

**UCSF**

**UC San Francisco Electronic Theses and Dissertations**

**Title**

Genetic Studies of Co-transcriptional pre-mRNA Splicing Regulation

**Permalink**

<https://escholarship.org/uc/item/9hw8c0pg>

**Author**

Nissen, Kelly Elizabeth

**Publication Date**

2017

Peer reviewed|Thesis/dissertation

Genetic Studies of Co-transcriptional pre-mRNA Splicing Regulation

by

Kelly E. Nissen

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Genetics

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

**Copyright 2017**

**by**

**Kelly E. Nissen**

## ACKNOWLEDGEMENTS

First and foremost, I want to thank my mentor and dissertation advisor Christine Guthrie. I have had the unique privilege of being Christine's last graduate student and she gave me tremendous freedom and support in charting my own scientific course. She has always treated me as a scientist, even when I did not feel that way about myself. To me, Christine is the epitome of a scientist and I will be forever grateful of her support and guidance.

The faculty at UCSF have also been instrumental in shaping my graduate training and I want to thank them. I especially want to recognize the immense contributions to my thesis committee members. First, Sandy Johnson, who counseled me into joining the Tetrad Program while I was still in college, has always had an open door to advise my work and career. He has advocated for my needs and pushed me to think independently. Next, Geeta Narlikar generously opened her lab meetings to me and provided critical insights into my primary research project. She supported me in coordinating its publication with external collaborators and helped me navigate the publication process. Hiten Madhani was an unofficial member of my thesis committee, my first rotation advisor, and he supported for my successful NSF Graduate Research Fellowship application. He and his lab also generously shared computational resources, *S. pombe* expertise and reagents. Nevan Krogan allowed me to perform part of my Guthrie lab rotation in his laboratory space and offered sage advice on our publication together, as well as connections to my collaborators. Moreover, other UCSF faculty have been outstanding rotation advisors and mentors, in particular, I'd like to thank Pat O'Farrell and Dave Toczyski. Barbara Panning did me a great honor by filling in on my qualifying exam committee on short notice and her contributions shaped my thesis work, and Wallace Marshall and Carol Gross have always been mentors that I could turn to when I felt lost.

Christine has a unique talent of bringing together phenomenal people towards a common goal and as the last student, I felt as if I “stood on the shoulders of giants.” I owe much of my success in graduate school to the support and guidance of members of the Guthrie Lab. I was drawn to join the Guthrie lab because of the “intellectual prowess,” as I dubbed it, that I saw the lab demonstrate as a group and that existed within each its members. In particular, I want to thank Kristin Patrick. Kristin was a joint post-doc in the Guthrie lab and Krogan lab, and she was my rotation supervisor. She then went on to become my in-lab mentor and a co-primary investigator on my main thesis project, which was born out of her own work. I looked up to Kristin my entire graduate experience, and hope that someday I can be as talented of a scientist and as clever, cool, and generous of a person as her. Additionally, the other postdocs in the Guthrie lab are phenomenal scientists and it was a pleasure learning from all of them. Thank you to Mike Marvin, my bay-mate and *S. pombe* ally; Sarah Ledoux, Jaci Greimann, and Megan Mayerle, for always being willing to answer questions and guide me through the Ph.D. process; Anne de Bruyn Kops, who has been with me in the Guthrie lab my entire time and always treated me with respect; Rebecca Holmes, Quinn Mitrovich, Corina Maeder, and Yangzi He; John Abelson, who is the biggest fan of science I know and always looks on the bright-side of life; Michael Dinglasan our lab technician for countless plates and helping me make special media under a tight time-frame for reviews. The graduate students in the Guthrie lab that I had the pleasure to overlap with and learn from never cease to amaze me. Thank you for all of your inspiration and guidance to Alex Plocik, Erica Moehle, Argenta Price, Bellos Hadjivassiliou, and Megan Bergkessel, my most recent collaborator.

Because I've learned how truly collaborative science can be, I owe a huge debt of gratitude to my collaborators. In particular, I would like to thank my friend and colleague Christina Homer for volunteering her computational expertise and helping me complete my thesis project. Thank you to Colm Ryan and Mike Shales for your computational support. Thank you to Tracy Johnson, Lauren Neves, and Tracy Kress for coordinating the publication of our work on H2A.Z and the Swr1 Complex and demonstrating the power of converging discoveries.

UCSF is an incredible environment and I have been surrounded by world-class colleagues and peers. I want to thank all of my incredible colleagues and friends from the 2009 Tetrad entering class. In particular, Phillip Dumesic, who provided critical feedback on my first publication and always made me feel like an equal; Isabel Nocedal, Dustin Dovala, and Trevor Sorrells for creating a fun and welcoming community in the North 3<sup>rd</sup> floor lab area; Josh Kane, who travelled with me to UCSF from the Freeling lab and never stopped feeling like a co-worker; Robert Stanley, Coral Zhou, and Alison Leaf for fun adventures in SF and friendship. Beyond my Tetrad class, I've made so many wonderful friends at UCSF, who continue to support me to this day. Thank you especially to Anna Reade, Theresa Berens, Josh Dunn, Michael Todhunter, Justin Farlow, Stefan Isaac, Silvi Rouskin; and the women of Women in Life Sciences at UCSF (WILS), particularly Linet Mera, Vanessa Van Voorhis, Candace Bitton, and Liron Noiman. Thank you also to my friends and mentors in other labs and programs, who enriched my daily experience at UCSF and for your continued support: Jordan Burke, Jessica Brown, Manon Eckhardt, Veronica Pessino, and Ali Taubes; members of the Madhani and Johnson labs.

I would not have survived this experience without the love and support of my community of friends and family. In particular, I want to thank Starlynn Clarke, Joanna Tong, Jackie Lambe,

Jenny McGarvey, Rachel Bramwell, Morad Alvarez, and Stephen Champ. I love you all, and you're the reason I made it through! Thank you for your unending patience and great adventures outside of the lab. I will be forever grateful to my parents, Ted and Nancy Nissen, who inspired me to pursue science, are my tireless cheerleaders, and supported me with love throughout this journey. Finally, I cannot express enough gratitude to Corey Siegel. Corey supported me during the last five years of my Ph.D. with an untold number of late night trips to lab, lifting heavy lab equipment when I was injured, and helping me with computational analysis of my thesis work. I could not have done this without you.

## **PUBLISHED WORK**

Chapter 1 is reproduced from work that is published in *Genes and Development* in collaboration with Kristin Patrick, Christina Homer, and Nevan Krogan's Lab:

Nissen KE, Homer CM, Ryan CJ, Shales M, Krogan NJ, Patrick KL, Guthrie C. 2017. The histone variant H2A.Z promotes splicing of weak introns. *Genes & development* **31**: 688-701.



# Genetic Studies of Co-transcriptional pre-mRNA Splicing Regulation

Kelly E. Nissen

## ABSTRACT

During eukaryotic gene expression, non-coding intron sequences are removed from the protein-coding exon sequences of pre-messenger RNA (pre-mRNA) by the process of splicing. Splicing is carried out by the spliceosome, a macromolecular ribozyme that assembles *de novo* on every intron. It has been established that the spliceosome assembles co-transcriptionally while the nascent pre-mRNA is attached to the elongating RNA polymerase II (RNAP II). However, mechanisms that enforce the coupling of splicing to transcription and the chromatin landscape remain poorly understood.

Here, I describe our work to gain further insight into the regulation of co-transcriptional pre-mRNA splicing by employing genetic and molecular biology approaches. Using the traditional splicing model organism *Saccharomyces cerevisiae*, we interrogated the mechanism of a rapid splicing block at the ribosomal protein genes (RPGs) that occurs in response to amino acid starvation stress. By creating chimeric gene constructs, we show that an RPG promoter sequence element, and not the intron sequence, is both necessary and sufficient to confer a transcript-specific splicing response. This study reveals a mechanism whereby in response to stress RNAP II becomes de-enriched near the 3' splice site and early splicing factor recruitment is delayed, consistent with an environmentally-responsive splicing-dependent polymerase checkpoint, and underscoring the highly coordinated nature of splicing and transcription.

We further expanded our studies of splicing regulation into *Schizosaccharomyces pombe*, an emerging model organism that shares many splicing features with mammalian systems, such as

abundant non-consensus splice sites, multi-intron containing genes, and conserved *trans*-acting splicing factors. In our work with the *S. pombe* system, we developed techniques for probing splicing factor association with chromatin. Starting from a genome-wide genetic screen, we characterized interactions between the spliceosome and Swr1 complex, which deposits the highly conserved histone variant H2A.Z. We identified H2A.Z's requirement in efficient splicing at introns with weak splice sites, supporting a growing hypothesis of chromatin-dependent splicing checkpoints. Taken together, these studies reveal an unappreciated role of the promoter and its proximal nucleosome composition in transcript-specific pre-mRNA splicing regulation. We expect these results to provide foundational insights into the interconnectedness of gene expression processes and have important implications in human disease.

## TABLE OF CONTENTS

<b>Introduction</b> .....	1
<b>Chapter 1: The Histone Variant H2A.Z Promotes Splicing of Weak Introns</b> .....	10
<b>Chapter 2: A Ribosomal Protein Gene Promoter Element Influences Co-transcriptional Splicing Regulation in Response to Amino Acid Starvation in <i>S. cerevisiae</i></b> .....	77
<b>Conclusion</b> .....	122
<b>Appendix</b> .....	131
Part A: What is the genetic relationship between <i>swr1</i> and <i>pht1</i> in influencing pre-mRNA splicing? .....	133
Part B: What residues in H2A.Z are required for its role in splicing regulation? .....	140
Part C: Splicing factor chromatin immunoprecipitation (ChIP) protocol optimization in <i>S. pombe</i> .....	147
<b>References</b> .....	157

## LIST OF TABLES

### Chapter 1:

Available online at [genesdev.cshlp.org](http://genesdev.cshlp.org)

Supplemental Table 1: Splicing Sub-complex Definitions.....	76
Supplemental Table 2: GO Term Analysis Data .....	76
Supplemental Table 3: Strain List.....	76
Supplemental Table 4: Processed Microarray Data.....	76

### Chapter 2:

Table 1: Strain List .....	121
----------------------------	-----

## LIST OF FIGURES

### Chapter 1:

Figure 1: The Swr1 Complex has strong genetic interactions across the spliceosome. ....	46
Figure 2: H2A.Z is required for pre-mRNA splicing. ....	48
Figure 3: Loss of H2A.Z impairs splicing of weak introns at the 5' ends of genes. ....	50
Figure 4: H2A.Z is enriched at splice sites. ....	52
Figure 5: Overexpression of <i>prp16</i> ATPase suppresses growth and splicing defects in <i>pht1Δ</i> and <i>swr1Δ</i> at restrictive temperatures. ....	54
Figure 6: Mutating <i>cis</i> -splicing sequences to consensus suppresses the splicing defect in <i>pht1Δ</i> at 16°C. ....	56
Figure 7: Model illustrating the effect of H2A.Z loss on splicing efficiency and suppression by overexpression of <i>prp16</i> and exacerbation by overexpression of <i>prp43</i> . ....	57
Supplemental Figure 1: Splicing cycle diagram ....	58
Supplemental Figure 2: Extended EMAP scores and <i>swr1Δ</i> serial dilutions ....	59
Supplemental Figure 3: Additional panels supporting Figure 2 ....	61
Supplemental Figure 4: Splicing factor gene expression in Swr1 Complex mutants ....	63
Supplemental Figure 5: Comparison between introns with splicing in the <i>pht1Δ</i> microarray at 16°C and introns found to employ alternative splice sites in <i>S. pombe</i> . ....	65
Supplemental Figure 6: Comparison of H2A.Z ChIP-seq replicates. ....	66
Supplemental Figure 7: Additional replicates of pile up graphs of H2A.Z ChIP-seq around the TSS, 5'SS, and 3'SS ....	67

Supplemental Figure 8: Pile up graphs of H2A.Z ChIP-seq around the TSS, 5'SS, and 3'SS for genes with introns affected by <i>pht1Δ</i> .....	68
Supplemental Figure 9: First exon lengths in <i>S. pombe</i> .....	70
Supplemental Figure 10: Additional replicates of average H2A.Z coverage analysis around TSS .....	71
Supplemental Figure 11: Additional replicates of average H2A.Z coverage analysis for introns in the top and bottom deciles of <i>cis</i> -splicing sequence strength.....	72
Supplemental Figure 12: H2A.Z-associated genes are lowly expressed .....	74
Supplemental Figure 13: <i>prp16</i> and <i>prp43</i> expression levels.....	75

## Chapter 2:

Figure 1: Schematics of gene structures (WT or chimeric) used.....	108
Figure 2: Effects of intron on expression, splicing efficiency, and splicing response to 3-AT..	109
Figure 3: Effects of promoter on expression, splicing efficiency, and splicing response to 3-AT. ....	111
Figure 4 (two pages): ChIP profiles for U1, U2, RNAP II and RNAP II in a $\Delta$ <i>ckb2</i> strain with or without 3-AT treatment.....	113
Figure 5: Aggregated ChIP-Seq profiles for U1, U2, RNAP II and RNAP II in a $\Delta$ <i>ckb2</i> strain with or without 3-AT treatment, comparing intron-containing RPG transcripts with all other intron containing transcripts. ....	116
Figure 6: Model.....	119

## Appendix:

Figure A.1: H2A.Z-dependent splicing defect partially suppressed in *pht1Δ swr1Δ* double mutant strain..... 138

Figure A.2: Acetylation cycling on Pht1 is important for splicing..... 145

Figure A.3: U1 and U2AF splicing factor ChIP on two highly expressed genes in *S. pombe* ... 156

## INTRODUCTION

### **Splicing is a central regulation point in gene expression**

The central dogma of molecular biology posits that genetic information flows from DNA to RNA and then to protein through the gene expression process. Gene expression is tightly regulated through each step (transcription, messenger RNA (mRNA) processing, export, translation) in both space and time to ensure that the correct messenger ribonucleoprotein (mRNP) is formed. Pre-messenger RNA (pre-mRNA) undergoes several processing steps to remove sequence information that is not desired for that specific mRNA, and to check that the transcript has the signals necessary for export and translation.

Splicing is a complex and dynamic processing pathway during gene expression that leads to the removal of intervening sequences called introns from the protein-coding exon sequences in pre-mRNA. Splicing is carried out by the spliceosome, a highly conserved, macromolecular ribozyme, in an ordered, step-wise manner for each splicing event (Staley and Guthrie 1998; Wahl et al. 2009). The spliceosome assembles primarily from five small nuclear ribonucleic particles (snRNPs) and is collectively composed 5 RNAs, and of over 100 proteins, depending on the species (Wahl et al. 2009; Cvitkovic and Jurica 2013). The splicing process begins with the U1 snRNP recognizing the 5' splice site (5'SS) and the branch point binding protein/U2AF recognizing the branch point (BP) assembly on the pre-mRNA intron, forming the pre-spliceosome. Next, the U2 snRNP replaces the branch point binding protein/U2AF at the BP within the intron. Subsequently, the U4/U6•U5 tri-snRNP is recruited to the intron and the U1 and U4 snRNPs are released from the assembly as a consequence of massive rearrangements, leading to a catalytically active spliceosome (Brow 2002). The first catalytic step then takes



place, in which an adenosine in the BP of the RNA, bulged out by the spliceosome, performs a nucleophilic attack on the 5' SS, thereby cleaving the 5' exon and creating a lariat intermediate. Next, in the second catalytic step, the 5' SS performs a nucleophilic attack on the 3' splice site (3' SS), which causes the ligation of the 5' and 3' exons. The remaining lariat intron and the spliced RNA are released from the spliceosome, which is disassembled and recycled to allow for further splicing reactions (Fabrizio et al. 2009).

Our understanding of the dynamics of spliceosome assembly and catalytic complexes has been greatly aided in the past two years by breakthrough structural studies in yeast and humans (Yan et al. 2015; Galej et al. 2016; Nguyen et al. 2016; Rauhut et al. 2016; Wan et al. 2016; Yan et al. 2016; Bertram et al. 2017a; Bertram et al. 2017b; Yan et al. 2017). These tremendous efforts have revealed snapshots of megadalton spliceosomes at various steps along the splicing pathway, revealing changes in protein composition and spatial arrangements of the active site, substrates and products. Rearrangements of the spliceosome during assembly and catalysis are driven by splicing DExD/H ATPases and are essential for the two-step trans-esterification reaction (Staley and Guthrie 1998; Cordin et al. 2012). These enzymes regulate the rate of spliceosomal rearrangements, which creates an opportunity for proofreading of the substrates at each step. Proofreading each step in the splicing reaction is a checkpoint for the spliceosome to distinguish introns with sub-optimal *cis*-splicing sequences (5' SS, BP, 3' SS), which may take more time to process or discard truly incorrect substrates (Horowitz 2011).

The regulation of spliceosome assembly and catalysis is tightly controlled to produce both canonical and alternative transcripts from a diverse set of intronic sequences, under a variety of environmental and intercellular signaling conditions (Kalsotra and Cooper 2011). Canonical splicing refers to splicing of a single intron from the most commonly used splice sites

for that intron while alternative splicing events can take many forms including exon skipping, intron retention, use of alternative splice sites, and others (Blencowe 2006). Alternative splicing is a mechanism for generating protein diversity in a cell-specific or context-specific manner, and a driver of evolution (Keren et al. 2010). A large number of human diseases are associated with mutations or mis-regulation of both canonical and alternative pre-mRNA splicing events, underscoring the pathway's essential role in gene expression (Fackenthal and Godley 2008; Wang et al. 2012; Vanichkina et al. 2017). Nevertheless, the mechanisms by which splicing decisions are regulated within the complex nuclear milieu remain poorly understood.

### **Splicing occurs co-transcriptionally**

Concealed in the central dogma pathway is the possibility for the underlying DNA sequences or environment to influence mRNA production. It has been well established that pre-mRNA splicing occurs co-transcriptionally, while the nascent pre-mRNA is still attached to the elongating RNA polymerase II (RNAP II) and that splicing is regulated by both chromatin structure and transcription. Throughout the duration of my graduate training, there have been numerous review articles written on the coordination between splicing, chromatin and transcription. Many of the individual studies of co-transcriptional splicing regulation look at how factors in one pathway influence the others, for example, how a chromatin factor influences either transcription or splicing. Putting a large field worth of experiments together yields a picture of extensive crosstalk between these pathways and insights into how modulating one factor can have both downstream and auto-regulatory effects. Yet, an overarching picture of all of the connections between chromatin, splicing and transcription are intertwined is still on the

horizon, and will require further *in vitro*, *in vivo*, and high-throughput genome-wide technique development.

One model to ensure co-transcriptional splicing is through direct communication between splicing and chromatin. Chromatin structure may act as a platform for directly recruiting splicing factors to the site of transcription, or conversely, spliceosome assembly on *cis*-splicing sequences may affect the recruitment of chromatin remodeling factors, called the “recruitment model”. In an alternative model, splicing and chromatin may communicate indirectly through the transcription complex, called the “kinetic model”. A large body of work, for which it is presently impossible to disentangle the direct or indirect nature of physical and genetic interactions, supports both of these models (Merkhofer et al. 2014; Herzel et al. 2017).

For example, splicing signals have been shown to regulate transcription *in vitro* and *in vivo*. *Cis*-splicing sequences, such as splice sites and full introns, have been shown to increase transcription rates via nuclear run-on experiments (Fong and Zhou 2001; Furger et al. 2002; Damgaard et al. 2008). Additionally, in budding yeast chromatin immunoprecipitation (ChIP) experiments, RNA polymerase II (RNAP II) had higher levels of association directly downstream of an intron, and this enrichment was lost when the upstream intron’s splice sites were mutated (Alexander et al. 2010). In this landmark experiment, when compensatory mutations in snRNPs were combined with splice site mutations, RNAP II ChIP signal was restored to a wild-type pattern, showing that the regulation of transcription by splicing is mediated through a direct interaction between splice site sequences and the spliceosome. However, the mechanism of how co-transcriptional spliceosome assembly on *cis*-splicing signals regulates transcription is unknown. Moreover, it is not well understood whether proteins are

necessary to ensure the spliceosome is recruited to the site of transcription in order to correctly form an mRNP.

It is well established that chromatin remodeling processes such as nucleosome remodeling and histone modification influence transcription and splicing (Workman and Kingston 1998; Li et al. 2007). For example, transcription dynamics can be regulated by nucleosomes; nucleosomes impose a barrier to Pol II elongation *in vitro* and knocking-down expression of histone proteins leads to increased transcription initiation *in vivo* (Han and Grunstein 1988; Hodges et al. 2009). The speed of transcription can also influence splicing, as shown by Pol II point mutant with a slow elongation rate causes an accumulation of intron-containing pre-mRNA and skipping of alternative exons in the human fibronectin gene (de la Mata et al. 2003; Braberg et al. 2013; Fong et al. 2014). Evidence for chromatin playing a direct role in co-transcriptional spliceosome association comes from mammalian studies which show two chromatin remodeling proteins, Chd1 and MRG15, bind specific histone modifications (H3K4me3 and H3K36me3, respectively) and splicing factors (SF3a and polypyrimidine tract binding protein, respectively), and enhance the co-transcriptional association of those splicing factors by ChIP (Sims et al. 2007; Luco et al. 2011). An example of both of these models is that HP1gamma histone protein in humans regulates alternative exon inclusion of the CD44 gene by bridging splicing factors U2AF65 and Prp8 with chromatin, and increasing RNAP II ChIP association (Saint-Andre et al. 2011). However, there is limited evidence for whether co-transcriptional spliceosome assembly on pre-mRNA recruits chromatin remodeling factors. I am excited to see what the future holds in this line of inquiry.

### **A yeast model for every splicing question**

Both the budding yeast *Saccharomyces cerevisiae* and the fission yeast *Schizosaccharomyces pombe* are fundamental model organisms for the study of molecular and cellular biology. As unicellular eukaryotes, they are powerful genetic and biochemical tools to identify and characterize factors in a variety of biological processes that are conserved from yeast to humans. *S. pombe* and *S. cerevisiae* diverged ~350 million years ago, and share approximately the same evolutionary distance between them as either of these yeasts do with mammals (Sipiczki 2000; Heckman et al. 2001). While both yeasts provide opportunities to interrogate the pre-mRNA splicing pathway, differences in their genomes can be taken advantage of to best address core splicing mechanism versus regulatory inputs of co-transcriptional splicing. Moreover, comparing knowledge between the two yeast systems can yield further insight into the evolutionary origins of splicing mechanisms and how splicing functions in humans.

Much of our understanding of pre-mRNA splicing mechanism has come from work in budding yeast, *Saccharomyces cerevisiae*. Over the last 30 years, many core splicing factors were identified by random mutagenesis genetic screens for temperature and cold sensitivity; suppressor screens also have been instrumental in identifying how these splicing factors work in a pathway and also helped define the roles of the splicing ATPases in splicing fidelity (Vijayraghavan et al. 1989; Noble and Guthrie 1996). Additionally, the biochemical mechanism of spliceosome assembly and splicing catalysis was largely worked out using *in vitro* splicing reactions from purified budding yeast extracts (Wassarman and Steitz 1992). The budding yeast spliceosome is easier to analyze and utilize *in vitro* because it has retained fewer splicing factors that are conserved in mammals (Wahl et al. 2009; Cvitkovic and Jurica 2013). Furthermore, less than 5% of the *S. cerevisiae* genome contains introns, which are over-represented in the highly

expressed class of genes, the ribosomal protein genes (RPGs) (Ares et al. 1999; Bon et al. 2003; Cherry et al. 2012). With few exceptions, this relatively small number of introns occurs as a single intron in genes, largely removing the opportunity for alternative splicing (Davis et al. 2000; Ast 2004). These introns also have *cis*-splicing sequences that are highly constrained, and thus rarely diverge from a consensus sequence, and make understanding rules of splicing mechanism easier to interrogate (Kupfer et al. 2004). Despite the field-changing advances to our understanding of pre-mRNA splicing brought about by using budding yeast, questions that pertain to splicing in the context of complex gene architectures are unable to be addressed, underscoring the need for another complementary model system to study splicing regulatory mechanisms.

Compared to *S. cerevisiae*, pre-mRNA splicing in *S. pombe* more closely resembles that in humans, making it a better-suited model organism for studying the splicing of genes with complex intron/exon architectures within genes, alternative splicing, and how conserved auxiliary splicing factors influence splicing outcomes. While maintaining nearly the same number of genes as *S. cerevisiae* (5054 genes versus 5821 genes in budding yeast), almost half of the genes in *S. pombe* contain introns, and of those, roughly half are multi-intron containing genes, comprising greater than 5000 introns (Hoffman et al. 2015). The introns in *S. pombe* have less constrained *cis*-splicing sequences across the entire intron, which allows for observed alternative splicing events in at least 2% of genes and in others during certain environmental conditions (Kupfer et al. 2004; Awan et al. 2013; Bitton et al. 2015; Stepankiw et al. 2015). For example, the fission yeast BP consensus sequence of CURAY is more similar to that of humans than the UACU AAC of *S. cerevisiae* (Kaufer and Potashkin 2000; Ast 2004; Kupfer et al. 2004). Also, there is a polypyrimidine tract in the majority of fission yeast introns, which can occur

upstream and/or downstream of the BP (Sridharan et al. 2011). This variety of intron distribution and sequence within a gene and throughout the genome makes *S. pombe* more reliant on the balance between splicing efficiency and splicing fidelity and provides opportunities to study splicing genome wide with statistically significant sampling of introns. Additionally, several important *trans*-acting splicing factors are present in *S. pombe* and humans that are absent from *S. cerevisiae*, such as the small subunit of U2AF, multiple SR proteins, and splicing protein kinase Prp4, all of which help the spliceosome identify non-consensus splice sites and influence alternative splicing decisions (Kaufer and Potashkin 2000). Overall, 38% of splicing factors in fission yeast are more similar in sequence to the human proteins than to their budding yeast homologs (Kaufer and Potashkin 2000). Taken together, the intron architecture and composition of *S. pombe* makes it an attractive system to study questions pertaining to the regulated usage of diverse splicing signals while *S. cerevisiae* presents opportunities to dissect core splicing mechanisms on strongly consensus model introns with a streamlined spliceosome.

In this work, my colleagues and I utilize both of these model yeast systems to understand different facets of co-transcriptional pre-mRNA splicing regulation, using genetic and molecular biology techniques. Underlying these studies is my interest in understanding the regulatory mechanisms that govern how cells make dynamic decisions about gene expression levels in response to environmental changes and disease. Here, I explore how cells harness the chromatin landscape and specific sequences of the promoter to enhance constitutive splicing efficiency of weak introns and rapidly control splicing in response to stress. In Chapter 1, we identify and characterize the role of the highly conserved histone variant, H2A.Z, as promoter-proximal beacon for co-transcriptional splicing regulation of 5' introns in lowly expressed genes with non-consensus splice sites. In Chapter 2, we determine that sequences in the promoter and not the

intron are required for a rapid accumulation of unspliced RNA during amino acid starvation stress, in coordination with changes in spliceosome recruitment and polymerase distribution. Additionally, I show unpublished studies related to work in Chapter 1 that includes a method for splicing factor chromatin immunoprecipitation in *S. pombe*, in the Appendix. Taken together, these studies highlight the role of the promoter region in connecting transcription and the chromatin landscape to co-transcriptional mRNA processing. By better understanding the inputs that regulate pre-mRNA splicing co-transcriptionally, we reveal handles for future work in manipulating splicing in human disease.



## **CHAPTER 1:**

The Histone Variant H2A.Z Promotes Splicing of Weak Introns

## The Histone Variant H2A.Z Promotes Splicing of Weak Introns

Kelly E. Nissen<sup>1</sup>, Christina M. Homer<sup>1</sup>, Colm J. Ryan<sup>2</sup>, Michael Shales<sup>3,4</sup>, Nevan J. Krogan<sup>3,4,5</sup>,  
Kristin L. Patrick<sup>1,6,\*</sup>, Christine Guthrie<sup>1,\*</sup>

1. Department of Biochemistry and Biophysics, University of California, San Francisco, California, United States of America
2. Systems Biology Ireland, University College Dublin, Belfield, Dublin, Ireland
3. Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California, United States of America
4. California Institute for Quantitative Biosciences, QB3, San Francisco, California, United States of America
5. J. David Gladstone Institutes, San Francisco, California, United States of America
6. Department of Microbial Pathogenesis and Immunology, Texas A&M Health Science Center, Bryan, Texas, United States of America

\*Corresponding Author

**Author Contributions:**

Kelly Nissen: conceived experiments, made strains and reagents, designed and performed experiments, analyzed data, wrote the paper

Christina Homer: designed and performed experiments, analyzed data

Colm Ryan: analyzed data

Michael Shales: provided computational support

Nevan Krogan: supervised the research

Kristin Patrick: conceived experiments, made strains and reagents, supervised the research, wrote the paper

Christine Guthrie: conceived experiments, supervised the research, wrote the paper

**Abstract:**

Multiple lines of evidence implicate chromatin in the regulation of pre-mRNA splicing. However, the influence of chromatin factors on co-transcriptional splice-site usage remains unclear. Here we investigated the function of the highly conserved histone variant H2A.Z in pre-mRNA splicing using the intron-rich model yeast, *Schizosaccharomyces pombe*. Using Epistatic Mini-Array Profiles (EMAPs) to survey the genetic interaction landscape of the Swr1 nucleosome remodeling complex, which deposits H2A.Z, we uncovered evidence for functional interactions with components of the spliceosome. In support of these genetic connections, splicing-specific microarrays show that H2A.Z and the Swr1 ATPase are required during temperature stress for the efficient splicing of a sub-set of introns. Notably, affected introns are enriched for H2A.Z occupancy, and more likely to contain non-consensus splice sites. To test the significance of the latter correlation, we mutated the splice sites in an affected intron to consensus and found this suppressed the requirement for H2A.Z in splicing of that intron. These data suggest that H2A.Z occupancy promotes co-transcriptional splicing of sub-optimal introns that may otherwise be discarded via proofreading ATPases. Consistent with this model, we show that over-expression of splicing ATPase Prp16 suppresses both the growth and splicing defects seen in the absence of H2A.Z.

## **Introduction:**

Functional crosstalk between chromatin, transcription and mRNA processing is well documented (de Almeida and Carmo-Fonseca 2014). Multiple lines of evidence support a model whereby pre-mRNA splicing is primarily co-transcriptional and splicing factors are recruited to, and act upon, nascent RNA while it is engaged with the RNA polymerase II (RNAP II) complex (Merkhofer et al. 2014). Chromatin structure is an important foundation for the coordination of splicing and transcription, providing both a recruitment platform for factors in both pathways and an environment to modulate the kinetics of transcription, which influences splice site recognition and the kinetics of splicing (Gomez Acuna et al. 2013). As such, ablating chromatin factors or altering their distribution leads to defects in splicing efficiency and changes in alternative splicing patterns (Naftelberg et al. 2015). However, the mechanisms by which chromatin influences splice-site choice are still poorly understood.

Splicing, the process of removing intronic sequences from pre-mRNA, is carried out by the spliceosome, a highly conserved macromolecular machine, in an intricate step-wise manner (Wahl et al. 2009) (Supplemental Figure S1). Spliceosome assembly and catalysis are tightly regulated to produce both canonical and alternative transcripts from a diverse set of *cis*- splicing sequences (Ast 2004). The splicing reaction is driven forward by ATPases that control the rate of spliceosomal rearrangements while providing an opportunity for proofreading of the substrates (Staley and Guthrie 1998; Koodathingal and Staley 2013). Proofreading allows the spliceosome to identify introns with sub-optimal *cis*-splicing sequences that may take more time to process or discard incorrect substrates (Staley and Guthrie 1998; Koodathingal and Staley 2013). In organisms with many diverse intronic signals, auxiliary splicing factors aid splice site selection, but precisely how additional gene expression factors, such as chromatin, participate remains

unclear (Chen and Manley 2009).

The distribution and modification pattern of nucleosomes serves not only as a marker of gene boundaries but also provides information across the gene body during gene expression (Campos and Reinberg 2009). Nucleosomes, by virtue of their enriched occupancy in exons, are thought to play a role in defining exon-intron boundaries and differential histone modifications have been shown to influence splicing efficiency and alternative splicing outcomes (Schwartz and Ast 2010). Reciprocally, splicing factors and intron/exon architecture have been shown to direct histone modifications and positioning (de Almeida et al. 2011; Kim et al. 2011; Bieberstein et al. 2012). In addition to histone modification, replacement of canonical histone proteins with histone variants is another way to add specialization to the nucleosome. H2A variant H2A.Z is conserved across 90% of eukaryotes with 60% sequence similarity (Zlatanova and Thakar 2008). Deposition of the H2A.Z/H2B dimer in exchange for a canonical H2A/H2B dimer is mediated by the Swr1 complex (Krogan et al. 2003; Kobor et al. 2004; Mizuguchi et al. 2004) and typically occurs at the +1 and -1 nucleosomes around the transcription start site (TSS) (Subramanian et al. 2015). Recently, H2A.Z has also been found to incorporate into gene bodies in metazoans (Weber et al. 2010; Tolstorukov et al. 2012). Notably, closely related variant histones in *Drosophila melanogaster* (H2A.V) and *Homo sapiens* (H2A.Bbd) are specifically enriched at intron-exon boundaries where they promote efficient pre-mRNA splicing (Weber et al. 2010; Tolstorukov et al. 2012).

