by the cycle of ubiquitination and de-ubiquitination? That deletion of *UBP8*, the H2B ubiquitin protease, exacerbated the splicing defect of *npl3*Δ, suggests the entire cycle promotes splicing. The ubiquitin moiety is a bulky modification, far larger than, for example, tri-methyl groups that create specific binding surfaces for chromatin-associating factors.[249] For that reason it does not seem well suited to be a landing pad that will directly recruit splicing factors to an actively transcribed locus. Rather, its role in changing the dynamics of RNAPII passage through nucleosomes seems a promising direction for model building. Results from biophysical studies on nucleosomes and nucleosomal arrays with ubiquitinated H2B have yielded conflicting results, with some showing

that ubiquitination increases chromatin compaction[151,250] and one suggesting it leads to decondensation.[251] Either scenario has the potential to impact how RNAPII traverses genes marked with ubiquitinated H2B. It has, in fact, been observed that ubiquitinated H2B aids in repression of silenced genes,[95,252] even while it is enriched within the coding region of highly expressed genes.[95,253] Its interplay with the de-ubiquitination machinery and with transcription factors such as Spt16 and Ctk1 during transcription elongation[107,254] suggests that this modification cycle occurs with the passage of RNAPII through every nucleosome. Therefore, ubiquitinated H2B must be a very dynamic modification, and its levels could quickly adapt to changes in the environment of the cell. Given that splicing is regulated in an equally dynamic fashion,[4,5] it perhaps makes sense for a "regulatory partnership" to have developed between H2B ubiquitination and splicing, both in budding yeast[26,104] and humans.[144,145]

The question of whether the splicing defect in strains lacking ubiquitinated H2B is due to RNAPII elongating too quickly is still an open one; one way to address this experimentally is to use an inducible gene, but none of the intron-containing genes in yeast rigorously fit into that category. It had been shown that the speed of RNAPII can impact the usage of alternative splice sites in metazoans, which inspired the question of whether this phenomenon occurs in some form in budding yeast. It was for this reason that I was interested in determining whether changes in RNAPII speed could impact splicing, since this could be investigated directly using a panel of RNAPII point mutants isolated by the Kaplan lab.[31,174,192]

RNA Polymerase II elongation

Our splicing microarray data suggest a model in which splicing efficiency is generally anti-correlated with transcription rate (*i.e.*, increased splicing efficiency when RNAPII rate is decreased, and decreased splicing efficiency when RNAPII rate is increased). Thus our data provide experimental support for the prediction (previously untested directly) that transcription rate and splicing efficiency would be generally anti-correlated. The resulting model fits with data from

metazoans, where it has been shown that changing RNAPII elongation rate near specific encoded splice sites affects the balance of the use of upstream and downstream splice sites. Because the spliceosome assembles in a stepwise fashion onto each intron, slowing transcription can afford more time for formation of the catalytically active machine before transcription of a downstream, stronger site. While budding yeast lack alternative splicing, it nonetheless follows that the efficiency of co-transcriptional splicing would be favored by allowing sufficient time for spliceosome assembly prior to transcription termination. Indeed, recent work suggests that RNAPII may slow down to favor co- versus post-transcriptional splicing.[34,36,138]

Are there conditions in which RNAPII elongation rate is specifically altered to impact splicing in a chromatin-directed way? It has recently been observed in human cells that splicing patterns during development respond to specific chromatin signatures that exert their effect on splicing by impacting RNAPII elongation.[255] These data provide a case-in-point study for a chromatin-RNAPII-splicing regulatory cascade during vertebrate development, and inspire the question of whether this phenomenon is a generally-applicable mechanism. The direct effect of changing RNAPII elongation rate on splicing in budding yeast[31] leaves open this possibility for genes with single introns as well.

Towards a "meta-model" of splicing within a chromatin environment

When I started on these projects, the idea of histone modification impacting splicing was a pipe-dream. By now, it is very clear that H2B ubiquitination[26,104,144,145] is not the only chromatin modification to have been tied in some way to splicing: the list now includes H3/H4 acetylation,[15] H2A.Z,[18] H3K4 methylation and H3K36 methylation (Chapter 2) in budding yeast; H3K36 methylation,[99,139,140,143,256-259] H3K9 methylation,[260] H3K4 methylation,[149] H4 deacetylation,[126] H2A.Bbd,[261] and H3K79 methylation[143] in metazoans. How do all these modifications contribute to splicing? The set of key outstanding questions posed earlier still stands:

- Do specific histone modifications directly recruit individual splicing factors?
- Do some modifications impact splicing indirectly, *i.e.*, through an effect on RNAPII?
- Does the logic of "chromatin – transcription" regulatory coupling apply to chromatin and splicing as well?

In some of the studies cited above, a direct physical association was observed and in others, this has been proposed. Some authors inferred an effect of polymerase dynamics that impact splicing. Some showed a causative, gene-specific effect of abrogating the modification on splicing, while others simply observed that the location of the modification is sensitive to intron-exon structure. My emerging sense is that the answers to the aforementioned questions are likely "in some cases," "in some cases," and "probably." The field of "coupling" is now in a remarkably similar position to where the "chromatin – transcription" field was 15 years ago: beginning to interpret the roles of myriad histone modifications together into a "meta-model" for how genes are transcribed. As the splicing field now approaches the need for a meta-model, I think it fair to hypothesize a complex and multi-factorial picture in which the spliceosome integrates multiple inputs from the chromatin landscape. Therefore, the landscape drawn in Chapter 1, Figure 6[26] is an apt one to generally describe this model, because it implies that at any given intron-containing gene there exists a dynamic, complex backdrop on which transcription and splicing occur. Because histone modifications and transcription are dynamic processes, the potential involvement of these factors may represent a way by which the cell rapidly adapts multiple steps of its gene expression program to changes in environmental stimuli. Because this chromatin landscape varies from gene to gene, this allows splicing of transcripts from different intron-containing genes to be coordinated with a different set of modifications.

Digging deeper into specific connections

An issue that deserves some thought as the field moves forward is how to bridge what is being learned about how splicing is coupled to chromatin and transcription with what is known about the mechanism of splicing. An intriguing possibility, first suggested by the Johnson lab for histone acetylation and deacetylation,[105] is that a specific modification during transcription could promote a specific spliceosome assembly step. Does each histone modification that has been connected to splicing make a contribution to a specific spliceosomal transition? This has been measured by chromatin immunoprecipitation (ChIP), which is an assay fraught with technical issues – poor spatial resolution, a signal-to-noise ratio difficult to work within, and the unpredictability of crosslinking. ChIP was originally used to conclude that completion of splicing could not possibly occur co-transcriptionally because late splicing factors are rarely observed by ChIP;[129] however, this conclusion was disproved by more recent reports using new and alternative assays.[34,35,262] The problem of what proteins can be crosslinked and immunoprecipitated at a given time during the splicing cycle makes any lack of signal un-trustworthy. It was this, and the desire to follow the individual catalytic steps of splicing as they occur co-transcriptionally that inspired the assay proposed in the Appendix (NETseq).

Much has been gleaned from *in vitro* systems about the mechanisms by which the spliceosome achieves its high fidelity.[263] Because completion of splicing often occurs co-transcriptionally,[34] these quality control mechanisms, must necessarily operate co-transcriptionally as well. Is there input from RNAPII or other co-transcriptional elements that promotes splicing fidelity? A transcription–splicing checkpoint has been proposed, during which RNAPII pauses briefly at the 3' splice site, and which may promote splicing fidelity.[36] As described above, the cycle of histone acetylation and deacetylation has been associated with specific spliceosomal rearrangements,[105] and leaves open the possibility that each step in splicing is associated with a chromatin-specific quality control mechanism.

It a general sense, the field has completely moved away from the textbook idea that splicing is a separate process that occurs on fully transcribed pre-mRNA. Splicing is generally co-transcriptional, and as I have shown, in some way integrates both the speed of RNAPII and the marks on the chromatin template that the polymerase transcribes. Do we have now a list of all the parts to build the model of how the integration occurs, or are we still at the tip of the iceberg? As the field progresses over the next decade, I am eager to see several specific directions that I believe will start to address the questions posed here: (1) a way to accurately model transcription-coupled splicing on physiological chromatin templates *in vitro*, in which to test specific predictions from modifying the chromatin structure or RNAPII, (2) a high-resolution structural understanding of spliceosome dynamics, and (3) novel high-resolution assays to measure progression through the splicing cycle *in vivo*. With these in hand, very specific questions can be asked about what precise step in splicing is affected when, for example, RNAPII is accelerated or slowed down, or when the histone ubiquitination cycle is impaired.

I know from ten years of experience playing in a symphony orchestra that many different players, with very different instruments, can come together on the same stage and play cohesively. In the nucleus, there is no conductor, but somehow, everyone plays from the same score. I think that tremendous progress has been recently made toward understanding how many different players there are in the gene expression symphony, and I have very high hopes that its score will be revealed in the not-too-distant future.

REFERENCES

1. Staley JP, Guthrie C. Mechanical Devices of the Spliceosome: Motors, Clocks, Springs, and Things. Cell 1998; 92:315–26.
2. Engebrecht JA, Voelkel-Meiman K, Roeder GS. Meiosis-specific RNA splicing in yeast. Cell 1991; 66:1257–68.
3. Munding EM, Igel AH, Shiue L, Dorighi KM, Trevino LR, Ares M. Integration of a splicing regulatory network within the meiotic gene expression program of Saccharomyces cerevisiae. Genes Dev 2010; 24:2693–704.
4. Bergkessel M, Whitworth GB, Guthrie C. Diverse environmental stresses elicit distinct responses at the level of pre-mRNA processing in yeast. RNA 2011; 17:1461–78.
5. Pleiss JA, Whitworth GB, Bergkessel M, Guthrie C. Rapid, transcript-specific changes in splicing in response to environmental stress. Mol Cell 2007; 27:928–37.
6. Wahl MC, Will CL, Lührmann R. The Spliceosome: Design Principles of a Dynamic RNP Machine. Cell 2009; 136:701–18.
7. Listerman I, Sapra AK, Neugebauer KM. Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. Nat Struct Mol Biol 2006; 13:815–22.
8. Kotovic KM, Lockshon D, Boric L, Neugebauer KM. Cotranscriptional Recruitment of the U1 snRNP to Intron-Containing Genes in Yeast. Mol Cell Biol 2003; 23:5768–79.
9. Lacadie SA, Rosbash M. Cotranscriptional Spliceosome Assembly Dynamics and the Role of U1 snRNA:5′ss Base Pairing in Yeast. Mol Cell 2005; 19:65–75.
10. Görnemann J, Kotovic KM, Hujer K, Neugebauer KM. Cotranscriptional Spliceosome Assembly Occurs in a Stepwise Fashion and Requires the Cap Binding Complex. Mol Cell 2005; 19:53–63.
11. Moore MJ, Schwartzfarb EM, Silver PA, Yu MC. Differential Recruitment of the Splicing

Machinery during Transcription Predicts Genome-Wide Patterns of mRNA Splicing. Mol Cell 2006; 24:903–15.

12. Nilsen TW. Spliceosome assembly in yeast: one ChIP at a time? Nat Struct Mol Biol 2005; 12:571–3.

13. Gregory PD, Schmid A, Zavari M, Lui L, BERGER SL, Hörz W. Absence of Gcn5 HAT activity defines a novel state in the opening of chromatin at the PHO5 promoter in yeast. Mol Cell 1998; 1:495–505.

14. Fuchs SM, Laribee RN, Strahl BD. Protein modifications in transcription elongation. BBA - Gene Regulatory Mechanisms 2009; 1789:26–36.

15. Gunderson FQ, Johnson TL. Acetylation by the Transcriptional Coactivator Gcn5 Plays a Novel Role in Co-Transcriptional Spliceosome Assembly. PLoS Genet 2009; 5:e1000682.

16. McKay SL, Johnson TL. An Investigation of a Role for U2 snRNP Spliceosomal Components in Regulating Transcription. PLoS ONE 2011; 6:e16077.

17. Hossain MA, Claggett JM, Nguyen T, Johnson TL. The cap binding complex influences H2B ubiquitination by facilitating splicing of the SUS1 pre-mRNA. RNA 2009; 15:1515–27.

18. Albulescu L-O, Sabet N, Gudipati M, Stepankiw N, Bergman ZJ, Huffaker TC, Pleiss JA. A Quantitative, High-Throughput Reverse Genetic Screen Reveals Novel Connections between Pre–mRNA Splicing and 5′ and 3′ End Transcript Determinants. PLoS Genet 2012; 8:e1002530.

19. Burckin T, Nagel R, Mandel-Gutfreund Y, Shiue L, Clark TA, Chong J-L, Chang T-H, Squazzo S, Hartzog G, Ares M. Exploring functional relationships between components of the gene expression machinery. Nat Struct Mol Biol 2005; 12:175–82.

20. Morris DP, Greenleaf AL. The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. J Biol Chem 2000; 275:39935–43.

21. Gornemann J, Barrandon C, Hujer K, Rutz B, Rigaut G, Kotovic KM, Faux C, Neugebauer KM, Séraphin B. Cotranscriptional spliceosome assembly and splicing are independent of the Prp40p WW domain. RNA 2011; 17:2119–29.

22. Yuryev A, Patturajan M, Litingtung Y, Joshi RV, Gentile C, Gebara M, Corden JL. The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. Proc Natl Acad Sci USA 1996; 93:6975–80.

23. Martinez-Contreras R, Cloutier P, Shkreta L, Fisette J-F, Revil T, Chabot B. hnRNP proteins and splicing control. Adv Exp Med Biol 2007; 623:123–47.

24. Kress TL, Krogan NJ, Guthrie C. A single SR-like protein, Npl3, promotes pre-mRNA splicing in budding yeast. Mol Cell 2008; 32:727–34.

25. Dermody JL, Dreyfuss JM, Villén J, Ogundipe B, Gygi SP, Park PJ, Ponticelli AS, Moore CL, Buratowski S, Bucheli ME. Unphosphorylated SR-like protein Npl3 stimulates RNA polymerase II elongation. PLoS ONE 2008; 3:e3273.

26. Moehle EA, Ryan CJ, Krogan NJ, Kress TL, Guthrie C. The Yeast SR-Like Protein Npl3 Links Chromatin Modification to mRNA Processing. PLoS Genet 2012; 8:e1003101.

27. Hérissant L, Moehle EA, Bertaccini D, Van Dorsselaer A, Schaeffer-Reiss C, Guthrie C, Dargemont C. H2B Ubiquitylation Modulates Spliceosome Assembly and Function in Budding Yeast. Biol Cell 2014; in press. doi:10.1111/boc.201400003.

28. la Mata de M, Alonso CR, Kadener S, Fededa JP, Blaustein M, Pelisch F, Cramer P, Bentley D, Kornblihtt AR. A slow RNA polymerase II affects alternative splicing in vivo. Mol Cell 2003; 12:525–32.

29. Howe KJ, Kane CM, Ares M. Perturbation of transcription elongation influences the fidelity of internal exon inclusion in Saccharomyces cerevisiae. RNA 2003; 9:993–1006.

30. Lacadie SA, Tardiff DF, Kadener S, Rosbash M. In vivo commitment to yeast cotranscriptional splicing is sensitive to transcription elongation mutants. Genes Dev 2006;

20:2055–66.

31. Braberg H, Jin H, Moehle EA, Chan YA, Wang S, Shales M, Benschop JJ, Morris JH, Qiu C, Hu F, et al. From Structure to Systems: High-Resolution, Quantitative Genetic Analysis of RNA Polymerase II. Cell 2013; 154:775–88.

32. Rhind N, Chen Z, Yassour M, Thompson DA, Haas BJ, Habib N, Wapinski I, Roy S, Lin MF, Heiman DI, et al. Comparative Functional Genomics of the Fission Yeasts. Science 2011; 332:930–6.

33. Jurica MS, Moore MJ. Pre-mRNA splicing: awash in a sea of proteins. Mol Cell 2003; 12:5–14.

34. Oesterreich FC, Preibisch S, Neugebauer KM. Global Analysis of Nascent RNA Reveals Transcriptional Pausing in Terminal Exons. Mol Cell 2010; 40:571–81.

35. Churchman LS, Weissman JS. Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature 2011; 469:368–73.

36. Alexander RD, Innocente SA, Barrass JD, Beggs JD. Splicing-Dependent RNA Polymerase Pausing in Yeast. Mol Cell 2010; 40:1–12.

37. Li B, Carey M, Workman JL. The Role of Chromatin during Transcription. Cell 2007; 128:707–19.

38. Singh R, Valcarcel J. Building specificity with nonspecific RNA-binding proteins. Nat Struct Mol Biol 2005; 12:645–53.

39. Long JC, Caceres JF. The SR protein family of splicing factors: master regulators of gene expression. Biochem J 2009; 417:15.

40. Dreyfuss G, Kim VN, Kataoka N. Messenger-RNA-binding proteins and the messages they carry. Nat Rev Mol Cell Biol 2002; 3:195–205.

41. Manley JL, Krainer AR. A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). Genes Dev 2010; 24:1073–4.

42. Lei EP, Krebber H, Silver PA. Messenger RNAs are recruited for nuclear export during transcription. Genes Dev 2001; 15:1771–82.

43. Bucheli ME, Buratowski S. Npl3 is an antagonist of mRNA 3' end formation by RNA polymerase II. EMBO J 2005; 24:2150–60.

44. Lund MK, Kress TL, Guthrie C. Autoregulation of Npl3, a yeast SR protein, requires a novel downstream region and serine phosphorylation. Mol Cell Biol 2008; 28:3873–81.

45. Wong CM, Qiu H, Hu C, Dong J, Hinnebusch AG. Yeast Cap Binding Complex Impedes Recruitment of Cleavage Factor IA to Weak Termination Sites. Mol Cell Biol 2007; 27:6520–31.

46. Gilbert W, Siebel CW, Guthrie C. Phosphorylation by Sky1p promotes Npl3p shuttling and mRNA dissociation. RNA 2001; 7:302–13.

47. Gilbert W, Guthrie C. The Glc7p nuclear phosphatase promotes mRNA export by facilitating association of Mex67p with mRNA. Mol Cell 2004; 13:201–12.

48. Wood A, Krogan NJ, Dover J, Schneider J, Heidt J, Boateng MA, Dean K, Golshani A, Zhang Y, Greenblatt JF, et al. Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. Mol Cell 2003; 11:267–74.

49. Hwang WW, Venkatasubrahmanyam S, Ianculescu AG, Tong A, Boone C, Madhani HD. A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. Mol Cell 2003; 11:261–6.

50. Daniel JA, Torok MS, Sun Z-W, Schieltz D, Allis CD, Yates JR, Grant PA. Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. J Biol Chem 2004; 279:1867–71.

51. Sanders SL, Jennings J, Canutescu A, Link AJ, Weil PA. Proteomics of the eukaryotic transcription machinery: identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry. Mol Cell Biol 2002; 22:4723–38.

52. Henry KW, Wyce A, Lo W-S, Duggan LJ, Emre NCT, Kao C-F, Pillus L, Shilatifard A, Osley MA, Berger SL. Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. Genes Dev 2003; 17:2648–63.

53. Tong AH, Evangelista M, Parsons AB, Xu H, Bader GD, Pagé N, Robinson M, Raghibizadeh S, Hogue CW, Bussey H, et al. Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science 2001; 294:2364–8.

54. Tong AHY, Lesage G, Bader GD, Ding H, Xu H, Xin X, Young J, Berriz GF, Brost RL, Chang M, et al. Global mapping of the yeast genetic interaction network. Science 2004; 303:808–13.

55. Wilmes GM, Bergkessel M, Bandyopadhyay S, Shales M, Braberg H, Cagney G, Collins SR, Whitworth GB, Kress TL, Weissman JS, et al. A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. Mol Cell 2008; 32:735–46.

56. Lee MS, Henry M, Silver PA. A protein that shuttles between the nucleus and the cytoplasm is an important mediator of RNA export. Genes Dev 1996; 10:1233–46.

57. Singleton DR, Chen S, Hitomi M, Kumagai C, Tartakoff AM. A yeast protein that bidirectionally affects nucleocytoplasmic transport. J Cell Sci 1995; 108:265–72.

58. Fortes P, Bilbao-Cortés D, Fornerod M, Rigaut G, Raymond W, Séraphin B, Mattaj IW. Luc7p, a novel yeast U1 snRNP protein with a role in 5′ splice site recognition. Genes Dev 1999; 13:2425–38.

59. Krebber H, Taura T, Lee MS, Silver PA. Uncoupling of the hnRNP Npl3p from mRNAs during the stress-induced block in mRNA export. Genes Dev 1999; 13:1994–2004.

60. Kim Guisbert K, Duncan K, Li H, Guthrie C. Functional specificity of shuttling hnRNPs revealed by genome-wide analysis of their RNA binding profiles. RNA 2005; 11:383–93.

61. Kobor MS, Venkatasubrahmanyam S, Meneghini MD, Gin JW, Jennings JL, Link AJ,

Madhani HD, Rine J. A Protein Complex Containing the Conserved Swi2/Snf2-Related ATPase Swr1p Deposits Histone Variant H2A.Z into Euchromatin. PLoS Biol 2004; 2:e131.

62. Krogan NJ, Keogh M-C, Datta N, Sawa C, Ryan OW, Ding H, Haw RA, Pootoolal J, Tong A, Canadien V, et al. A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Mol Cell 2003; 12:1565–76.

63. Mizuguchi G, Shen X, Landry J, Wu W-H, Sen S, Wu C. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science 2004; 303:343–8.

64. Shi X, Finkelstein A, Wolf AJ, Wade PA, Burton ZF, Jaehning JA. Paf1p, an RNA polymerase II-associated factor in Saccharomyces cerevisiae, may have both positive and negative roles in transcription. Mol Cell Biol 1996; 16:669–76.

65. Mueller CL, Jaehning JA. Ctr9, Rtf1, and Leo1 Are Components of the Paf1/RNA Polymerase II Complex. Mol Cell Biol 2002; 22:1971–80.

66. Squazzo SL, Costa PJ, Lindstrom DL, Kumer KE, Simic R, Jennings JL, Link AJ, Arndt KM, Hartzog GA. The Paf1 complex physically and functionally associates with transcription elongation factors in vivo. EMBO J 2002; 21:1764–74.

67. Krogan NJ, Kim M, Ahn SH, Zhong G, Kobor MS, Cagney G, Emili A, Shilatifard A, Buratowski S, Greenblatt JF. RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach. Mol Cell Biol 2002; 22:6979–92.

68. Miller T, Krogan NJ, Dover J, Erdjument-Bromage H, Tempst P, Johnston M, Greenblatt JF, Shilatifard A. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. Proc Natl Acad Sci USA 2001; 98:12902–7.

69. Grant PA, Duggan LJ, Côte J, Roberts SM, Brownell JE, Candau R, Ohba R, Owen-Hughes T, Allis CD, Winston F, et al. Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA

(Spt/Ada) complex. Genes Dev 1997; 11:1640–50.

70. Pijnappel WW, Schaft D, Roguev A, Shevchenko A, Tekotte H, Wilm M, Rigaut G, Séraphin B, Aasland R, Stewart AF. The S. cerevisiae SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. Genes Dev 2001; 15:2991–3004.

71. Robzyk K, Recht J, Osley MA. Rad6-dependent ubiquitination of histone H2B in yeast. Science 2000; 287:501–4.

72. Kim J, Roeder RG. Direct Bre1-Paf1 complex interactions and RING finger-independent Bre1-Rad6 interactions mediate histone H2B ubiquitylation in yeast. J Biol Chem 2009; 284:20582–92.

73. Vitaliano-Prunier A, Menant A, Hobeika M, Géli V, Gwizdek C, Dargemont C. Ubiquitylation of the COMPASS component Swd2 links H2B ubiquitylation to H3K4 trimethylation. Nat Cell Biol 2008; 10:1365–71.

74. Krogan NJ, Dover J, Wood A, Schneider J, Heidt J, Boateng MA, Dean K, Ryan OW, Golshani A, Johnston M, et al. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol Cell 2003; 11:721–9.

75. Ng HH, Robert F, Young RA, Struhl K. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol Cell 2003; 11:709–19.

76. Wood A, Schneider J, Dover J, Johnston M, Shilatifard A. The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. J Biol Chem 2003; 278:34739–42.

77. Ng HH, Dole S, Struhl K. The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B. J Biol Chem 2003; 278:33625–8.

78. Roguev A, Schaft D, Shevchenko A, Pijnappel WW, Wilm M, Aasland R, Stewart AF. The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. EMBO J 2001; 20:7137–48.

79. Sun Z-W, Allis CD. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature 2002; 418:104–8.

80. Dover J, Schneider J, Tawiah-Boateng MA, Wood A, Dean K, Johnston M, Shilatifard A. Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J Biol Chem 2002; 277:28368–71.

81. van Leeuwen F, Gafken PR, Gottschling DE. Dot1p modulates silencing in yeast by methylation of the nucleosome core. Cell 2002; 109:745–56.

82. Lacoste N, Utley RT, Hunter JM, Poirier GG, Côte J. Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. J Biol Chem 2002; 277:30421–4.

83. Briggs SD, Xiao T, Sun Z-W, Caldwell JA, Shabanowitz J, Hunt DF, Allis CD, Strahl BD. Gene silencing: trans-histone regulatory pathway in chromatin. Nature 2002; 418:498.

84. Ng HH, Xu R-M, Zhang Y, Struhl K. Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79. J Biol Chem 2002; 277:34655–7.

85. Ingvarsdottir K, Krogan NJ, Emre NCT, Wyce A, Thompson NJ, Emili A, Hughes TR, Greenblatt JF, Berger SL. H2B ubiquitin protease Ubp8 and Sgf11 constitute a discrete functional module within the Saccharomyces cerevisiae SAGA complex. Mol Cell Biol 2005; 25:1162–72.

86. Shukla A, Stanojevic N, Duan Z, Sen P, Bhaumik SR. Ubp8p, a histone deubiquitinase whose association with SAGA is mediated by Sgf11p, differentially regulates lysine 4 methylation of histone H3 in vivo. Mol Cell Biol 2006; 26:3339–52.

87. Lee KK, Florens L, Swanson SK, Washburn MP, Workman JL. The deubiquitylation

activity of Ubp8 is dependent upon Sgf11 and its association with the SAGA complex. Mol Cell Biol 2005; 25:1173–82.

88. Lee KK, Swanson SK, Florens L, Washburn MP, Workman JL. Yeast Sgf73/Ataxin-7 serves to anchor the deubiquitination module into both SAGA and Slik(SALSA) HAT complexes. Epigenetics Chromatin 2009; 2:2.

89. Daniel JA, Grant PA. Multi-tasking on chromatin with the SAGA coactivator complexes. Mutat Res 2007; 618:135–48.

90. Siebel CW, Guthrie C. The essential yeast RNA binding protein Npl3p is methylated. Proc Natl Acad Sci USA 1996; 93:13641.

91. Collins SR, Kemmeren P, Zhao X-C, Greenblatt JF, Spencer F, Holstege FCP, Weissman JS, Krogan NJ. Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae. Mol Cell Proteomics 2007; 6:439–50.

92. Pleiss JA, Whitworth GB, Bergkessel M, Guthrie C. Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. PLoS Biol 2007; 5:e90.

93. Xiao T, Kao C-F, Krogan NJ, Sun Z-W, Greenblatt JF, Osley MA, Strahl BD. Histone H2B ubiquitylation is associated with elongating RNA polymerase II. Mol Cell Biol 2005; 25:637–51.

94. Nickel BE, Allis CD, Davie JR. Ubiquitinated histone H2B is preferentially located in transcriptionally active chromatin. Biochemistry 1989; 28:958–63.

95. Batta K, Zhang Z, Yen K, Goffman DB, Pugh BF. Genome-wide function of H2B ubiquitylation in promoter and genic regions. Genes Dev 2011; 25:2254–65.

96. Clark TA, Sugnet CW, Ares M. Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. Science 2002; 296:907–10.

97. Loomis RJ, Naoe Y, Parker JB, Savic V, Bozovsky MR, Macfarlan T, Manley JL,

Chakravarti D. Chromatin Binding of SRp20 and ASF/SF2 and Dissociation from Mitotic Chromosomes Is Modulated by Histone H3 Serine 10 Phosphorylation. Mol Cell 2009; 33:450–61.

98. Huang Y, Li W, Yao X, Lin Q-J, Yin J-W, Liang Y, Heiner M, Bin Tian, Hui J, Wang G. Mediator Complex Regulates Alternative mRNA Processing via the MED23 Subunit. Mol Cell 2012; 45:1–11.

99. Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T. Regulation of Alternative Splicing by Histone Modifications. Science 2010; 327:996–1000.

100. Hossain MA, Rodriguez CM, Johnson TL. Key features of the two-intron Saccharomyces cerevisiae gene SUS1 contribute to its alternative splicing. Nucleic Acids Res 2011; 39:8612–27.

101. Rodriguez-Navarro S, Fischer T, Luo M-J, Antúnez O, Brettschneider S, Lechner J, Pérez-Ortín JE, Reed R, Hurt E. Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. Cell 2004; 116:75–86.

102. Shen EC, Stage-Zimmermann T, Chui P, Silver PA. The yeast mRNA-binding protein Npl3p interacts with the cap-binding complex. J Biol Chem 2000; 275:23718–24.

103. Schulze JM, Jackson J, Nakanishi S, Gardner JM, Hentrich T, Haug J, Johnston M, Jaspersen SL, Kobor MS, Shilatifard A. Linking Cell Cycle to Histone Modifications: SBF and H2B Monoubiquitination Machinery and Cell-Cycle Regulation of H3K79 Dimethylation. Mol Cell 2009; 35:626–41.

104. Shieh GS, Pan C-H, Wu J-H, Sun Y-J, Wang C-C, Hsiao W-C, Lin C-Y, Tung L, Chang T-H, Fleming AB, et al. H2B ubiquitylation is part of chromatin architecture that marks exon-intron structure in budding yeast. BMC Genomics 2011; 12:627.

105. Gunderson FQ, Merkhofer EC, Johnson TL. Dynamic histone acetylation is critical for cotranscriptional spliceosome assembly and spliceosomal rearrangements. Proc Natl Acad

Sci USA 2011; 108:2004–9.