Our initial efforts to examine how the chromatin landscape influences splicing utilized *Saccharomyces cerevisiae* (Wilmes et al. 2008). Although genome-wide genetic interaction mapping uncovered many connections between splicing and other gene expression pathways, the scope of this analysis was limited by the fact that the budding yeast genome contains only ~250

introns and most adhere to strong consensus *cis*-splicing signals. In order to extend this analysis to a more metazoan-like system, we turned to the fission yeast, *Schizosaccharomyces pombe*, the genome of which encodes over 5,000 introns that employ a diverse array of *cis*-splicing signals and a spliceosome that shares more conservation with metazoans (Wood et al. 2002; Kuhn and Kaufer 2003; Sridharan et al. 2011). Furthermore, a sub-set of fission yeast introns undergo alternative splicing in the form of intron retention in response to certain environmental conditions (Awan et al. 2013; Bitton et al. 2015; Stepankiw et al. 2015). Thus, using the fission yeast system, we can better interrogate how a diversity of intron strengths and splicing decisions may be impacted by chromatin.

Building upon work from Patrick et al., in which a genome-wide genetic screen in *S. pombe* revealed extensive connections between splicing and chromatin factors, here we have identified a role for the Swr1 chromatin remodeling complex and the histone variant H2A.Z in splicing weak introns in *S. pombe* (Patrick et al. 2015). Specifically, introns with sub-optimal architecture and non-consensus *cis*-splicing sequences require H2A.Z for their efficient removal. The magnitude of this effect is further enhanced during temperature stress. Our data show that H2A.Z is enriched in intron-containing genes, located primarily within the +1 nucleosome, near the intron/exon boundaries of weak introns. Overexpression of Prp16, a splicing ATPase with proofreading activity, suppresses the temperature-dependent growth defect and non-consensus intron retention defect of strains lacking H2A.Z or the Swr1 ATPase. Additionally, mutating non-consensus splice sites of a model intron was sufficient to relieve the requirement for H2A.Z in promoting splicing. Taken together, these data suggest a crucial role for H2A.Z in “marking” weak introns, allowing for additional inspection by the spliceosome and prevention of spurious discard.

## Results:

### The Swr1 Complex has Strong Negative Genetic Interactions with the Spliceosome

An Epistatic Mini-Array Profile (EMAP), a genome-wide genetic screen for interactions of gene products, in *S. pombe*, previously uncovered a number of strong genetic interactions between splicing factors and the Swr1 nucleosome remodeling complex (Patrick et al. 2015). To further understand this observation, we performed a permutation analysis on the *S. pombe* EMAP to ask which Gene Ontology (GO) processes and complexes are enriched for genetic interactions with the Swr1 complex. Indeed, we found that “mRNA metabolic process” is amongst the top hits for this GO processes analysis (Figure 1A). To extend the resolution of the GO complex term “spliceosome,” we manually curated sub-complexes of the spliceosome and included them with the GO term complex analysis (Wahl et al. 2009; Cvitkovic and Jurica 2013)(Supplemental Table S1). This permutation analysis of GO complexes identified an enrichment in negative interactions between the Swr1 complex and several early splicing complexes, such as: splicing complex A, SF3A, splicing complex B, U1 snRNP and U2 snRNP. These sub-complex definitions are not mutually exclusive, as SF3A is associated with the U2 snRNP and both U1 snRNP and U2 snRNP are part of splicing complexes A and B. While we observed interactions across the splicing cycle, the strongest enrichment is with the early assembling spliceosome factors of the U1 snRNP and U2 snRNP, which promote the recognition of the 5' splice site (5'SS) and branch point (BP) (Supplemental Figure S1).

The majority of splicing factors in fission yeast are essential and are perturbed in the EMAP screen as “Decreased Abundance by mRNA Perturbation”(DAMP) alleles, where an antibiotic resistance cassette has been inserted into the gene between the stop codon and the 3' UTR to reduce the gene's expression (Schuldiner et al. 2005). However, subsequent reports show



that DAmP alleles of RNA binding proteins can be overexpressed (Lund et al. 2008; Patrick et al. 2015). Given the varying phenotypic strength reported for DAmP alleles, we were encouraged to see that the magnitudes of genetic interaction scores are consistent across double mutants between the Swr1 complex and the splicing factor DAmP strains across the spliceosome (Figure 1B, Supplemental Figure S2A). Additionally, not all splicing factor DAmP alleles yielded strong genetic interactions (Supplemental Figure S2B).

To more closely examine the interactions of the Swr1 complex with splicing factors at multiple steps of the splicing cycle, we manually generated double mutant strains through crosses and performed serial dilution growth assays (Figure 1C). Focusing on *pht1*, the gene encoding the variant histone H2A.Z and primary substrate of the Swr1 complex, and *swr1*, the gene encoding the ATPase and organizational platform of the Swr1 complex, we crossed deletions of each of these chromatin factors to alleles of splicing factors that have roles throughout the splicing cycle, including: *usp107<sup>DAmP</sup>*, a U1 snRNP factor; *sap14 $\Delta$* , a U2 snRNP factor; *dim1<sup>DAmP</sup>*, a tri-snRNP factor; *cwf15<sup>DAmP</sup>*, a Nineteen complex factor; *prp22<sup>DAmP</sup>*, a DEAH-box ATP required for mRNA release after the second catalytic step of the splicing reaction; and *prp43<sup>DAmP</sup>*, a DEAD-box ATPase required for spliceosome disassembly after proofreading-induced discard or the completion of the splicing reaction (Ohi et al. 2002; Carnahan et al. 2005; Newo et al. 2007; Koodathingal et al. 2010; Shao et al. 2012; Cvitkovic and Jurica 2013). As predicted by the negative genetic interaction scores in the EMAP, combining the *pht1 $\Delta$*  strain with these six splicing factor alleles results in varying levels of synthetic sickness in the double mutant strains at 30°C. The serial dilution profiles of *swr1 $\Delta$*  and splicing factor double mutant strains were similarly synthetic sick at 30°C (Supplemental Figure S2C).

## **The Swr1 Complex is required for pre-mRNA splicing of a sub-set of introns**

Given the strong genetic interactions between the spliceosome and the Swr1 complex, we wanted to investigate what molecular changes were responsible for the observed phenotypes. To test whether changes to splicing were contributing to the Swr1 complex/spliceosome double mutant growth defects, we performed genome-wide splicing-specific microarray analysis (Figure 2A). These microarrays contain oligonucleotide probes that are specific for exon, intron and spliced exon-exon junction regions for the vast majority of intron-containing genes in the *S. pombe* genome, allowing us to compare total mRNA, pre-mRNA, and mature mRNA between two strains (Lipp et al. 2015; Patrick et al. 2015). From these individual measurements, we generate a “Splicing Index” score, which reports the ratio of pre-mRNA to mature mRNA multiplied by the level of total mRNA to normalize changes in transcription. We compared total RNA from an isogenic wild-type strain to a variety of single and double mutant strains of the Swr1 complex and the spliceosome.

In order to understand if the synthetic sickness we observed correlated with a molecular splicing defect, we first analyzed the splicing efficiencies of *pht1*Δ and *usp107<sup>DAmP</sup>*, and the double mutant strain *pht1*Δ *usp107<sup>DAmP</sup>* (Figure 2B). We were particularly interested in a U1 snRNP factor, such as *usp107*, because of the enrichment of splicing complex A and U1 snRNP in our permutation analysis of complexes. *usp107<sup>DAmP</sup>* is overexpressed by RT-qPCR, possibly creating a dominant negative allele (Supplemental Figure S3A). The double mutant *pht1*Δ *usp107<sup>DAmP</sup>* strain had an exacerbated splicing defect beyond either single mutant strain with 567 introns retained, suggesting that fission yeast are especially reliant on Pht1 for splicing in the context of an already defective spliceosome. However, because the double mutant *pht1*Δ

*usp107<sup>DAmp</sup>* strain is sicker than the single mutants, we cannot rule out the possibility that poor growth indirectly accounts for the exacerbation of intron accumulation.

These observations led us to speculate that Pht1 and/or Swr1 may be particularly important for splicing when cells are grown under non-ideal conditions. Because many strains harboring mutations in splicing factors are sensitive to changes in temperature, likely due to defects in the kinetics of spliceosomal rearrangements, we repeated the above serial dilution growth assays at 16°C and 37°C (Rosbash et al. 1981; Noble and Guthrie 1996) (Figure 1C). Intriguingly, the *pht1Δ* strain has previously been described as cold-sensitive at 16°C, while the *swr1Δ* strain has been shown to be heat-sensitive at 37°C (Ahmed et al. 2007) (Supplemental Figure S3B). This observation indicates that H2A.Z may be important for recruiting or stabilizing gene expression complexes such as the spliceosome, during thermal stress. Indeed, at 16°C, the double mutants of *usp107<sup>DAmp</sup>*, *sap14Δ*, *dim1<sup>DAmp</sup>*, or *prp43<sup>DAmp</sup>* with *pht1Δ* are synthetic lethal, and *cwf15<sup>DAmp</sup>* and *prp22<sup>DAmp</sup>* are strongly synthetic sick. The serial dilution profiles of *swr1Δ* and splicing factor double mutant strains were similarly exacerbated compared to either single mutant, but under heat stress at 37°C (Supplemental Figure S2C). The observation that growing these double mutants under temperature stress exacerbates growth defects supports the hypothesis that the interplay between the Swr1 complex and spliceosome is particularly important for splicing when yeast are grown in non-ideal kinetic conditions.

To test this hypothesis, we grew the *pht1Δ* and *swr1Δ* single mutant strains under temperature stress of 16°C and 37°C, respectively, and looked for an intron retention defect beyond the level observed during growth at 30°C (Figure 2C). Indeed, we observed a significant increase in the number of affected introns when *pht1Δ* cells were grown at 16°C, with 883 introns affected (compared to 37 introns at 30°C). A similar trend was observed for *swr1Δ* at

37°C with 155 introns affected compared to 4 introns at 30°C (Supplemental Figure S3C). While the loss of the Swr1 ATPase has been shown to dramatically decrease the level of H2A.Z in chromatin, Swr1-independent incorporation and a residual fraction of deposited H2A.Z may account for this moderate intron retention profile in *swr1Δ* compared to *pht1Δ* (Papamichos-Chronakis et al. 2011; Sadeghi et al. 2011). We confirmed the intron retention defect in the *pht1Δ* strain at 16°C by measuring spliced and un-spliced RNA from several affected introns using RT-qPCR (Figure 2D).

To further examine the role of the Swr1 complex in pre-mRNA splicing, we analyzed a published RNA-seq data set of *swc5Δ*, a deletion strain of a component of the Swr1 complex responsible for the histone replacement activity along with *swr1* (Wu et al. 2005; Clement-Ziza et al. 2014). We compared *swc5Δ* datasets to wild-type datasets, and observed a statistically significant increase in intronic reads normalized to total reads for each dataset in the *swc5Δ* strain, indicating a global increase in retained introns, and further supporting a role for Swr1 complex in promoting pre-mRNA splicing in fission yeast (Supplemental Figure S3D).

To address the concern that changes in global transcription may be responsible for the intron accumulation defects observed, we utilized published *pht1Δ* and *swr1Δ* expression microarrays, and *swc5Δ* RNA-seq (all three performed at 30°C) to look for splicing factor genes with greater than two-fold expression level changes. We did not find consistent overlap of gene expression changes between these three datasets. Additionally, we selected eight splicing factor genes from these published datasets and performed RT-qPCR comparing WT to *pht1Δ* strains at 16°C (Supplemental Figure S4). We observed no changes in expression in these splicing factor genes greater than two-fold in the *pht1Δ* strain background. Taken together, these data strongly

suggest that changes in global transcription are not responsible for the intron accumulation defects that we observe (Kim et al. 2009; Clement-Ziza et al. 2014; Larson et al. 2016).

### **H2A.Z-affected introns are more likely to be 5' in the gene and contain non-consensus *cis*-splicing sequences**

Because we observed intron specificity in our splicing-specific microarray of *pht1* $\Delta$  at 16°C, we asked if the sub-set of *pht1* $\Delta$ -affected introns contained *cis*-splicing signals that made them more sensitive to loss of H2A.Z. To address this question, we analyzed different intron features using a logistic regression model to see if there was a correlation between *cis*-splicing signals and intron retention in the *pht1* $\Delta$  strain at 16°C (Lipp et al. 2015). Because H2A.Z is known to be enriched in promoter regions of genes, we first tested whether the distance of an intron from the TSS correlated with intron retention in the absence of *pht1* at 16°C (Subramanian et al. 2015) (Figure 3A). Interestingly, we observed that the likelihood of intron retention in *pht1* $\Delta$  at 16°C was inversely proportional to proximity to the TSS, meaning that introns located in the 3' end of a gene were less likely to accumulate in *pht1* $\Delta$  cells at 16°C. The increased sensitivity of introns at the 5' ends of genes to the ablation of *pht1* is consistent with the enrichment of H2A.Z in the promoter region of genes.

We next examined correlations between intron retention and other architectural features of introns, including intron length, distance between the BP and the 3' splice site (3'SS), and presence of a polypyrimidine tract. Previous work in *S. pombe* has shown that deviation from average length or absence of a polypyrimidine tract is associated with lower splicing efficiency (Sridharan et al. 2011). Indeed, we found that introns are more likely to accumulate in the absence of H2A.Z if they are longer, have a longer distance between the BP and 3'SS, and lack a

polypyrimidine tract (Figure 3A), suggesting that Pht1 is particularly important for splicing pre-mRNAs with sub-optimal intron architectures.

To determine whether affected introns were more or less likely to be enriched for particular *cis*-splicing sequences, we looked at the 5'SS, BP, and 3'SS of introns that accumulated in *pht1* $\Delta$  cells at 16°C. We based this analysis on the fact that although *cis*-splicing sequences in *S. pombe* are less constrained than those in *S. cerevisiae*, there is still an overall consensus sequence for each feature (Kuhn and Kaufner 2003; Stepankiw et al. 2015). Our analysis revealed that particular nucleotide deviations from the consensus *cis*-splicing sequences (5'SS, BP, 3'SS) are associated with significantly increased probability that an intron will be retained in *pht1* $\Delta$  at 16°C (Figure 3B). As these deviations occur at a variety of nucleotides across all three primary *cis*-splicing signals, this result is consistent with our observation that *pht1* $\Delta$  has negative genetic interactions with genes that contribute to recognition of each of these signals.

Given the enrichment of *pht1* $\Delta$ -affected introns containing non-consensus *cis*-splicing sequences, we wondered if these introns were more likely to use alternative splice sites. The Pleiss group has recently revealed numerous instances of alternative splice site usage in *S. pombe* (Stepankiw et al. 2015). We overlapped the list of *pht1* $\Delta$ -affected introns with a published list of introns utilizing alternative 5'SSs and/or 3'SSs (Supplemental Figure S5). Because the alternative introns were identified by lariat-sequencing, the list of introns using alternative 3'SS may also reflect alternative BPs. We found a significant enrichment of introns employing alternative splice sites in the sub-set of *pht1* $\Delta$ -affected introns, the majority of which utilized alternative 5'SSs rather than alternative 3'SSs (BPs), suggesting that Pht1 may play an important role in 5'SS choice.

Next, we asked whether the 5' bias of sub-optimal introns affected in *pht1Δ* was indicative of a general distribution pattern of introns with non-consensus *cis*-splicing sequences in the *S. pombe* genome. In order to address this question, we compared the 5' SS distance from the TSS with the 5' SS or BP sequence strength for each intron-containing gene in the *S. pombe* genome. Using published position-weighted matrix scores to classify *cis*-splicing signal strength, we analyzed the 5' SS and BP sequence strength for all introns within 4000bp of the transcription start site (99% of annotated introns) (Stepankiw et al. 2015). We then plotted 5' SS distances from the TSS for the sub-set of introns with the highest 10% and lowest 10% of 5' SS or BP scores (Figure 3C). Importantly, introns with 5' SS or BP sequences that fell into the bottom decile for sequence strength score were more likely to be found within the first 500bp from the TSS compared to introns with 5' SS or BP sequences in the top decile of sequence strength, implying that introns with weak *cis*-splicing sequences are more likely to be at the 5' end of genes genome-wide.

Taken together, these data indicate that H2A.Z is required for efficient splicing of a sub-set of introns that are “weak,” characterized by sub-optimal architectures, non-consensus *cis*-splicing sequences, and enrichment for alternative 5' SS usage. Furthermore, introns in the 5' ends of genes are more likely to have non-consensus *cis*-splicing sequences and be retained in the absence of H2A.Z, consistent with H2A.Z's known enrichment near the TSS (Subramanian et al. 2015).

### **H2A.Z is enriched near introns with non-consensus *cis*-splicing sequences**

In order to influence splicing, one would predict that H2A.Z would be enriched near *S. pombe* intron-exon boundaries. To test this prediction, we used ChIP-seq to probe H2A.Z

association genome-wide. We combined one replicate of our ChIP-seq with two biological replicates from a recently published dataset for analysis (Clement-Ziza et al. 2014). We confirmed that the replicates were correlated highly enough to be analyzed together by plotting the average coverage of each kilobase of the genome for each dataset and calculating the Pearson correlation pairwise between each replicate, showing a high correlation of all three replicates ( $R>0.8$ ) (Supplemental Figure S6).

We first aligned all of the TSSs or intron-exon boundaries in the genome and then plotted the ChIP-seq reads that fell around each feature over the background whole cell extract (WCE) coverage at that same position (Figure 4A, Supplemental Figure S7). As expected, we observed strong enrichment of H2A.Z around the TSS, with the strongest placement around the +1 nucleosome, followed by the -1 nucleosome, consistent with previous observations (Zofall et al. 2009).

Strikingly, we also observed a strong association of H2A.Z directly upstream of both the 5'SS and 3'SS. (Figure 4A, Supplemental Figure S7). Similar H2A.Z association patterns for the TSS, 5'SS, and 3'SS are also observed when looking only at genes containing introns with splicing defects in the absence of *pht1* at 16°C (Supplemental Figure S8). These data suggest that H2A.Z is located in physical proximity to intron-exon boundaries in order to directly influence splicing. Moreover, the median distance from the TSS of the first intron is 249bp in *S. pombe*, further correlating the location of many introns (n=2552 introns) with the +1 nucleosome (Supplemental Figure S9). However, we cannot separate the bias for H2A.Z enrichment and introns located near the TSS without further molecular experiments.



Because we observed an enrichment of H2A.Z near intron-exon boundaries, we next asked if H2A.Z deposition is biased towards intron-containing genes. We calculated H2A.Z coverage measured by ChIP-seq across every gene body and compared the average levels between genes with and without introns, normalized to WCE and the average coverage of both data sets. We found significantly higher levels of H2A.Z association on intron-containing genes, providing strong correlative evidence of a functional relationship between H2A.Z and the spliceosome (Figure 4B, Supplemental Figure S10A).

To strengthen our hypothesis that the presence of H2A.Z promotes the splicing of proximal introns, we next analyzed whether *pht1Δ*-affected introns are more likely to be associated with H2A.Z. We calculated H2A.Z coverage across a 300bp window around the TSS (150bp up and downstream), and then compared the average H2A.Z association levels between genes with and without *pht1Δ*-affected introns, normalized as above. As expected, there is a statistically significant enrichment of H2A.Z around the TSS for introns that have splicing defects in a *pht1Δ* background (Figure 4C, Supplemental Figure S10B). Importantly, these data suggest that the presence of H2A.Z near the TSS is predictive of a weak intron, and further support the claim that H2A.Z-containing nucleosomes promote splicing of nearby introns.

Given that H2A.Z is required for efficient splicing of sub-optimal introns, next we asked whether introns with weak *cis*-splicing sequences are more likely to coincide with H2A.Z-containing nucleosomes. We sampled introns from the top and bottom deciles of both 5' SS and BP strength scores and asked whether they had significantly different levels of normalized average H2A.Z coverage across the intron. We observed significantly higher levels of H2A.Z association on introns in the bottom decile of 5' SS strength scores compared to the top decile. However, there was not a significant difference in H2A.Z coverage between the top and bottom

deciles of BP strength scores. Weak BPs lacking significant H2A.Z coverage may be due to a smaller dataset of BP strength scores analyzed than of 5'SS strength scores. These data demonstrate that introns coincident with H2A.Z association are more likely to have weak 5'SSs (Figure 4D, Supplemental Figure S11).

Previous work in *S. pombe* showed a correlation between H2A.Z association and low gene expression (Zofall et al. 2009). To confirm this observation genome-wide, we compared the expression levels of H2A.Z-associated genes to the expression levels of genes without H2A.Z and found that H2A.Z-associated genes tend to be more lowly expressed ( $p < 0.001$ ) (Supplemental Figure S12). Intriguingly, it has been shown that introns in lowly-expressed genes in *S. pombe* are more likely to undergo aberrant splicing (Stepankiw et al. 2015).

Taken together, these analyses demonstrate that H2A.Z tends to be positioned in close proximity to weak introns in lowly-expressed, intron-containing genes, raising the possibility that it marks these genes to promote co-transcriptional intron recognition and splicing fidelity.

### **Overexpression of Prp16 suppresses H2A.Z-dependent splicing defects**

Given the spatial overlap of H2A.Z and weak introns, we next wondered if H2A.Z has any molecular connections to parts of the spliceosome involved in identifying and proofreading *cis*-splicing signals. Splicing ATPases known to participate in proofreading include *prp28*, *prp11* (Prp5 in *S. cerevisiae*), *prp16*, *prp22*, and *prp43* (Koodathingal and Staley 2013). In the *S. pombe* EMAP, *prp28*<sup>DAmP</sup> and *prp11*<sup>DAmP</sup> do not show any strong genetic interactions with the Swr1 complex, possibly due to insufficient perturbation of expression by the DAmP alleles. However, we observed synthetic sickness in double mutants of both *prp22*<sup>DAmP</sup> or *prp43*<sup>DAmP</sup> and *pht1Δ* by serial dilution growth assays at 30°C, which were exacerbated at 16°C, consistent with

the strong negative interactions across the Swr1 complex with these two splicing mutants in the EMAP (Figure 1C).

Because the synthetic sick interactions with *prp43*<sup>DAmP</sup> and the Swr1 complex were particularly strong, we further characterized the *prp43*<sup>DAmP</sup> allele by examining its expression level by RT-qPCR at 30°C (Supplemental Figure S13A). *prp43* is overexpressed in the *prp43*<sup>DAmP</sup> strain, by more than three times the level of wild-type. As Prp43 promotes the discard of sub-optimal substrates, the synthetic sickness exacerbated at 16°C in the double mutant *pht1Δ prp43*<sup>DAmP</sup> strain might be explained by increased spurious discard of weak introns (Mayas et al. 2010).

Because the *prp16*<sup>DAmP</sup> strain was not present in the *S. pombe* EMAP analysis, we generated double mutants with the *pht1Δ* and *swr1Δ* strains by hand and performed serial dilution growth assays across different temperatures to look for genetic interactions (Figure 5A). Prp16 is a splicing DEAH-box helicase required for transition from the first catalytic step to the second catalytic step of splicing and shown *in vitro* to possess proofreading activity, acting at the 5'SS and BP (Horowitz 2011). Strikingly, the *pht1Δ prp16*<sup>DAmP</sup> and *swr1Δ prp16*<sup>DAmP</sup> double mutant strains suppressed both the *pht1Δ* cold-sensitivity and the *swr1Δ* heat-sensitivity. RT-qPCR revealed that *prp16* transcript levels are at wild-type level in *pht1Δ* but more than three fold above the wild-type level in *prp16*<sup>DAmP</sup> and *pht1Δ prp16*<sup>DAmP</sup> strains 30°C and 16°C (Figure S13A, Supplemental Figure S13B).

Because overexpression of *prp16* suppresses both *pht1Δ* and *swr1Δ* temperature-dependent growth defects, we next tested if the double mutants also suppress the intron accumulation defects of the *pht1Δ* and *swr1Δ* at 16°C and 37°C, respectively. We used the splicing-specific microarray platform to assess splicing efficiency of *prp16*<sup>DAmP</sup> and *pht1Δ*

*prp16<sup>DAmP</sup>* at 16°C and *prp16<sup>DAmP</sup>* and *swr1Δ prp16<sup>DAmP</sup>* at 37°C (Figure 5B). The *prp16<sup>DAmP</sup>* strain by itself had a mild splicing defect at 16°C with 190 introns affected at 37°C with 83 introns affected, which is not as strong as either *pht1Δ* or *swr1Δ*, respectively. However, overexpression of *prp16* in strains lacking *pht1* or *swr1* led to a significant reduction in the total number of introns with splicing defects. In *pht1Δ prp16<sup>DAmP</sup>*, there was a partial reduction in the total number of introns affected, from 883 in *pht1Δ* to 500, consistent with the partial suppression of the cold-sensitive growth defect.

Importantly, the vast majority of the sub-set of 883 introns retained in *pht1Δ* at 16°C are rescued in the *prp16<sup>DAmP</sup>* strain (97%) and the *pht1Δ prp16<sup>DAmP</sup>* strain (85%) at 16°C. We further analyzed the sub-set of introns with significant splicing changes in strains overexpressing *prp16* at 16°C, and found 167 introns in the *prp16<sup>DAmP</sup>* strain and 370 introns in the *pht1Δ prp16<sup>DAmP</sup>* strain that are distinct from the introns that accumulate in the absence of *pht1*. This sub-set of *prp16<sup>DAmP</sup>*-affected introns is too small to perform intron feature analysis upon, with the method that we employed. These data suggest that excess Prp16 is not favorable for constitutive splicing of all introns. In the *swr1Δ prp16<sup>DAmP</sup>* strain, we detected no splicing events with significant changes from wild-type, indicating that overexpression of *prp16* restored splicing of all 155 introns that accumulated in *swr1Δ* at 37°C. However, because of improved growth in the double mutants, we cannot rule out the possibility of indirect suppression of intron accumulation.

Because the splicing of weak introns in *pht1Δ* at 16°C can be rescued by overexpression of proofreading ATPase Prp16, we propose that non-consensus introns require additional inspection by the spliceosome to be properly processed in sub-optimal environmental conditions. Furthermore, because overexpression of Prp43 exacerbates the *pht1Δ*-dependent splicing defect,

these data suggest that H2A.Z may have a role in preventing spurious discard of sub-optimal substrates by promoting the splicing reaction.

### **Restoration of consensus *cis*-splicing sequences suppresses H2A.Z-dependent splicing defect**

Next, we sought to directly test the requirement for H2A.Z in promoting splicing of “weak” introns. If H2A.Z indeed promotes splicing of introns with non-consensus *cis*-splicing sequences, then we would predict it to be dispensable for splicing when the intron’s *cis*-splicing sequences are strengthened to consensus.

To address this question, we chose a model intron from the single-intron-containing-gene SPBPJ4664.05, which is found at the 5’ end of the gene and accumulates in the absence of H2A.Z by microarray (Splicing Index = 1.187;  $p < 0.0006$ ). Importantly, SPBPJ4664.05 is enriched for H2A.Z deposition by ChIP-seq and is lowly expressed (Marguerat et al. 2012). The SPBPJ4664.05 intron 1 was tagged with an auxotrophic marker and its non-consensus 5’SS and BP sequences were mutated to the *S. pombe* consensus sequences (Figure 6A). These constructs were integrated into the genome, replacing the endogenous alleles in wild-type and *pht1Δ* strain backgrounds. Consistent with the microarray conditions, cells were grown at 16°C and RT-qPCR was performed to assess splicing efficiency of SPBPJ4664.05 intron 1.

Importantly, we found that strengthening the *cis*-splicing sequences of SPBPJ4664.05 intron 1 to consensus in the *pht1Δ* background is sufficient to restore splicing efficiency to the wild-type level (Figure 6B). Together with our genome-wide microarray data and global H2A.Z ChIP, these results strongly suggest that weak introns require H2A.Z for efficient removal by the spliceosome.

## **Discussion:**

Proper recognition and removal of introns requires regulation at a number of steps. How cells integrate multiple inputs to make co-transcriptional splicing decisions is still an area of active investigation. In this work, we demonstrate a requirement for the histone variant H2A.Z and the Swr1 nucleosome remodeling complex in the efficient splicing of weak introns. We show that H2A.Z is enriched at intron-exon boundaries near promoters of lowly-expressed genes. Additionally, introns located at the 5' end of *S. pombe* genes are more likely to contain non-consensus *cis*-splicing signals. Correlation between H2A.Z occupancy and non-consensus introns suggests that the histone variant itself may mark introns with weak splicing signals to aid in their recognition and removal. Consistent with this idea, we show that overexpression of a proofreading ATPase, Prp16, rescues growth and splicing defects in the absence of H2A.Z-containing nucleosomes. We propose a model whereby the presence H2A.Z proximal to introns alters the kinetics of co-transcriptional intron recognition and proofreading, thereby promoting the splicing of introns with weak splice sites.

## **H2A.Z promotes recognition and removal of weak, promoter-proximal introns**

The genome of *S. pombe* is rich with thousands of *cis*-splicing signals that are not constrained to a consensus sequence (Wood et al. 2002; Kuhn and Kaufer 2003). Such diversity in sequences implies a need for multiple layers of regulation for every splicing decision. The work presented here strongly implicates H2A.Z and the Swr1 complex in facilitating these decisions in fission yeast. Our results are consistent with published studies from the Pleiss group in *S. pombe* and the Stevens group in *S. cerevisiae*, in which multiple components of the Swr1

complex were identified in genome-wide screens for splicing efficiency changes of a model gene construct (Sorenson and Stevens 2014; Larson et al. 2016).

One way that the Swr1 complex may promote splicing is through the recruitment of splicing factors to the nascent pre-mRNA during spliceosome assembly and/or catalysis. Our work shows that H2A.Z, in addition to being enriched around the TSS, is also more likely to be found in nucleosomes flanking weak introns in lowly-expressed genes, spatially correlating H2A.Z deposition and splicing at genes with weak gene expression signals. In humans, SF3B was shown to weakly co-immunoprecipitate with H2A.Z, suggesting that there may be physical interactions underlying these genetic observations (Tolstorukov et al. 2012). In addition, mass spectrometry experiments of the Swr1 complex in *S. pombe* and *S. cerevisiae* show other early splicing factors are present at sub-complex levels, particularly splicing factors for which there are strong negative genetic interactions in both the *S. pombe* and *S. cerevisiae* EMAPs (Shevchenko et al. 2008; Wilmes et al. 2008; Kim et al. 2009). Further biochemistry is needed to examine the nature of these genetic interactions.

Alternatively, and not mutually exclusively, H2A.Z may help coordinate splicing and transcription by influencing the kinetics of RNAP II initiation and elongation through promoter-proximal introns. To further understand H2A.Z's role in the co-transcriptional recruitment of the spliceosome and/or dynamics of transcription, we performed ChIP-qPCR with splicing factors (U1 and U2AF) and RNAP II (holoenzyme and Ser5-P) at multiple genes of interest that we observed to have both H2A.Z deposition and intron retention in *pht1Δ* at 16°C (data not shown). However, because of the correlation between H2A.Z occupancy and low levels of gene expression, we could not consistently detect either spliceosome proteins or RNAP II at levels over background at genes with *pht1Δ*-affected introns. These results are consistent with previous

reports that RNAP II has a low ChIP signal in the *htz1*Δ (Adam et al. 2001).

Recent work from the Johnson group shows that in the absence of H2A.Z the Ser2-P population of RNAP II, representing elongating RNAP II, accumulates at the 5' end of the gene bodies (Neves et al. 2017). They also observe significant perturbations in snRNP association at introns with non-consensus BPs, suggesting a defect in spliceosome assembly, rearrangements or discard. Taken together with our data, these observations demonstrate that H2A.Z's influence on transcription is correlated with its role in splicing of non-consensus introns, possibly through changes in the kinetics of spliceosome rearrangements. Observations made regarding RNAP II elongation in *S. cerevisiae* are likely applicable to the fission yeast system. Previous analyses of our EMAPs showed that Swr1 Complex components as a group had highly-correlated, global genetic interaction profiles in both *S. pombe* and *S. cerevisiae*, strongly suggesting that this protein complex plays similar roles in the two yeast species.

### **H2A.Z likely influences splicing fidelity**

The splicing reaction proceeds in a complex, step-wise manner, driven forward by ATPases that regulate the rate of spliceosomal rearrangements, allowing proofreading of the substrates (Staley and Guthrie 1998; Koodathingal and Staley 2013). Given our observation that weak, non-consensus introns are sensitive to loss of H2A.Z, we were curious as to whether expression of proofreading ATPases—the bastions of splicing fidelity—could compensate for loss of H2A.Z at introns with weak and/or alternative *cis*-splicing signals. Indeed, we observed strong suppression of the *pht1*Δ-dependent splicing defect in cells overexpressing the splicing fidelity factor DEAH-box ATPase Prp16. This requirement for additional Prp16 is underscored in our experiments performed at 16°C, where the loss of *pht1* likely contributes to



slow/inefficient RNAPII elongation that is further exacerbated at low temperature. Under these conditions, overexpression of Prp16 may rescue splicing by increasing the association frequency of Prp16 with the spliceosome, thereby increasing the probability that Prp16 can sample and engage the *cis*-splicing signals of weak, promoter-proximal introns (Figure 7). By increasing the likelihood that Prp16 is available through an increase in its local concentration, both first and second step catalysis can be pushed forward, even in the absence of H2A.Z to mark weak introns and coordinate co-transcriptional intron recognition and removal. This role for Prp16 is consistent with a report from the Cheng group (Tseng et al. 2011), which showed that Prp16 has an ATP-independent role in promoting the first catalytic step of splicing, and with multiple reports from the Staley group that have characterized Prp16's roles in kinetically proofreading the second step (Koodathingal et al. 2010), and allowing for sampling of non-consensus splice sites by the spliceosome (Semlow et al. 2016)

Such a model predicts that in a wild-type cell, splicing of sub-optimal substrates requires additional attention from the spliceosome, which is afforded through coordination between H2A.Z and the spliceosome. By allowing the spliceosome to try and fail several times on sub-optimal substrates before the first catalytic step is completed, the cell can avoid discarding substrates with weak splice sites. Consistent with this idea, overexpression of Prp43 in the *pht1Δ* strain causes synthetic sickness, possibly through increased association of Prp43 leading to discard of already poorly-spliced, weak introns. Rescuing the growth and splicing defects of *pht1Δ* by overexpressing Prp16 suggests that H2A.Z has a role in promoting the splicing of weak introns by prompting additional inspection by the spliceosome, thus preventing spurious discard of non-consensus splice site-containing introns. This model is bolstered by our observation that

mutating an intron with a non-consensus 5' SS and BP sequences to consensus suppresses the requirement for H2A.Z in promoting efficient splicing.

Taken together with results from the Johnson group, our genetic and molecular data implicate the H2A.Z histone variant as a signal that marks weak introns to promote co-transcriptional splicing. By uncovering the complexities of how chromatin influences splicing fidelity and efficiency, we can begin to interrogate how uncoupling these coordinated processes may result in cellular dysfunction and disease.

## **Materials and Methods:**

### **Strains**

A list of strains used in this study is available in Supplemental Table S3.

### **E-MAP Analysis**

E-MAP data and their generation were as previously described (Patrick et al., 2015, Ryan et al., 2012). Evaluation of the links between the Swr1 complex and GO processes or complexes was performed as in Patrick et al. using the same dataset. Additional complexes related to splicing were defined by manual curation and are available in Supplemental Table S1.

### **RNA Isolation and qRT-PCR**

Strains were grown in YES5 overnight at 30°C and then diluted to 0.1 OD<sub>600</sub> the next morning to allow for growth to mid log-phase. Temperature shifts occurred when strains had reached 0.3-0.5 OD<sub>600</sub>. 37°C shifts took place for 2 hours and 16°C shifts for 9 hours. Cultures were harvested by centrifugation between 0.5–0.8 OD<sub>600</sub>. Pellets were washed once in water and then flash-frozen in liquid nitrogen and stored at -80°C. RNA was isolated using hot acid phenol followed by isopropanol precipitation as previously described (Bergkessel et al. 2011).

For RT-qPCR, RNA was treated with DNase I (Fermentas) and then MMLV1- RT (Promega) with dN9 primers (0.5mg/mL) to generate cDNA. Remaining RNA was then removed from cDNA by NaOH treatment and then the cDNA was purified using a spin kit (Zymo). qPCR was performed using a CFX96 Real Time system (BioRad).

### **Splicing-specific Microarray Data Collection and Feature Analysis**

Collection and analysis of splicing-specific microarray data was performed as previously described (Lipp et al. 2015; Patrick et al. 2015). Logistic regression analysis and subsequent intron feature analysis was performed as previously described (Lipp et al. 2015).

### **ChIP-seq Library Construction**

ChIP was performed as previously described (Patrick et al. 2015). After a chromatin immunoprecipitation reaction, the sequencing library preparation was performed as previously described (Dumesic et al. 2015), with the following modifications: The DNA yield of a single chromatin immunoprecipitation reaction—or 100ng WCE DNA— was used as starting material. Library concentration was determined by Qubit (Invitrogen).

### **Sequencing Data Processing, Read Alignment and Analysis**

Alignment of reads, generation of BAM files, and generation of bedgraph files were performed as previously described (Dumesic et al. 2015). Bedgraph files were used for analysis of signal around intron-exon junctions as described below. Bedgraph files were visualized using the Integrative Genomics Viewer 2.0.30 (Broad Institute). The *S. pombe* genome from Pombase.org was used as the reference.

Additional ChIP-seq data of wild-type and *pht1*-3xFLAG samples were obtained from ArrayExpress with identifier E-MTAB-2650 (Clement-Ziza et al. 2014). These data were treated as additional replicates and processed in the same manner as the library generated in this study. To confirm that replicates were similar enough to be used together in analysis, the average

coverage of each kilobase of the genome for each dataset was plotted. The Pearson correlation was calculated for each replicate compared to another replicate, using these average kilobase coverage values.

### **Meta-Intron-Exon Junction and Meta-TSS Graphs**

Each intron-exon junction, either 5'SS, 3'SS or both, for every intron in the genome, as well as the subset of introns with splicing defects in the *pht1Δ* microarray at 16°C, was aligned and the pile up of ChIP signal 500bp upstream and downstream of this feature was analyzed. The total coverage around this type of feature was then normalized on a per-base basis to that of the total coverage of the input sample at these same locations. To control for differences in sequencing depth, each sample was further normalized to read depth such that it equaled read depth in the input sample. Meta-TSS plots were generated in a similar manner by aligning the TSS for each gene in the genome and analyzing ChIP signal 500bp upstream and downstream of the feature, normalized to input and coverage of both datasets. The significance of ChIP-enrichment in these pile-up graphs was determined using the Kolmogorov-Smirnov (K-S) test compared to input sample signal distribution.

### **ChIP-seq Coverage Analysis**

For each gene in the genome, the number of reads in H2A.Z ChIP-seq datasets aligning to the gene body or the intron were counted. The same process was applied to the input sample dataset. Fold enrichment was calculated for each gene by dividing ChIP counts by input dataset counts, normalized to the total read counts for both ChIP and input samples:  $(\text{ChIP dataset reads for gene} / \text{Input dataset reads for gene}) * (\text{Input datasets total reads} / \text{ChIP dataset total reads})$ .

The same analysis was applied for each intron and TSS region in the genome, with TSS region defined as 150bp up- and downstream from each TSS. Mean enrichment of ChIP signal over input was compared between genes containing introns whose splicing was altered by the presence/absence of H2A.Z, between introns that are sensitive to the presence/absence of H2A.Z, or between the TSS region of genes that contain an intron sensitive to the presence/absence of H2A.Z. This analysis was repeated on introns that were found to be in the top and bottom deciles of 5'SS and BP scores in the *S. pombe*, as previously reported (Stepankiw et al. 2015).

### **Analysis of Intron Position and Strength in *S. pombe* Genome**

Intron locations were obtained from Pombase ([www.pombase.org](http://www.pombase.org)). *S. pombe* 5'SS and BP strength scores were previously reported (Stepankiw et al. 2015). These data consist of position weight matrix scores for the majority of annotated 5'SSs (5294) and BPs (3076). Some introns do not have scores for these *cis*-splicing sequences because the methods employed were not able to positively identify these features. The distance of each intron from the TSS of its gene was graphed on a theoretical gene body. Distributions across a theoretical gene body were also graphed for introns in the top and bottom decile of 5'SS and BP strength scores. These significance of intron distribution difference was determined by the K-S test.

## Supplemental Materials and Methods:

### Chromatin Immunoprecipitation (ChIP)

100mL cultures were grown to between 0.5–0.8 OD<sub>600</sub>, crosslinked with 1% formaldehyde (Sigma) for 15 minutes, and then quenched for 5 minutes with 125mM glycine (Sigma). Crosslinked cells were isolated by centrifugation and washed twice with 4°C TBS. Pellets were aliquoted into 50 OD<sub>600</sub> units of cells, flash frozen in liquid nitrogen and stored at -80°C. Aliquots were thawed on ice and resuspended in 1 mL ChIP lysis buffer (50mM HEPES-KOH pH 7.5, 140mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, and 1x EDTA-free Complete protease inhibitor (PI; Roche) Protease inhibitors (2mM PMSF, 20µg/mL leupeptin, 20µg/mL pepstatin, 20µg/mL aprotonin, 20µg/mL antipain). Lysate was generated using a Mini-Beadbeater (Biospec Products) for 8 cycles of 1.5 minutes beadbeating followed by 2 minutes on ice. DNA was isolated by centrifugation. DNA was sheared in a Bioruptor waterbath sonicator (Diagenode) with 6 cycles on high (10 min total, 30 sec ON, 1 min OFF). After sonication, cell debris was removed by centrifugation and the supernatant was brought up to 1mL in ChIP lysis buffer. 100µL of the lysate was then removed and saved as an input. 10ug of antibody against M2 FLAG (F3165, Sigma) was added to the remaining 900µL of lysate and this reaction was incubated for 2 hours at 4°C. Prior to addition, 30µL of packed Protein G DYNAbeads (Invitrogen) were washed for 5 minutes each at 4°C: twice with 1mL TBS and then twice in 1mL ChIP lysis buffer, and then resuspended in 100µL ChIP lysis buffer. This 100µL of washed beads was add to the reaction and incubated for 2 hours at 4°C for immunoprecipitation. Beads were then washed for 5 minutes each at 4°C: twice with 1mL ChIP lysis buffer, twice with 1mL high salt ChIP lysis buffer (ChIP lysis buffer with 500mM NaCl), and once with TE at

room temperature. Crosslinking was reversed in 1% SDS at 65°C overnight in the presence of Proteinase K (30µL of 20mg/mL). SDS was removed from DNA using NucleoSpin Gel Clean-up Columns with 1250µL NTB (Macherey-Nagel). The resulting DNA was analyzed by high throughput sequencing or qPCR.

### **Comparing ChIP-seq to PomBase Gene Expression Levels**

BAM files were used to call peaks from ChIP-seq data. Peaks were called using the Danpos algorithm, “dpeak”, run on three replicates of ChIP samples (two previously published datasets and one generated in this report) normalized to three replicates of input sample (where each input corresponds to each ChIP sample). Peaks were filtered for quality, requiring a value of more than 50000 in total signal for the peak, a value of more than 100 in height for the peak, and a p-value of less than  $10^{-100}$  corresponding to the comparison of peak height to background signal. These high quality peaks were used for downstream analysis. Gene annotations were obtained from Pombase ([www.pombase.org](http://www.pombase.org)).

We then compared the RNA expression level of each gene, as previously determined during vegetative growth (Marguerat et al. 2012), to the expression levels of genes that we found to have an H2A.Z ChIP-seq peak inside the gene body or within 150bp of the gene. The significance of this comparison was determined by using a two-tailed t-test.

### **RNA-seq Analysis**

RNA-seq data for three replicates of wild-type and *swc5Δ* samples were obtained from ArrayExpress with identifier E-MTAB-2640 (Clement-Ziza et al. 2014). The ABI SoLiD data were aligned using SHRIMP 2.2.3 program package’s aligner “gmapper” using default



parameters in color space. These datasets were poly-A selected and, to control for any artifacts caused by that technique, the amount of rRNA remaining in each sample was analyzed. While there is no significant difference, there is a trend towards more rRNA in wild type samples (data not shown), indicating that the effects discussed in the results section are not due to an artifact of poly-A selection.

Because read counts across individual introns were too low to perform statistical analysis for individual splicing events, the sum of reads falling in all introns in a sample were calculated. Intronic reads were defined as any sequencing read overlapping with an intron. The average of these values was calculated for 3 replicates of wild-type samples and 3 replicates of *swc5Δ* samples. 95% confidence intervals were calculated using bootstrapping. Bootstrap analysis was used to test the difference between the average intron reads in wild type datasets and *swc5Δ* datasets.  $p < 0.05$  was used as the significance threshold.

### **Comparing Microarray to Alternative Introns**

Published data of *S. pombe* alternative introns were obtained (Stepankiw et al. 2015). These data were identified by lariat branches and included chromosome number, 5'SS, BP and 3'SS position, as well as whether the 5'SS or 3'SS used was upstream or downstream of the annotated position. The authors describe their method of defining alternative 3'SSs as selecting the first AG dinucleotide at least 5 nucleotides downstream of branch read. We culled this list so that it only contained introns measured on the splicing-specific microarrays that we used in the main text. We then partitioned the list into introns with alternative 5'SS usage, alternative 3'SS usage, and introns with alternative 5'SS and/or 3'SS usage. We then compared these three sub-populations of alternative introns to the list of introns with significant splicing index changes in

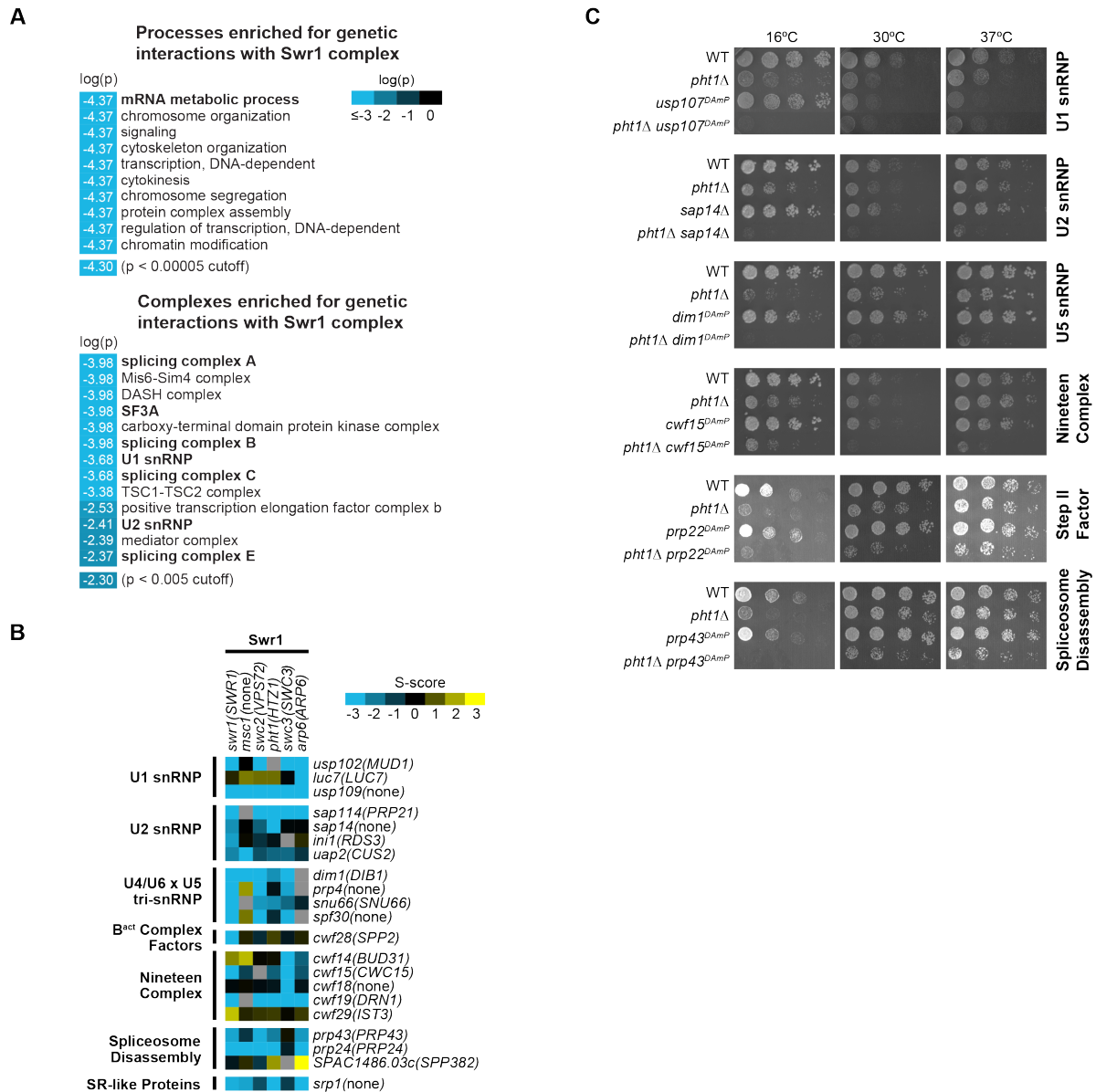
*pht1*Δ at 16°C as determined by the splicing-specific microarray. We used contingency tables followed by a chi-squared test to determine the significance of the overlap between these data sets.

**Acknowledgments:**

We wish to thank J. Pleiss for sharing his design of the *S. pombe* splicing-specific microarrays and data on intron *cis*-splicing sequence strength, and H. Madhani for collaborating on the ChIP-seq analysis. We are grateful to S. Grewal for sharing *S. pombe* strains and to J. Lipp, C. Siegel, and R. Bramwell for assistance with computational analysis. Thank you to P. Dumesic and members of the Guthrie and Madhani labs for their critical reading of this manuscript and thoughtful discussions.

**Funding:**

KEN was supported by a Graduate Research Fellowship from the National Science Foundation. CMH was funded by the National Institutes of Health grant number F30HL120496-02. CJR is a Sir Henry Wellcome Fellow jointly funded by Science Foundation Ireland, the Health Research Board and the Wellcome Trust (Grant Number 103049/Z/13/Z) under the SFI-HRB-Wellcome Trust Biomedical Research Partnership. NJK was funded by the National Institutes of Health grant numbers GM84279, GM081879, GM098101, GM084448 ([www.nih.gov](http://www.nih.gov)). KLP was funded by the American Cancer Society, grant number 119093-PF-10-182-01-TBE ([www.cancer.org](http://www.cancer.org)). CG is an American Cancer Society Research Professor of Molecular Genetics and was funded by the National Institutes of Health grant number GM21119 ([www.nih.gov](http://www.nih.gov)). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.



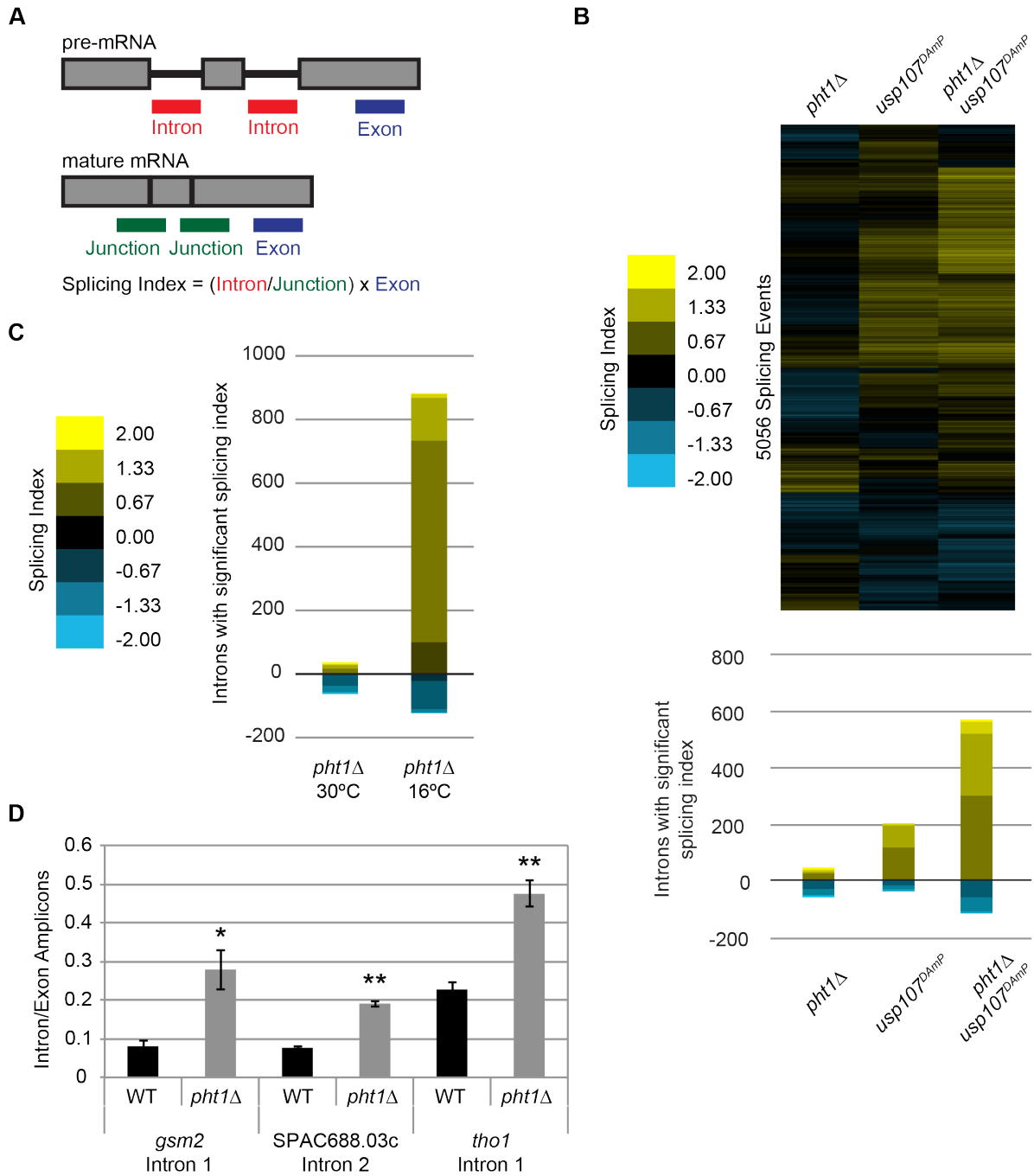
**Figure 1: The Swr1 Complex has strong genetic interactions across the spliceosome.**

(A) Processes and complexes enriched for significant negative genetic interactions with the Swr1 Complex in the *S. pombe* E-MAP. Processes defined by *S. pombe* GO Slim database (PomBase). Complexes are from the *S. pombe* cellular component GO list (PomBase), with splicing sub-complexes manually defined as in Table S1. Complexes involved in chromatin remodeling are

not shown. Significance defined as Bonferonni corrected p value  $<0.00005$  for processes and  $<0.005$  for complexes.

(B) Heat map of genetic interaction S-scores between Swr1 complex factors and splicing factors. Splicing factors shown have at least one S-score either  $> 2.0$  or  $< -2.5$ . Grey squares indicate no data.

(C) Serial dilutions (1:5) of WT, the *pht1* $\Delta$  mutant, splicing factors representing different splicing sub-complexes, and their double mutants with *pht1* $\Delta$ , grown at 16°C, 30°C, and 37°C.



**Figure 2: H2A.Z is required for pre-mRNA splicing.**

(A) Cartoon depicting splicing-specific microarray probes.

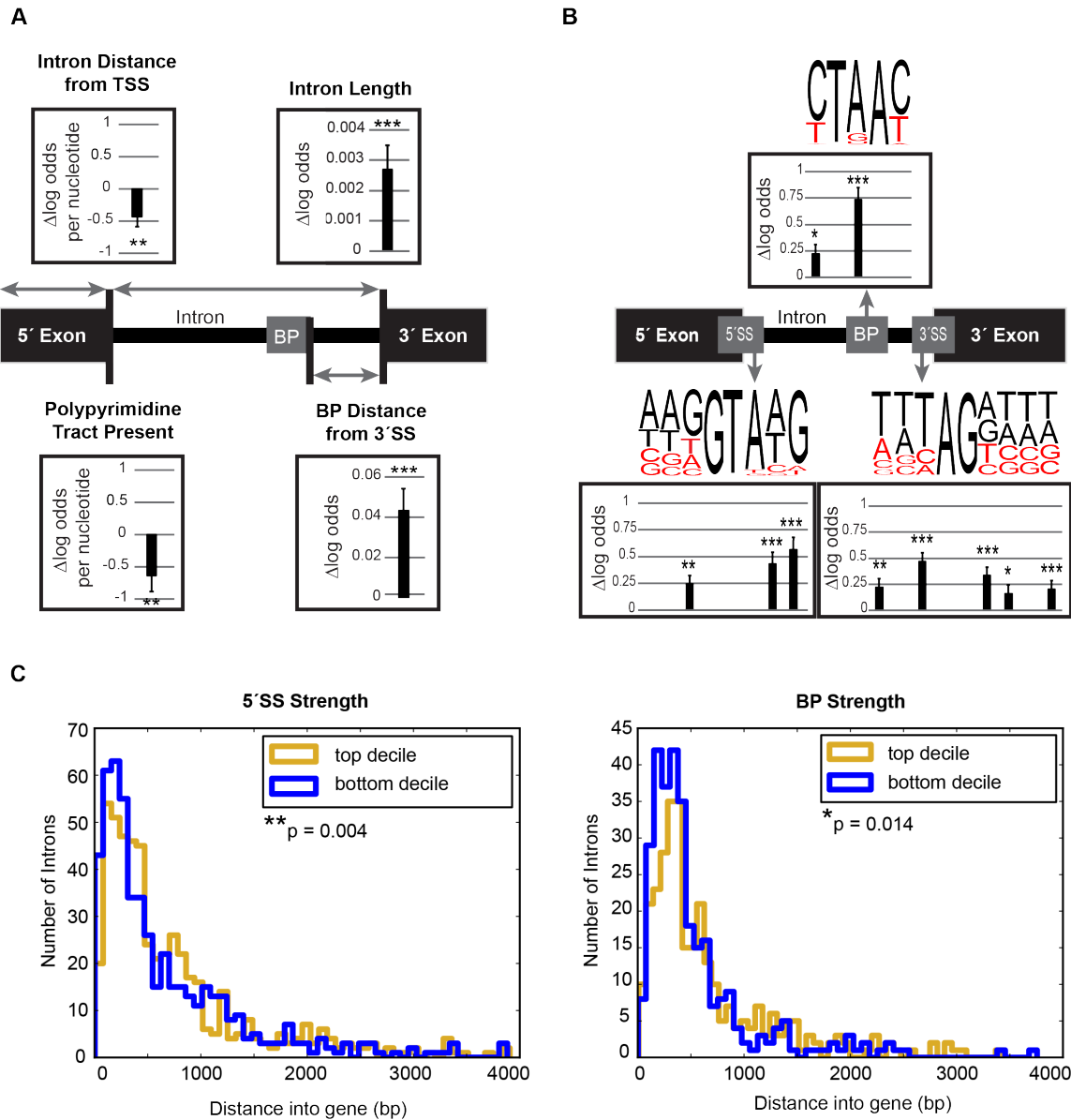
(B) Splicing profile of *pht1*Δ and *usp107*<sup>DAmP</sup> single mutants and *pht1*Δ *usp107*<sup>DAmP</sup> double mutant strains. EMAP S-score for *pht1*Δ *usp107*<sup>DAmP</sup> is 0.455335 (neutral). cDNA from each

single and double mutant strain was competitively hybridized on the splicing-specific microarray against that from WT. A Splicing Index value was calculated for each intron by normalizing the  $\log_2$ -ratio of the intron change to junction change, multiplied by the exon change. The heat map shows the Splicing Index score for most introns in *S. pombe* of the indicated strain compared to WT. Gene order along the y-axis is the same for all arrays. Stacked bar plots below show the number of introns with a Splicing Index score greater than 0.3. Color scale shows distribution of the severity of splicing defect.

(C) Stacked bar plots of Splicing Index scores for *pht1* $\Delta$  strain. Method is the same as Figure 2B except that cultures were grown to mid-log phase at 30°C and shifted to 16°C for 9 hours where indicated. 30°C data is the same as Figure 2B.

(D) RT-qPCR validation of *pht1* $\Delta$  at 16°C microarray for 3 introns comparing WT to *pht1* $\Delta$  at 16°C. Cultures were grown to mid-log phase at 30°C and shifted to 16°C for 9 hours. Bars represent ratios of intronic to exonic RNA signals. Error bars represent S.E.M. of 3 biological replicates. \* $p < 0.05$ ; \*\* $p < 0.01$  calculated by one-tailed t-test





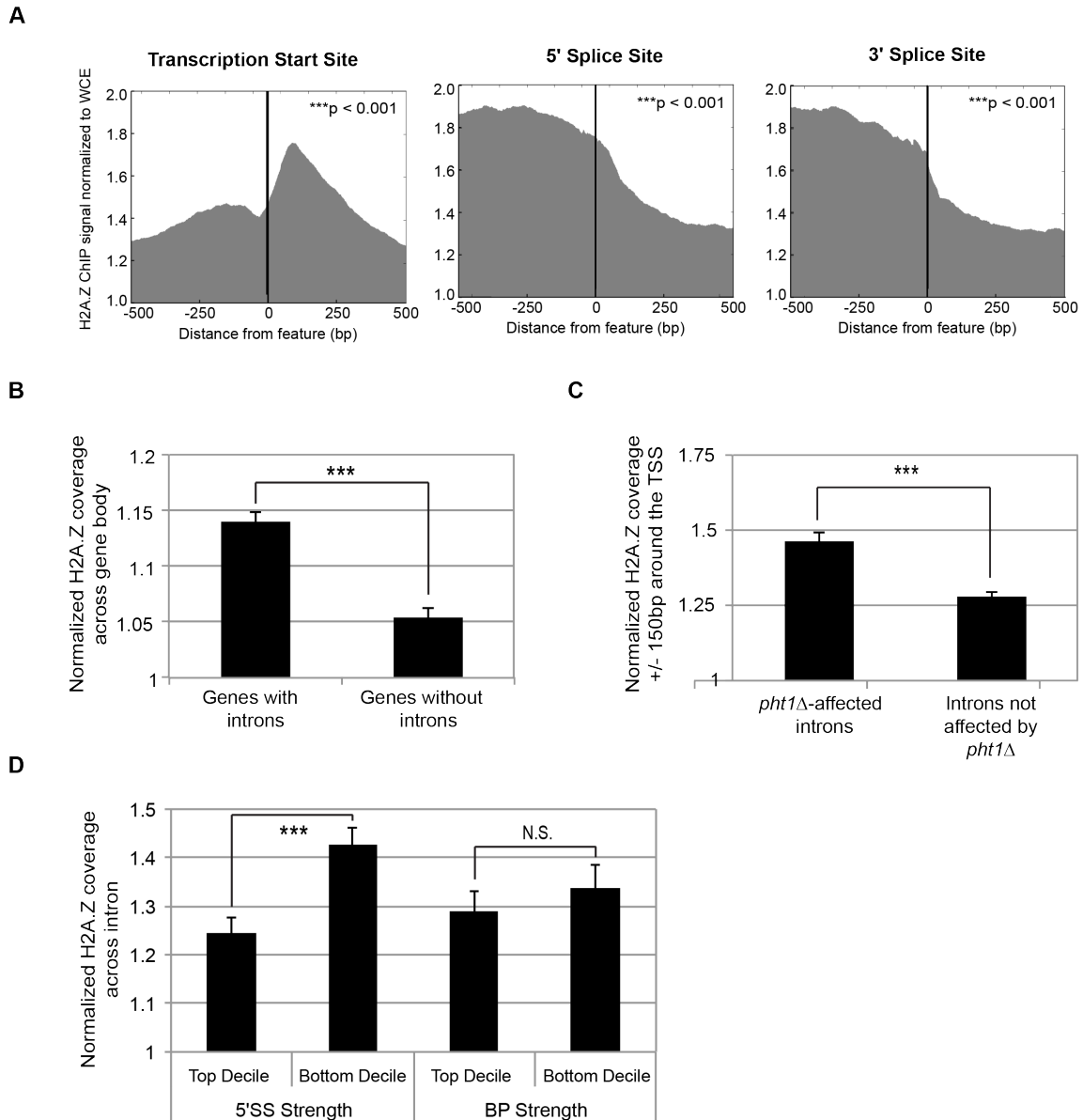
**Figure 3: Loss of H2A.Z impairs splicing of weak introns at the 5' ends of genes.**

(A) Schematic representation of an *S. pombe* intron with arrows indicating features of intron architecture. Bar graph shows the probability of indicated feature increasing the likelihood of intron retention represented as log odds or log odds per nucleotide (positive log odds at a given feature indicate that increased length or frequency correlate with intron retention.) Affected

introns were defined as having a corrected p-value of  $< 0.05$ . Error bars represent S.E.M. of a logistic regression model (Lipp et al. 2015). \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

(B) Schematic representation of the *S. pombe* 5' SS, BP and 3' SS, with consensus nucleotides indicated in black and non-consensus nucleotides in red. Probability of a specific nucleotide change increasing the likelihood of retention is represented as log odds (positive log odds at a given position indicate that non-consensus nucleotides correlate with intron retention). Statistics are the same as above.

(C) Histogram of 5' SS and BP strength in *S. pombe* introns, plotted by their distance in base pairs from the TSS. Introns in the bottom decile of 5' SS and BP strength are enriched in the 5' ends of genes. p-values are from Kolmogorov-Smirnov (K-S) test.