106. Vitaliano-Prunier A, Babour A, Hérissant L, Apponi L, Margaritis T, Holstege FCP, Corbett AH, Gwizdek C, Dargemont C. H2B ubiquitylation controls the formation of export-competent mRNP. Mol Cell 2012; 45:132–9.

107. Fleming AB, Kao C-F, Hillyer C, Pikaart M, Osley MA. H2B Ubiquitylation Plays a Role in Nucleosome Dynamics during Transcription Elongation. Mol Cell 2008; 31:57–66.

108. Bhatt DM, Pandya-Jones A, Tong A-J, Barozzi I, Lissner MM, Natoli G, Black DL, Smale ST. Transcript Dynamics of Proinflammatory Genes Revealed by Sequence Analysis of Subcellular RNA Fractions. Cell 2012; 150:279–90.

109. Tilgner H, Knowles DG, Johnson R, Davis CA, Chakrabortty S, Djebali S, Curado J, Snyder M, Gingeras TR, Guigó R. Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. Genome Res 2012; 22:1616–25.

110. Ameur A, Zaghlool A, Halvardson J, Wetterbom A, Gyllensten U, Cavelier L, Feuk L. Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. Nat Struct Mol Biol 2011; 18:1435–40.

111. Guthrie C, Fink GR. Guide to Yeast Genetics and Molecular and Cell Biology. London: Academic Press: Methods in Enzymology; 2002.

112. Janke C, Magiera MM, Rathfelder N, Taxis C, Reber S, Maekawa H, Moreno-Borchart A, Doenges G, Schwob E, Schiebel E, et al. A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast 2004; 21:947–62.

113. Tong AHY, Boone C. Synthetic genetic array analysis in Saccharomyces cerevisiae. Methods Mol Biol 2006; 313:171–92.

114. Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G, Chan ET, Christie KR,

Costanzo MC, Dwight SS, Engel SR, et al. Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res 2012; 40:D700–5.

115. Pu S, Wong J, Turner B, Cho E, Wodak SJ. Up-to-date catalogues of yeast protein complexes. Nucleic Acids Res 2009; 37:825–31.

116. Berriz GF, King OD, Bryant B, Sander C, Roth FP. Characterizing gene sets with FuncAssociate. Bioinformatics 2003; 19:2502–4.

117. Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 2011; 27:431–2.

118. Schmitt ME, Brown TA, Trumpower BL. A rapid and simple method for preparation of RNA from Saccharomyces cerevisiae. Nucleic Acids Res 1990; 18:3091–2.

119. DeRisi JL, Iyer VR, Brown PO. Exploring the metabolic and genetic control of gene expression on a genomic scale. Science 1997; 278:680–6.

120. de Hoon MJL, Imoto S, Nolan J, Miyano S. Open source clustering software. Bioinformatics 2004; 20:1453–4.

121. Saldanha AJ. Java Treeview--extensible visualization of microarray data. Bioinformatics 2004; 20:3246–8.

122. Amberg DC, Goldstein AL, Cole CN. Isolation and characterization of RAT1: an essential gene of Saccharomyces cerevisiae required for the efficient nucleocytoplasmic trafficking of mRNA. Genes Dev 1992; 6:1173–89.

123. Knop M, Siegers K, Pereira G, Zachariae W, Winsor B, Nasmyth K, Schiebel E. Epitope tagging of yeast genes using a PCR-based strategy: more tags and improved practical routines. Yeast 1999; 15:963–72.

124. Suganuma T, Workman JL. Signals and Combinatorial Functions of Histone Modifications. Annu Rev Biochem 2011; 80:473–99.

125. Carrillo Oesterreich F, Bieberstein N, Neugebauer KM. Pause locally, splice globally.

Trends in Cell Biology 2011; 21:1–8.

126. Hnilicová J, Hozeifi S, Dušková E, Icha J, Tománková T, Staněk D. Histone Deacetylase Activity Modulates Alternative Splicing. PLoS ONE 2011; 6:e16727.

127. Luco RF, Alló M, Schor IE, Kornblihtt AR, Misteli T. Epigenetics in Alternative Pre-mRNA Splicing. Cell 2011; 144:16–26.

128. Niño CA, Hérissant L, Babour A, Dargemont C. mRNA Nuclear Export in Yeast. Chem Rev 2013; 113:8523−8545.

129. Tardiff DF, Lacadie SA, Rosbash M. A genome-wide analysis indicates that yeast pre-mRNA splicing is predominantly posttranscriptional. Mol Cell 2006; 24:917–29.

130. Babour A, Dargemont C, Stutz F. Ubiquitin and assembly of export competent mRNP. BBA - Gene Regulatory Mechanisms 2012; 1819:521–30.

131. Ares M, Grate L, Pauling MH. A handful of intron-containing genes produces the lion's share of yeast mRNA. RNA 1999; 5:1138–9.

132. Lenstra TL, Benschop JJ, Kim T, Schulze JM, Brabers NACH, Margaritis T, van de Pasch LAL, van Heesch SAAC, Brok MO, Groot Koerkamp MJA, et al. The specificity and topology of chromatin interaction pathways in yeast. Mol Cell 2011; 42:536–49.

133. Margaritis T, Oreal V, Brabers N, Maestroni L, Vitaliano-Prunier A, Benschop JJ, van Hooff S, van Leenen D, Dargemont C, Géli V, et al. Two Distinct Repressive Mechanisms for Histone 3 Lysine 4 Methylation through Promoting 3′-End Antisense Transcription. PLoS Genet 2012; 8:e1002952.

134. Oeffinger M, Wei KE, Rogers R, DeGrasse JA, Chait BT, Aitchison JD, Rout MP. Comprehensive analysis of diverse ribonucleoprotein complexes. Nat Methods 2007; 4:951–6.

135. Tuck AC, Tollervey D. A Transcriptome-wide Atlas of RNP Composition Reveals Diverse Classes of mRNAs and lncRNAs. Cell 2013; 154:996–1009.

136. Heintz D, Gallien S, Wischgoll S, Ullmann AK, Schaeffer C, Kretzschmar AK, Van Dorsselaer A, Boll M. Differential membrane proteome analysis reveals novel proteins involved in the degradation of aromatic compounds in Geobacter metallireducens. Mol Cell Proteomics 2009; 8:2159–69.

137. Pabis M, Neufeld N, Steiner MC, Bojic T, Shav-Tal Y, Neugebauer KM. The nuclear cap-binding complex interacts with the U4/U6.U5 tri-snRNP and promotes spliceosome assembly in mammalian cells. RNA 2013; 19:1054–63.

138. Aitken S, Alexander RD, Beggs JD. Modelling Reveals Kinetic Advantages of Co-Transcriptional Splicing. PLoS Comp Biol 2011; 7:e1002215.

139. Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, Ahringer J. Differential chromatin marking of introns and expressed exons by H3K36me3. Nat Genet 2009; 41:376–81.

140. Schwartz S, Meshorer E, Ast G. Chromatin organization marks exon-intron structure. Nat Struct Mol Biol 2009; 16:990–5.

141. Spies N, Nielsen CB, Padgett RA, Burge CB. Biased Chromatin Signatures around Polyadenylation Sites and Exons. Mol Cell 2009; 36:245–54.

142. Dhami P, Saffrey P, Bruce AW, Dillon SC, Chiang K, Bonhoure N, Koch CM, Bye J, James K, Foad NS, et al. Complex Exon-Intron Marking by Histone Modifications Is Not Determined Solely by Nucleosome Distribution. PLoS ONE 2010; 5:e12339.

143. Huff JT, Plocik AM, Guthrie C, Yamamoto KR. Reciprocal intronic and exonic histone modification regions in humans. Nat Struct Mol Biol 2010; 17:1495–9.

144. Jung I, Kim SK, Kim M, Han YM, Kim YS, Kim D, Lee D. H2B monoubiquitylation is a 5'-enriched active transcription mark and correlates with exon-intron structure in human cells. Genome Res 2012; 22:1026–35.

145. Zhang Z, Jones A, Joo HY, Zhou D, Cao Y, Chen S, Erdjument-Bromage H, Renfrow M, He H, Tempst P, et al. USP49 deubiquitinates histone H2B and regulates cotranscriptional

pre-mRNA splicing. Genes Dev 2013; 27:1581–95.

146. Spingola M, Grate L, Haussler D, Ares M. Genome-wide bioinformatic and molecular analysis of introns in Saccharomyces cerevisiae. RNA 1999; 5:221–34.

147. Colot HV, Stutz F, Rosbash M. The yeast splicing factor Mud13p is a commitment complex component and corresponds to CBP20, the small subunit of the nuclear cap-binding complex. Genes Dev 1996; 10:1699–708.

148. Hossain MA, Chung C, Pradhan SK, Johnson TL. The Yeast Cap Binding Complex Modulates Transcription Factor Recruitment and Establishes Proper Histone H3K36 Trimethylation during Active Transcription. Mol Cell Biol 2013; 33:785–99.

149. Sims RJ III, Millhouse S, Chen C-F, Lewis BA, Erdjument-Bromage H, Tempst P, Manley JL, Reinberg D. Recognition of Trimethylated Histone H3 Lysine 4 Facilitates the Recruitment of Transcription Postinitiation Factors and Pre-mRNA Splicing. Mol Cell 2007; 28:665–76.

150. Dujardin G, Lafaille C, Petrillo E, Buggiano V, Acuña LIG, Fiszbein A, Herz MAG, Moreno NN, Muñoz MJ, Alló M, et al. Transcriptional elongation and alternative splicing. BBA - Gene Regulatory Mechanisms 2013; 1829:134–40.

151. Chandrasekharan MB, Huang F, Sun Z-W. Ubiquitination of histone H2B regulates chromatin dynamics by enhancing nucleosome stability. Proc Natl Acad Sci USA 2009; 106:16686–91.

152. Longtine MS, McKenzie A, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR. Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast 1998; 14:953–61.

153. Gwizdek C, Hobeika M, Kus B, Ossareh-Nazari B, Dargemont C, Rodriguez MS. The mRNA nuclear export factor Hpr1 is regulated by Rsp5-mediated ubiquitylation. J Biol Chem 2005; 280:13401–5.

154. Ossareh-Nazari B, Bonizec MEL, Cohen M, Dokudovskaya S, Delalande FCO, Schaeffer C, Van Dorsselaer A, Dargemont C. Cdc48 and Ufd3, new partners of the ubiquitin protease Ubp3, are required for ribophagy. EMBO Rep 2010; 11:548–54.

155. Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, Devos D, Suprapto A, Karni-Schmidt O, Williams R, Chait BT, et al. The molecular architecture of the nuclear pore complex. Nature 2007; 450:695–701.

156. Iglesias N, Tutucci E, Gwizdek C, Vinciguerra P, Dach Von E, Corbett AH, Dargemont C, Stutz F. Ubiquitin-mediated mRNP dynamics and surveillance prior to budding yeast mRNA export. Genes Dev 2010; 24:1927–38.

157. Gokce E, Shuford CM, Franck WL, Dean RA, Muddiman DC. Evaluation of Normalization Methods on GeLC-MS/MS Label-Free Spectral Counting Data to Correct for Variation during Proteomic Workflows. J Am Soc Mass Spectrom 2011; 22:2199–208.

158. Miguet L, Lennon S, Baseggio L, Traverse-Glehen A, Berger F, Perrusson N, Chenard M-P, Galoisy A-C, Eischen A, Mayeur-Rousse C, et al. Cell-surface expression of the TLR homolog CD180 in circulating cells from splenic and nodal marginal zone lymphomas. Leukemia 2013; 27:1748–50.

159. Gwizdek C, Iglesias N, Rodriguez MS, Ossareh-Nazari B, Hobeika M, Divita G, Stutz F, Dargemont C. Ubiquitin-associated domain of Mex67 synchronizes recruitment of the mRNA export machinery with transcription. Proc Natl Acad Sci USA 2006; 103:16376–81.

160. Merz C, Urlaub H, Will CL, Lührmann R. Protein composition of human mRNPs spliced in vitro and differential requirements for mRNP protein recruitment. RNA 2006; 13:116–28.

161. Beltrao P, Cagney G, Krogan NJ. Quantitative genetic interactions reveal biological modularity. Cell 2010; 141:739–45.

162. Schuldiner M, Collins SR, Thompson NJ, Denic V, Bhamidipati A, Punna T, Ihmels J, Andrews B, Boone C, Greenblatt JF, et al. Exploration of the function and organization of

the yeast early secretory pathway through an epistatic miniarray profile. Cell 2005; 123:507–19.

163. Pan X, Yuan DS, Xiang D, Wang X, Sookhai-Mahadeo S, Bader JS, Hieter P, Spencer F, Boeke JD. A robust toolkit for functional profiling of the yeast genome. Mol Cell 2004; 16:487–96.

164. Collins SR, Roguev A, Krogan NJ. Quantitative genetic interaction mapping using the E-MAP approach. Meth Enzymol 2010; 470:205–31.

165. Schuldiner M, Collins SR, Weissman J, Krogan NJ. Quantitative genetic analysis in Saccharomyces cerevisiae using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. Methods 2006; 40:344–52.

166. Collins SR, Miller KM, Maas NL, Roguev A, Fillingham J, Chu CS, Schuldiner M, Gebbia M, Recht J, Shales M, et al. Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature 2007; 446:806–10.

167. Fiedler D, Braberg H, Mehta M, Chechik G, Cagney G, Mukherjee P, Silva AC, Shales M, Collins SR, van Wageningen S, et al. Functional Organization of the S. cerevisiae Phosphorylation Network. Cell 2009; 136:952–63.

168. Aguilar PS, Frohlich F, Rehman M, Shales M, Ulitsky I, Olivera-Couto A, Braberg H, Shamir R, Walter P, Mann M, et al. A plasma-membrane E-MAP reveals links of the eisosome with sphingolipid metabolism and endosomal trafficking. Nat Struct Mol Biol 2010; 17:901–8.

169. Greger IH, Proudfoot NJ. Poly(A) signals control both transcriptional termination and initiation between the tandem GAL10 and GAL7 genes of Saccharomyces cerevisiae. EMBO J 1998; 17:4771–9.

170. Kaplan CD, Holland MJ, Winston F. Interaction between transcription elongation factors and mRNA 3'-end formation at the Saccharomyces cerevisiae GAL10-GAL7 locus. J Biol

Chem 2005; 280:913–22.

171. Winston F, Sudarsanam P. The SAGA of Spt proteins and transcriptional analysis in yeast: past, present, and future. Cold Spring Harb Symp Quant Biol 1998; 63:553–61.

172. Shaw RJ, Reines D. Saccharomyces cerevisiae transcription elongation mutants are defective in PUR5 induction in response to nucleotide depletion. Mol Cell Biol 2000; 20:7427–37.

173. Archambault J, Friesen JD. Genetics of eukaryotic RNA polymerases I, II, and III. Microbiol Rev 1993; 57:703–24.

174. Kaplan CD, Jin H, Zhang IL, Belyanin A. Dissection of Pol II Trigger Loop Function and Pol II Activity–Dependent Control of Start Site Selection In Vivo. PLoS Genet 2012; 8:e1002627.

175. Cramer P, Bushnell DA, Kornberg RD. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science 2001; 292:1863–76.

176. Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science 2001; 292:1876–82.

177. Kaplan CD. Basic mechanisms of RNA polymerase II activity and alteration of gene expression in Saccharomyces cerevisiae. Biochim Biophys Acta 2013; 1829:39–54.

178. Roguev A, Bandyopadhyay S, Zofall M, Zhang K, Fischer T, Collins SR, Qu H, Shales M, Park HO, Hayles J, et al. Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. Science 2008; 322:405–10.

179. Wang D, Bushnell DA, Westover KD, Kaplan CD, Kornberg RD. Structural Basis of Transcription: Role of the Trigger Loop in Substrate Specificity and Catalysis. Cell 2006; 127:941–54.

180. Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JL, Toufighi K, Mostafavi S, et al. The genetic landscape of a cell. Science 2010; 327:425–31.

181. Benschop JJ, Brabers N, van Leenen D, Bakker LV, van Deutekom HW, van Berkum NL, Apweiler E, Lijnzaad P, Holstege FC, Kemmeren P. A consensus of core protein complex compositions for Saccharomyces cerevisiae. Mol Cell 2010; 38:916–28.

182. Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia W-J, Anderson S, Yates J, Washburn MP, et al. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell 2005; 123:581–92.

183. Keogh M-C, Kurdistani SK, Morris SA, Ahn SH, Podolny V, Collins SR, Schuldiner M, Chin K, Punna T, Thompson NJ, et al. Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell 2005; 123:593–605.

184. Spencer F, Gerring SL, Connelly C, Hieter P. Mitotic chromosome transmission fidelity mutants in Saccharomyces cerevisiae. Genetics 1990; 124:237–49.

185. Ohkuni K, Kitagawa K. Endogenous transcription at the centromere facilitates centromere activity in budding yeast. Curr Biol 2011; 21:1695–703.

186. Kettenberger H, Armache KJ, Cramer P. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. Mol Cell 2004; 16:955–65.

187. Edwards AM, Kane CM, Young RA, Kornberg RD. Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro. J Biol Chem 1991; 266:71–5.

188. Tan Q, Prysak MH, Woychik NA. Loss of the Rpb4/Rpb7 subcomplex in a mutant form of the Rpb6 subunit shared by RNA polymerases I, II, and III. Mol Cell Biol 2003; 23:3329–38.

189. Chen HT, Warfield L, Hahn S. The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. Nat Struct Mol Biol 2007; 14:696–703.

190. Eichner J, Chen HT, Warfield L, Hahn S. Position of the general transcription factor TFIIF

within the RNA polymerase II transcription preinitiation complex. EMBO J 2010; 29:706–16.

191. Ghazy MA, Brodie SA, Ammerman ML, Ziegler LM, Ponticelli AS. Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II. Mol Cell Biol 2004; 24:10975–85.

192. Kaplan CD, Larsson K-M, Kornberg RD. The RNA Polymerase II Trigger Loop Functions in Substrate Selection and Is Directly Targeted by α-Amanitin. Mol Cell 2008; 30:547–56.

193. Malagon F, Kireeva ML, Shafer BK, Lubkowska L, Kashlev M, Strathern JN. Mutations in the Saccharomyces cerevisiae RPB1 gene conferring hypersensitivity to 6-azauracil. Genetics 2006; 172:2201–9.

194. Bentley DL. Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. Curr Opin Cell Biol 2005; 17:251–6.

195. la Mata de M, Muñoz MJ, Alló M, Fededa JP, Schor IE, Kornblihtt AR. RNA Polymerase II Elongation at the Crossroads of Transcription and Alternative Splicing. Genetics Research International 2011; 2011:1–9.

196. Mason PB, Struhl K. Distinction and Relationship between Elongation Rate and Processivity of RNA Polymerase II In Vivo. Mol Cell 2005; 17:831–40.

197. Conesa C, Acker J. Sub1/PC4 a chromatin associated protein with multiple functions in transcription. RNA Biology 2010; 7:287–90.

198. Garcia A, Collin A, Calvo O. Sub1 associates with Spt5 and influences RNA polymerase II transcription elongation rate. Mol Biol Cell 2012; 23:4297–312.

199. Knaus R, Pollock R, Guarente L. Yeast SUB1 is a suppressor of TFIIB mutations and has homology to the human co-activator PC4. EMBO J 1996; 15:1933–40.

200. Wu WH, Pinto I, Chen BS, Hampsey M. Mutational analysis of yeast TFIIB. A functional relationship between Ssu72 and Sub1/Tsp1 defined by allele-specific interactions with

TFIIB. Genetics 1999; 153:643–52.

201. Rosonina E, Willis IM, Manley JL. Sub1 functions in osmoregulation and in transcription by both RNA polymerases II and III. Mol Cell Biol 2009; 29:2308–21.

202. Tavenet A, Suleau A, Dubreuil G, Ferrari R, Ducrot C, Michaut M, Aude J-C, Dieci G, Lefebvre O, Conesa C, et al. Genome-wide location analysis reveals a role for Sub1 in RNA polymerase III transcription. Proc Natl Acad Sci USA 2009; 106:14265–70.

203. Sikorski TW, Ficarro SB, Holik J, Kim T, Rando OJ, Marto JA, Buratowski S. Sub1 and RPA Associate with RNA Polymerase II at Different Stages of Transcription. Mol Cell 2011; 44:397–409.

204. Pinto I, Ware DE, Hampsey M. The yeast SUA7 gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection in vivo. Cell 1992; 68:977–88.

205. Pinto I, Wu WH, Na JG, Hampsey M. Characterization of sua7 mutations defines a domain of TFIIB involved in transcription start site selection in yeast. J Biol Chem 1994; 269:30569–73.

206. He Y, Fang J, Taatjes DJ, Nogales E. Structural visualization of key steps in human transcription initiation. Nature 2013; 495:481–6.

207. Goel S, Krishnamurthy S, Hampsey M. Mechanism of start site selection by RNA polymerase II: interplay between TFIIB and Ssl2/XPB helicase subunit of TFIIH. J Biol Chem 2012; 287:557–67.

208. Giardina C, Lis JT. DNA melting on yeast RNA polymerase II promoters. Science 1993; 261:759–62.

209. Kuehner JN, Brow DA. Quantitative analysis of in vivo initiator selection by yeast RNA polymerase II supports a scanning model. J Biol Chem 2006; 281:14119–28.

210. Arribere JA, Gilbert WV. Roles for transcript leaders in translation and mRNA decay

revealed by transcript leader sequencing. Genome Res 2013; 23:977–87.

211. Rojas-Duran MF, Gilbert WV. Alternative transcription start site selection leads to large differences in translation activity in yeast. RNA 2012; 18:2299–305.

212. Perales R, Bentley D. "Cotranscriptionality": The Transcription Elongation Complex as a Nexus for Nuclear Transactions. Mol Cell 2009; 36:178–91.

213. Ip JY, Schmidt D, Pan Q, Ramani AK, Fraser AG, Odom DT, Blencowe BJ. Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. Genome Res 2011; 21:390–401.

214. Cramer P, Pesce CG, Baralle FE, Kornblihtt AR. Functional association between promoter structure and transcript alternative splicing. Proc Natl Acad Sci USA 1997; 94:11456–60.

215. Harel-Sharvit L, Eldad N, Haimovich G, Barkai O, Duek L, Choder M. RNA polymerase II subunits link transcription and mRNA decay to translation. Cell 2010; 143:552–63.

216. Trcek T, Larson DR, Moldon A, Query CC, Singer RH. Single-molecule mRNA decay measurements reveal promoter- regulated mRNA stability in yeast. Cell 2011; 147:1484–97.

217. Hazelbaker DZ, Marquardt S, Wlotzka W, Buratowski S. Kinetic Competition between RNA Polymerase II and Sen1-Dependent Transcription Termination. Mol Cell 2013; 49:55–66.

218. Roguev A, Wiren M, Weissman JS, Krogan NJ. High-throughput genetic interaction mapping in the fission yeast Schizosaccharomyces pombe. Nature methods 2007; 4:861–6.

219. Ryan CJ, Roguev A, Patrick K, Xu J, Jahari H, Tong Z, Beltrao P, Shales M, Qu H, Collins SR, et al. Hierarchical modularity and the evolution of genetic interactomes across species. Mol Cell 2012; 46:691–704.

220. Butland G, Babu M, Diaz-Mejia JJ, Bohdana F, Phanse S, Gold B, Yang W, Li J, Gagarinova AG, Pogoutse O, et al. eSGA: E. coli synthetic genetic array analysis. Nat Methods 2008; 5:789–95.

221. Typas A, Nichols RJ, Siegele DA, Shales M, Collins SR, Lim B, Braberg H, Yamamoto N, Takeuchi R, Wanner BL, et al. High-throughput, quantitative analyses of genetic interactions in E. coli. Nat Methods 2008; 5:781–7.

222. Bassik MC, Kampmann M, Lebbink RJ, Wang S, Hein MY, Poser I, Weibezahn J, Horlbeck MA, Chen S, Mann M, et al. A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility. Cell 2013; 152:909–22.

223. Laufer C, Fischer B, Billmann M, Huber W, Boutros M. Mapping genetic interactions in human cancer cells with RNAi and multiparametric phenotyping. Nat Methods 2013; 10:427–31.

224. Lin YY, Kiihl S, Suhail Y, Liu SY, Chou YH, Kuang Z, Lu JY, Khor CN, Lin CL, Bader JS, et al. Functional dissection of lysine deacetylases reveals that HDAC1 and p300 regulate AMPK. Nature 2012; 482:251–5.

225. Roguev A, Talbot D, Negri GL, Shales M, Cagney G, Bandyopadhyay S, Panning B, Krogan NJ. Quantitative genetic-interaction mapping in mammalian cells. Nat Methods 2013; 10:432–7.

226. Gaj T, Gersbach CA, Barbas CF3. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol 2013; 31:397–405.

227. van de Peppel J, Kemmeren P, Van Bakel H, Radonjic M, van Leenen D, Holstege FC. Monitoring global messenger RNA changes in externally controlled microarray experiments. EMBO Rep 2003; 4:387–93.

228. Guthrie C, Fink GR. Guide to Yeast Genetics and Molecular and Cell Biology. Academic Press; 2002.

229. Pikielny CW, Rosbash M. mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. Cell 1985; 41:119–26.

230. Brugiolo M, Herzel L, Neugebauer KM. Counting on co-transcriptional splicing.

F1000Prime Rep 2013; 5.

231. Khodor YL, Rodriguez J, Abruzzi KC, Tang C-HA, Marr MT, Rosbash M. Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in Drosophila. Genes Dev 2011; 25:2502–12.

232. Mischo HE, Proudfoot NJ. Disengaging polymerase: Terminating RNA polymerase II transcription in budding yeast. BBA - Gene Regulatory Mechanisms 2013; 1829:174–85.

233. Dye MJ, Proudfoot NJ. Terminal exon definition occurs cotranscriptionally and promotes termination of RNA polymerase II. Mol Cell 1999; 3:371–8.

234. Vagner S, Rüegsegger U, Gunderson SI, Keller W, Mattaj IW. Position-dependent inhibition of the cleavage step of pre-mRNA 3'-end processing by U1 snRNP. RNA 2000; 6:178–88.

235. Rigo F, Martinson HG. Functional Coupling of Last-Intron Splicing and 3'-End Processing to Transcription In Vitro: the Poly(A) Signal Couples to Splicing before Committing to Cleavage. Mol Cell Biol 2008; 28:849–62.

236. Martins SB, Rino J, Carvalho T, Carvalho C, Yoshida M, Klose JM, de Almeida SF, Carmo-Fonseca M. Spliceosome assembly is coupled to RNA polymerase II dynamics at the 3' end of human genes. Nat Struct Mol Biol 2011; 18:1115–23.

237. Noble SM, Guthrie C. Identification of novel genes required for yeast pre-mRNA splicing by means of cold-sensitive mutations. Genetics 1996; 143:67–80.

238. Chanfreau G, Noble SM, Guthrie C. Essential yeast protein with unexpected similarity to subunits of mammalian cleavage and polyadenylation specificity factor (CPSF). Science 1996; 274:1511–4.

239. Zhang F, Yu X. WAC, a Functional Partner of RNF20/40, Regulates Histone H2B Ubiquitination and Gene Transcription. Mol Cell 2011; 41:384–97.

240. Jaehning JA. The Paf1 complex: Platform or player in RNA polymerase II transcription? BBA - Gene Regulatory Mechanisms 2010; 1799:379–88.

241. Hebbes TR, Thorne AW, Crane-Robinson C. A direct link between core histone acetylation and transcriptionally active chromatin. EMBO J 1988; 7:1395–402.

242. Eperon LP, Graham IR, Griffiths AD, Eperon IC. Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? Cell 1988; 54:393–401.

243. Lewicki BT, Margus T, Remme J, Nierhaus KH. Coupling of rRNA transcription and ribosomal assembly in vivo. J Mol Biol 1993; 231:581–93.

244. Pandya-Jones A, Black DL. Co-transcriptional splicing of constitutive and alternative exons. RNA 2009; 15:1896–908.

245. Brody Y, Neufeld N, Bieberstein N, Causse SZ, Böhnlein E-M, Neugebauer KM, Darzacq X, Shav-Tal Y. The In Vivo Kinetics of RNA Polymerase II Elongation during Co-Transcriptional Splicing. PLoS Biol 2011; 9:e1000573.

246. Rhee HS, Pugh BF. Genome-wide structure and organization of eukaryotic pre-initiation complexes. Nature 2012; 483:1–8.

247. Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science 2008; 322:1845–8.

248. Crick F. On protein synthesis. Symp Soc Exp Biol 1958; 12:138–63.

249. Eissenberg JC. Structural biology of the chromodomain: Form and function. Gene 2012; 496:69–78.

250. Davis FP, Braberg H, Shen MY, Pieper U, Sali A, Madhusudhan MS. Protein complex compositions predicted by structural similarity. Nucleic Acids Res 2006; 34:2943–52.

251. Fierz B, Chatterjee C, McGinty RK, Bar-Dagan M, Raleigh DP, Muir TW. Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction. Nat Chem Biol 2011; 7:113–9.

252. Turner SD, Ricci AR, Petropoulos H, Genereaux J, Skerjanc IS, Brandl CJ. The E2

Ubiquitin Conjugase Rad6 Is Required for the ArgR/Mcm1 Repression of ARG1 Transcription. Mol Cell Biol 2002; 22:4011–9.

253. Schulze JM, Hentrich T, Nakanishi S, Gupta A, Emberly E, Shilatifard A, Kobor MS. Splitting the task: Ubp8 and Ubp10 deubiquitinate different cellular pools of H2BK123. Genes Dev 2011; 25:2242–7.

254. Wyce A, Xiao T, Whelan KA, Kosman C, Walter W, Eick D, Hughes TR, Krogan NJ, Strahl BD, Berger SL. H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex. Mol Cell 2007; 27:275–88.

255. Schor IE, Fiszbein A, Petrillo E, Kornblihtt AR. Intragenic epigenetic changes modulate NCAM alternative splicing in neuronal differentiation. EMBO J 2013; 32:2264–74.

256. Simon JM, Hacker KE, Singh D, Brannon AR. Variation in chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with widespread RNA processing defects. Genome Res 2013; 24:241–50.