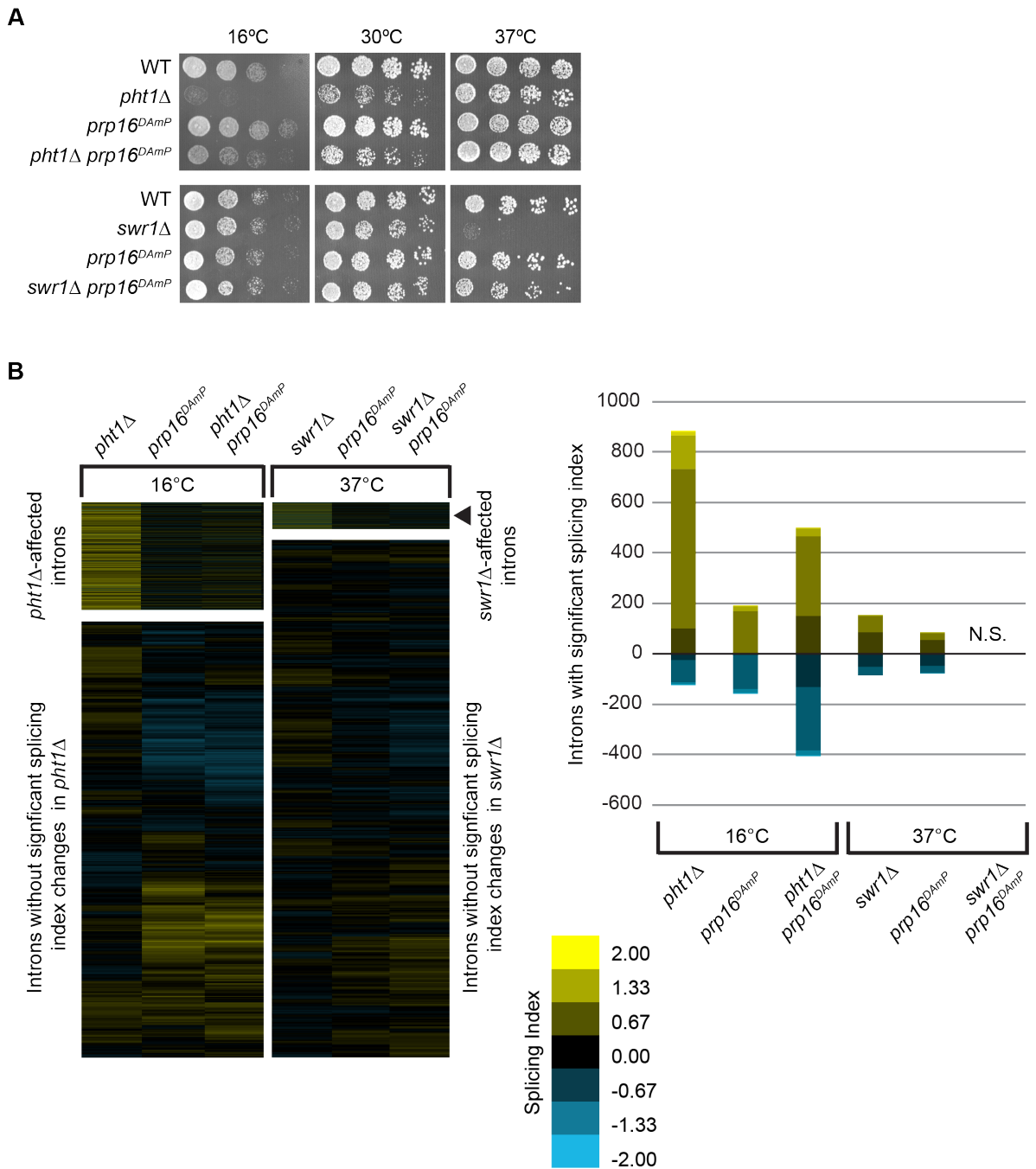
**Figure 4: H2A.Z is enriched at splice sites.**

(A) Pile up of H2A.Z ChIP-seq reads around all TSSs, 5'SSs, and 3'SSs, normalized to input sample. Graphs represent a single replicate from this study. p-values are from K-S test for data compared to a uniform distribution.

(B) Bar plots representing the average level of H2A.Z coverage measured by ChIP-seq across genes with or without introns, normalized to whole cell extract (WCE), and the coverage of both data sets. Error bars represent S.E.M. \*\*\* $p < 0.001$  calculated by t-test.

(C) Bar plots representing the average level of H2A.Z coverage measured by ChIP-seq within +/- 150bp around the TSS for genes with or without introns with splicing defects in the *pht1Δ* microarray at 16°C, normalized to WCE and the coverage of both data sets. Error bars represent S.E.M. \*\*\* $p < 0.001$  calculated by t-test.

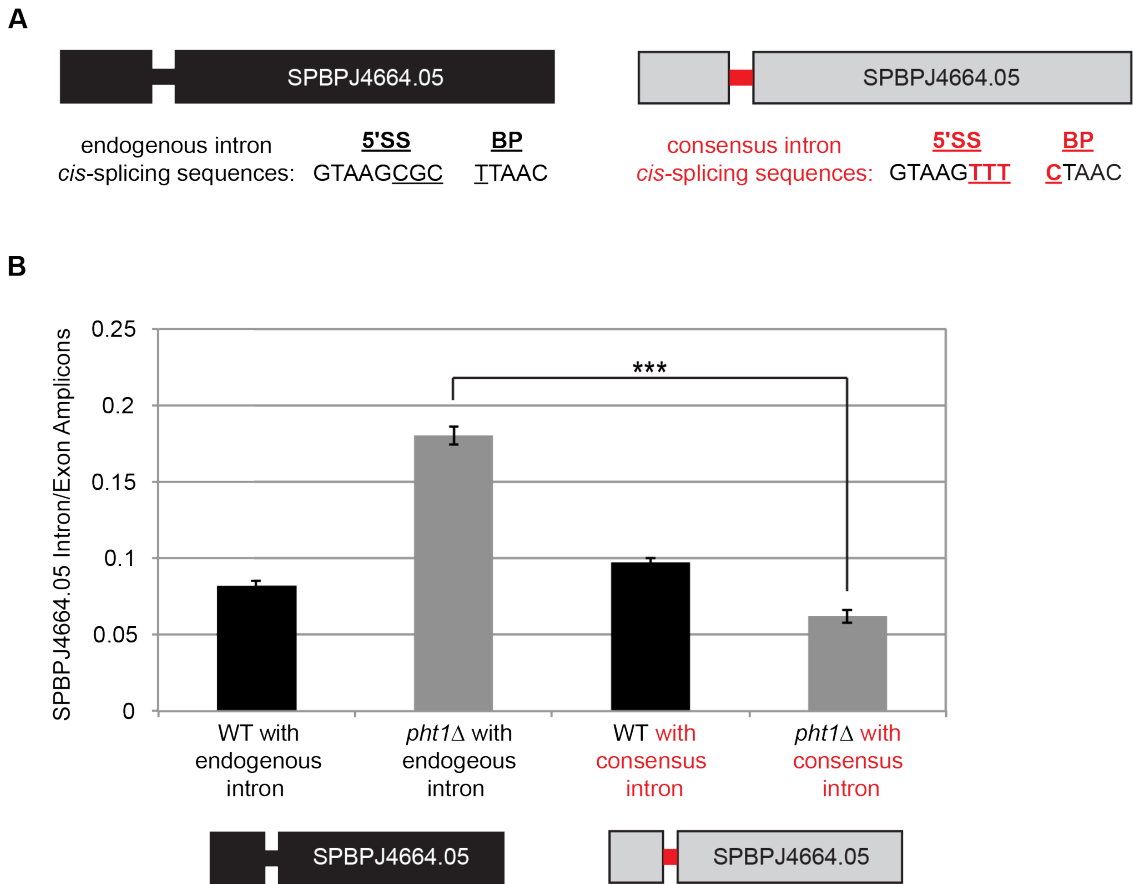
(D) Bar plots representing the average level of H2A.Z coverage measured by ChIP-seq across introns in the top and bottom deciles of 5'SS and BP strength, normalized to WCE and the coverage of both data sets. Error bars represent S.E.M. \*\*\* $p < 0.001$  calculated by t-test. N.S. = not significant.



**Figure 5: Overexpression of *prp16* ATPase suppresses growth and splicing defects in *pht1Δ* and *swr1Δ* at restrictive temperatures.**

(A) Serial dilutions (1:5) of WT, single and double mutant strains grown at 16°C, 30°C, and 37°C.

(B) Heatmaps of the splicing profiles of the *pht1Δ* or *swr1Δ*, *prp16<sup>DAmP</sup>*, and their respective double mutant strains. Cultures were grown to mid-log phase at 30°C and shifted to 16°C for 9 hours or 37°C for 2 hours where indicated; cDNA from each single and double mutant strain was competitively hybridized on the splicing-specific microarray against that from WT. Stacked bar plot shows the number of introns with a Splicing Index score greater than 0.3. Color scale shows distribution of the severity of splicing defect. N.S. = not significant

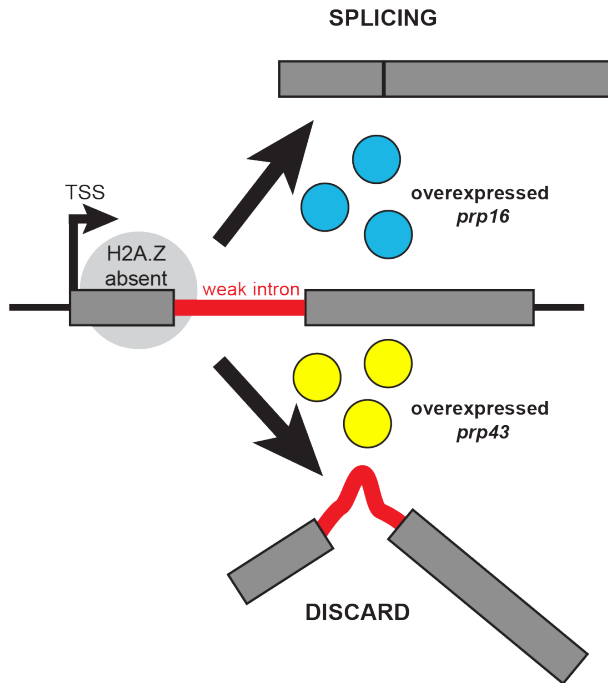


**Figure 6: Mutating *cis*-splicing sequences to consensus suppresses the splicing defect in *pht1*Δ at 16°C.**

(A) Cartoon depicting architecture and 5'SS and BP sequences of intron 1 of model gene SPBPJ4664.05. Mutation of *cis*-splicing sequence nucleotides to the consensus sequence is indicated in red.

(B) Bar graphs representing splicing efficiency of SPBPJ4664.05 intron 1 in WT and *pht1*Δ strains, with and without mutations to 5'SS and BP as indicated above. Cultures were grown to mid-log phase at 30°C and shifted to 16°C for 9 hours. Bars represent ratio of intronic to exonic signal measured by RT-qPCR. Error bars represent S.E.M. of 3 biological replicates. \*\*\* $p < 0.001$  calculated by one-tailed t-test.

A

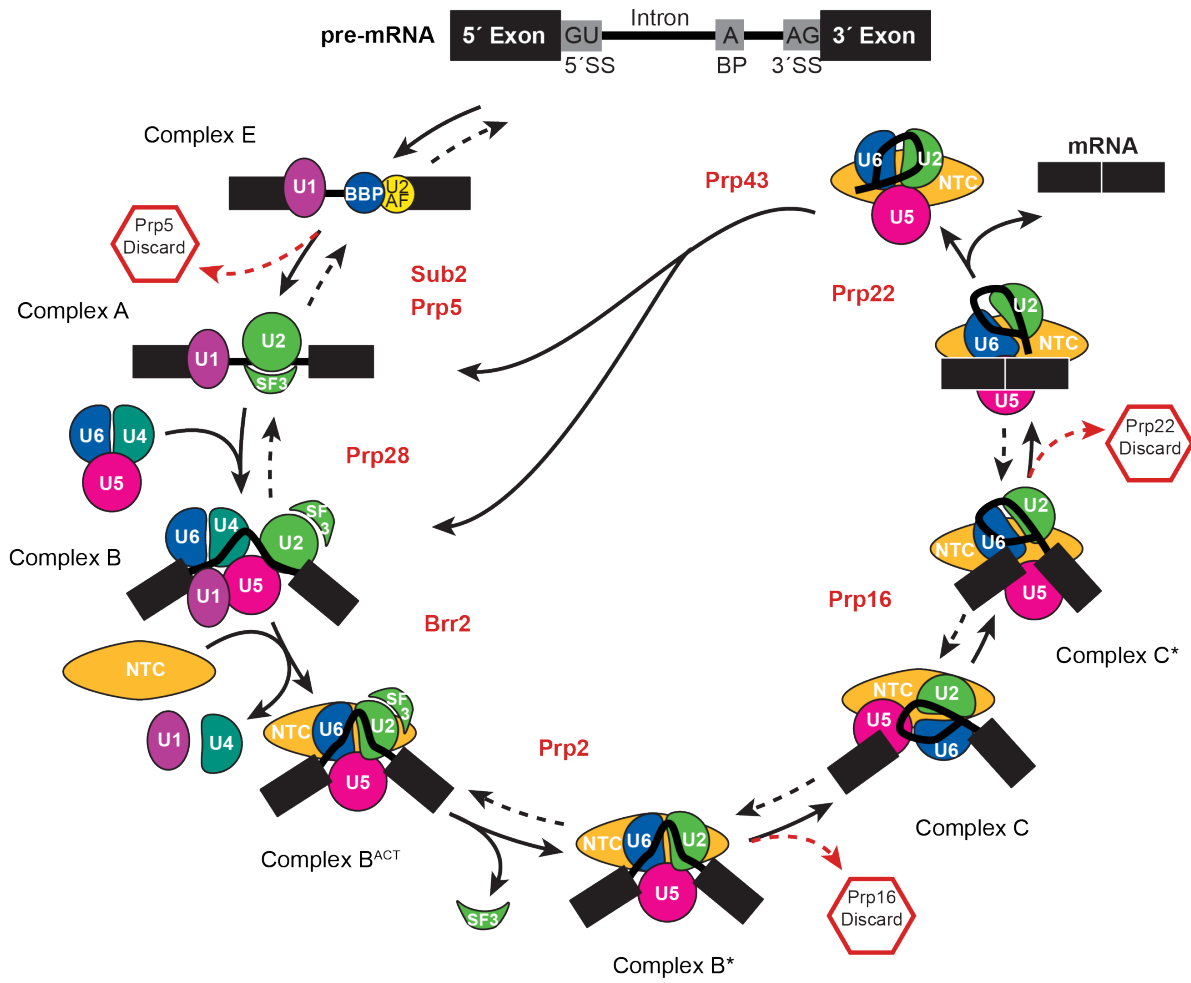


**Figure 7: Model illustrating the effect of H2A.Z loss on splicing efficiency and suppression by overexpression of *prp16* and exacerbation by overexpression of *prp43*.**

(A) Cartoon depicting a gene with a weak intron near the +1 nucleosome. When H2A.Z is absent, splicing of this weak intron does not proceed efficiently. Overexpression of Prp16 ATPase recues splicing efficiency, possibly through altering the kinetics of splicing catalysis. Through mass action, overexpressed Prp16 is more likely to be bound to the spliceosome, preventing discard by Prp43, promoting weak splice site sampling and/or driving splicing catalysis forward. Overexpression of Prp43 on the other hand preferentially acts to discard the spliceosomes on the weak substrate, which are possibly stalled due to the absence of H2A.Z.



A



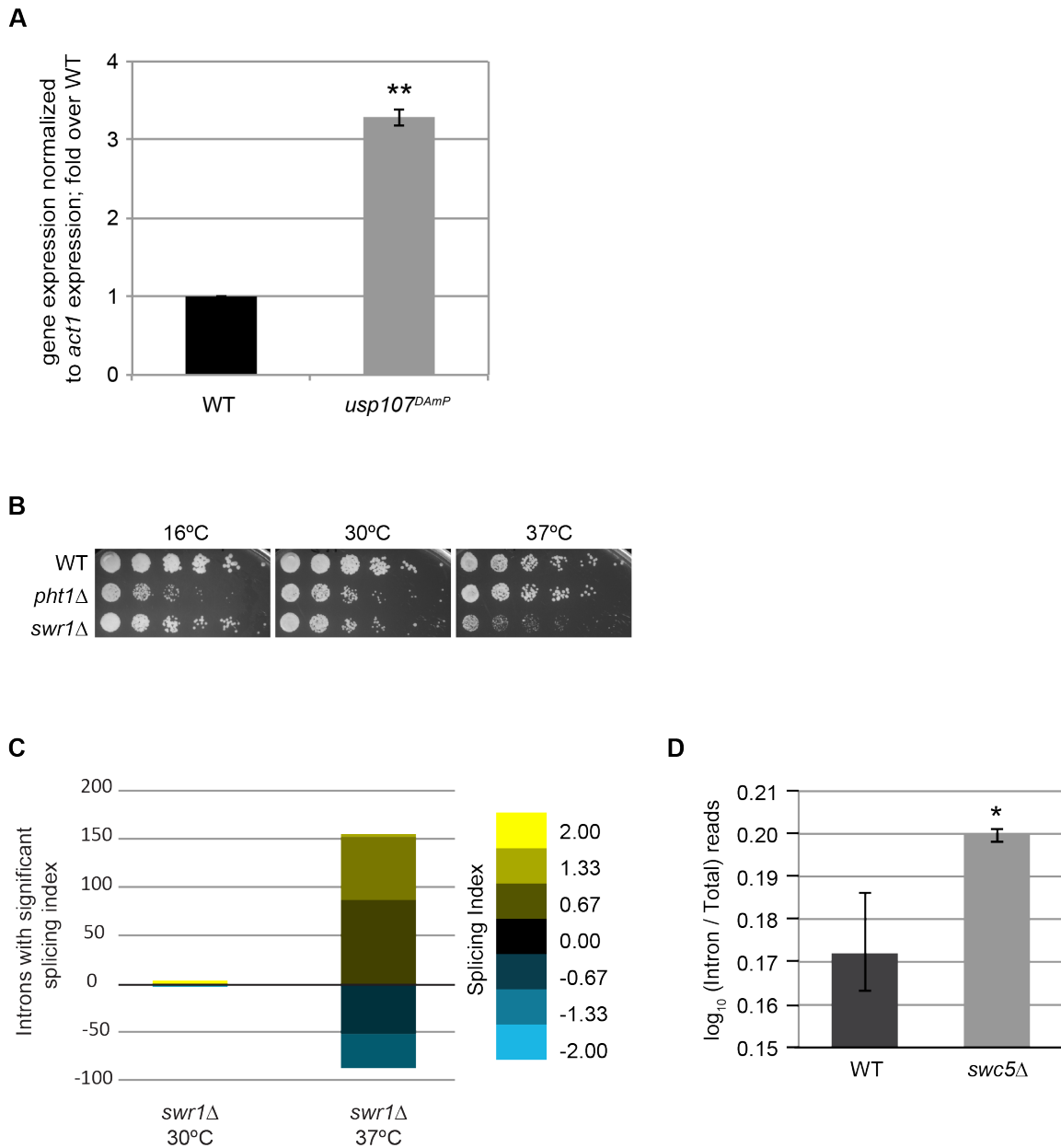
**Supplemental Figure 1: Splicing cycle diagram**

(A) Schematic drawing of the splicing cycle highlighting proofreading steps.



(B) Serial dilutions (1:5) of WT, *pht1* $\Delta$  and *swr1* $\Delta$  mutants, splicing factors, and their respective double mutants showing a neutral genetic interaction, grown at 16°C, 30°C, and 37°C.

(C) Serial dilutions (1:5) of WT, the *swr1* $\Delta$  mutant, splicing factors representing different splicing sub-complexes, and their double mutants with *swr1* $\Delta$ , grown at 16°C, 30°C, and 37°C.



**Supplemental Figure 3: Additional panels supporting Figure 2**

(A) RT-qPCR of *usp107<sup>DAmp</sup>* mRNA levels at 30°C. Bars represent fold change of mutant strain signal over WT, normalized to *act1* mRNA levels. Error bars represent S.E.M. of 3 biological replicates. \*\* $p < 0.01$  calculated by one-tailed t-test.

(B) Serial dilutions (1:5) of WT, *pht1Δ* and *swr1Δ* strains grown at 16°C, 30°C, and 37°C.

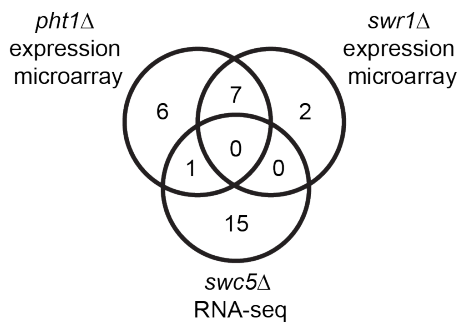
(C) Stacked bar plot of Splicing Index scores for *swr1Δ* mutant strain. Cultures were grown to mid-log phase at 30°C and shifted to 37°C for 2 hours where indicated. Stacked bar plot shows the number of introns with a Splicing Index score greater than 0.3. Color scale shows distribution of the severity of splicing defect.

(D) Analysis of RNA-seq data of Swr1 Complex factor *swc5Δ* shows a splicing defect (Clement-Ziza et al. 2014). Average intron read count compared to total reads for each dataset are plotted. Error bars represent 95% confidence interval.

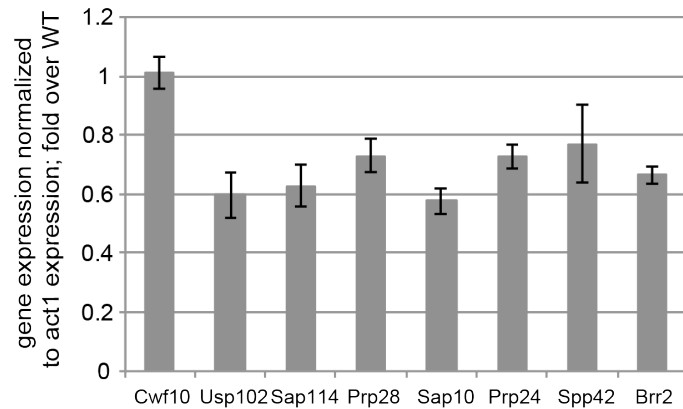
A

<i>pht1Δ</i> expression microarray		<i>swr1Δ</i> expression microarray		<i>swc5Δ</i> RNA-seq	
Gene	LogFC average	Gene	LogFC average	Gene	average RPKM fold change over WT
<i>brr2</i>	0.21	<i>brr2</i>	0.34	<i>cwf10</i>	0.36
<i>spp42</i>	0.30	<i>spp42</i>	0.38	<i>uap2</i>	0.38
<i>prp16</i>	0.36	<i>prp24</i>	0.42	<i>cwf18</i>	0.54
<i>usp109</i>	0.37	<i>sap10</i>	0.44	<i>prp3</i>	0.57
<i>prp24</i>	0.38	<i>prp28</i>	0.45	<i>prp17</i>	0.63
<i>cwf22</i>	0.41	<i>prp38</i>	2.06	<i>snu66</i>	0.64
<i>cwf10</i>	0.46	<i>usp102</i>	2.16	<i>cwf12</i>	0.69
<i>sap10</i>	0.46	<i>sap14</i>	2.22	<i>cwf28</i>	1.75
<i>cwf14</i>	0.46	<i>snu13</i>	2.85	<i>prp31</i>	2.14
<i>prp28</i>	0.47			<i>lin1</i>	2.16
<i>sap114</i>	0.48			<i>prp1</i>	2.26
<i>cdc5</i>	0.48			<i>ubp10</i>	2.42
<i>usp102</i>	2.03			<i>rna4</i>	3.34
<i>snu13</i>	2.07			<i>cwf4</i>	3.35
				<i>cwf29</i>	4.53
				<i>uap56</i>	7.87

B



C



### Supplemental Figure 4: Splicing factor gene expression in Swr1 Complex mutants

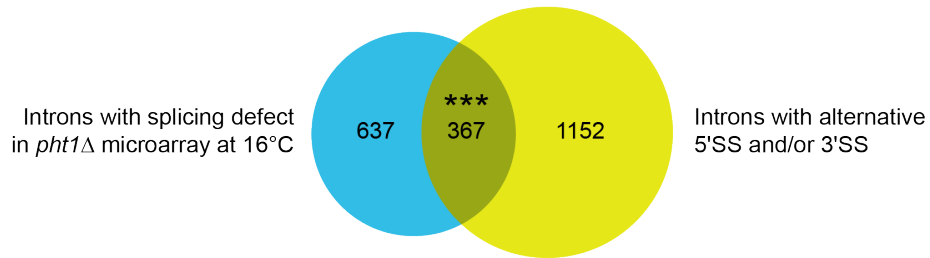
(A) Table of splicing factor genes with greater than 2-fold expression changes in a published *pht1Δ* expression microarray, *swr1Δ* expression microarray (Kim et al. 2009) and *swc5Δ* RNA-seq (Clement-Ziza et al. 2014). All published data were collected at 30°C.

(B) Venn diagram of genes listed in (A).

(C) RT-qPCR of splicing factor gene expression in *pht1Δ* at 16°C. Bars represent fold change of *pht1Δ* signal over WT for splicing factor gene mRNA level normalized to *act1* mRNA level.

Error bars represent root sum of squares of 4 biological replicates.

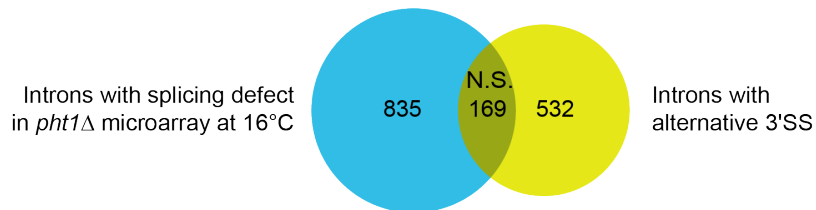
**A**



**B**



**C**

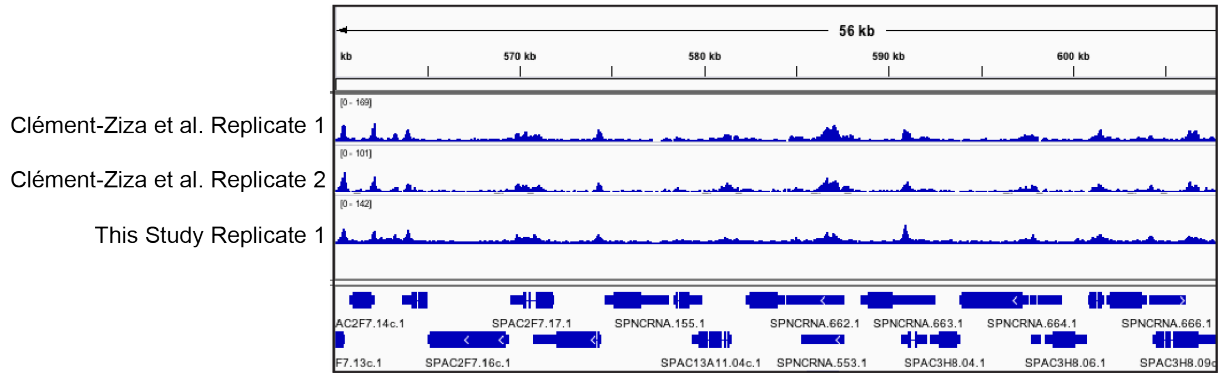


**Supplemental Figure 5: Comparison between introns with splicing in the *pht1Δ* microarray at 16°C and introns found to employ alternative splice sites in *S. pombe***

Venn diagram of introns with splicing defects in the *pht1Δ* microarray at 16°C and a published list of introns using alternative splice sites in *S. pombe* determined via lariat sequencing (Stepankiw et al. 2015). (A) Intron with alternative 5'SSs and/or 3'SSs, (B) introns with alternative 5'SSs only, and (C) introns with alternative 3'SSs only. \*\*\* $p < 0.001$  calculated by chi-squared test of a 2x2 contingency table.



A



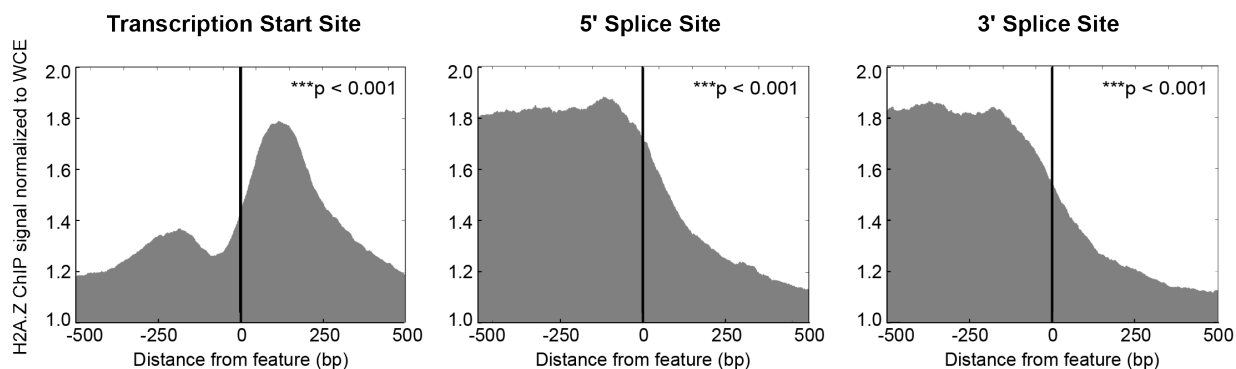
Comparison	R Value (RPKM)
Clément-Ziza Replicate 1 vs. Clément-Ziza Replicate 2	0.939
Clément-Ziza Replicate 1 vs. This Study Replicate 1	0.886
Clément-Ziza Replicate 2 vs. This Study Replicate 1	0.847

**Supplemental Figure 6: Comparison of H2A.Z ChIP-seq replicates**

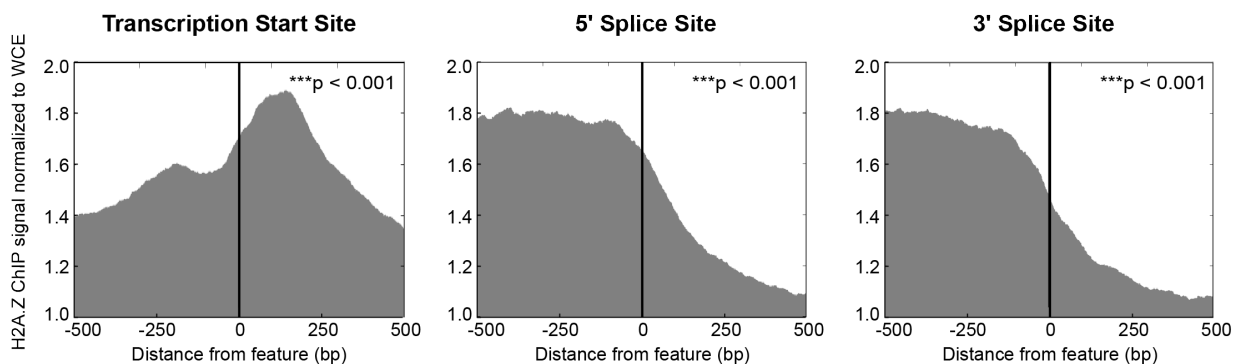
(A) Representative ChIP-seq pht1-3XFLAG binding profile of a sample collected in this study and two published replicates used for analysis throughout this study (Clément-Ziza et al. 2014). R values in the table below compare the average read coverage in each dataset for each kilobase of the genome.

**A**

**Clément-Ziza Replicate 1**



**Clément-Ziza Replicate 2**

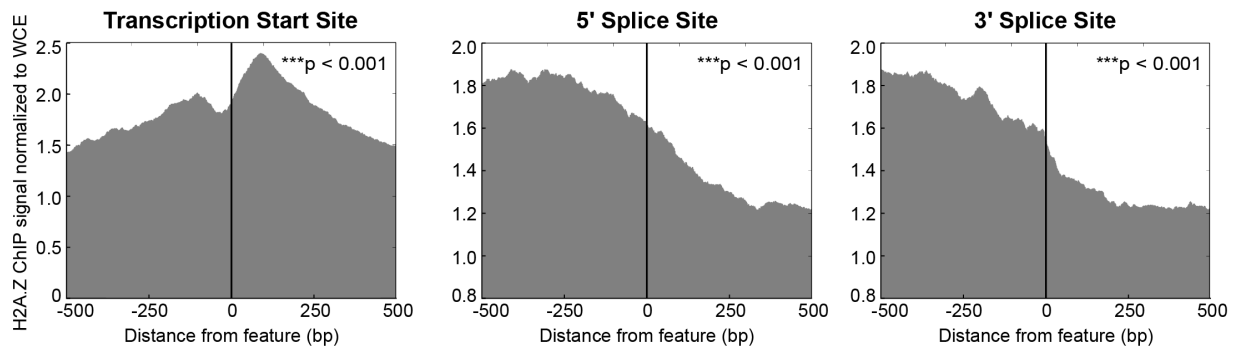


**Supplemental Figure 7: Additional replicates of pile up graphs of H2A.Z ChIP-seq around the TSS, 5'SS, and 3'SS**

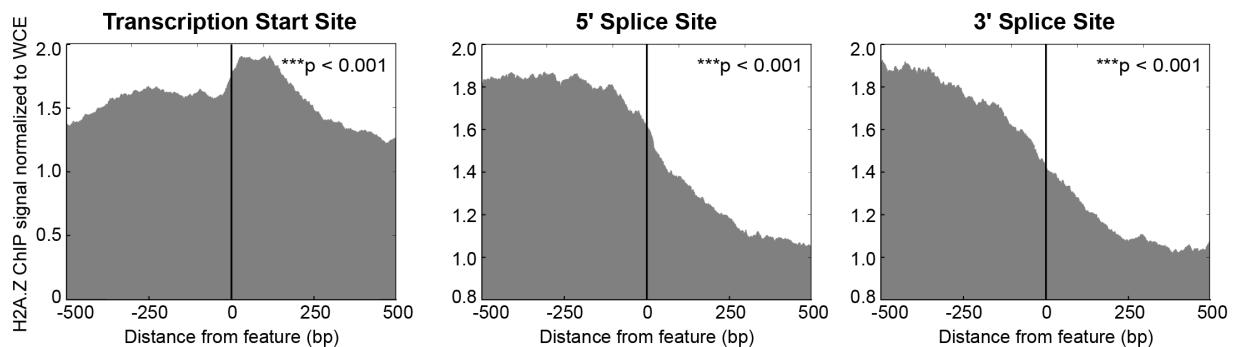
(A) Pile up of H2A.Z ChIP-seq reads at all TSSs, 5'SSs, and 3'SSs, normalized to input sample. Graphs represent data from two published replicates (Clément-Ziza et al. 2014). p-values were calculated by the K-S test for the data as compared to a uniform distribution.