257. Andersson R, Enroth S, Rada-Iglesias A, Wadelius C, Komorowski J. Nucleosomes are well positioned in exons and carry characteristic histone modifications. Genome Res 2009; 19:1732–41.

258. Hon G, Wang W, Ren B. Discovery and Annotation of Functional Chromatin Signatures in the Human Genome. PLoS Comp Biol 2009; 5:e1000566.

259. Nahkuri S, Taft RJ, Mattick JS. Nucleosomes are preferentially positioned at exons in somatic and sperm cells. Cell Cycle 2009; 8:3420–4.

260. Saint-André V, Batsché E, Rachez C, Muchardt C. Histone H3 lysine 9 trimethylation and HP1γ favor inclusion of alternative exons. Nat Struct Mol Biol 2011; 18:1–9.

261. Tolstorukov MY, Goldman JA, Gilbert C, Ogryzko V, Kingston RE, Park PJ. Histone Variant H2A.Bbd Is Associated with Active Transcription and mRNA Processing in Human Cells. Mol Cell 2012; 47:1-12.

262. Alexander R, Beggs JD. Cross-talk in transcription, splicing and chromatin: who makes the first call? Biochem Soc Trans 2010; 38:1251.

263. Koodathingal P, Staley JP. Splicing fidelity: DEAD/H-box ATPases as molecular clocks. RNA Biology 2013; 10:1073–9.

APPENDICES

ESS1

Question: Does *ESS1* play a role in splicing?

Background:

This experiment came out of a discussion with Steven Hanes at the ASBMB meeting at Granlibakken entitled "Transcriptional Regulation of Chromatin and RNA Polymerase II." I was presenting work on Npl3 and its genetic interaction landscape. Dr. Hanes was looking at the transcription and chromatin genes that genetically interact with *npl3*Δ (See Chapter 1, Figures 1 and 2), and noticed that a mutant that he has been studying, *ess1*-H164R, also exhibits genetic interactions with many of the same genes. *ESS1*, a gene so named because it is essential,[1] is a prolyl isomerase that acts on the Ser5-Pro6 bond CTD of RNA Polymerase II (RNAPII).[2] In turn, isomerization of the CTD controls de-phosphorylation of Ser5 by Ssu72[3-6] and is required for methylation of Histone H3 on K4.[7] Through these connections and others, *ESS1* has been implicated in multiple aspects of the RNAPII transcription cycle,[7] specifically initiation,[8,9] elongation,[8] and termination.[2,3,9,10]

Npl3 itself has been implicated in transcription elongation and termination, as mutations in *NPL3* cause improved elongation and transcriptional read-through[11-13] (also see Appendix – Transcription Termination). Npl3 also has connections to CTD phosphorylation: Npl3 preferentially binds Ser2-phosphorylated CTD over its unphosphorylated form.[12] The similar genetic interactions and similar localization to elongating RNAPII, led me to ask whether, like Npl3,[14] Ess1 might be involved in splicing.

Experimental procedure:

I received two *ess1*-H164R strains from Dr. Hanes, with 2 matched WT strains:

EMy476: *ess1*-H164R MATα

EMy477: *ESS1* WT MATα

EMy478: *ess1*-H164R MAT**a**

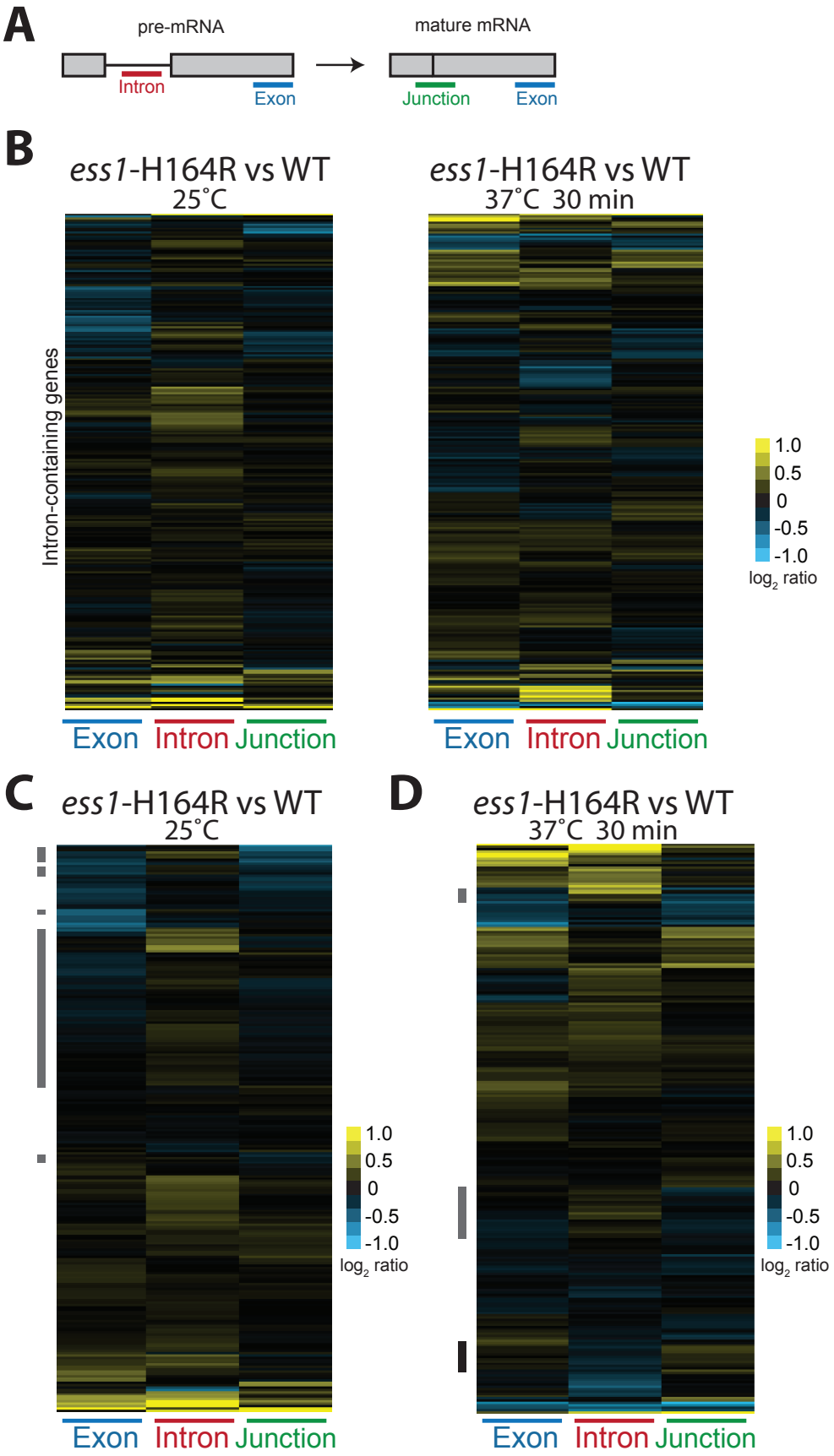
EMy479: *ESS1* WT MAT**a**

These strains were grown up and collected at logphase after growth at 25˚C, as well as after a 30-minute temperature shift to 37˚C. RNA was extracted and cDNA was made, labeled and competitively hybridized to our in-house splicing microarrays.[15] The results from both mating types for each condition were averaged together and are shown in Figure 1. These data are available in electronic form in the ESS1 folder in Erica's Documents.

Results:

The different mating type strains yielded very similar microarray data (data not shown). Increases and decreases in the abundance of the exon probe suggest some mild transcriptional changes occurring in the *ess1*-H164R mutant (Figure 1). This is consistent with Ess1's role in various transcriptional processes. On the backdrop of these changes, some interesting patterns can be discerned that are indicative of mild splicing phenotypes. At 25˚C, a subset of genes exhibit the characteristic pattern of a decrease in splicing efficiency: accumulation of the intron and a decrease in the abundance of the exon-exon junction feature (Figure 1C and 1E). At 37˚C, the *ess1*-H164R strain also contains some genes whose splicing efficiency may be decreased, as well as those whose splicing appears mildly improved (Figure 1D and 1F). Follow-up studies would be required to nail down a direct functional connection between Ess1 and splicing.

Interestingly, the phenotypes from cells grown at 25˚C are distinct from those that have been shifted to 37˚C for 30 minutes. Whether this is due to different *ESS1*-dependent functions at these two temperatures is unknown.


A
pre-mRNA
Intron
Exon
mature mRNA
Junction
Exon
B
ess1-H164R vs WT
25˚C
ess1-H164R vs WT
37˚C 30 min
Intron-containing genes
Exon
Intron
Junction
1.0
0.5
0
-0.5
-1.0
log2 ratio
C
ess1-H164R vs WT
25˚C
D
ess1-H164R vs WT
37˚C 30 min

E

ess1-H164R vs WT

25˚C

F

ess1-H164R vs WT

37˚C 30 min

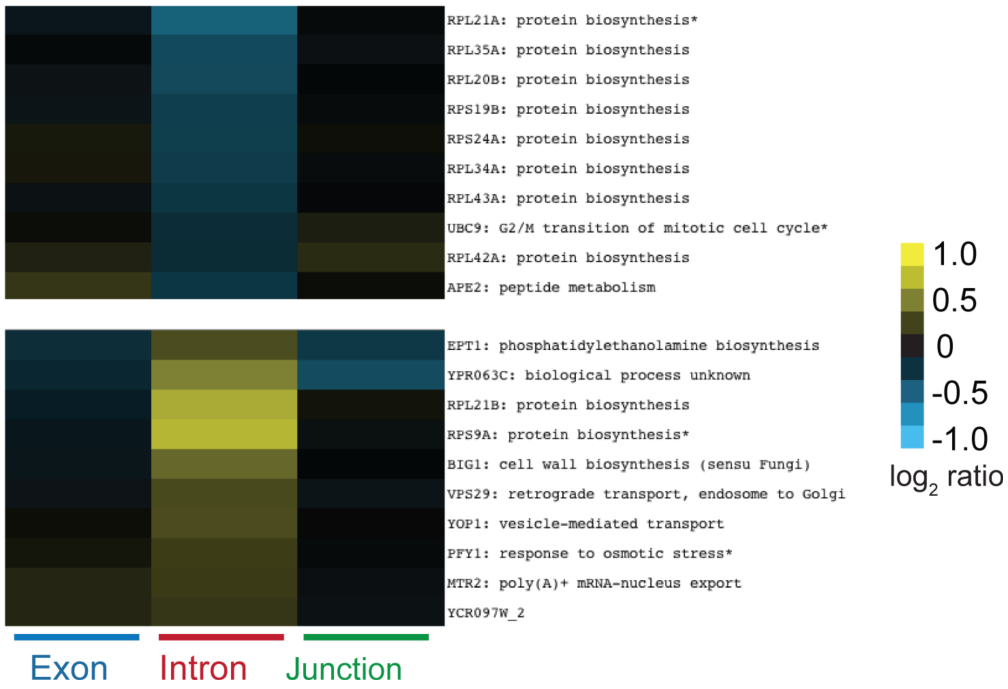


Figure 1. Splicing phenotypes of *ess1*-H164R mutant

(A.) Schematic of probes contained on the microarray

(B.) The *ess1*-H164R strain and isogenic wild-type strains grown at the indicated temperatures were hybridized to the microarray. Shown is a heat map of $\log_2$ ratios. Gene order is the same for both arrays.

(C.) Re-clustering of the *ess1*-H164R array at 25˚C from (B.). The grey lines to the left of the heatmap indicate intron-containing genes that may have a splicing defect, for which there is an accumulation of intron and a concomitant decrease in the abundance of the exon-exon junction.

(D.) Re-clustering of the *ess1*-H164R array at 37˚C from (B.). The grey lines to the left of the heatmap indicate intron-containing genes that may have a splicing defect, for which there is an accumulation of intron and a concomitant decrease in the abundance of the exon-exon junction. The black line to the left of the heatmap indicates intron-containing genes that may have improved splicing, for which there is a decrease in the abundance of intron and concomitant increase of the exon-exon junction.

(E.) Subset of genes from (C.).

(F.) Subset of genes from (D.).

REFERENCES

1. Hanes SD, Shank PR, Bostian KA. Sequence and mutational analysis of ESS1, a gene essential for growth in Saccharomyces cerevisiae. Yeast 1989; 5:55–72.
2. Hani J, Schelbert B, Bernhardt A, Domdey H, Fischer G, Wiebauer K, Rahfeld JU. Mutations in a peptidylprolyl-cis/trans-isomerase gene lead to a defect in 3'-end formation of a pre-mRNA in Saccharomyces cerevisiae. J Biol Chem 1999; 274:108–16.
3. Singh N, Ma Z, Gemmill T, Wu X, DeFiglio H, Rossettini A, Rabeler C, Beane O, Morse RH, Palumbo MJ, et al. The Ess1 Prolyl Isomerase Is Required for Transcription Termination of Small Noncoding RNAs via the Nrd1 Pathway. Mol Cell 2009; 36:255–66.
4. Xiang K, Nagaike T, Xiang S, Kilic T, Beh MM, Manley JL, Tong L. Crystal structure of the human symplekin–Ssu72– CTD phosphopeptide complex. Nature 2010; 467:729–33.
5. Werner-Allen JW, Lee CJ, Liu P, Nicely NI, Wang S, Greenleaf AL, Zhou P. cis-Proline-mediated Ser(P)5 Dephosphorylation by the RNA Polymerase II C-terminal Domain Phosphatase Ssu72. J Biol Chem 2011; 286:5717–26.
6. Bataille AR, Jeronimo C, Jacques P-É, Laramée L, Fortin M-È, Forest A, Bergeron M, Hanes SD, Robert F. A Universal RNA Polymerase II CTD CycleIs Orchestrated by Complex Interplays between Kinase, Phosphatase, and Isomerase Enzymes along Genes. Mol Cell 2012; 45:158–70.
7. Ma Z, Atencio D, Barnes C, DeFiglio H, Hanes SD. Multiple Roles for the Ess1 Prolyl Isomerase in the RNA Polymerase II Transcription Cycle. Mol Cell Biol 2012; 32:3594–607.
8. Wilcox CB, Rossettini A, Hanes SD. Genetic interactions with C-terminal domain (CTD) kinases and the CTD of RNA Pol II suggest a role for ESS1 in transcription initiation and elongation in Saccharomyces cerevisiae. Genetics 2004; 167:93–105.
9. Krishnamurthy S, Ghazy MA, Moore C, Hampsey M. Functional Interaction of the Ess1 Prolyl Isomerase with Components of the RNA Polymerase II Initiation and Termination

Machineries. Mol Cell Biol 2009; 29:2925–34.

10. Morris DP, Phatnani HP, Greenleaf AL. Phospho-carboxyl-terminal domain binding and the role of a prolyl isomerase in pre-mRNA 3'-End formation. J Biol Chem 1999; 274:31583–7.

11. Bucheli ME, Buratowski S. Npl3 is an antagonist of mRNA 3' end formation by RNA polymerase II. EMBO J 2005; 24:2150–60.