**A**

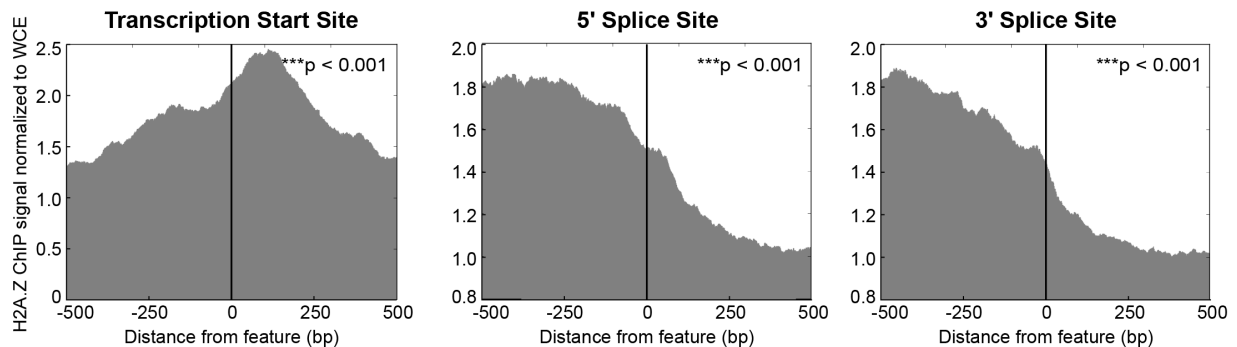
**This Study Replicate**



**Clément-Ziza Replicate 1**



**Clément-Ziza Replicate 2**

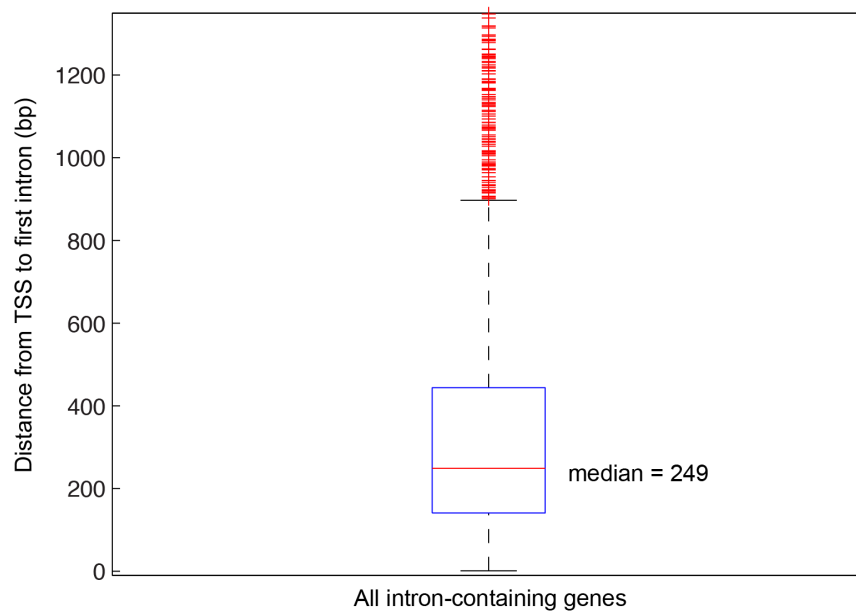


**Supplemental Figure 8: Pile up graphs of H2A.Z ChIP-seq around the TSS, 5'SS, and 3'SS for genes with introns affected by *pht1Δ***

(A) Pile up of H2A.Z ChIP-seq reads around the TSS, 5'SS, and 3'SS, for genes with introns showing splicing defects in *pht1Δ* at 16°C microarray, normalized to input sample. Graphs

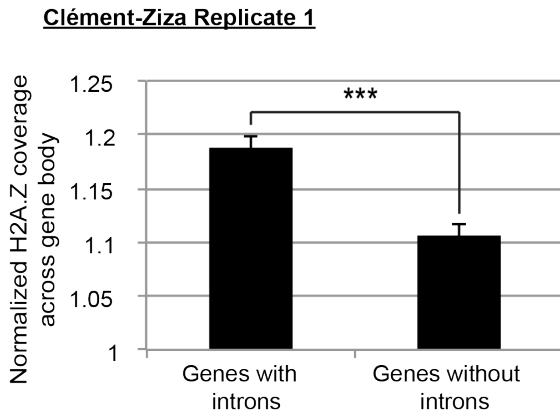
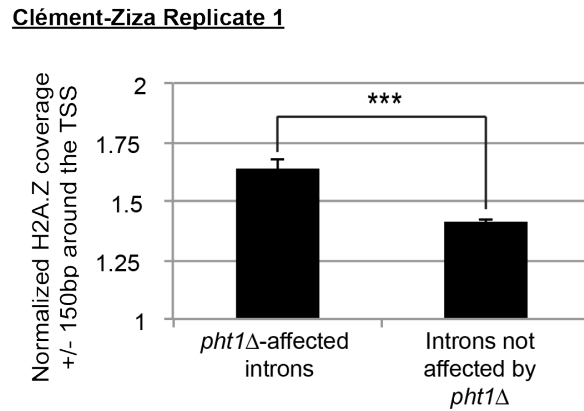
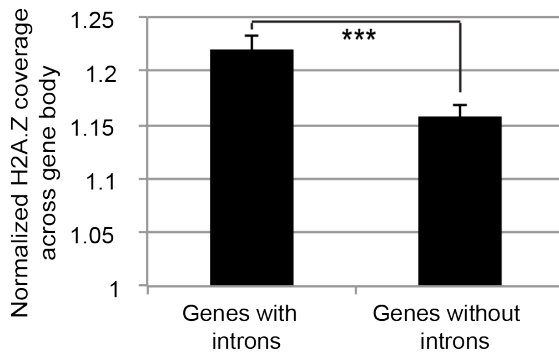
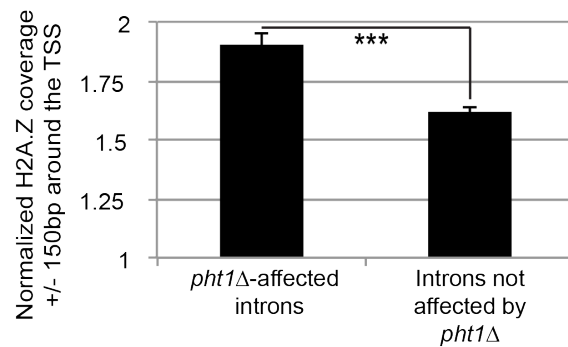
represent data from this study and two published replicates (Clement-Ziza et al. 2014). p-values were calculated by the K-S test for the data as compared to a uniform distribution.

**A**



**Supplemental Figure 9: First exon lengths in *S. pombe***

(A) Box plot of first exon lengths in *S. pombe* (Pombase). Median = 249bp, Maximum = 6549bp, n=2552 introns.

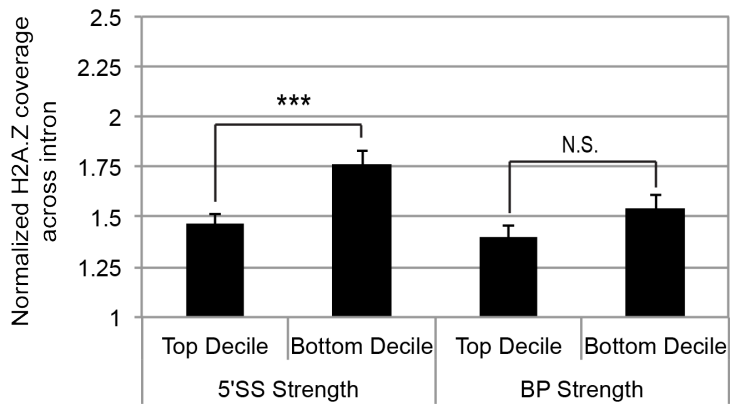
**A****B****Clément-Ziza Replicate 2****Clément-Ziza Replicate 2**

### Supplemental Figure 10: Additional replicates of average H2A.Z coverage analysis around TSS

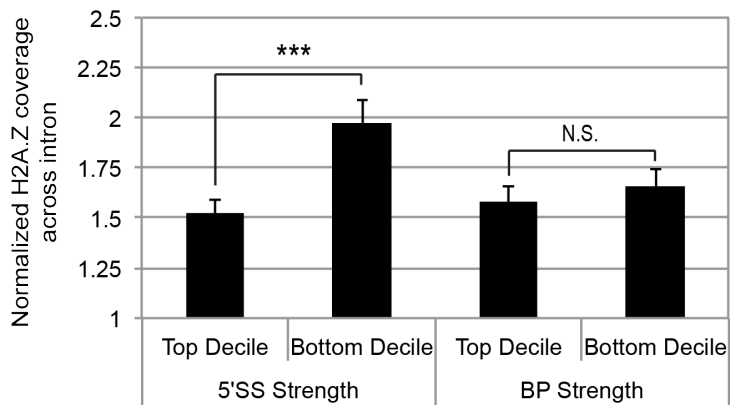
Average level of H2A.Z coverage measured by ChIP-seq (A) across genes with or without introns, or (B) within +/- 150bp around the TSS for introns with or without splicing defects in the *pht1*Δ microarray at 16°C, normalized to whole cell extract (WCE) and the coverage of both data sets. Graphs represent data from two published replicates (Clement-Ziza et al. 2014). Error bars represent S.E.M. \*\*\*p < 0.001 calculated by t-test.

A

**Clément-Ziza Replicate 1**



**Clément-Ziza Replicate 2**



**Supplemental Figure 11: Additional replicates of average H2A.Z coverage analysis for introns in the top and bottom deciles of *cis*-splicing sequence strength**

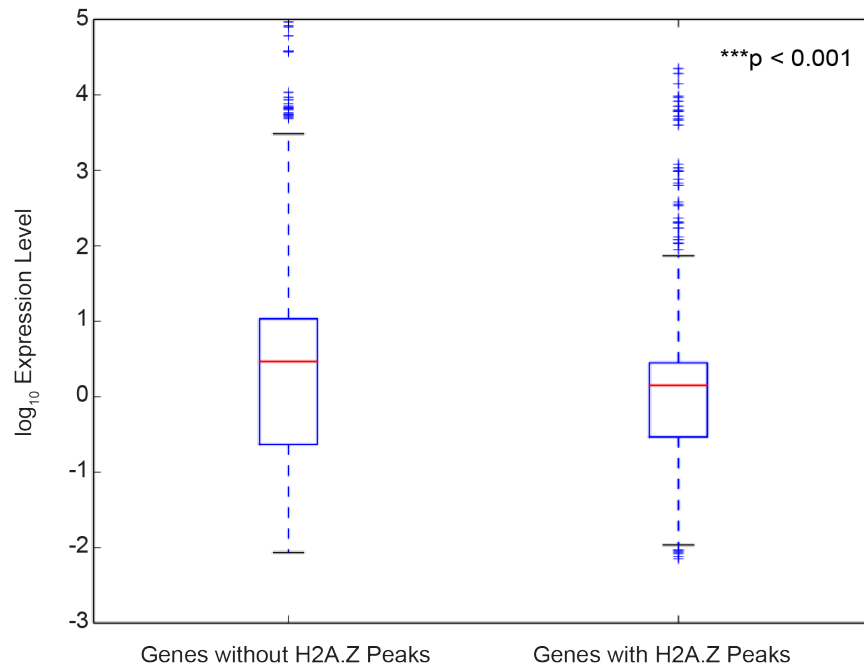
(A) Average level of H2A.Z coverage measured by ChIP-seq across introns in the top and bottom deciles of 5'SS and BP strength, normalized to WCE and the coverage of both data sets.

Graphs represent data from two published replicates (Clément-Ziza et al. 2014). Error bars

represent S.E.M. \*\*\* $p < 0.001$  calculated by t-test. N.S. = not significant. 5'SS and BP strength as previously defined (Stepankiw et al. 2015).



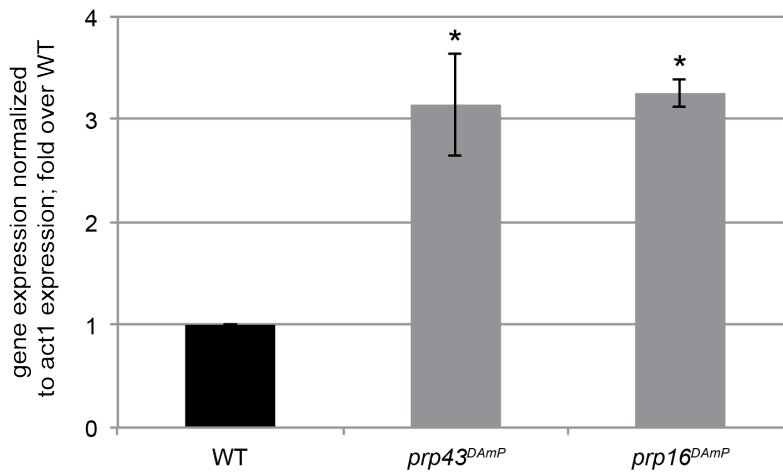
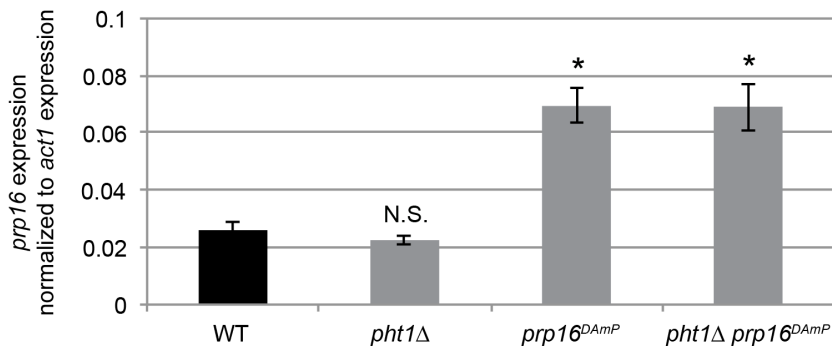
**A**



**Supplemental Figure 12: H2A.Z-associated genes are lowly expressed**

(A) Box plot of gene expression values for genes that have at least one associated H2A.Z ChIP-seq peak (peaks located within the gene body or 150bp of start or finish of the gene) compared to genes that do not have H2A.Z ChIP-seq peaks associated.

Gene expression values as previously defined (Marguerat et al. 2012). \*\*\*p < 0.001 calculated by t-test.

**A****B**

### Supplemental Figure 13: *prp16* and *prp43* expression levels

(A) RT-qPCR of *prp43<sup>DAmP</sup>* and *prp16<sup>DAmP</sup>* mRNA levels at 30°C. Bars represent fold change of mutant strain signal over WT, normalized to *act1* mRNA levels. Error bars represent S.E.M. of 3 biological replicates. \* $p < 0.05$  calculated by one-tailed t-test.

(B) RT-qPCR of *prp16* mRNA levels at 16°C normalized to *act1* mRNA levels. Error bars represent S.E.M. of 3 biological replicates. \* $p < 0.05$  calculated by one-tailed t-test. N.S. = not significant.

## **Supplemental Tables:**

Available online at [genesdev.cshlp.org](http://genesdev.cshlp.org)

### **Supplemental Table 1: Splicing Sub-complex Definitions**

Manually curated list of splicing sub-complexes and their component genes by *S. pombe* systematic name.

### **Supplemental Table 2: GO Term Analysis Data**

Full dataset of permutation analysis on the *S. pombe* EMAP for main text Figure 1A.

### **Supplemental Table 3: Strain List**

Detailed list of all strains described in this paper.

### **Supplemental Table 4: Processed Microarray Data**

Processed microarray data for splicing specific microarrays presented in the main text. “logFC” value refers to the “Splicing Index” score discussed in the main text and “adj.P.Val.” value was used for significance cut-off at  $p < 0.05$ . Analysis output from methods as previously described (Lipp et al. 2015).

## **CHAPTER 2:**

A Ribosomal Protein Gene Promoter Element Influences Co-transcriptional Splicing Regulation  
in Response to Amino Acid Starvation in *S. cerevisiae*

**Author Contributions:**

Megan Bergkessel: conceived experiments, designed and performed experiments, made strains and reagents, computational data analysis, wrote and edited the draft

Kelly Nissen: designed and performed experiments, wrote and edited the draft

Jaclyn Greimann: made strains and reagents

Christine Guthrie: conceived experiments, supervised the research, wrote and edited the draft

**Background:**

Megan Bergkessel initiated this project to gain further insight into the mechanism of an environmentally regulated pre-mRNA splicing phenomena completed during her tenure in the Guthrie Lab and she published the progress she made as a chapter in her graduate thesis (Bergkessel 2010). This project was subsequently driven forward for several years by a post-doctoral fellow, Jaclyn Greimann, but remained unpublished when she departed the lab. Given my interest in co-transcriptional splicing regulation mechanisms, I offered to help Megan complete this project by applying my knowledge of advances to the field and improved methodologies for probing genome-wide questions. I collaborated with Megan to revise initial RT-qPCR and ChIP experiments she had performed with more sensitive methods of data collection, as well as extend genome-wide analyses to more factors. The results of this effort allowed us to draw more significant and broad-reaching conclusions, as well as probe the mechanism in more detail.

**Abstract:**

In *S. cerevisiae*, mRNAs for most of the intron-containing ribosomal protein genes (RPGs) have been shown to rapidly accumulate in an unspliced state in response to amino acid starvation stress (Pleiss et al. 2007a). This block to normal splicing progression can be initiated by the addition of the drug 3-amino-1,2,4-triazole (3-AT), which mimics histidine starvation. To further understand the mechanism of the amino acid starvation-induced stall to co-transcriptional splicing, herein called the “3-AT response”, we interrogated the sequence requirements of introns and promoters of RPGs and non-RPGs by creating chimeric genetic constructs that swapped these sequences between endogenous loci. Surprisingly, we found that the identity of the intron is neither necessary nor sufficient for mediating the 3-AT response. However, the promoter sequence of an intron-containing RPG was both necessary and sufficient to confer a rapid splicing response to 3-AT treatment. Additionally, we show that constitutive splicing efficiency is influenced by the promoter. Given the promoter dependence of the 3-AT response and the intimate connection between splicing and transcription, we next examined the distribution patterns of RNA polymerase II (RNAP II) and early splicing factors of the U1 and U2 snRNPs, by ChIP-seq. We found that RNAP II signal was lower over the introns of RPGs but present and similar or higher levels at the 3' ends of these genes during 3-AT treatment. U1 signal came up later in the gene sequence and persisted aberrantly to the 3' ends of genes during 3-AT treatment, while U2 signal reached much lower levels than it did on RPGs in the absence of 3-AT. Finally, we interrogated the RNAPII association distribution in the absence casein kinase 2 (CK2) factor CKB2, a regulatory factor that is necessary for pre-mRNA accumulation during amino acid starvation stress (Bergkessel et al. 2011). We observed that the 3' end accumulation of RNAPII observed following 3-AT treatment is dependent on CKB2, suggesting

that its relevant activity could be co-transcriptional. Taken together, these results suggest RPG promoters contain sequence elements that are required for the pre-mRNA accumulation seen in response to 3-AT and that this response involves alterations to the co-transcriptional recruitment of splicing factors.



## **Introduction:**

Splicing is a complex, step-wise process, orchestrated by rearrangements of multipartite sub-complexes collectively composed of 5 snRNAs and over 100 proteins (Wahl et al. 2009). The U1 snRNP is recruited first to the 5' splice site of the intron followed by the U2 snRNP, which recognizes the branch point. Next, the U4/U6 • U5 tri-snRNP is recruited and the now massive assembly of splicing sub-complexes on the intron undergoes vast rearrangements that result in the release of U1 and U4 and lead to a catalytically active spliceosome. Splicing proceeds through two trans-esterification reaction steps separated by further rearrangements and protein composition changes. The activity of splicing factors, particularly the DExD/H-box ATPases that drive spliceosomal rearrangements, can be regulated by post-translational modifications, interactions with various nucleotide tri-phosphates and protein-protein/protein-RNA interactions (Staley and Guthrie 1998; Koodathingal and Staley 2013). The abundant opportunities to fine-tune splicing through altering spliceosome assembly dynamics and protein composition are only beginning to be unraveled.

In addition to regulatory interactions within the spliceosome, spliceosome assembly and catalysis occur largely co-transcriptionally, while the nascent pre-mRNA is still associated with RNA polymerase II (RNAP II) (Merkhofer et al. 2014). This high degree of coordination between splicing and transcription provides the opportunity for further regulatory interactions between the spliceosome, RNAP II and the local chromatin environment. In landmark papers from the Beggs and Neugebauer labs, RNAP II elongation dynamics were shown to be dependent on splicing, such that a pause occurs both at the 3' end of the intron, and at the 3' end of the gene, just to transcription termination (Alexander et al. 2010; Carrillo Oesterreich et al. 2010), suggesting that there are splicing and/or processing checkpoints for RNAP II. This RNAP

II elongation behavior is hypothesized to ensure that the splicing reaction completes co-transcriptionally. Changes to RNAP II elongation dynamics reciprocally influence splicing outcomes and can be further modulated by local nucleosome changes and post-translational modifications, such as C-terminal domain (CTD) phosphorylation states (Herzel et al. 2017). Moreover, co-transcriptional gene expression processes, such as the capping, splicing, and packaging for export of the mRNA, can significantly affect the expression of a gene, in an environmentally responsive manner (Kim Guisbert et al. 2007; Pleiss et al. 2007a; Munding et al. 2010; Hossain et al. 2016). Understanding the mechanistic underpinnings of regulation occurring at the level of co-transcriptional processing is complicated because multiple processes are interdependent. How signals from the spliceosome, RNAP II, chromatin landscape and other co-transcriptional gene expression processes are integrated to ensure tight regulation of transcript quality remains an active field of investigation.

In *S. cerevisiae*, the vast majority of expressed transcripts come from the ribosomal protein genes (RPGs), which are tightly regulated in response to environmental stress (Warner 1999). In addition to their high levels of expression, RPGs are biased in the distribution of introns in the budding yeast genome, in that 101 of the 296 introns are found in this sub-set of genes (Ares et al. 1999; Davis et al. 2000; Bon et al. 2003; Cherry et al. 2012). Given the additional energetic cost of splicing vast quantities of RPG pre-mRNA, it follows that splicing may also be tightly regulated in response to environmental stress. Indeed, work from the Guthrie lab has previously shown that when budding yeast are treated with an amino acid starvation-mimicking drug, 3-amino-1,2,4-triazole (3-AT), pre-mRNA splicing was rapidly and specifically inhibited for most intron-containing RPGs, causing a transient accumulation of unspliced transcripts (Pleiss et al. 2007a). In a follow-up study, this splicing inhibition response

(3-AT response) was robust in a variety of amino-acid limiting stresses and requires the activity of casein kinase II (CK2) via regulatory protein CKB2 (Bergkessel et al. 2011). This observation suggests that while alterations to rates of surveillance and decay could contribute to the magnitude of the 3-AT effect for some transcripts, an alteration to the co-transcriptional recruitment or activity of spliceosome on RPG transcripts also likely contributes. CK2 has numerous putative substrates involved in transcription and splicing, so its regulatory subunit is an attractive candidate as a mediator of changes to coordination between transcription and splicing (Allen 2008). However, the mechanism by which RPG splicing is regulated in response to amino acid starvation stress remains largely unclear.

In order to better understand possible mechanisms by which RPG introns accumulate during 3-AT treatment, we asked which *cis*-acting sequences are required for the effect. By constructing chimeric transcripts with swapped introns and promoters between 3-AT-responsive and -unresponsive genes and monitoring pre-mRNA accumulation, we found that the promoter of an intron-containing RPG was necessary and sufficient for the 3-AT splicing response, but that the intron identity had no effect. We further examined the dynamics of spliceosome recruitment and RNAP II elongation using ChIP-seq to ask whether 3-AT affected the distribution of U1, U2, and RNAP II on intron-containing genes. We found that early splicing factor recruitment, as well as splicing checkpoint behavior of the polymerase, were both affected, and that CKB2 affected the polymerase distribution especially in the presence of 3-AT. Taken together these results suggest that co-transcriptional recruitment of splicing factors is affected by 3-AT, with the transcript specificity of the response imposed by sequences in the promoter.

## Results:

### Chimeras integrated into their chromosomal loci test effects of cis-sequences on splicing

In order to test whether *cis*-sequences contribute to the high level of splicing efficiency and specificity of the 3-AT response on RPGs, we generated strains of yeast with chimeric, integrated constructs that swapped introns and promoters between responsive and non-responsive intron-containing genes (Fig. 1) and monitored splicing efficiency during a 3-AT treatment time-course (Fig. 2 and 3). By examining both the promoter and intron, which are landmark *cis*-sequences in the transcription and splicing cycles, we hoped to narrow the point in gene expression that leads to 3-AT splicing response. Genes were chosen based on their responsiveness to 3-AT as determined in a previously published splicing-specific microarray (Bergkessel et al. 2011). Intron swaps were generated using a modification of the “*Delitto perfetto*” method (Fiedler et al. 2009) and promoter swaps were generated by a gene replacement strategy, inserting the new promoter and an antibiotic resistance cassette 5' of the promoter in place of the old promoter. To determine whether the intron is necessary for rapid pre-mRNA accumulation during amino acid starvation stress, we replaced the intron of the NMD2 gene, for which splicing is not responsive to 3-AT, with an intron from a gene which is highly-expressed and strongly accumulates pre-mRNA after 3-AT treatment, RPS9B. To test whether the intron is sufficient to ectopically confer the 3-AT splicing response in a non-responsive intron-containing gene, we replaced the intron of RPL24B, another highly expressed and strongly 3-AT responsive RPG, with the intron of NMD2. To interrogate the requirement of RPG-specific promoter sequences in the 3-AT response, we also replaced the promoter in front of RPL24B with two promoters from non-intron containing genes that drive expression levels in the range of RPGs. We swapped the promoter from RPS5 gene, an intron-less RPG that is constitutively highly

expressed under favorable growth conditions, and the promoter from the CUP1 gene, which can be driven to high levels of expression by the addition of copper sulfate to the growth media. To test the sufficiency of an RPG promoter in conferring a splicing response to amino acid starvation stress in a non-responsive intron-containing gene, we replaced the NMD2 endogenous promoter with the promoter from RPL24B. The promoter sequences defined for RPL24B and RPS5 are broad and contain the entire genomic region between the transcription start site of interest and the neighboring upstream transcript's boundaries by manual curation of published data sets (David et al. 2006).

### **Intron identity affects steady-state splicing efficiency, but not responsiveness to 3-AT**

The chimeric constructs were tested against wild-type for their splicing response to 3-AT treatment by RT-qPCR, as described in the Materials and Methods. In brief, the strains containing either the wild type or chimeric sequences were grown in culture to mid-log phase in a synthetic media lacking histidine. An aliquot was removed as a zero time point prior to 3-AT, followed by additional aliquots taken at 4 and 8 minutes after 3-AT addition. Total RNA was extracted and prepared into cDNA for RT-qPCR analysis for each relevant transcript. Primers were designed to the exon sequence and either the endogenous or chimeric intron sequence. Because the 3-AT response is characterized by a loss in splicing efficiency and accumulation of unspliced pre-mRNA, we calculated steady state transcript levels and splicing efficiencies prior to 3-AT addition (Fig. 2A and 2B respectively) as well as the change in the percent of unspliced transcript, relative to time zero, at 4 and 8 minutes after 3-AT addition (Figure 2C).

The intron sequence did appear to have some effect on the steady state splicing efficiency and even possibly on steady state transcript levels. Specifically, replacing the NMD2 intron with

the much longer RPS9B intron appeared to decrease its splicing efficiency, while replacing the RPL24B intron with the much shorter NMD2 intron appeared to increase its splicing efficiency (Fig. 2B). However, this decrease in percent unspliced transcript seemed to come at a cost of lower overall transcript levels for the RPL24B-NM2 chimeric transcript compared to the RPL24B wild type transcript (Fig. 2A).

While swapping introns between RPG and non-RPG transcripts did appear to subtly affect steady-state splicing, chimeric intron swap constructs revealed that RPG intron sequences are neither necessary nor sufficient to confer pre-mRNA accumulation following 3-AT treatment (Figure 2C). The splicing of chimeric RPL24B transcript containing the NMD2 intron behaved more similar to that of the wild-type RPL24B transcript than the wild-type NMD2 transcript, suggesting that the intron identity is not necessary for the 3-AT splicing response. Similarly, when comparing the chimeric transcript of NMD2 containing the RPS9B intron to the wild-type NMD2 transcript and the wild-type RPS9B transcript, we found that like wild-type NMD2 (and unlike wild-type RPS9B), no significant change in the chimera's splicing efficiency occurred following 3-AT treatment, suggesting that an RPG intron is not sufficient to confer this response. This is consistent with the findings of a high-throughput screen in which each yeast intron was inserted into a fluorescent reporter gene, and no difference in behavior was observed between reporters with RPG introns and those with non-RPG introns, under a variety of stress conditions (Yofe et al. 2014).

### **Promoter identity affects both steady-state splicing efficiency and 3-AT responsiveness**

In contrast to the intron sequence, the results of the promoter swap experiments show that an intron-containing RPG-specific promoter element is both necessary and sufficient to confer a

splicing response to 3-AT treatment and that steady-state splicing efficiency is also influenced by the promoter (Figure 3). Replacing the native RPL24B promoter with either the RPS5 promoter or the CUP1 promoter resulted in high levels of transcript expression (Figure 3 panel A). However, the native RPL24B promoter yielded higher levels of constitutive splicing efficiency on the RPL24B transcript than the two other non-native promoters (Figure 3B). Furthermore, the chimeric RPL24B constructs driven by RPS5 and CUP1 promoters did not accumulate increased levels of unspliced pre-mRNA in response to 3-AT treatment, suggesting that a sequence element in the promoter of an intron-containing RPG is necessary for both efficient constitutive splicing of a highly expressed transcript and a splicing response to amino acid starvation stress.

Finally, we determined that a sequence element of an RPG promoter, in this case RPL24B's promoter, is sufficient to impart an intron accumulation response after 3-AT treatment to a non-RPG transcript (Figure 3 Panel C). The wild-type NMD2 transcript does not change splicing efficiency in response to 3-AT treatment, but the chimeric transcript of NMD2 driven by the RPL24B promoter accumulates additional unspliced transcript, especially by the 8 minute time point (Figure 3 Panel C). Again, we observed that a non-native promoter (the RPL24B promoter) driving expression resulted in a lower level of constitutive splicing efficiency compared to that seen for the wild-type NMD2 transcript, possibly because expression levels under the RPL24B promoter were much higher.

### **ChIP-Seq shows RPG-specific changes to distributions of U1 snRNP, U2 snRNP, and RNA polymerase II in response to 3-AT**

Given the importance of promoter sequence, we wondered whether co-transcriptional recruitment of early splicing factors might be affected by 3-AT. To address this question, we

performed ChIP-seq and compared distributions of U1 snRNP, U2 snRNP, and RNAP II in the presence and absence of 3-AT genome wide (Figure 4 and Figure 5). To measure the U1 and U2 snRNP association, we HA tagged U1 factor Prp42 and U2 factor Lea1. To measure RNAP II, we used an antibody raised against core polymerase factor Rpb3. ChIP enrichment was calculated as a ratio of normalized sequencing reads per 100 base pair (bp) segment of the genome for the immunoprecipitated DNA versus input DNA (whole cell extract). We monitored two biological replicates for each ChIP reaction, with and without 3-AT treatment for 4 minutes prior to crosslinking, and the values plotted are the average of both experiments.

First, we show a two-page of example of ChIP distributions of U1, U2 and RNAP II on RPGs and non-RPGs, with and without 3-AT treatment, including genes used in the intron and promoter swap experiments above (Figure 4). Next, for each ChIP experiment, we show the average behavior across all intron-containing RPGs and intron-containing non-RPGs for which transcription start sites and polyadenylation sites were defined, that had only one intron, with either 3-AT treatment or a mock water control (Figure 5). In order to generalize across multiple genes, we averaged together the values for the same positions in each gene relative to the transcription start, intron boundaries, and polyadenylation sites (Figure 5 panel A). The error bars are the standard error of the means across all the genes in the category.