12. Dermody JL, Dreyfuss JM, Villén J, Ogundipe B, Gygi SP, Park PJ, Ponticelli AS, Moore CL, Buratowski S, Bucheli ME. Unphosphorylated SR-like protein Npl3 stimulates RNA polymerase II elongation. PLoS ONE 2008; 3:e3273.

13. Lund MK, Kress TL, Guthrie C. Autoregulation of Npl3, a yeast SR protein, requires a novel downstream region and serine phosphorylation. Mol Cell Biol 2008; 28:3873–81.

14. Kress TL, Krogan NJ, Guthrie C. A single SR-like protein, Npl3, promotes pre-mRNA splicing in budding yeast. Mol Cell 2008; 32:727–34.

15. Pleiss JA, Whitworth GB, Bergkessel M, Guthrie C. Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. PLoS Biol 2007; 5:e90.

NETseq

This is an extended experiment that I proposed to undertake in collaboration with Stirling Churchman, but I did not finish it, and it makes for an ideal "future direction." I am including it here on the off-chance that someone will be inspired to pick it up. Experimental details are available in my NETseq notebook.

Purpose: To characterize the influence of polymerase dynamics and chromatin on co-transcriptional splicing, by defining the factors and conditions that determine *where* RNA Polymerase II is during the transcription of a pre-mRNA when the two biochemical steps of co-transcriptional splicing occur.

Part 1

Develop and validate a high-throughput sequencing pipeline that will measure where RNA Polymerase II (RNAPII) is along a gene when the 1st and 2nd steps of splicing occur. Use this pipeline to determine for which genes the (a.) first or (b.) first *and* second steps of splicing occur co-transcriptionally. Determine what *cis* elements best correspond to where RNAPII is when the steps of splicing occur. Characterize how tight the distribution of splice events is across a locus and how this distribution varies among spliced genes.

Part 2

Use this assay on strains with specific alterations in the co-transcriptional environment to dissect the *cis* elements and *trans* factors that regulate the 1st and 2nd steps of splicing. Identify factors that influence the frequency with which the two steps of splicing occur co-transcriptionally. Identify factors that influence the timing of when the 1st and 2nd steps of splicing occur along a gene. In the context of FAST and SLOW polymerases, do the steps of splicing occur at different locations along genes?

Background:

There is growing interest in understanding the mechanisms by which co-transcriptional events are coordinated with each other. Assays have been developed to look at co-transcriptional splicing that have improved this understanding: the three most often discussed are from the Beggs, Neugebauer, and Black labs. The Beggs lab has developed an inducible reporter assay to study the dynamics of (a.) splicing factor association with nascent transcripts, (b.) RNAPII elongation, and (c.) splicing catalysis after induction of transcription. With this system, they see evidence suggesting that RNAPII pauses at the 3' splice site, which depends on an intact 3'SS-U2 interaction.[1] The Neugebauer lab uses steady state microarray measurements of chromatin-associated transcripts to infer dynamic information about RNAPII speed. These experiments also suggested that RNAPII pauses; however this pause occurs at a more or less defined distance from the 3' end, and is specific for genes with short second exons. These experiments also led to measurements of "co-transcriptional splicing efficiency," *i.e.*, the fraction of transcripts from a given locus that have undergone intron removal prior to release from the chromatin.[2] The Black lab induces transcription of pro-inflammatory genes and isolates chromatin-associated transcripts to achieve a dynamic picture of when, temporally, introns are removed from transcripts. From these experiments, they see evidence suggesting that unspliced transcripts can be held at the 3' end after polyadenylation and that completion of splicing likely occurs before release into the nucleoplasm.[3,4]

The conclusions of these experiments add important nuances to the picture we previously had from straightforward chromatin immunoprecipitation assays of splicing factors[5] of how the connection between transcription and co-transcriptional splicing is controlled. However, to really understand the complex interplay between splicing and various aspects of the co-transcriptional environment, it is not sufficient to just know the nature of the final spliced transcript, a limitation of current assays. Rather, we need a way to assess what factors affect the co-transcriptional regulation of individual steps of splicing. In addition, by specifying transcripts as co-transcriptional (*i.e.*, associated with RNAPII) rather than chromatin-associated, we will focus only on transcripts that are

actively being transcribed, rather than those that may be fully transcribed but are being held at the 3' end. Lastly, next generation sequencing allows deep coverage at single nucleotide resolution. The experiments proposed here aim to address these issues.

Proposed experiments:

The assay builds on a robust technique called NETseq, developed at UCSF.[6] NETseq maps the location of RNAPII genome-wide using purification of the ternary complex (RNAPII, chromatin, and RNA) followed by next-generation sequencing of the 3' ends of each RNA. By virtue of the fact that each 3' end represents the last nucleotide of a nascent transcript that RNAPII has transcribed, NETseq offers a view into the distribution of polymerases across genes.

The proposed outline of this splicing-coupled NETseq is shown in Figure 2, and should theoretically allow pairing of the information about where RNAPII is to whether each transcript is spliced. Briefly, following isolation of nascent transcripts and ligation of a linker on the 3' end is a debranching step. This should theoretically be performed using purified debranchase enzyme. The transcripts then need to be prepared for paired end sequencing. The spectrum of nascent transcripts we would expect to isolate are shown in Figure 3.

There are two ways that splicing information can be determined from this type of experiment: from the 3' reads or from the 5' reads. In the Churchman dataset, a few genes showed evidence of the first step (free exon 1) and second step (lariat product) of splicing (see Supplement of Churchman *et al.*[7]) Furthermore, in an analysis performed by Vida Ahyong in her rotation in the lab, there was also evidence of lariat intermediate: 3' ends just downstream of the 3'SS were enriched for having 5' ends just short of the branch site A, suggesting that reverse transcription of those isolated RNAs had been blocked by the lariat branch (no debranching step had been performed). For further information, see Vida's group meeting on the Guthrie lab server. In the case of quickly-splicing genes, or genes on which RNAPII slows or pauses in exon 2,[2,8] the 3' read may be a reasonable way of capturing splicing information. Increasing the read length from 30nt to 200nt will increase the number of reads

that can be analyzed this way. However, the number of reads that informed on splicing in this manner was very low, and unfortunately, these reads aren't associated with information about the location of RNAPII.

The goal for the protocol proposed here is to use information from the 5' end, such that *most* transcripts are associated with splicing information. A prerequisite for this is that a 5' sequencing read covers the 5'UTR and the intron (or second exon in the case of completed splicing). Using two separate transcription start site mapping studies,[9,10] I calculated that the introns of ~75% of intron-containing genes start within 200nt of the transcription start; this means that sequencing 200bp from the 5' end should include sequencing through the first exon (Figure 4). The read will either continue into an intron (indicating the nascent transcript is unspliced) or into the second exon (indicating the nascent transcript had been fully spliced) (Figure 3). Any lariat-exon intermediates, which are the products of the 1st step of splicing, will have a 5' end beginning at the first nucleotide of the intron if they have been debranched. Statistically, any enrichment of transcripts with the 5' end being the start of the intron will represent lariat intermediates. By also sequencing from the 3' end, splicing information will be paired with the location of that individual polymerase.

In the published NETseq assay, a circularization step allows intra-molecular ligation of an adapter to the 5' end. The longer sequence reads proposed here, as well as the distribution of lengths of transcripts in *S. cerevisiae* require the sequencing library to be as long as possible. The maximum length of a sequencing library is ~800bp, which means that the library preparation protocol will need a shortening step. A potential solution is to treat with limiting DNaseI after circularization, followed by re-circularization and size-selection on a gel.

The 30nt reads used earlier did not always allow unique mapping to highly similar paralogs,[6] for example, the ribosomal protein genes, which exhibit limited sequence divergence, but which are a very biologically relevant class of spliced genes in budding yeast. Increasing the read length will address this issue.

Because of the massively parallel nature of high-throughput sequencing, each gene will contain a distribution of polymerase locations, each associated with fraction unspliced, partially spliced and fully spliced.

Hypothetical Results:

One result from the genome-wide dataset would be a nucleotide map of the location where the first and second steps of splicing begin to occur – this will result in a map that is different for each gene. A possible single-gene result is shown in Figure 5. In the example shown, the first and second steps of splicing are separated in space. If the splice events occur with a tight spatial distribution, one could infer that there is a specific signal directing the act of splicing. If the splicing events are widely distributed across a gene, then this might indicate that splicing is a more stochastic event *in vivo*.

Meta-gene models should allow identification of which *cis*-element best correlate with appearance of splice intermediates and products.

By performing these splicing-coupled NETseq experiments on the polymerase speed mutants, discussed in Chapter 3, it should be possible to resolve whether a simple timing mechanism or a signal/location mechanism is at play. The hypothesis is that if RNAPII is slowed, but splicing takes the same amount of time, both steps of splicing will occur more 5' than in a wild-type strain (Compare Figures 6 and 7). In addition, by analyzing mutants or deletions in candidate chromatin genes, it should be possible to begin to define elements of the chromatin landscape that direct when and where co-transcriptional splicing occurs.

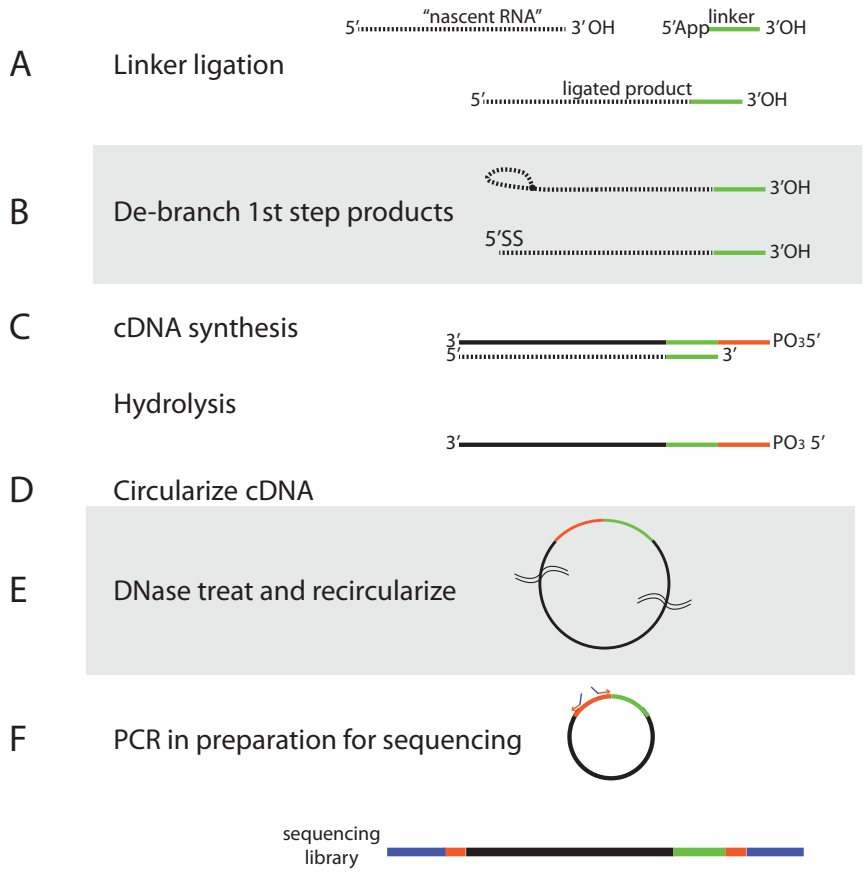


Figure 2. Proposed protocol for adapting NETseq to study splicing

Based on the protocol from Churchman *et al.,*[7] this protocol occurs after isolation of nascent RNAs that co-precipitate with RNAPII. The steps highlighted in grey are proposed additions to the original NETseq protocol.

(A.) Linker ligation and gel purification. The goal is to keep full transcripts, so the limiting hydrolysis step from Churchman *et al.*[7] is left out.

(B.) Lariat debranching *in vitro* using debranchase enzyme

(C.) cDNA synthesis using primer complementary to 3' linker and containing sequences necessary for paired-end Illumina sequencing

(D.) cDNA circularization

(E.) DNaseI treatment and recircularization to isolate cDNAs under 800bp

(F.) PCR to make double-stranded cDNA to generate final library

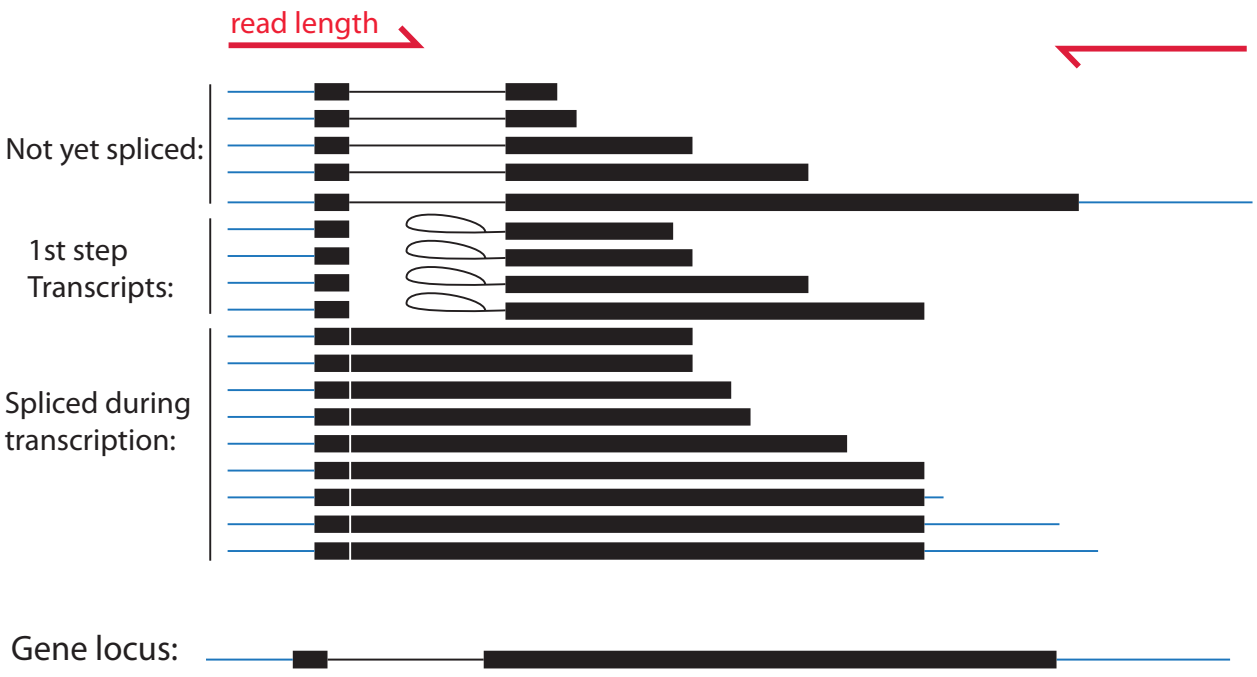


Figure 3. Schematic of pool of transcripts

The left red arrow represents the length of the 5' sequencing read; this read will encounter an intron (transcript not yet spliced) or the second exon (transcript was spliced during transcription). The right red arrow represents the length of the 3' sequencing read; this read may or may not encounter an intron. Thin blue lines denote 5' and 3' UTRs. Thin black lines denote introns. Thick black lines denote exons.

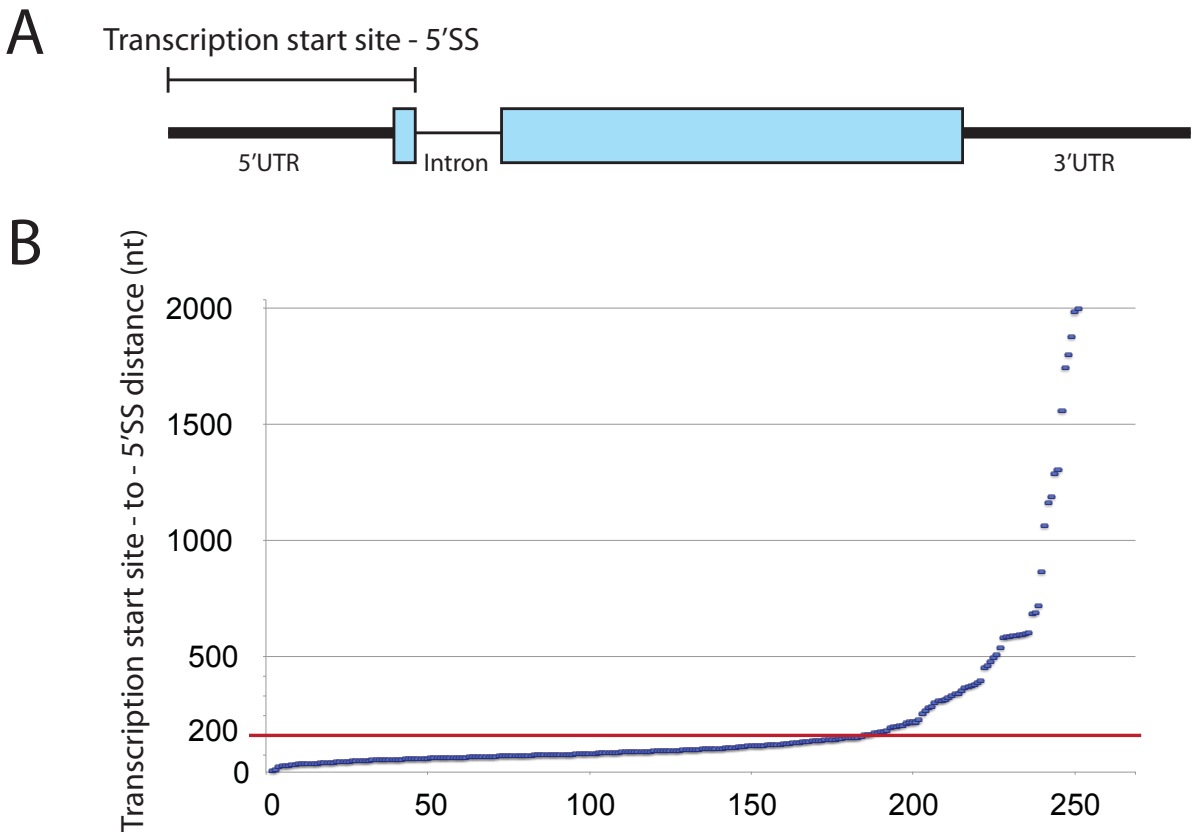


Figure 4. Distribution of length of TSS-5'SS distance of intron-containing genes

Many introns in *S. cerevisiae* contain introns toward the 5' end, which allows for retrieval of splicing information from 5' end sequencing. The length of Illumina sequencing read will change the number of intron-containing genes that can be analyzed with this technique.

(A.) Schematic of the distance being measured

(B.) Graph of the distribution of lengths using transcription start site information from Xu *et al.*[10] Genes are ordered based on this distance. The red bar at 200nt is the current reasonable limit for length of sequencing read from Illumina; 193 intron-containing genes can theoretically be analyzed using this assay.

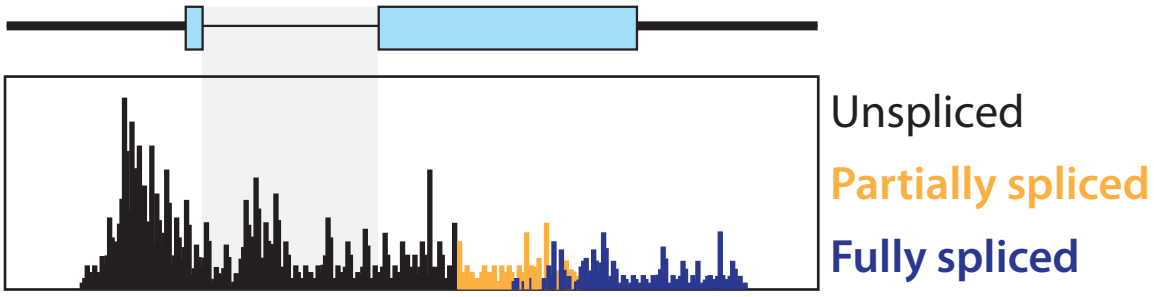


Figure 5. Hypothetical single gene data

Histogram mapping the 3' ends of nascent transcripts. The 3' ends of un-spliced and spliced transcripts are yellow and blue. These data would suggest that as RNAPII approaches the 3' end of this gene, all transcripts undergo the first and second steps of splicing, and have completed splicing before release from RNAPII.

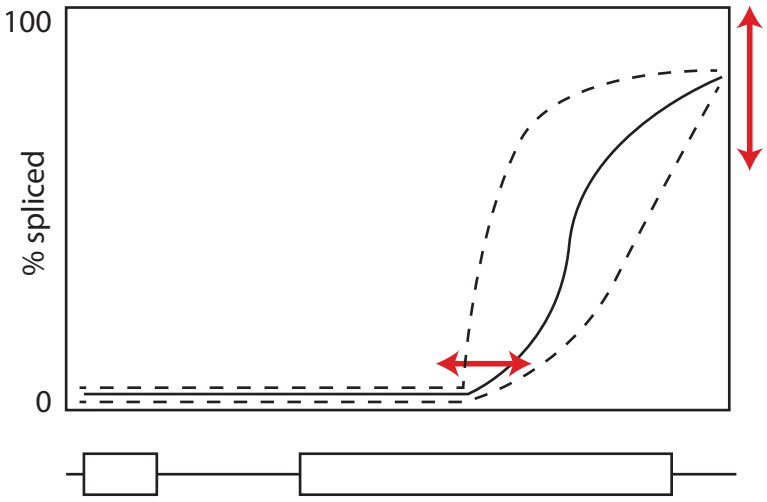


Figure 6. Hypothetical data showing cumulative splicing measurements

Vertical arrow on the right indicates the percent of transcripts that are spliced prior to release from RNAPII. Horizontal arrow indicates the location that spliced transcripts first appear in a gene. Shape of curve may be specific for individual genes, or could indicate whether there are location-specific regulators of splicing.

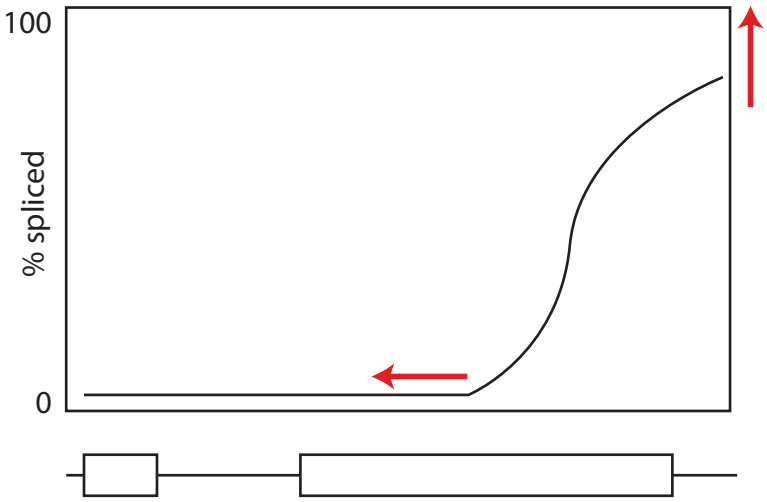


Figure 7. Possible results of slowing RNAPII

Slowing RNAPII could cause a 5' shift in the location that splicing is completed and an increase in the degree of co-transcriptional splicing. If a *cis*-acting signal defines where splicing occurs, one might not expect these differences.

REFERENCES

1. Alexander RD, Innocente SA, Barrass JD, Beggs JD. Splicing-Dependent RNA Polymerase Pausing in Yeast. Mol Cell 2010; 40:1–12.
2. Oesterreich FC, Preibisch S, Neugebauer KM. Global Analysis of Nascent RNA Reveals Transcriptional Pausing in Terminal Exons. Mol Cell 2010; 40:571–81.
3. Bhatt DM, Pandya-Jones A, Tong A-J, Barozzi I, Lissner MM, Natoli G, Black DL, Smale ST. Transcript Dynamics of Proinflammatory Genes Revealed by Sequence Analysis of Subcellular RNA Fractions. Cell 2012; 150:279–90.
4. Pandya-Jones A, Bhatt DM, Lin CH, Tong AJ, Smale ST, Black DL. Splicing kinetics and transcript release from the chromatin compartment limit the rate of Lipid A-induced gene expression. RNA 2013; 19:811–27.
5. Nilsen TW. Spliceosome assembly in yeast: one ChIP at a time? Nat Struct Mol Biol 2005; 12:571–3.
6. Churchman LS, Weissman JS. Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature 2011; 469:368–73.
7. Churchman LS, Weissman JS. Native Elongating Transcript Sequencing (NET-seq). Curr Protocols Mol Biol 2001; 98:4.14.1-4.14.17.
8. Alexander R, Beggs JD. Cross-talk in transcription, splicing and chromatin: who makes the first call? Biochem Soc Trans 2010; 38:1251.
9. Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M. The Transcriptional Landscape of the Yeast Genome Defined by RNA Sequencing. Science 2008; 320:1344–9.
10. Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Münster S, Camblong J, Guffanti E, Stutz F, Huber W, Steinmetz LM. Bidirectional promoters generate pervasive transcription in yeast. Nature 2009; 457:1033–7.

TRANSCRIPTION TERMINATION

Question: Do the genes that genetically interact with Npl3 regulate transcription termination?

Background:

The *npl3*Δ strain exhibits positive and negative genetic interactions with multiple factors involved in chromatin and transcription (See chapter 1 for details). Because Npl3 has been shown to play a role in regulating the switch from transcription elongation to termination[1-4] it was possible that these genetically interacting factors may play a role in that process as well.

Experimental Approaches:

A panel of ß-galactosidase reporters had been constructed in which the intron of RP51 (RPS17A/B)[5] was interrupted with a transcription terminator sequence.[6] The termination signals in yeast are fairly degenerate; however, the termination site for the *ADH2* gene has been mapped and it is known that it requires specific elements that are present within the reporter construct.[6] ß-galactosidase activity in the extract from a given strain was calculated for both the WT reporter (pHZ18Δ2[6]) and the reporter with the *ADH2* terminator (pL101[6]) (Figure 8). This ratio was compared to the ratio for each mutant versus a matched WT strain. In WT cells, the WT reporter is well spliced and translated, but pL101 undergoes transcription termination in the intron and very little ß-galactosidase activity results. In a strain with defective transcription termination, readthrough occurs and an increase in ß-galactosidase activity is measured.

Results:

Strains with mutations in *NPL3* are known to have defects in transcription termination, resulting in readthrough past termination signals.[4] Our results with the ß-galactosidase reporters

revealed that when measuring the ratio of pL101/pHZ18Δ2, the *npl3*Δ strain exhibited an increase in this ratio compared to WT (Figure 9A), suggesting a defect in termination. We next used these reporters in strains that exhibited suppressive genetic interactions with *npl3*Δ (*i.e.,* deletion of these factors improved the *npl3*Δ growth defect). The genetics suggested that these genes/proteins might have opposing functions. However, we were surprised to find, like *npl3*Δ, deletion of *CHD1*, *CKB2*, *SET2*, and *EAF3* showed an increase in this ratio. Furthermore, when combining two deletions that each have a readthrough phenotype, one might expect an exacerbation; however, when analyzing the double mutants, we observed suppression of the *npl3*Δ readthrough phenotype.

In Chapter 1, we focused on the connection between Npl3 and the H2B ubiquitination machinery – we observed that by growth, abrogation of H2B ubiquitination exacerbated the growth defect of the *npl3*Δ strain, and deletion of factors responsible for removing the mark suppressed. Using the ß-galactosidase reporters, we saw that *bre1*Δ and *htb1K123R* caused improved transcription termination and *ubp8*Δ caused worse transcription termination. As above, the phenotypes of these single mutants are counter-intuitive considering their genetic interactions with *npl3*Δ. When looking at the double mutants with *npl3*Δ, there was no change with respect to the *npl3*Δ mutant alone, suggesting that these factors are not working together to regulate termination of this reporter gene. Lastly, if levels of ubiquitinated H2B are important for the behavior of this reporter we would expect *bre1*Δ to be epistatic to *ubp8*Δ, as neither the *bre1*Δ and the *bre1*Δ*upb8*Δ strains should have ubiquitinated H2B. Therefore, consistent with expectations, the double mutant strain phenocopied the *bre1*Δ strain (Figure 10).

While the reporter construct was designed to assay transcription termination, it is an unusual situation for the transcription machinery to encounter. Assuming that the 5' splice site is bound by U1, the splicing reaction will be in competition with the termination machinery, a scenario that one can imagine only occurs when something has gone wrong during transcription. Furthermore, I was interested in the interplay between chromatin and splicing, and while plasmids are generally

chromatinized in yeast, there was no guarantee that this chromatin would appropriately recapitulate that at endogenous loci. I made an attempt to validate the results from the reporter assay on the endogenous *GAL10/7* locus. Readthrough qPCR primers were designed such that the amplicon would span the intergenic space between *GAL10* and *GAL7*. To correct for altered expression of GAL10 in the analyzed strains, the abundance of this amplicon was normalized to an amplicon that falls at the 3' end of the *GAL10* open reading frame. As a positive control, *rna14-64*,[7] a mutant in a component of the cleavage and polyadenylation factor, was analyzed – as expected, this mutant exhibits a readthrough phenotype (Figure 11): an increase in the ratio of readthrough product/expression. Furthermore, in agreement with the ß-galactosidase reporter assay, loss of *NPL3* caused a readthrough phenotype. However, the genetic interacting partners of *npl3*Δ did not exhibit changes in readthrough when deleted individually. The double mutants of these with *npl3*Δ did not have any effect on the *npl3*Δ readthrough phenotype. I additionally performed this experiment to measure readthrough at the intron-containing RPS9B and made the same conclusion (data not shown). These data suggest either that these chromatin factors affect termination, but not at *GAL10* or *RPS9B*, or that they are not involved in termination at all.

These results further draw into question what the ß-galactosidase reporters are actually reporting on. There is the intriguing possibility that because of the unusual order of *cis* elements (5'SS – terminator – 3'SS), that these phenotypes are in some way reporting on a connection between splicing and transcription termination. In fact, in metazoans, in which exon definition occurs, the terminal exon is defined by an interaction between the spliceosome and the 3' end processing machinery.[8] Whether this coupling exists in *S. cerevisiae* is unknown. However, if this phenomenon does occur, perhaps mutation of one of the chromatin factors discussed here impacts this splicing/3' end processing interaction.