We found that the peak of U1 snRNP (Prp42) chromatin association signal on RPGs is shifted downstream and is generally lower after 3-AT treatment. This downstream shift of U1 accumulation is seen in the example RPGs in Figure 4 A, B, and D, as well as in Figure 5 transcript region 6, where the levels of U1 are higher at the 3' end of the RPGs in 3-AT treatment. This indicates that 3-AT might affect steps even prior to U1 recruitment, and cause U1 recruitment to be less efficient, and thus not be well recruited until the polymerase is further



downstream. Consistent with our results above that show a role for a promoter element in the 3-AT response, it has been shown in humans that U1 starts to be recruited directly to the RNAP II, further suggesting possibility of non-intron signal affecting U1 recruitment patterns (Morris and Greenleaf 2000). There is a lower level of U1 snRNP association in non-RPGs across the gene body as well, and it is further lowered in 3-AT treatment, as see in the aggregate graph of Figure 5. This slight difference is hard to distinguish on the example genes of Figure 4. For the non-RPGs, the beginnings and endings of the Prp42 peak are largely unchanged with and without 3-AT treatment (Figure 5 transcript region 1, 2, 6, and 7).

In contrast to U1, the U2 snRNP (Lea1) association peaks are very similar in their shapes and positions between the 3-AT treated and untreated samples, but the peak heights are smaller in 3-AT treatment, indicating a lower level of U2 association across the gene body. The magnitude of this effect is larger for RPG introns. Additionally, the loss of U2 association in 3-AT treatment is more easily visible in the aggregate genome-wide graphs for non-RPGs because of the smaller magnitude (Figure 5). One interpretation of these data is that delayed recruitment of U1 snRNP fails to efficiently recruit U2 co-transcriptionally, thus lowering overall U2 signal. Another possibility is that U2 recruitment is affected by a separate mechanism that is not directly dependent on the U1 status.

The RNAP II distribution was also altered in the presence of 3-AT (Figure 4 and Figure 5). Strikingly, in the RPGs, there is a peak of RNAP II around the 3' end of the intron in the untreated sample that is abolished in 3-AT treatment. This behavior is easily observable in the intron region of the three RPG example genes in Figure 4 and in the aggregate analysis of Figure 5 at transcript regions 3, 4, and 5, which span the intron and beginning of the second exon. This local peak around the end of the intron is not observed in non-RPGs with or without 3-AT

treatment. Such a peak is consistent with a "splicing checkpoint" in transcription, observed by Jean Beggs' group as a buildup of RNAP II by ChIP around the 3' splice/beginning of the second exon, that is dependent on Prp5 and Cus2, U2 snRNP-associated splicing factors (Alexander et al. 2010). If after 3-AT treatment, U1 recruitment is late and U2 recruitment fails to occur co-transcriptionally, this checkpoint must be much less enforced. Interestingly, by the terminal exon, RNAP II levels have been restored to wild type levels in 3-AT treatment on the RPGs, such that the effect of 3-AT primarily affects polymerase accumulation around the intron. Again, we observe a peak of RNAP II prior to the end of the transcript, consistent with the Neugebauer lab's observation of 3' end accumulation of RNAP II in co-transcriptionally spliced genes (Carrillo Oesterreich et al. 2010). This distribution change of RNAP II across RPGs in 3-AT treatment could contribute to intron accumulation by skipping a splicing dependent checkpoint for polymerase in the gene body but allowing transcripts to be terminated normally, and therefore exported from the nucleus unspliced, rather than causing the whole polymerase-transcript-splicing factor machinery to stall and wait for proper recruitment of splicing factors at the end of the intron.

### **Deletion of CKB2 alters pattern of RNA polymerase accumulation on RPGs after 3-AT treatment**

Because we previously saw that a CK2 regulatory subunit, CKB2, is required for the RPG splicing response to 3-AT, we examined the distribution of RNAP II in the presence and absence of 3-AT in a  $\Delta ckb2$  strain (Figure 4 and Figure 5) (Bergkessel et al. 2011). We performed ChIP-seq for Rpb3, as above, on a single biological sample of the  $\Delta ckb2$  strain, because of time constraints. Nevertheless, the results are interesting. The most striking difference

between the behavior of RNAP II in the  $\Delta$ ckb2 strain compared to the wild-type strain is that RNAP II levels are much lower at the 3 ends of RPGs following 3-AT treatment (Figure 5, transcript region 6 and 7). As with wild type described above, in aggregate (Figure 5) and in the three individual RPG examples shown (Figure 4), RNAP II levels in  $\Delta$ ckb2 under 3-AT treatment are lower in the intron. The effects on non-RPGs are more subtle, and the effects of the CKB2 deletion in the absence of 3-AT are also subtle.

## **Discussion:**

In previous work, we showed that pre-mRNAs for RPG transcripts rapidly accumulate in response to amino acid starvation stress, and that this accumulation depends on the casein kinase 2 regulatory subunit, CKB2 (Pleiss et al. 2007a; Bergkessel 2010; Bergkessel et al. 2011). In the current work, we attempt to further constrain the possible mechanisms that could contribute to the observed pre-mRNA accumulation, and the roles that CKB2 might play. We show that accumulation of pre-mRNA in response to 3-AT depends on *cis*-sequences in the promoter of an RPG transcript and not on intron sequence, while both the promoter and intron identity influence steady-state splicing efficiency (Figure 2 and 3). Given that promoter sequences have been proposed to influence early splicing factor recruitment in higher eukaryotes, we examined the distributions of association of U1, U2 and RNAP II across intron-containing genes using CHIP-Seq, in the presence and absence of 3-AT. We found that 3-AT treatment biases RNAP II and U1 distribution toward the 3' ends of RPGs and decreases the association of U2. These effects were much more pronounced for RPGs than for non-RPGs (Figure 4 and 5). Finally, we showed that deletion of CKB2 lowers the association of RNAP II with the 3' ends of RPGs in the presence of 3-AT, suggesting that the impact of the CKB2 deletion on pre-mRNA accumulation is due to a co-transcriptional regulatory role.

### **The 3-AT response is due to a change in splicing, not just a change in decay**

One possibility that has been suggested is that the observed accumulation of pre-mRNA in response to 3-AT treatment is due exclusively to inhibition of decay processes, either nuclear or cytoplasmic. This model requires that a steady-state level of surveillance is constitutively responsible for removing RPG pre-mRNAs that have failed to splice, and that this surveillance

activity is inhibited by the 3-AT response, thus allowing for the constitutive levels of splicing failure on RPG transcripts to be detected. This model is based in part on work suggesting that cytoplasmic decay of transcripts is inhibited in response to osmotic stress (Molin et al. 2009; Romero-Santacreu et al. 2009), and also on findings that nuclear decay can change in a transcript-specific manner in response to glucose withdrawal (Bresson et al. 2017). However, we believe that several lines of evidence, including the new data presented here, argue against this model for the 3-AT responsive accumulation of RPG pre-mRNAs.

First, we have previously shown that the responses of RPG transcripts to osmotic stress and glucose withdrawal are distinct from the response to amino acid starvation, so whichever mechanisms act under those conditions are not acting equivalently under the amino acid starvation condition (Bergkessel et al. 2011). Additionally, we have previously shown that deletions of many different nuclear and cytoplasmic decay factors fail to phenocopy the accumulation of RPG pre-mRNAs that we see after 3-AT treatment, suggesting that inhibition of these factors cannot fully explain the 3-AT response (Pleiss et al. 2007a). This evidence is essentially negative however - it shows that the 3-AT response does not match and is not explained by characterized decay mechanisms, but it does not rule out the possibility that an as-yet undescribed decay pathway is being inhibited in the presence of 3-AT in order to allow RPG pre-mRNA accumulation.

The new ChIP-Seq data presented here positively shows that co-transcriptional recruitment of splicing factors is altered by 3-AT treatment, in a manner that is consistent with the transient and partial blockage to splicing that is observed. Additionally, the finding that the promoter is necessary and sufficient for conferring a 3-AT responsive pre-mRNA accumulation to a spliced transcript further supports the idea that the mechanism responsible occurs co-

transcriptionally. We do not dispute the possibility that changes to nuclear or cytoplasmic decay rates could in fact modulate the size of the new pool of unspliced pre-mRNA that results from this decrease in splicing factor recruitment, but we think it is a critical point that the pool of unspliced RPG pre-mRNA that could be affected by changing decay pathways comes into existence through a change in the behavior of the splicing machinery.

### **A model for the mechanism of the 3-AT splicing block**

While we have shown that the 3-AT response in a wild-type strain background correlates with changes in U1, U2, and RNAP II distribution across intron-containing RPGs and that the accumulation of RPG pre-mRNA in response to 3-AT is lost if the promoter is replaced or if CKB2 is deleted, connecting these observations to yield a deep mechanistic understanding of how co-transcriptional splicing can be altered in response to environmental signals will require much more work. Nevertheless, we think it useful to propose a model that connects the current observations and identifies specific questions for further exploration.

Many others have related peaks in RNAP II ChIP signal to pauses in elongation, and have used the position of splicing factor ChIP signal peaks along genes as proxies for the time of recruitment (Carrillo Oesterreich et al. 2010). Although it is true that a ChIP experiment provides a single static snapshot of the population of factor positions (Bieberstein et al. 2014), we think that these approximations are useful and appropriate in our case. Our previous microarray data show that the total amount of RPG transcript does not decrease in the time frame of our ChIP experiments, so we assume that the rate of transcription initiation remains relatively constant, and under this assumption, the magnitude of the RNAP II ChIP signal should be inversely proportional to elongation speed, and the position of splicing factor peaks should relate to the

time of their recruitment. In our untreated samples, we see small peaks of RNAP II signal near the end of the intron and near the end of the transcript, especially for the RPG transcripts (Figure 5C, red). These positions have previously been proposed as “checkpoints”, where RNAP II can slow in order to allow for proper splicing and/or other transcript processing, by the Beggs and Neugebauer labs respectively (Figure 6, top panel). In 3-AT treated samples, by contrast, we see much lower RNAP II signal throughout the intron, which we suggest could indicate a bypass of the checkpoint, proposed by the Beggs lab, at the 3' end of the intron. At the same time, we see that the increase in U1 signal is generally shifted somewhat later in the transcript by 3-AT treatment, which could be consistent with bypassing a RNAP II checkpoint that otherwise would provide additional time for U1 recruitment closer to the 5' end of the gene. In contrast to U1, U2 signal shows an overall decrease for RPG genes but no shift in the position of the peak. This could suggest that transcripts that recruit U1 too late simply fail to recruit U2 co-transcriptionally. A failure to recruit later splicing factors could also lead to prolonged retention of U1, which we see as increased U1 signal at the 3' ends of genes (Figure 6, middle panel). Previous work has suggested that while post-transcriptional splicing is possible, it is much less efficient than co-transcriptional splicing (Tardiff et al. 2006), so any transcripts that fail to recruit U2 co-transcriptionally and retain U1 at the end of the gene would be expected to splice very inefficiently, if at all.

Based on single-molecule fluorescence in-situ hybridization experiments performed against the intron sequences of two different RPG transcripts, in which a large increase in intron signal was observed in the cytoplasm after 3-AT treatment (Bergkessel 2010). We think it likely that most of the RPG transcripts that fail to splice co-transcriptionally are nevertheless exported from the nucleus (Figure 6, middle panel). Interestingly, we see that a peak of RNAP II signal

that occurs at the 3' ends of RPG transcripts is present regardless of whether 3-AT treatment has occurred (Figure 5C). In contrast to the behavior observed over the intron regions, the RNAPII signal is if anything slightly higher over the 3' ends of RPGs in the presence of 3-AT compared to its absence (Figure 4A, B, and D). If the 3'-end-of-gene "checkpoint" detected by both ChIP and nascent transcript sequencing by the Neugebauer lab relates to processing transcripts for export, the fact that this checkpoint is relatively unperturbed by 3-AT could be consistent with ongoing export of transcripts, whether they are spliced or not.

Interestingly, the most obvious impact of CKB2 deletion on RNAP II ChIP profile across RPGs was a loss of the peak at the 3' ends of genes during 3-AT treatment (Figure 5D). We do not yet have ChIP data for U1 and U2 in the  $\Delta ckb2$  strain, so we cannot comment on whether CKB2 directly influences splicing factor recruitment. But one possible interpretation of our RNAP II ChIP data for this strain is that export of the unspliced transcript to the cytoplasm is affected. Instead, the unspliced transcript could be rapidly degraded by the nuclear decay machinery, or could dissociate from the site of transcription and go on to be spliced post-transcriptionally (Figure 6, bottom panel). We note that in our previously reported microarray data, a loss of total transcript was evident for the RPGs in the  $\Delta ckb2$  strain earlier in 3-AT treatment than was seen for the wild-type, possibly supporting the idea that the unspliced transcripts are aberrantly targeted for nuclear decay in this strain, but more work is needed.

Major questions remain regarding how and whether our correlated observations can be linked mechanistically. For example, how might a promoter sequence be mechanistically linked to an environmentally-responsive change in U1 recruitment and splicing checkpoint enforcement? Connections between promoters and splicing have been investigated extensively, and decades ago it was observed that transcripts produced from genes swapped with RNA



polymerase III promoters are not spliced (Sisodia et al. 1987). Several mechanistic links between promoters and splicing have been described that could in theory provide for environmental sensing, but more work will be needed in order to determine which might be important in the case of the 3-AT response. One possibility is that the promoter sequence itself directly recruits splicing factors that affect transcription and splicing kinetics from the point of transcription initiation onward (Rosonina et al. 2005). Numerous examples exist in humans where alternative splicing patterns are dependent on recruitment factors to the promoter (Monsalve et al. 2000; Huang et al. 2012) (Cramer et al. 1997), but many of these alternative splicing factors do not exist in yeast. Alternatively, post-translational modifications to the polymerase, such as acetylation or phosphorylation, could be set at the promoter and subsequently influence polymerase behavior throughout the gene. These two possibilities have been proposed to explain promoter-dependent changes in alternative splicing in higher eukaryotes (Kornblihtt 2005), and the phosphorylation state of the RNAP II CTD has been shown to correlate with the checkpoint pausing observed at the 3' ends of introns (Alexander et al. 2010). In fact, CKB2 (but not CKB1) is reported to physically interact with both the large subunit of RNAP II and the RPG transcription factors FHL1 and IFH1, potentially providing a link between the requirement for intron-containing RPG promoters and CKB2 in the 3-AT response, but this link will require testing (Rudra et al. 2007). Finally, different promoter elements can influence polymerase kinetics by affecting chromatin structures such as histone modifications and variant histones, which indirectly influence splicing (Huff et al. 2010; Neves et al. 2017), Chapter 1. The chromatin modifications of RPG promoters are known to be affected by environmental signals, and CK2 has recently been shown to directly phosphorylate histone H2A, leading to changes in

multiple other histone modifications (Basnet et al. 2014), suggesting another potential type of link among promoter sequence, CK2 activity, and splicing.

Which roles of CK2 are relevant in the 3-AT response remains an outstanding question. As noted above, CK2 has been reported to phosphorylate RNAP II and histone H2A, but these are just two of dozens of reported substrates that have roles in transcription, splicing, and mRNA export. On the one hand, this makes CK2 a very attractive candidate for a regulator that could sensitively and coordinately adjust rates of multiple interrelated processes in order to change transcriptional output in response to changes in environmental conditions. On the other hand, this promiscuous activity makes it very difficult to even definitively identify *bona fide* substrates, let alone determine how they might all interact in phosphorylated versus unphosphorylated forms. Nevertheless, a few putative CK2 substrates bear mentioning. A screen for CK2 substrates in human cells, carried out by Jasmina Allen in Kevan Shokat's lab, identified the conserved splicing ATPases Prp28, Prp5, Prp2, and Prp43 (Allen 2008). Prp5 has been shown to play a role in the transcription checkpoint occurring at the 3' ends of introns (Chathoth et al. 2014), and Prp28 also plays a role in rearrangements of early splicing factors (Price et al. 2014). Prp43 has putative roles in spliceosome disassembly, destabilizing the interaction between U2 and the intron (Fourmann et al. 2016), but also interestingly is involved in rRNA processing and ribosome biogenesis (Bohnsack et al. 2009). Alterations in any or all of these activities could be imagined to play a role in the 3-AT response. Finally, CKB2 (but not CKB1) has also been shown to physically interact with Nab2, the nuclear polyA binding protein that is involved in mRNA export, but also has interactions with splicing and nuclear decay pathways (Soucek et al. 2016). Given our observation that deletion of CKB2 results in a change in RNAP II behavior near the 3' ends of RPGs in the presence of 3-AT, it is tempting to speculate that a change in the

phosphorylation status of Nab2 could have impacts on the partitioning of transcripts to export or nuclear decay, in a manner that is sensitive to splicing status.

### **What is the purpose of this response?**

While many mechanistic details remain to be worked out, an important broader question is what function this response might serve. What advantage might there be to exporting unspliced RPG transcripts to the cytoplasm as part of an amino acid starvation response, as opposed to downregulating RPG expression by some other means? This is a difficult question to answer definitively, but imagining the possibilities can help contextualize the mechanistic questions that remain.

First, we note that yeast cells are clearly capable of rapidly and dramatically downregulating total levels of RPG transcripts, but this is not the response that is observed during amino acid starvation. Instead, we saw extremely rapid loss of RPG transcripts following glucose withdrawal or treatment with drugs that inhibit the TORC1 complex (Bergkessel et al. 2011), and recently, this rapid downregulation of RPG expression in response to glucose withdrawal was also linked to the transcript-specific activity of the nuclear exosome (Bresson et al. 2017). The case of amino acid starvation, cells may not need to strongly shut down ribosome biogenesis, but only transiently repress expression and make RNAP II more available for expressing amino acid biosynthetic genes. Perhaps this can be achieved efficiently by relieving splicing checkpoint enforcement to allow RNAP II to rapidly complete expression of RPGs, and then export them to keep them from occupying nuclear machinery. This strategy could potentially also allow promoter histone modifications and transcription factor recruitment to be

maintained for rapid resumption of ribosome biogenesis following upregulation of amino acid biosynthetic genes.

While we have examined it in the context of shutting down ribosome biogenesis in a stress response, such a mechanism could also presumably function to help coordinate the establishment of the nuclear organization that supports high-level ribosome biogenesis during upregulation. Indeed, a high-resolution study of temporal organization of gene expression during the yeast metabolic cycle provides some evidence for this notion: at the time point immediately preceding a coordinated burst of ribosome production that is a hallmark of the yeast metabolic cycle, the ribosomal protein gene introns were observed to transiently accumulate, and expression of CK2 was also high at this time point (Kuang et al. 2014). Perhaps disposing of RPG transcripts by sending them unspliced to the cytoplasm is a way to prime the nuclear ribosome biogenesis machine without producing stoichiometrically unbalanced ribosome components or starting production before the correct metabolic conditions have been established.

### **Future Directions**

Many of the hypotheses presented in this discussion could be directly tested. Immediate next steps could be repeating the RNAP II ChIP-Seq experiment in the  $\Delta$ ckb2 strain, and testing whether splicing factor recruitment is affected in the  $\Delta$ ckb2 strain. Additional ChIP experiments to measure chromatin modifications and recruitment of CK2 and RPG-specific transcription factors to RPGs and RPG promoters of interest, in the presence and absence of 3-AT, could also be informative. Further experimentation might involve testing mutants for putative CK2 substrates to see if they are defective for the 3-AT response, measuring phosphorylation levels of putative CK2 substrates during the 3-AT response and determining whether the Ckb2 regulatory

subunit confers specificity toward likely relevant substrates. Already, our studies of the 3-AT responsive accumulation of pre-mRNA in budding yeast have suggested that co-transcriptional splicing of the highly expressed, essential RPG transcripts can be adjusted based on environmental conditions, and that this adjustment involves changes in the behavior of RNAP II and splicing factor recruitment. Ultimately, we hope that further study of the splicing response to this perturbation will yield fundamental insights into how co-transcriptional processes can be coordinately adjusted to sensitively optimize environmental responses in time and space.

## **Materials and Methods:**

Detailed information about the methods of strain creation can be found in Megan Bergkessel's dissertation (Bergkessel 2010) and the laboratory notebooks of Jaclyn Greimann.

### **Strains**

A list of strains used in this study is available in Table 1.

### **RNA Isolation for RT-qPCR**

Yeast were grown in a synthetic media lacking histidine overnight at 30°C and then diluted to 0.1 OD<sub>600</sub> in 50mL the next morning. For copper induction of CUP1, copper sulfate was added to a final concentration of 25mM to the 50mL culture. Cultures were then grown to mid log-phase (0.3-0.7 OD<sub>600</sub>) and an untreated 15mL aliquot was removed. 3-Amino-1,2,4-triazole (3-AT) (Sigma) was added to a final concentration of 50mM, and cultures were returned to 30°C. 15mL aliquots were removed from the treated culture at 4 minutes and 8 minutes. Cells were rapidly harvested by vacuum onto a nitrocellulose membrane, flash-frozen in liquid nitrogen, and stored at -80°C. RNA was isolated using hot acid phenol followed by isopropanol precipitation as previously described (Bergkessel et al. 2011).

### **qRT-PCR**

For RT-qPCR, RNA was treated with TURBO DNA-free kit (Ambion) and then iScript cDNA Synthesis kit (Bio-Rad) was used to generate cDNA. qPCR was performed using a CFX96 Real Time system (BioRad) using iTaq Mastermix (Bio-Rad).

qRT-PCR Data Analysis. For each primer pair, Ct values were converted to relative quantities by comparison to a standard curve generated by serial dilutions of the appropriate (i.e. WT or chimeric) genomic DNA. Although these quantities are still not absolute, because the intron and exon are present stoichiometrically in the genomic DNA, this comparison to a genomic DNA standard curve allows the intron and exon quantities to be compared to each other in order to estimate the percentage of each transcript that is unspliced at each time point. At least three biological replicates were analyzed for each strain and time point. Error bars report the standard error of the mean and the significance of the effect of 3-AT on the percentage of transcript that remained unspliced was analyzed using the aov function provided in the stats package for R.

### **Chromatin Immunoprecipitation (ChIP)**

Yeast were grown in a synthetic media lacking histidine overnight at 30°C and then diluted to 0.1 OD<sub>600</sub> the next morning. 500mL cultures were grown to between 0.3–0.7 OD<sub>600</sub> and then split. Half of the culture was treated with 3-AT to a final concentration of 50mM and the other half was treated with water.

Cells were crosslinked with 1% formaldehyde (Sigma) for 15 minutes, and then quenched for 5 minutes with 125mM glycine (Sigma). Crosslinked cells were isolated by centrifugation in 50mL aliquots and washed twice with 4°C TBS. Pellets were flash frozen in liquid nitrogen and stored at -80°C. Each aliquot was thawed on ice and resuspended in 1 mL ChIP lysis buffer (50mM HEPES-KOH pH 7.5, 140mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, and 1x EDTA-free Complete protease inhibitor (PI; Roche) Protease

inhibitors (2mM PMSF, 20µg/mL leupeptin, 20µg/mL pepstatin, 20µg/mL aprotonin, 20µg/mL antipain). Lysate was generated using a Mini-Beadbeater (Biospec Products) for 6 cycles of 1.5 minutes beadbeating followed by 1.5 minutes on ice after each cycle. DNA was isolated by centrifugation. DNA was sheared in a Bioruptor waterbath sonicator (Diagenode) with 6 cycles on high (10 min total, 30 sec ON, 1 min OFF). After sonication, cell debris was removed by centrifugation and the supernatant was brought up to 1mL in ChIP lysis buffer. 40µL of the lysate was then removed and saved as an input. The remaining 960µL of lysate was split into fifths of 192µL. One fifth was brought up to 500µL and 3ug of antibody against Rpb3 (1Y26, BioLegend) was added. The remaining four-fifths were brought up to 1000µL and 7ug of antibody against HA (12CA5, Roche) was added. Reactions were incubated with nutation for 2 hours at 4°C. Next, 30µL of packed Protein G DYNAbeads (Invitrogen) per reaction were washed for 5 minutes each at 4°C: twice with 1mL TBS and then twice in 1mL ChIP lysis buffer, and then resuspended in 100µL ChIP lysis buffer. 100µL of washed beads was add to each reaction and incubated for 2 hours at 4°C for immunoprecipitation. Beads were then washed for 5 minutes each at 4°C: twice with 1mL ChIP lysis buffer, twice with 1mL high salt ChIP lysis buffer (ChIP lysis buffer with 500mM NaCl), and once with TE at room temperature. Crosslinking was reversed in 1% SDS at 65°C overnight in the presence of Proteinase K (30µL of 20mg/mL). SDS was removed from DNA using NucleoSpin Gel Clean-up Columns with 1250µL NTB (Macherey-Nagel). The resulting DNA was recombined from the split aliquots based on antibody used for both treated and untreated samples and analyzed by high throughput sequencing.

### **ChIP-seq Library Construction**



After a chromatin immunoprecipitation reaction, the sequencing library prepared using NEBNext Ultra II kit (Illumina). The DNA yield of a single ChIP reaction was used as starting material. Library concentration was determined by Qubit (Invitrogen).

### **ChIP-seq Analysis**

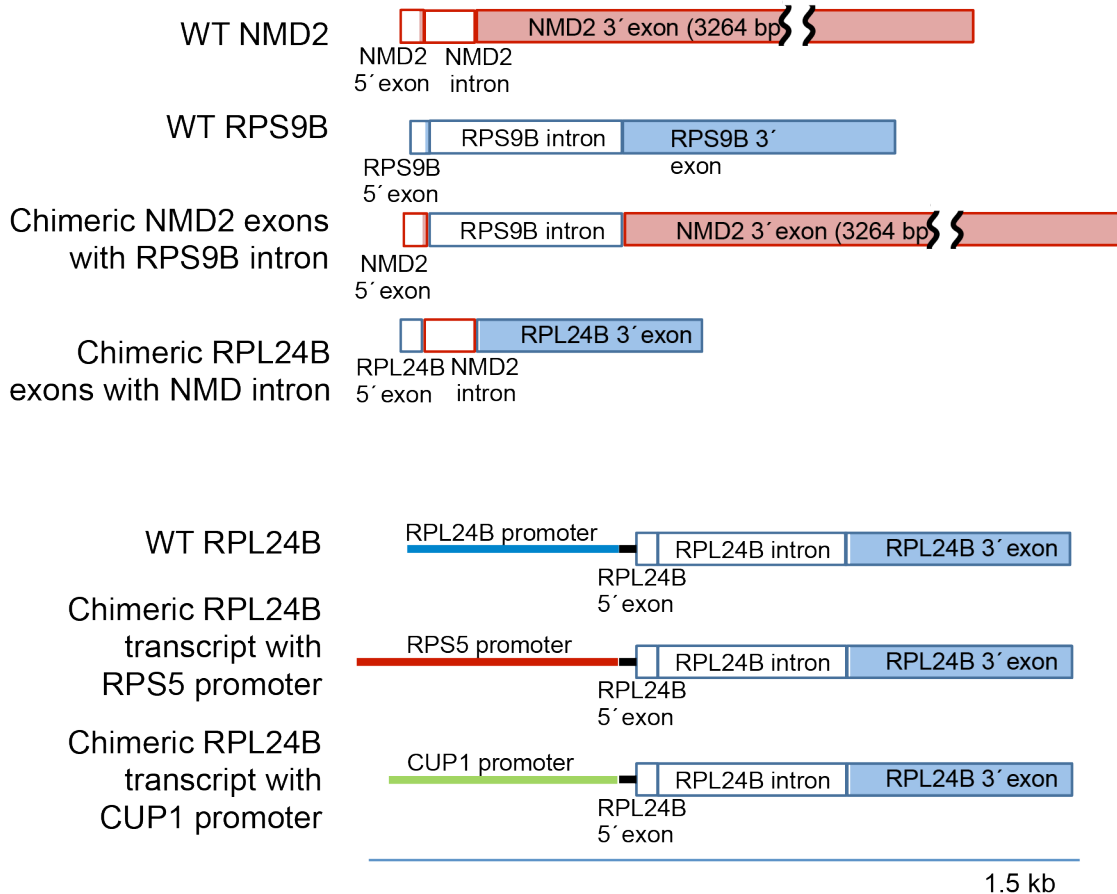
All sequencing was performed by the Center for Advanced Technologies (CAT) at UCSF using the Illumina HiSeq 2500 platform. 10-15 million reads of 50 base pairs (bp) each were collected for each sample. Base-calling and de-multiplexing were performed by the Illumina HiSeq Control Software (HCS). The resulting FASTQ files were concatenated into one file per sample and filtered and trimmed by quality score per base using the Trimmomatic software package with the following parameters: LEADING:27 TRAILING:27 SLIDINGWINDOW:4:20 MINLEN:35 (Bolger et al. 2014). Surviving reads were mapped to the *Saccharomyces cerevisiae* S288C reference genome release R64-2-1\_20150113, obtained via the SGD website (Cherry et al. 2012; Engel et al. 2014) using the Bowtie package with the -n 2 and --best arguments (Langmead et al. 2009). Mapped reads were sorted, indexed and converted to binary format using the samtools package (Van Alst et al. 2009). Two different .gff annotation files were used. First, a .gff file was generated which divided the entire genome into sequential 100bp regions, and the number of reads mapping to each 100bp region was calculated using the easyRNASeq package from the Bioconductor project in R (Delhomme et al. 2012). Additionally, a .gff annotation file was generated in which seven different 100bp regions were defined for each gene containing a single intron, with defined transcription start and polyadenylation (polyA) sites. The list of genes included 89 intron-containing ribosomal protein genes (RPGs) and 175 other intron-containing genes. The regions annotated for each gene were: 1) 50bp upstream and 50bp

downstream of the transcription start site; 2) bp 50-150 of the transcript; 3) first 100bp of the intron; 4) last 100bp of the intron; 5) first 100bp of second exon; 6) 100bp upstream of polyA site; 7) 100bp downstream of polyA site. The number of reads mapping to each region for each gene was calculated using the easyRNASeq package from the Bioconductor project in R (Delhomme et al. 2012). Reads per region for each of these two annotation sets were calculated for each pull-down and also for the corresponding total DNA sample from each culture, and normalized to library sizes. The ratio of reads for the pull-down sample compared to the corresponding total DNA sample was calculated for each region for each sample as a measure of pull-down enrichment. Enrichment values were log<sub>2</sub>-transformed and averaged for the two replicates of each pull-down. The full dataset will be made available through the NCBI GEO repository.

The average enrichment values for each 100bp window across the genome for each factor in each treatment were converted to .wig files and visualized using the MochiView software package (Homann and Johnson 2010). For individual gene traces, the genomic region boundaries surrounding the gene of interest were chosen to extend from the local minimum of Rpb3 pull-down enrichment upstream of the gene (in the minus 3-AT sample) to the local minimum of Rpb3 pull-down enrichment downstream of the gene. The upstream local minimum position was set to 0. The enrichment values were adjusted such that the lowest observed value in the region for each factor was set to 0.

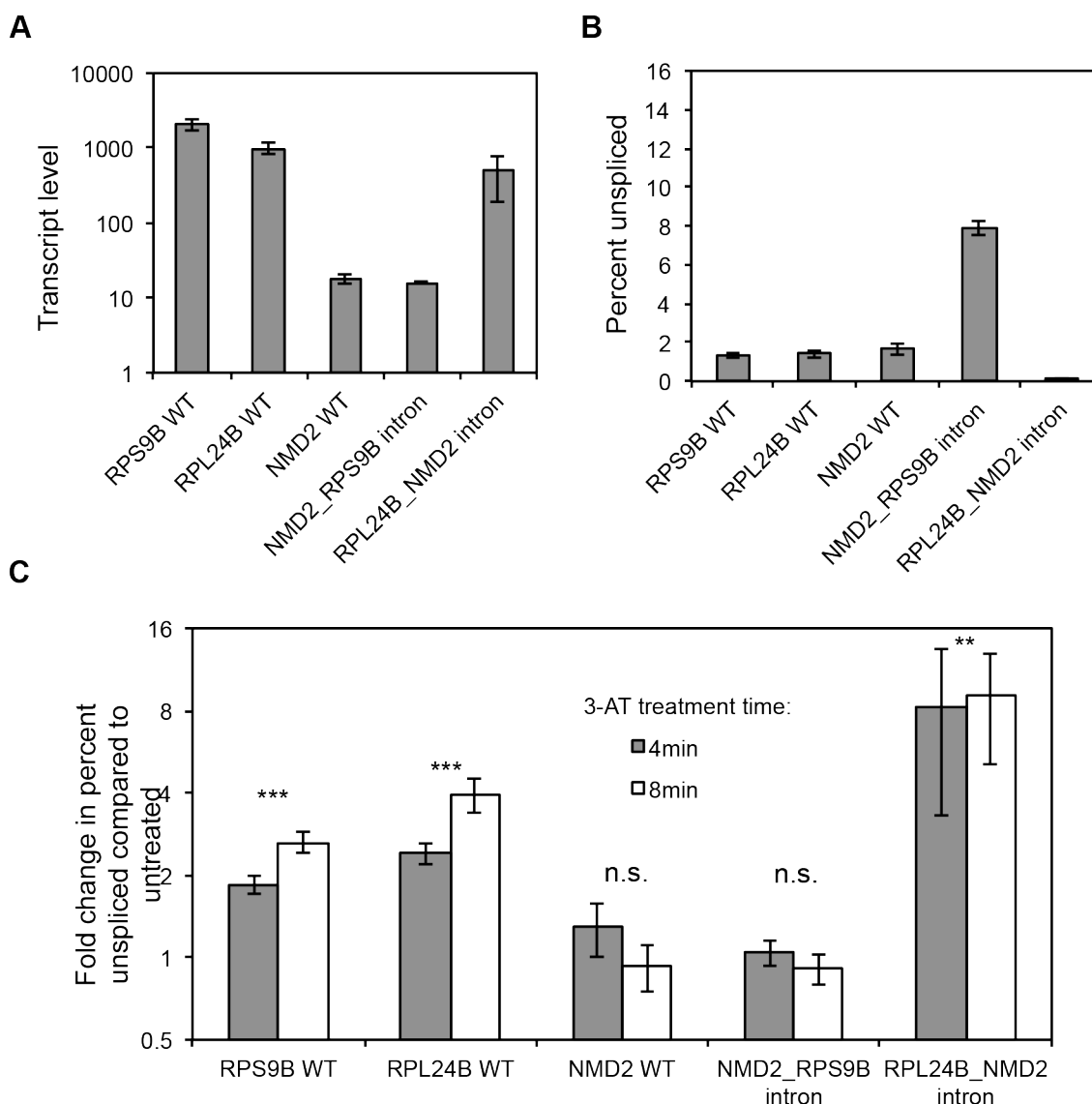
For the aggregated data plots, enrichment values for each factor, treatment, and gene region were averaged across either all RPGs or all non-RPGs in the data set. Error bars report the standard error of the mean.

**A**



**Figure 1: Schematics of gene structures (WT or chimeric) used.**

(A) Schematics of genes involved in chimeric constructs tested in this study. Block regions represent transcribed sequence, while lines represent untranscribed sequences. Colored lines correspond to sequence defined as promoter sequence for the purposes of this study. Shaded block regions represent coding sequence, while white block regions represent untranslated sequence – either part of the 5' UTR or the intron (or both). The 3' untranslated regions are not depicted – the end of the block in each case represents the end of the coding sequence for the transcript. All constructs are integrated at their genomic loci. Promoter swaps are marked by the kanMX drug resistance cassette upstream of the promoter.



**Figure 2: Effects of intron on expression, splicing efficiency, and splicing response to 3-AT.**

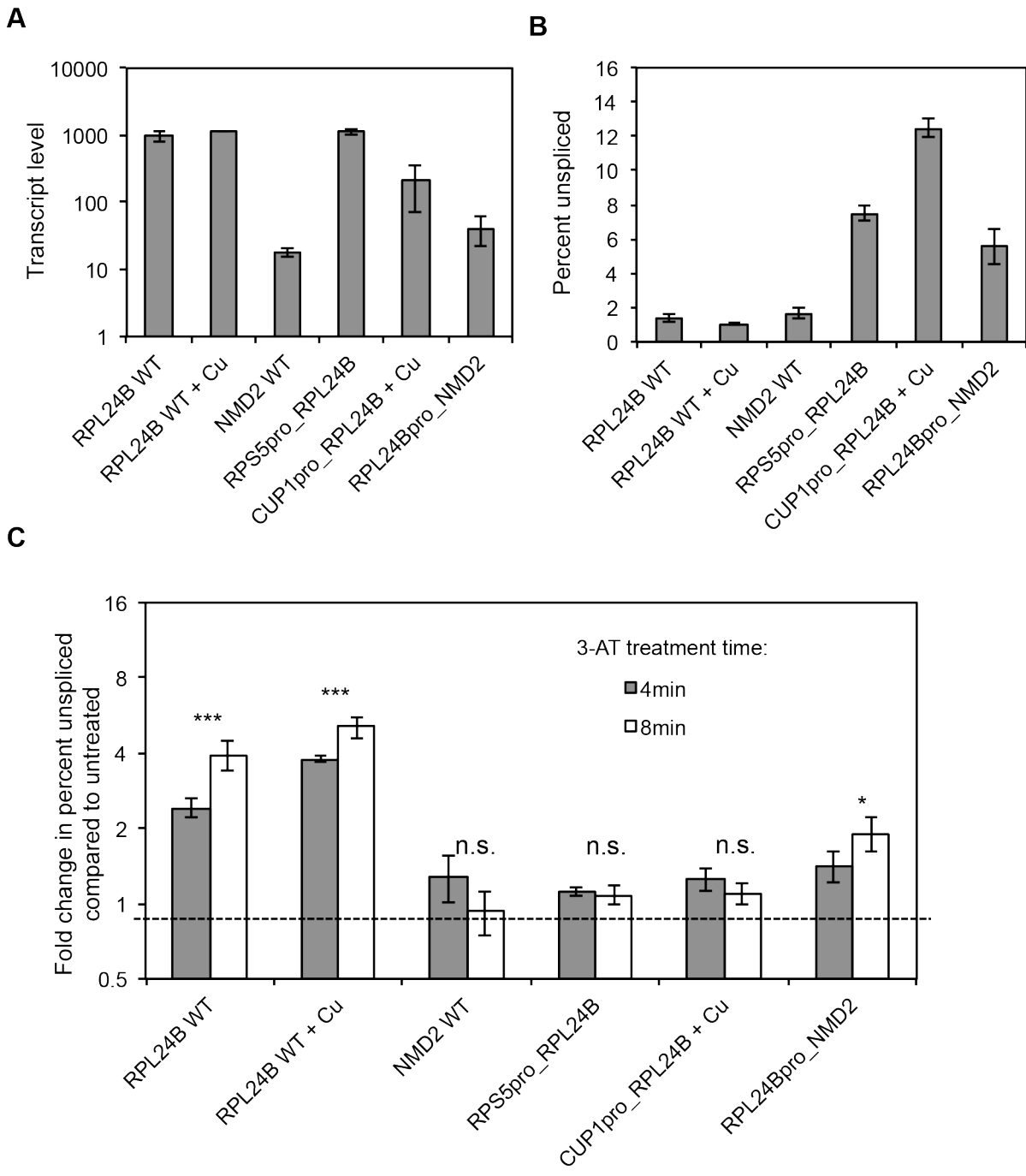
(A) Transcript levels in the absence of 3-AT as measured by primers directed against the second exon. All three “WT” transcripts were measured from the same wild-type culture, while the chimeric transcripts were each measured in their own relevant strain. Note log scale.

(B) Percent unspliced transcript in the absence of 3-AT.

(C) Change in percent unspliced transcript in response to 3-AT at 4 and 8 minutes of treatment.

One-way ANOVA was used to determine whether the change in percent unspliced was

significant: \*\*\*  $p < .001$ ; \*\*  $p < .01$



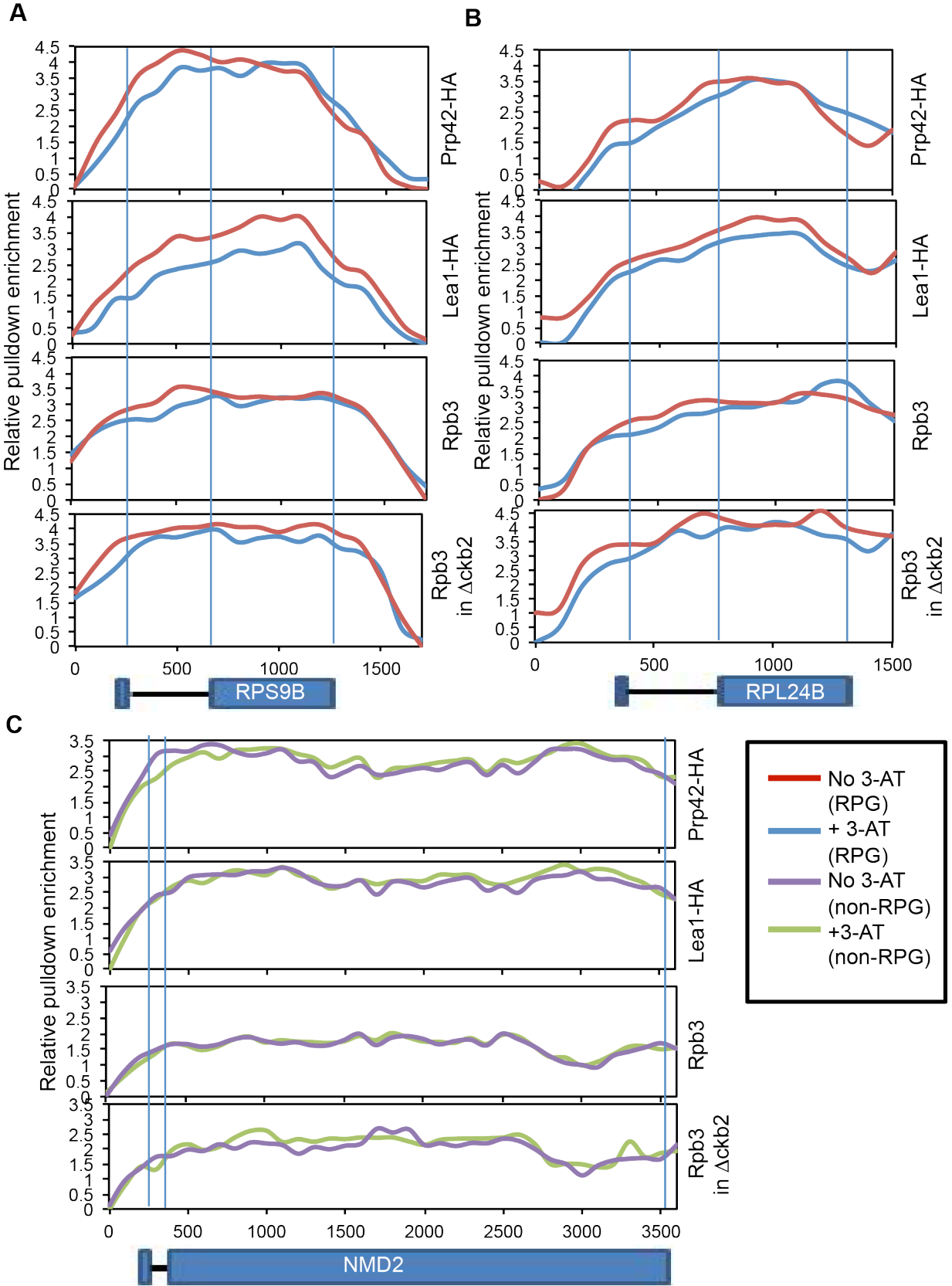
**Figure 3: Effects of promoter on expression, splicing efficiency, and splicing response to 3-AT.**

(A) Transcript levels in the absence of 3-AT as measured by primers directed against the second exon. The RPL24B WT and NMD2 WT transcripts were measured from the same wild-type culture, and are the same data depicted in Figure 2 but are repeated here for comparison purposes. The RPL24B WT + Cu sample is the wild type strain grown with copper sulfate added to the growth media for direct comparison with the CUP1pro\_RPL24B + Cu samples. The chimeric transcripts were each measured in their own relevant strain. Note log scale.

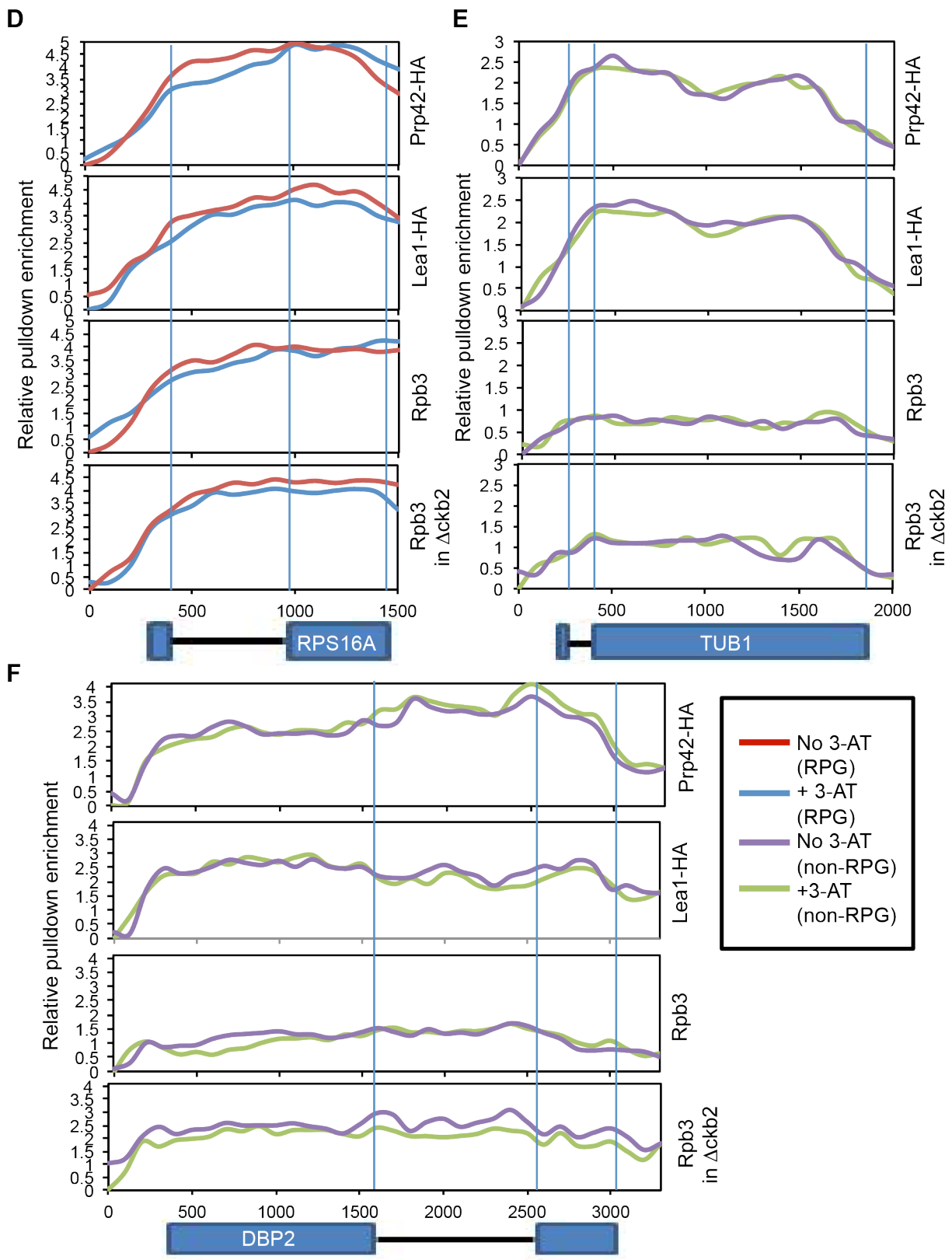
(B) Percent unspliced transcript in the absence of 3-AT.

(C) Change in percent unspliced transcript in response to 3-AT at 4 and 8 minutes of treatment.

One-way ANOVA was used to determine whether the change in percent unspliced was significant: \*\*\*  $p < .001$ ; \*  $p < .05$







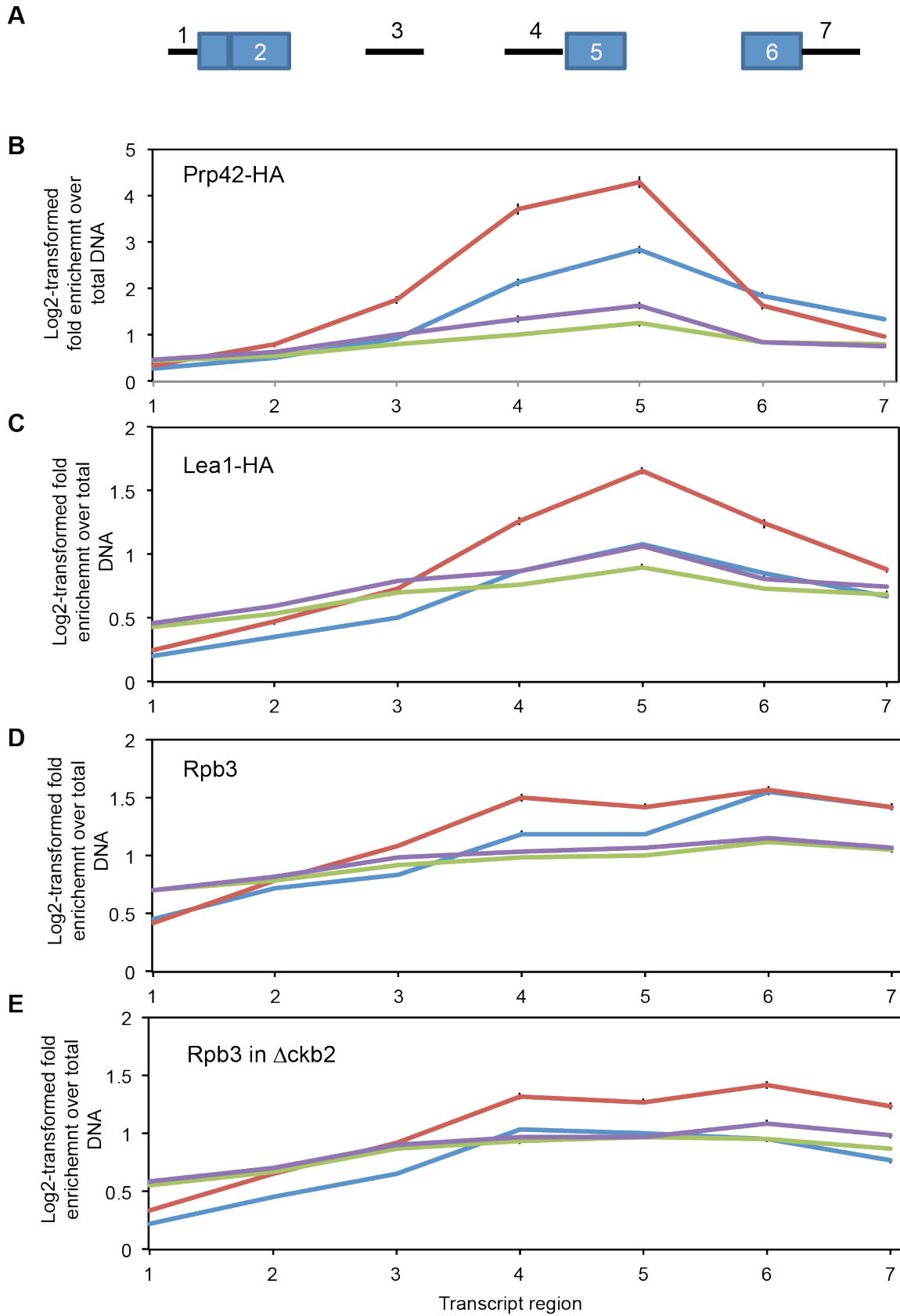
**Figure 4 (two pages): ChIP profiles for U1, U2, RNAP II and RNAP II in a  $\Delta$ ckb2 strain with or without 3-AT treatment.**

(A) Profiles for the genomic region surrounding RPS9B. The locations of the RPS9B exons and intron are indicated below the profiles. The pull-down factor and strain (wild type unless otherwise indicated) are shown to the right of the profiles. Profiles for the minus 3-AT samples are in red, and for the plus 3-AT samples are in blue. Values represent the log<sub>2</sub>-transformed enrichment of the pull-down sample compared to the total input DNA sample, scaled such that the lowest local value is set to 0 (see Materials and Methods for more detail). Thin vertical lines indicate the positions of the start and end of the intron and the end of the second exon.

(B) Profiles for the genomic region surrounding RPL24B, as depicted in (A).

(C) Profiles for the genomic region surrounding NMD2. Minus 3-AT samples are indicated in purple, while plus 3-AT samples are indicated in green, in order to emphasize that this gene is not a ribosomal protein gene, while those in (A) and (B) are. The genes in (A), (B), and (C) are genes that were used in the intron and promoter swap experiments.

(D-F) Profiles for the genomic regions surrounding RPS16A, TUB1, and DBP2 respectively, as described for (A-C).



**Figure 5: Aggregated ChIP-Seq profiles for U1, U2, RNAP II and RNAP II in a  $\Delta$ ckb2 strain with or without 3-AT treatment, comparing intron-containing RPG transcripts with all other intron containing transcripts.**

(A) Schematic of the genomic regions compared in this analysis. For each RPG (n=89) or non-RPG (n=175) intron-containing transcript, pull-down enrichment values for seven different 100bp windows across the transcript were calculated. These windows are: 1) 50bp upstream and 50bp downstream of the transcription start site; 2) bp 50-150 of the transcript; 3) first 100bp of the intron; 4) last 100bp of the intron; 5) first 100bp of second exon; 6) 100bp upstream of polyA site; 7) 100bp downstream of polyA site. For some transcripts, some of these windows will overlap, but this aggregate analysis is intended to capture overall genomic trends, as a complement to the individual gene traces presented in Figure 4.

(B) Aggregate analysis of U1 (Prp42-HA) pull-down enrichment. RPG transcripts in the absence of 3-AT are indicated by red, and in the presence of 3-AT by blue. Non-RPG transcripts in the absence of 3-AT are indicated by purple, and in the presence of 3-AT are indicated by green. Error bars indicate the standard error of the mean across all of the genes included in the category. For RPG transcripts, the peak of U1 signal appears to be overall somewhat lower in the presence of 3-AT, as well as later in the gene sequence. For non-RPGs, levels are perhaps slightly lower but the position of the peak relative to the gene sequence is not affected by 3-AT.

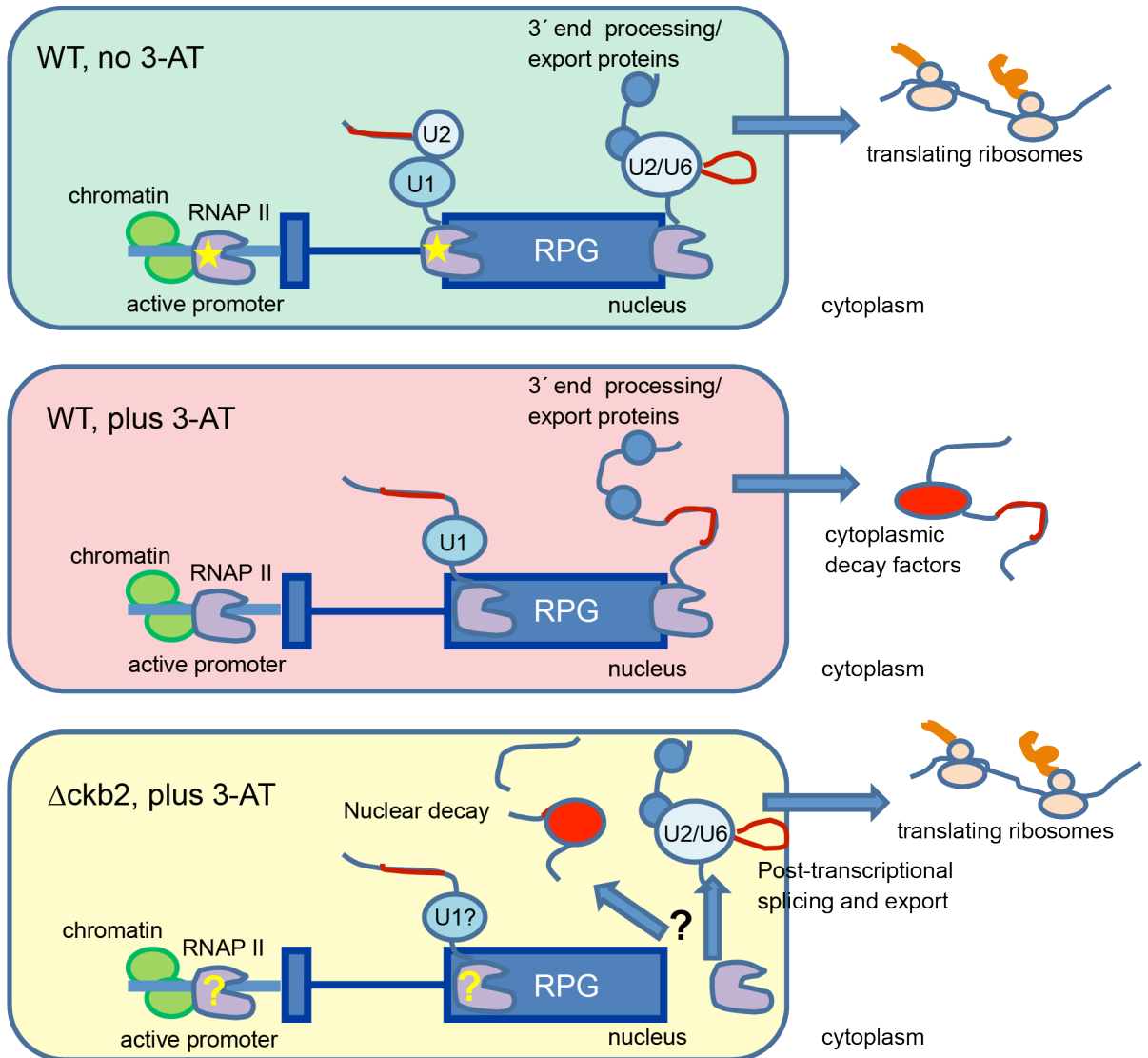
(C) Aggregate analysis of U2 (Lea1-HA) pull-down enrichment, as described for (B). Levels of U2 recruitment are slightly lower for non-RPGs and much lower for RPGs, but the peak of U2 signal is not shifted in position.

(D) Aggregate analysis of RNAP II (Rpb3) pull-down enrichment, as described for (B). For RPGs, RNAP II signal exhibits a small peak at the end of the intron and also at the end of the

transcript. Compared to the no 3-AT condition, addition of 3-AT diminished signal at the end of the intron, but not at the end of the transcript. Little effect of 3-AT was observed for non-RPGs.

(E) Aggregate analysis of RNA RNAP II (Rpb3) pull-down enrichment, in a  $\Delta$ ckb2 strain, as described for (B). Compared to the WT strain, this strain failed to show accumulation of RNAP II signal at the end of RPG transcripts in the presence of 3-AT.

A



**Figure 6: Model.**

(A) In a wild-type cell growing exponentially without stress, ribosomal protein genes are very highly transcribed and efficiently spliced. Their high transcription levels depend in part on modification to the chromatin in their promoter regions and recruitment to their promoters of strong transcriptional activators. Our results indicate that some promoter element may also help facilitate the efficient splicing of these genes, perhaps by recruiting an unknown factor that either

binds or post-translationally modifies RNAP II to affect its behavior. Polymerase may pause near the 3' end of the intron to ensure that splicing factors can be recruited and function. Splicing is largely complete by the time RNAP II reaches the end of the transcript, and it may pause again to allow for recruitment and assembly of export factors, which then facilitate export of the transcript to the cytoplasm and efficient translation. When 3-AT is added, cells must transiently repress ribosome biogenesis in order to re-direct resources toward amino acid biosynthesis. However, it may be more efficient for them to maintain the active promoter architecture and instead prevent splicing of RPG transcripts. The putative promoter-facilitated modification that enhances splicing efficiency is no longer made, and polymerase fails to pause at the 3' end of the intron. U1 and U2 are not recruited as efficiently, and more transcripts remain unspliced when RNAP II reaches the ends of genes. RNAP II still pauses at the ends of genes, and the unspliced transcripts are still packaged and exported but fail to be translated and are instead targeted for cytoplasmic decay due to their retained introns. In the absence of CKB2, polymerase fails to pause at the 3' end of RPG genes, possibly indicating a failure to package unspliced transcripts for export. These unspliced RPGs may then be spliced post-transcriptionally or rapidly degraded by the nuclear surveillance machinery, as pre-mRNA accumulation is not detectable in the  $\Delta ckb2$  strain.

ID	Strain	Genotype	Notes	Source
yMB1	wildtype	BY4742 MATalpha HIS3 leu2Δ0 lys2Δ0 ura3Δ0		Bergkessel et al. 2011
yMB11	Δckb2	Δckb2::NAT MATalpha	A	Fiedler et al. 2009
yMB9	RPL24Bpromoter_NMD2	RPL24Bpromoter_NMD2::KAN MATalpha HIS3 ade2::Gal_Isce1	B	This Study
yMB10	RPL24B_NMD2intron	RPL24B_NMD2intron::KAN MATalpha HIS3ade2::Gal_Isce1	C	This Study
yMB7	CUP1promoter_RPL24B	CUP1promoter_RPL24B::KAN MATalpha HIS3ade2::Gal_Isce1		Bergkessel, M. 2010 thesis
yMB6	RPS5promoter_RPL24B	RPS5promoter_RPL24B::KAN MATalpha HIS3 ade2::Gal_Isce1		Bergkessel, M. 2010 thesis
yMB4	NMD2_RPS9Bintron	NMD2_RPS9Bintron::KAN MATa HIS3 leu2Δ0 LYS2 ura3Δ0 ade2::Gal_Isce1	C	Bergkessel, M. 2010 thesis
yJCG021	Prp42-HA	yMB1; PRP42-HA::KAN MATa HIS3 leu2Δ0 LYS2 ura3Δ0 met15Δ0	D	This Study
yJCG035	Lea1-HA	yMB1; LEA1-HA::KAN MATa HIS3 leu2Δ0 LYS2 ura3Δ0 met15Δ0	D	This Study

Notes:

- A Fiedler D, Braberg H, Mehta M, Chechik G, Cagney G et al. (2009) Functional organization of the *S. cerevisiae* phosphorylation network. *Cell* 136: 952-963.
- B Promoter sequence was amplified by PCR and integrated upstream of locus by homologous recombination.
- C Intron swaps were performed using a modification of the "*Delitto Perfetto*" method (Storici et al. 2006).
- D Made by homologous recombination of a 3X-HA tag followed by the kanMX6 marker immediately downstream of the transcript stop codon.

## Table 1: Strain List

Detailed list of all strains described in this paper.



## CONCLUSION

To date, mechanisms of pre-mRNA splicing have largely been explored *in vitro*, but over my time in the Guthrie lab, the splicing field has grown tremendously in exploring *in vivo* considerations for how the cellular and external environments affect splicing outcomes. In this body of work, my colleagues and I explored ways in which splicing can be transcript-, gene-, and environmentally- specific. We showed that not only the local, co-transcriptional chromatin landscape (Chapter 1), but also the growth conditions outside the cell that signal gene expression changes (Chapter 2) influence spliceosome assembly and catalysis. Taken together, these studies reveal an under-appreciated role for the promoter region in the regulation of splice site selection and splicing catalysis. Furthermore, our work highlights a growing view that each splicing outcome is context specific, both by taking into account the state of the gene at the time of expression and integrating with a variety of nuclear regulatory input signals.

Such variability in splicing efficiency and regulatory capacity represents a new paradigm in the intricacies of eukaryotic gene expression and underscores the challenges that lie ahead in future studies of co-transcriptional splicing mechanisms. When I began my investigations, much attention was being given to a now well-accepted hypothesis that the chromatin state on a gene could regulate splicing decisions through directing RNA polymerase II dynamics (or kinetics) and/or modulating splicing factor recruitment (Merkhofer et al. 2014). Since this time, our lab and many others have collected evidence to support this hypothesis, but the exact molecular mechanisms underlying the coupled interactions between splicing, RNAP II and the chromatin landscape remain unclear. Creating a full picture of the interactions between splicing and other gene expression processes is a monumental task that I anticipate will be slowly accomplished

through future characterizations of individual interactions and the development of high-resolution *in vivo* and *in vitro* techniques.

### **Co-transcriptional Splicing Regulation Harnesses the Promoter**

We examined two different aspects of co-transcriptional splicing regulation using two different yeast systems, and serendipitously both studies yielded insights into the role of the promoter region. In Chapter 1 we interrogated the molecular connection underlying a genetic interaction between the spliceosome and a chromatin remodeling complex known for localizing around the promoters of a subset of genes in fission yeast. We showed that the histone variant H2A.Z localized around promoters of a subset of genes that are lowly expressed and enriched for promoter proximal introns with non-consensus splice sites. The statistically significant bias of H2A.Z containing promoters to coincide with weak introns suggests reciprocal enforcement of H2A.Z deposition and splice site degeneracy over evolutionary time. Perhaps having a promoter region decorated with H2A.Z allows splice sites in proximal introns to drift. Alternatively, perhaps introns in genes with lowly expressing promoters co-opted H2A.Z as a signal for enhancing splicing factor recruitment through direct interactions or indirectly through modulating RNAP II dynamics. Overall, this study revealed a novel role for promoter identity in splicing fidelity and possibly the evolution of splice site strength.

In Chapter 2, we interrogated an uncharacterized mechanism of regulated pre-mRNA accumulation in response to amino acid starvation stress in budding yeast. We found that a promoter sequence from an intron-containing ribosomal protein gene, a class of genes over-represented with introns, was both necessary and sufficient to impart a splicing response to

amino acid starvation stress to intron-containing genes. Through these experiments we observed that it was not simply promoter strength in driving transcript expression, but a particular, yet to be defined, sequence in the promoter that influences splicing in rapid response to extracellular cues, on a transcript-specific basis. Just as promoter sequences can recruit different transcription factors, so too does promoter identity influence splicing through coupling of RNAP II and transcription factors to the chromatin landscape. Our work highlights two ways in which the role of the promoter influences pre-mRNA splicing to be both gene-, and even transcript-specific, allowing for splicing outcome differences each round of transcription, and thereby broadening the role of the promoter in gene expression and revealing the importance of environmental context in each splicing decision.