It is important to note that it has been previously reported that deletion of *NPL3* has the opposite phenotype to the one described here. Wong *et al.*[4] found using an endogenous deletion of

the GAL10 terminator (*GAL10*-Δ56) that the *npl3*Δ strain exhibits improved growth on galactose, a phenotype indicative of *improved* termination. Consistent with these results, the *npl3*Δ strain allowed early termination at the *RNA14* gene, a gene that auto-regulates using several early and weak termination sites. A potentially important difference in the two studies is that Wong *et al*. have analyzed this strain in the context of defective or weak terminators, but the terminators in the ß-galactosidase assay and at endogenous genes are strong.

Nevertheless, the phenotypes we found here were quite robust, and are very possibly indicative of some very interesting connections among chromatin, splicing, and transcription termination.

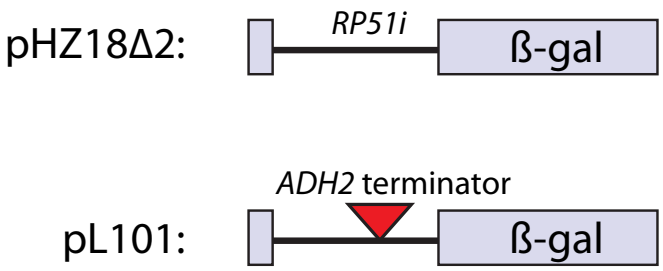


Figure 8. Schematic of ß-galactosidase reporter constructs used to assay transcription termination

The parent plasmid, pHZ18Δ2 contains the LacZ gene that is responsible for ß-galactosidase activity, interrupted by the intron of RP51. Upon being spliced, a functional LacZ transcript is made. In the termination reporter, the intron is interrupted by the terminator sequence of *ADH2*.

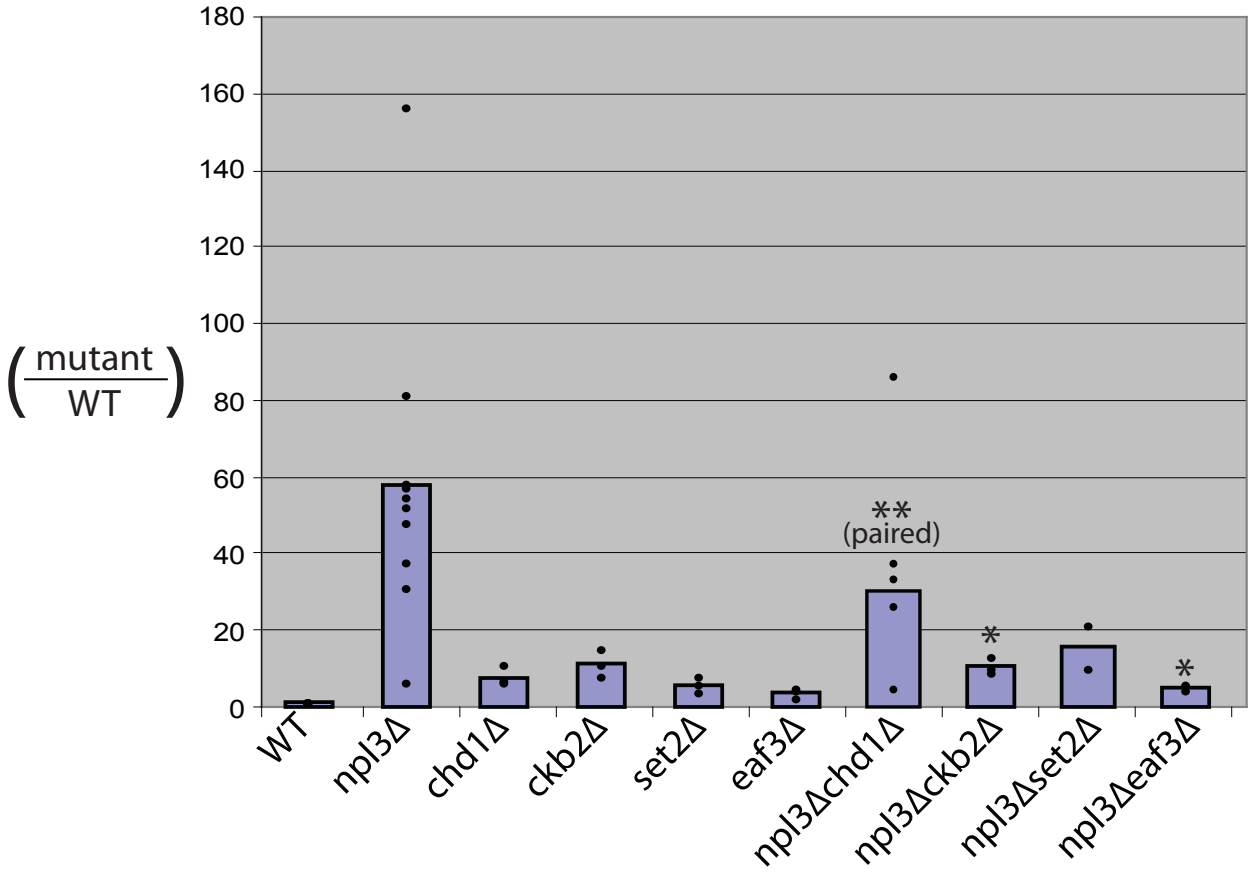


Figure 9. ß-galactosidase reporters illuminate strains with terminator readthrough.

Shown is the ratio of ß-galactosidase activity from the pL101 reporter to that from the pHZ18Δ2 reporter. These ratios are normalized to WT. Height of bars indicates average; dots indicate individual measurements. **paired $p<0.01$ by paired student t-test. * $p<0.05$ by unpaired student t-test.

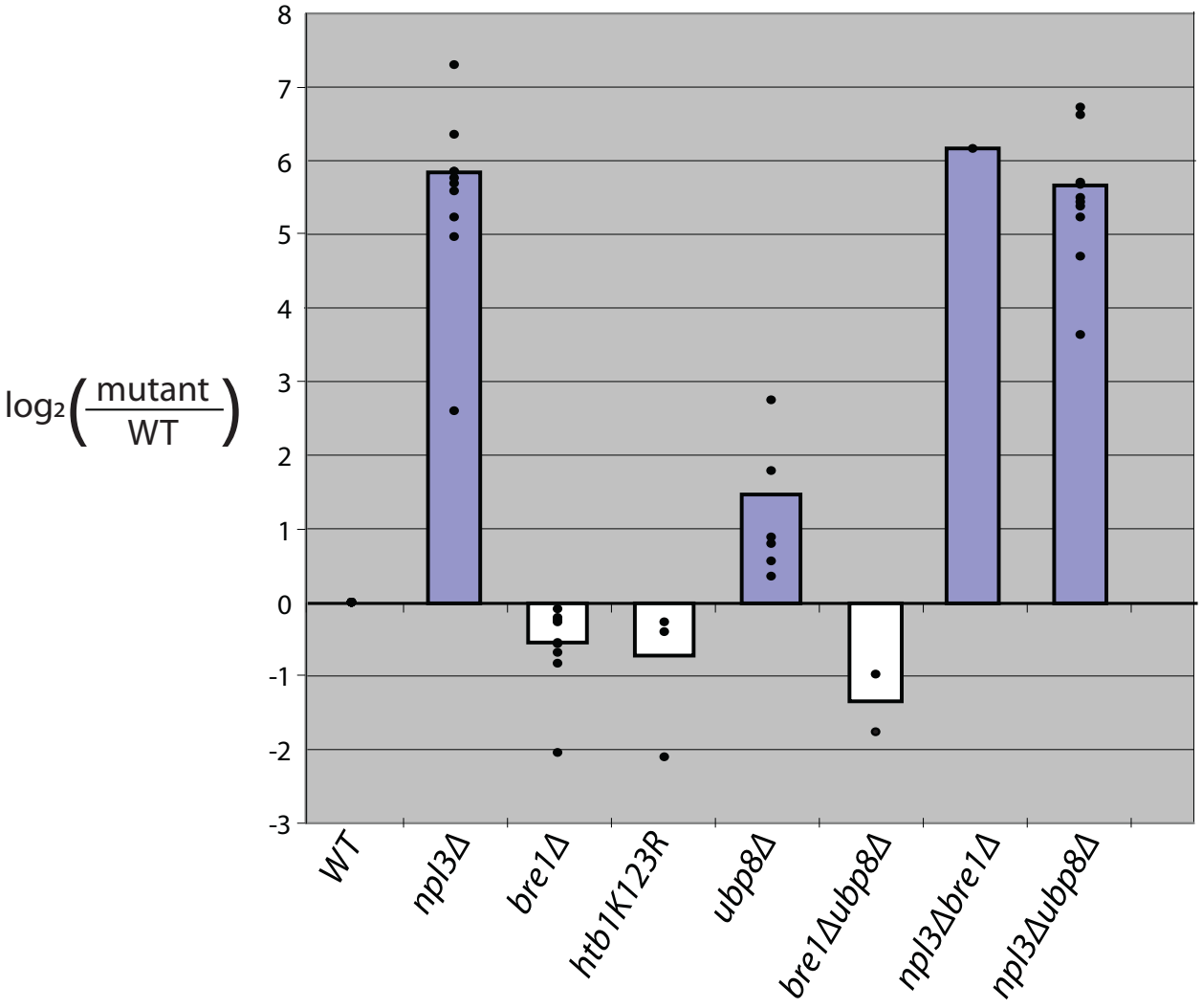


Figure 10. H2B ubiquitination machinery influence readthrough in ß-galactosidase reporters.

Shown is the ratio of ß-galactosidase activity from the pL101 reporter to that from the pHZ18Δ2 reporter. These ratios are normalized to WT and plotted on a $\log_2$ scale. Blue bars indicate increased readthrough was measured; white bars indicate decreased readthrough was measured. Height of bars indicates average; dots indicate individual measurements.

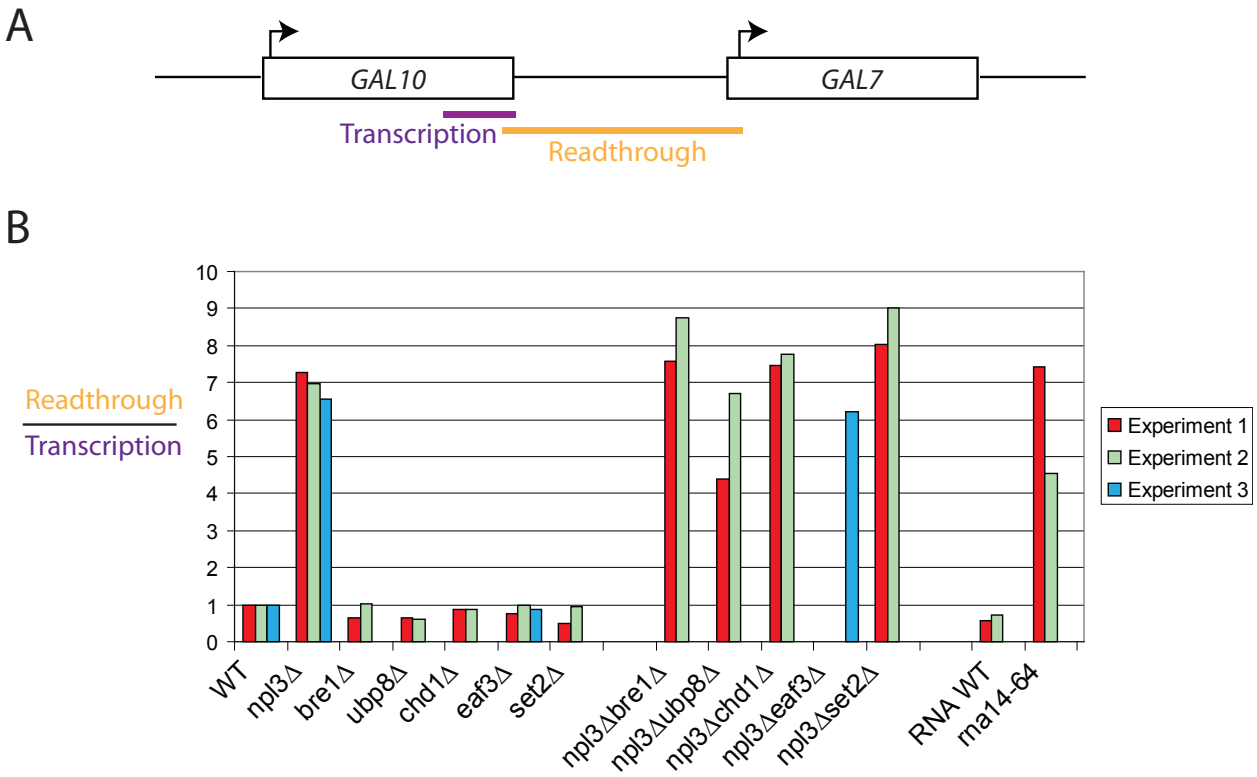


Figure 11. qPCR Readthrough assay at endogenous *GAL10/7* locus

(A.) Schematic of locus and location of qPCR amplicons. Arrows indicate direction of transcription.

(B.) Shown is the ratio of the abundance of readthrough product to the abundance of GAL10 transcription. The ratios are normalized to WT. Results of three independent experiments are shown.

REFERENCES

1. Bucheli ME, Buratowski S. Npl3 is an antagonist of mRNA 3' end formation by RNA polymerase II. EMBO J 2005; 24:2150–60.
2. Dermody JL, Dreyfuss JM, Villén J, Ogundipe B, Gygi SP, Park PJ, Ponticelli AS, Moore CL, Buratowski S, Bucheli ME. Unphosphorylated SR-like protein Npl3 stimulates RNA polymerase II elongation. PLoS ONE 2008; 3:e3273.
3. Wong CM, Tang HMV, Kong KYE, Wong GWO, Qiu H, Jin DY, Hinnebusch AG. Yeast arginine methyltransferase Hmt1p regulates transcription elongation and termination by methylating Npl3p. Nucleic Acids Res 2010; 38:2217–28.
4. Wong CM, Qiu H, Hu C, Dong J, Hinnebusch AG. Yeast Cap Binding Complex Impedes Recruitment of Cleavage Factor IA to Weak Termination Sites. Mol Cell Biol 2007; 27:6520–31.
5. Teem JL, Rosbash M. Expression of a beta-galactosidase gene containing the ribosomal protein 51 intron is sensitive to the rna2 mutation of yeast. Proc Natl Acad Sci USA 1983; 80:4403.
6. Hyman LE, Seiler SH, Whoriskey J, Moore CL. Point mutations upstream of the yeast ADH2 poly(A) site significantly reduce the efficiency of 3'-end formation. Mol Cell Biol 1991; 11:2004–12.
7. Hammell CM, Gross S, Zenklusen D, Heath CV, Stutz F, Moore C, Cole CN. Coupling of termination, 3' processing, and mRNA export. Mol Cell Biol 2002; 22:6441–57.
8. Dye MJ, Proudfoot NJ. Terminal exon definition occurs cotranscriptionally and promotes termination of RNA polymerase II. Mol Cell 1999; 3:371–8.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Erica A. Moehle
Signature

2/7/14
Date