### **Technical Challenges of Studying pre-mRNA splicing in *S. pombe***

Branching out into *S. pombe* afforded the Guthrie lab new experimental opportunities for understanding a more mammalian context for pre-mRNA splicing, but presented steep technical challenges. A large portion of my time was committed to developing methods to overcome technical challenges posed by molecular and genetic experiments in fission yeast, primarily in collaboration with a post-doctoral fellow, Kristin Patrick. In the Appendix, I describe a hard-fought method for chromatin immunoprecipitation (ChIP) of splicing factors that was optimized for *S. pombe*. While I ultimately overcame this technical hurdle, and utilized the method for two collaborative manuscripts ((Patrick et al. 2015) and the forthcoming manuscript from Chapter 2), I was unable to apply splicing factor ChIP to answer questions for my main project regarding H2A.Z and splicing in Chapter 1. Fortunately, our coordinated investigations with Tracy Johnson's group at UCLA allowed splicing factor ChIP to be applied to questions of H2A.Z's

role in spliceosome assembly, revealing defects in early spliceosome recruitment and/or catalysis, consistent with our findings (Neves et al. 2017).

Additionally, working with an essential and complex macromolecular machine such as the spliceosome presented its own unique set of challenges in *S. pombe*. Fission yeast are highly reliant on splicing for gene expression in that over half of the genes in the genome contain introns (Wood et al. 2012). A strong need for efficient splicing has been balanced over evolutionary time with the ability to recognize and splice non-consensus splice sites, limited our ability to mutate the snRNAs. Both Dr. Patrick and I attempted to mutate U2 and U1, respectively, multiple times, and our strains would not tolerate any snRNA mutations that were predicted to either weaken or strengthen interactions between the spliceosome and the consensus splice site sequence (data not shown). Additionally, the vast majority of the double mutant strains that we created between a splicing factor and a chromatin factor were too sick to work with, limiting our ability to ask questions about specific genetic interactions between splicing and chromatin (data not shown). Unlike *S. cerevisiae*, the community of fission yeast researchers does not have as extensive a collection of point mutants in splicing factors, limited existing genetic studies, and supporting the rationale for creating a DAmP allele collection of all *S. pombe* splicing factors (Patrick et al. 2015). Finally, many of the DAmP alleles of splicing factors that Dr. Patrick created, which were thought to decrease expression, actually resulted in increased transcript ((Lund et al. 2008; Patrick et al. 2015), Chapter 1, additional data not shown). These overexpression alleles have become a new, if unexpected, tool for future genetic studies of pre-mRNA splicing in *S. pombe*. Also, a significant number of the DAmP alleles did not significantly change expression levels, particularly for splicing factors that are central to the spliceosome (data not shown). It is important to keep in mind that a lack of a significant change

in transcription does not rule out a change in protein levels but there is no way high-throughput method to monitor protein levels from the DAmP alleles beyond native immunoprecipitation, as adding a protein tag essentially recreates a DAmP allele. Taken together, the essential nature and massive, multi-protein composition of the spliceosome make it a more challenging pathway to study genetically in fission yeast.

An early question we faced as a sub-group of *S. pombe* researchers in the Guthrie lab was deciding amongst genome-wide techniques to characterize changes in pre-mRNA splicing. At the time, splicing-specific microarrays were in common use in the *S. cerevisiae* splicing field, having been pioneered in part by Jeff Pleiss, when he was a post-doctoral fellow in the Guthrie lab (Pleiss et al. 2007b). Dr. Pleiss gifted access to a splicing-specific microarray that he designed for *S. pombe* to us, and I set about designing computational analysis in order to utilize the microarrays. Concurrently, RNA-seq was becoming a standard technique for characterizing the transcriptome, in part because it offered advantages to look at a variety of RNA species (spliced, unspliced, partially spliced, the splicing states of two consecutive introns, etc.), while splicing specific microarrays directly compared on type of RNA species between two biological samples. Ultimately, we wanted to make a direct comparison between wild-type and a variety of mutants, and this was easily achieved with the splicing-specific microarrays. However, while RNA-seq did not limit characterization to pre-determined short sequences, it was disadvantageous for *S. cerevisiae* because of the small number of introns in the budding yeast genome would require extreme sequencing depth to generate enough observations of unspliced pre-mRNA for statistically meaningful conclusions about splicing (personal communication with Christina Homer). *S. pombe* has at least 5000 introns compared to *S. cerevisiae*'s ~250 introns (Spingola et al. 1999; Wood et al. 2012). Given the cost to optimize the necessary sequencing

depth for *S. pombe* at the time, we were not confident in RNA-seq's ability to provide superior insights above and beyond the ~5000 splicing events measured by the splicing-specific microarray platform. In addition, Dr. Pleiss included alternative splicing events that his research group computationally predicted, but that had not been observed *in vivo*. Given the sequencing depth required to study alternative isoforms in *S. pombe*, which were predicted to be a minor species, the microarray platform offered greater sensitivity for monitoring changes in these alternative splicing events at the same cost. Thus, we decided to use splicing specific microarrays to characterize genome-wide splicing changes between two biological samples in *S. pombe*. In work not shown here, but fundamental for driving forward publications by Kristin Patrick (Patrick et al. 2015) and Michael Marvin (Lipp et al. 2015), I created normalization and analysis algorithms that were used early on in the microarray analysis pipeline. It is my sincerest hope that splicing-specific microarrays will continue to be a tool for directly comparing splicing outcomes between organisms with limited numbers of introns.

A large technical challenge for continuing to use *S. pombe* to study splicing is lack of an *in vitro* splicing system. Several other laboratories have attempted to create an *in vitro* splicing assay in *S. pombe*, but to date the furthest step in the splicing pathway able to be monitored is U2 snRNP association (Huang et al. 2002). A collaborative project to develop an *in vitro* splicing assay for *S. pombe* was initiated within the Guthrie lab, using my splicing-specific microarray data and analysis as a basis for transcript selection, but the project was never completed.

Finally, the biggest overall technical challenge for the field of co-transcriptional splicing regulation is the lack of an *in vitro* co-transcriptional splicing system to test mechanistic predictions, agnostic of model organism system. Several *in vitro* co-transcriptional splicing assays have been described at conferences, but to my knowledge, none have yet been published.

Despite these steep technical challenges, the power of *S. pombe* to ask questions of co-transcriptional splicing regulation in a more mammalian-like context, coupled with the awesome power of yeast genetics, and in concert with other labs developing both techniques and resources, presents ample opportunities for impactful future inquiries.

## **Future Directions**

In Chapter 1, we showed that the highly conserved histone variant H2A.Z influences splicing efficiency of promoter-proximal introns with weak *cis*-splicing sequences in *S. pombe*. However, the exact molecular mechanism of how H2A.Z participates in splicing remains unclear. Work from Tracy Johnson's lab in *S. cerevisiae*, published in coordination with Chapter 1, reveal changes in RNAP II ChIP accumulation profiles that correlate with the H2A.Z dependent splicing defect, suggesting that H2A.Z indirectly influences splicing through changes to initiation or elongation dynamics (Neves et al. 2017). It is not yet known if H2A.Z's modulation of RNAP II initiation and elongation dynamics is sufficient to enhance splicing of weak introns. Moreover, we do not have the tools to understand whether this change in RNAP II ChIP signal reflects a stall or a pausing checkpoint for either the polymerase or the spliceosome. Ultimately, all of these questions will eventually require the development of an *in vitro* co-transcriptional splicing assay or higher resolution, real-time *in vivo* techniques for monitoring chromatin, RNAP II and splicing together.

To further probe the mechanism of H2A.Z's role in splicing, I began a set of experiments to probe protein-protein and protein-RNA between the Swr1 complex and the spliceosome using mass spectrometry. While mass spectrometry data sets for H2A.Z and the Swr1 complex in several organisms exists, the sensitivity of the technique has improved greatly over my tenure on

the Guthrie lab, Indeed, these early studies camouflage weak, sub-complex level interactions between splicing factors and Swr1 complex factors in *S. pombe* and *S. cerevisiae* (Shevchenko et al. 2008; Kim et al. 2009). Human studies have also shown immunoprecipitation between splicing factors and H2A.Z (Tolstorukov et al. 2012). By using crosslinking followed by mass spec and *in vitro* binding studies, it may be possible to understand whether the Swr1 complex and H2A.Z recruit, or are recruited by, splicing factors.

In this body of work, we explored only a single node of interactions from the *S. pombe* splicing EMAP, between the spliceosome and one genetically interacting complex, the Swr1 complex (Patrick et al. 2015). Not only does this powerful genetic screen contain abundant unexplored interactions that are negative between splicing and chromatin, as well as between splicing and other cellular pathways, but also no study to date has explored positive genetic interactions from this screen. Moreover, because many of the splicing factors in the *S. pombe* EMAP are perturbed as DAmP alleles, I believe that a reanalysis of the EMAP could be prudent to account for our observation that many of the splicing factor DAmP alleles are over-expressed ((Patrick et al. 2015), Chapter 1). Lastly, there are also many more genetic interactions between the spliceosome and a variety of cellular pathways to be explored in the *S. cerevisiae* EMAP with or without comparison to *S. pombe* (Wilmes et al. 2008).

Finally, in the vein of thinking about coupling between gene expression processes, I am enthralled by the possibility of H2A.Z deposition or positioning being influenced by splicing. Indeed, I created several *S. pombe* strains to ChIP H2A.Z in a splicing factor defective background to see if H2A.Z deposition and localization depended on a functional spliceosome. While I was unable to complete experiments to test the hypothesis of coupling between H2A.Z and the catalytic activity of the spliceosome, in the time since I proposed to address that



question, work by other groups has provided supporting evidence for cross-talk between splicing and chromatin (de Almeida and Carmo-Fonseca 2014). For example, splice site mutations affect the tri-methylation of histone H3 lysine 36 in humans and yeast (de Almeida et al. 2011; Kim et al. 2011). Additionally, splicing can direct histone H3 lysine 4 tri-methylation in yeast (Bieberstein et al. 2012). I believe that goal of characterizing reciprocal interactions between splicing and other gene expression processes remains important for unraveling the interconnectedness of nuclear pathway relationships.

## **APPENDIX**

In this appendix, I present additional lines of experimental inquiry that were not included in the published manuscript of Chapter 1.

**Author Contributions:**

Kelly Nissen: conceived experiments, made strains and reagents, designed and performed experiments, analyzed data, wrote and edited the draft

Kristin Patrick: conceived experiments, made strains and reagents, designed and performed experiments

Christine Guthrie: conceived experiments, supervised the research, wrote and edited the draft

## **Part A: What is the genetic relationship between *swr1* and *pht1* in influencing pre-mRNA splicing?**

### **Background:**

Jasper Rine's group first hypothesized that in the absence of H2A.Z, the Swr1 complex could still perform first half of its two-step histone variant replacement mechanism, the H2A-H2B nucleosome dimer eviction step, and that this partial mechanism would disrupt promoter function by compromising the chromatin organization (Halley et al. 2010). Thus, the phenotype of cells lacking H2A.Z may represent two defects: the lack of H2A.Z's function and the disruption of nucleosomes by the Swr1 complex. It follows that in the absence of H2A.Z removing a functional Swr1 complex could suppress a subset of phenotypes. Indeed, suppression of temperature and cold-sensitive defects, as well as gene expression and genome instability defects, have been observed in *S. cerevisiae* in strains lacking HTZ1 (H2A.Z) and SWR1 (Swr1 ATPase) (Halley et al. 2010; Morillo-Huesca et al. 2010). In contrast, in the absence of Swr1 in *S. cerevisiae*, Htz1 incorporation is markedly reduced but not fully eliminated (Krogan et al. 2003; Kobor et al. 2004; Mizuguchi et al. 2004). In order to gain insight into the distinct roles of *pht1* and *swr1* in pre-mRNA splicing in *S. pombe*, I looked for epistatic or suppressive interactions between these two alleles by creating a double mutant *pht1* $\Delta$  *swr1* $\Delta$  strain and characterizing the phenotype in growth and splicing assays relative to the single mutants.

### **Experimental Approach:**

#### **Growth Assays**

Strains were grown in YES5 overnight at 30°C and then diluted to 0.1 OD<sub>600</sub> the next morning to allow for growth to mid log-phase. Mid-log phase cultures were re-diluted to 0.1 OD<sub>600</sub> and serially diluted 1:5 across a 96-well plate. Approximately 5µl of cultures were transferred to YES5 plates using a frogging tool and plates were grown at 30°C for 2 days, 37°C for 2 days or 16°C for 9 days. Images were then collected.

### **RNA Isolation and qRT-PCR**

Strains were grown in YES5 overnight at 30°C and then diluted to 0.1 OD<sub>600</sub> the next morning to allow for growth to mid log-phase. Temperature shifts occurred when strains had reached 0.3-0.5 OD<sub>600</sub>. 37°C shifts took place for 2 hours and 16°C shifts for 9 hours. Cultures were harvested by centrifugation between 0.5–0.8 OD<sub>600</sub>. Pellets were washed once in water and then flash-frozen in liquid nitrogen and stored at -80°C. RNA was isolated using hot acid phenol followed by isopropanol precipitation as previously described (Bergkessel et al. 2011).

For RT-qPCR, RNA was treated with DNase I (Fermentas) and then MMLV1- RT (Promega) with dN9 primers (0.5mg/mL) to generate cDNA. Remaining RNA was then removed from cDNA by NaOH treatment and then the cDNA was purified using a spin kit (Zymo). qPCR was performed using a CFX96 Real Time system (BioRad).

### **Splicing-specific Microarray Data Collection and Feature Analysis**

Collection and analysis of splicing-specific microarray data was performed as previously described (Lipp et al. 2015; Patrick et al. 2015).

### **Results and Discussion:**

First, I generated a double mutant strain of *pht1Δ swr1Δ* by crossing the two single mutants *pht1Δ* and *swr1Δ* together and dissecting tetrads. I collected three different double mutant isolates (1A, 3C, 4D) from different tetrad tetrads to act as biological replicates.

In order to characterize the genetic interaction between the two deletion alleles, I performed serial dilution growth assays at 16°C, 30°C, and 37°C and compared the single mutants to the double mutants and a wild-type strain (Figure A.1 Panel A). As predicted from similar experiments in *S. cerevisiae*, the *pht1Δ swr1Δ* double mutant strain suppresses both the growth defects of *pht1Δ* at 16°C and of *swr1Δ* at 37°C seen in Chapter 1 (Halley et al. 2010). To note, in *S. cerevisiae*, *htlzΔ* is heat sensitive and *swr1Δ* is cold sensitive, and the basis for this opposite phenotype in *S. pombe* (that *pht1Δ* is cold sensitive and *swr1Δ* is heat sensitive) is unknown. The suppression of the single mutant growth defects by the double mutant suggests that each single mutant has a complex phenotype, made up in part of the lack of function by the missing gene product and an aberrant function of the remaining gene product.

Next, in order to assess how *pht1Δ swr1Δ* influenced splicing efficiency, I performed RT-qPCR on two loci that I determined previously have an intron accumulation splicing defect at 16°C in *pht1Δ* as an initial screen (Figure A.1 Panel B). In this experiment, I compared three biological isolates of *pht1Δ swr1Δ* to single replicates of *pht1Δ* and *swr1Δ*, for which I had already observed a defect at 16°C for *pht1Δ* only, and no significant splicing defect at 30°C (data not shown). For both loci, Spbp22h7.06 and Spac12g12.12, the *pht1Δ swr1Δ* double mutants partially suppressed the 16°C splicing defect of *pht1Δ*. At 30°C, the double mutants behave similarly to both *pht1Δ* and *swr1Δ*, which do not show a splicing efficiency difference compared to wild-type (data not shown). Because this experiment was intended to be a quick screen for changes in splicing efficiency using two reporter genes, I did not test 37°C with RT-

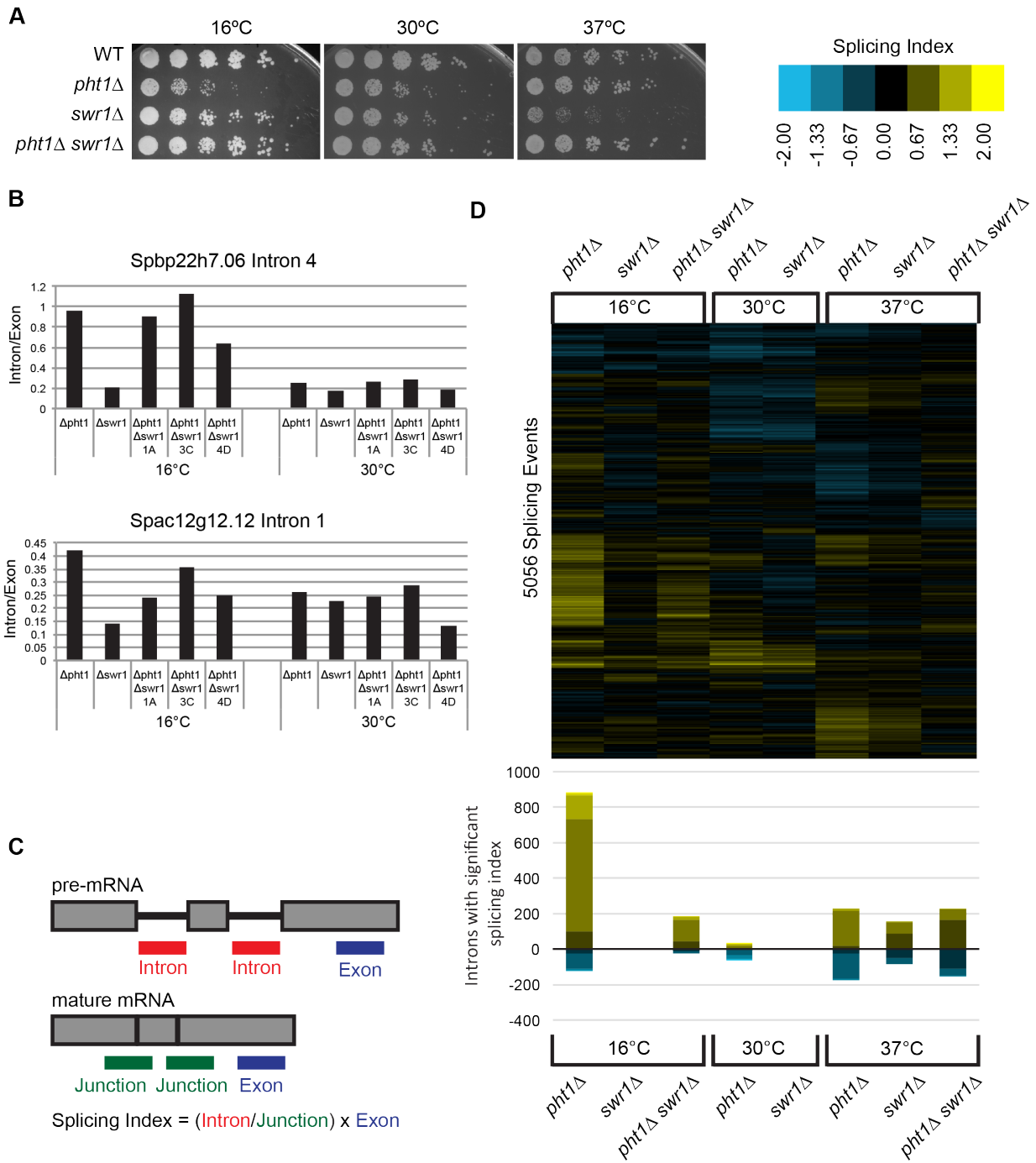
qPCR. The partial suppression of the *pht1Δ* splicing defect at 16°C by the *pht1Δ swr1Δ* double mutants is consistent with a model whereby some portion of the splicing efficiency change in *pht1Δ* is due to aberrant Swr1 complex function and some is due to loss of H2A.Z function.

Finally, given the partial suppression of the splicing defect in *pht1Δ* at 16°C on two reporter loci, I turned to splicing-specific microarrays to characterize the genome wide splicing profile of the *pht1Δ swr1Δ* double mutants at 16°C and 37°C (Figure A.1 Panels C and D). Previously, I had observed a significant splicing defect at 16°C in *pht1Δ* across a large number of introns, and a few introns significantly affected in *swr1Δ* at 37°C (Chapter 1) that correlated with the growth assay results. Additionally, at 30°C, there are very few introns with significant changes in their splicing index in either *pht1Δ* or *swr1Δ* strains, correlating with a relatively minor growth defects in Figure A1 Panel A, and these splicing profiles are shown for comparison. At 37°C, where *swr1Δ* is temperature sensitive, the single *pht1Δ* mutant mimics the *swr1Δ* splicing profile, but has stronger defects at the same introns affected, indicating that the introns affected are sensitive Swr1 complex deposition of H2A.Z. At 37°C, the double mutant *pht1Δ swr1Δ* the splicing defect in the introns affected in the single mutants is mostly suppressed, and now a variety of other introns are mildly splicing defective. At 16°C, the double mutant *pht1Δ swr1Δ* shows a partial suppression of the *pht1Δ* genome wide splicing profile, in which many introns are still affected in their splicing efficiency, but to a less extreme extent at each intron. This partial suppression suggests there is a role for H2A.Z in splicing that is partially independent of possible disruption of chromatin by Swr1 complex in its absence and that the splicing defect at 16°C of *pht1Δ* is not an indirect affect.

Taken together, the double mutant *pht1Δ swr1Δ* suppresses the growth defects partially suppresses the splicing defects of the single *pht1Δ* and *swr1Δ* mutants. That the growth defect

was fully suppressed by the double mutant strain at both heat and cold stress, but that the splicing defect was only partially suppressed indicates that splicing efficiency changes are not solely responsible for growth defects. These results are consistent with the mechanism proposed by the Rine group whereby in the absence of Pht1, the Swr1 complex still binds to the chromatin and evicts the canonical H2A-H2B dimer without an H2A.Z-H2B dimer to replace it with (Halley et al. 2010). This results in unstable chromatin, and because of H2A.Z's localization to promoters, disruption in transcription and splicing for a large number of genes. In contrast, when Swr1 is deleted, the Swr1 complex never forms and H2A.Z has a reduced capacity to incorporate in the chromatin, leading to aberrant incorporation sites and in turn, co-transcriptional splicing defects. This proposed *swr1Δ* scenario results in a different transcription and splicing defect than in the *pht1Δ*, which is reflected in different splicing defect severities and opposite temperature sensitivities. When the two mutations are combined in a double mutant strain, the Swr1 complex never forms which relieves the defect in chromatin structure caused by partial nucleosome remodeling, but the loss of H2A.Z function at its dependent loci persists. Thus, in the 16°C microarray result, the introns affected by the *pht1Δ swr1Δ* mutant reflect the true splicing defect when all H2A.Z and Swr1 complex function is removed, confirming a role in splicing is based on the function of H2A.Z and not an indirect consequence of the loss of chromatin integrity in its absence. More biochemical work is needed to fully unravel the molecular basis of this fascinating reciprocal suppression and test this model.





**Figure A.1: H2A.Z-dependent splicing defect partially suppressed in *pht1*Δ *swr1*Δ double mutant strain**

(A) Serial dilutions (1:5) of WT, *pht1*Δ, *swr1*Δ, and double mutant *pht1*Δ *swr1*Δ strains, grown at 16°C, 30°C, and 37°C.

(B) RT-qPCR validation of *pht1Δ* splicing defect suppression on two introns in three isolates of *pht1Δ swr1Δ* double mutant strain compared to *pht1Δ* and *swr1Δ* single mutant strains at 16°C and 30°C. Cultures were grown to mid-log phase at 30°C and shifted to 16°C for 9 hours. Bars represent ratios of intronic to exonic RNA signals.

(C) Cartoon depicting splicing-specific microarray probes.

(D) Splicing profile of *pht1Δ*, *swr1Δ* and *pht1Δ swr1Δ* strains at 16°C, 30°C and 37°C. cDNA from each strain listed was competitively hybridized on the splicing-specific microarray against that from WT. A Splicing Index value was calculated for each intron by normalizing the log<sub>2</sub>-ratio of the intron change to junction change, multiplied by the exon change. The heat map shows the Splicing Index score for most introns in *S. pombe* of the indicated strain compared to WT. Gene order along the y-axis is the same for all arrays. Stacked bar plots below show the number of introns with a Splicing Index score greater than 0.3. Color scale shows distribution of the severity of splicing defect.

## **Part B: What residues in H2A.Z are required for its role in splicing regulation?**

### **Background:**

This study came out of a thesis committee meeting discussion in which Geeta Narlikar posed the question of defining the functional residues in H2A.Z required for its affect on co-transcriptional pre-mRNA splicing regulation. There are several surfaces and modification sites known to differ between H2A.Z and canonical H2A that could account for the mechanism of how H2A.Z interacts with the spliceosome and RNAP II. These differences include: an extension of the acidic patch on H2A.Z, which interacts with H2B, promoters internucleosome interactions, and formation of the 30nm fibre confirmation of DNA, thereby altering chromosome folding and impacting transcription (Zhou et al. 2007); a destabilized docking surface with H3; and a metal ion binding site that may create a surface for specific protein interactions, such as with a zinc-finger domain (Suto et al. 2000). Additionally, there is a conserved array of four lysines in the N terminus that may be acetylated by the NuA4 acetyltransferase (Babiarz et al. 2006; Keogh et al. 2006; Millar et al. 2006). In *S. pombe*, H2A.Z is acetylated on four lysine residues (K5, K7, K12, and K16) by acetyltransferase Mst1 (Kat5) (Kim et al. 2009). While the function of acetylation on H2A.Z is not fully understood, it has been shown to be important for gene expression, genome stability, and chromosome condensation (Krogan et al. 2003; Kim et al. 2009; Halley et al. 2010).

Rather than undertaking a new mutagenesis screen, I opted to search the literature for point mutants in *ph1* and phenotypes in point mutants in H2A.Z homologs in other organisms, because the amino acid sequence of H2A.Z is highly conserved between different organisms (Baxevanis et al. 1995). To this end, Michael Keogh's group had recently shown that mutants in

*pht1* that disrupt the acetylation of the N terminus mimic the *pht1Δ* strain in the *S. pombe* EMAP and in expression microarrays (Kim et al. 2009). This strong phenocopy implies that preventing Pht1 acetylation disrupts the function of Swr1 complex pathway, and that the acetylation cycle plays a pivotal role in H2A.Z's functions.

I wondered if acetylation or the other N-terminal residues mutated in these strains were required for H2A.Z's role in splicing. I contacted Dr. Keogh to obtain the *pht1* acetylation mutant strains in 2012. The strains included a deletion of the N-terminal 38 amino acids, which contained all four potentially acetyltable lysines, as well as strains where these four lysines were mutated to either an acetylation mimic (lysine to glutamine) or were not acetyltable (lysine to arginine). Unfortunately, due to Hurricane Sandy flooding his lab around the time of our communication, I temporarily abandoned this line of inquiry with the *pht1* acetylation mutant strains. At the Cold Spring Harbor Eukaryotic mRNA Processing Meeting 2015, Charles Query suggested a similar H2A.Z mutagenesis experiment as Dr. Narlikar, and I mentioned my interest in Dr. Keogh's strains. Fortuitously, Dr. Query inherited the freezer containing these strains when Dr. Keogh left Albert Einstein College of Medicine, and he worked with Dr. Keogh to identify and send me these *pht1* acetylation mutant strains.

## **Experimental Approach:**

### **Growth Assays**

Strains were grown in YES5 overnight at 30°C and then diluted to 0.1 OD<sub>600</sub> the next morning to allow for growth to mid log-phase. Mid-log phase cultures were re-diluted to 0.1 OD<sub>600</sub> and serially diluted 1:5 across a 96-well plate. Approximately 5μl of cultures were

transferred to YES5 plates using a frogging tool and plates were grown at 30°C for 2 days, 37°C for 2 days or 16°C for 9 days. Images were then collected. 1 replicate shown for 3 biological replicates collected.

### **RNA Isolation**

Strains were grown in YES5 overnight at 30°C and then diluted to 0.1 OD<sub>600</sub> the next morning to allow for growth to mid log-phase. Temperature shifts occurred when strains had reached 0.3-0.5 OD<sub>600</sub>. 16°C shifts for 9 hours. Cultures were harvested by centrifugation between 0.5–0.8 OD<sub>600</sub>. Pellets were washed once in water and then flash-frozen in liquid nitrogen and stored at -80°C. RNA was isolated using hot acid phenol followed by isopropanol precipitation as previously described (Bergkessel et al. 2011).

### **Splicing-specific Microarray Data Collection and Feature Analysis**

Collection and analysis of splicing-specific microarray data was performed as previously described (Lipp et al. 2015; Patrick et al. 2015).

### **Results and Discussion:**

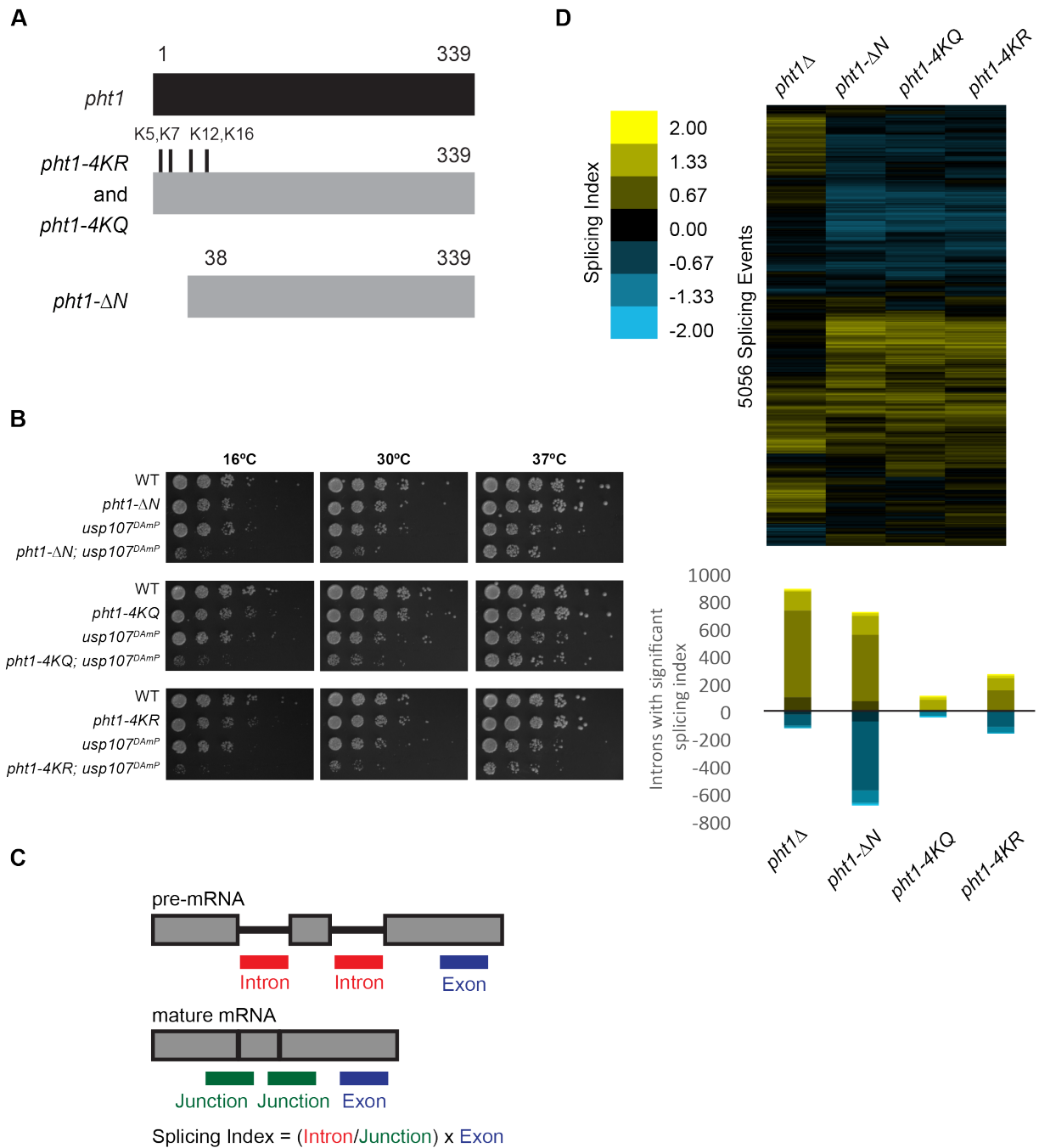
First, I tested the *pht1* acetylation mutant strains from Dr. Keogh for genetic interactions with the spliceosome by crossing them to the Guthrie lab's collection of splicing factor DAmP strains (Figure A.2 Panel A) (Patrick et al. 2015). Because the *pht1* acetylation mutants and splicing factor DAmP mutants were both entered into the *S. pombe* EMAP as query strains, they were not previously tested against each other (Kim et al. 2009; Patrick et al. 2015). Given that the *pht1* acetylation mutant strains strongly phenocopied the *pht1*Δ strain interaction profile in

the *S. pombe* EMAP, it is not surprising that we observed that they mimic *pht1Δ* in genetic interaction with splicing factors in serial growth assays (compare to Chapter 1 Figure 1) (Figure A.2 panel B). Indeed, the double mutants of U1 factor *usp107<sup>DAmP</sup>* and *pht1-ΔN*, *pht1-4KR* and *pht1-4KQ*, are synthetic sick at 16°C, 30°C, and 37°C (Ohi et al. 2002; Cvitkovic and Jurica 2013). Crosses between the *pht1* acetylation mutants and other splicing factors revealed similar behavior (data not shown). The phenocopy of the genetic interactions suggests that the cycling of acetylation on Pht1 is required for influencing splicing. Similar to *pht1Δ*, the *pht1* acetylation mutants are all slightly cold sensitive. This suggests that the essential function(s) of Pht1 depend on its ability to be acetylated.

Next, I examined the *pht1* acetylation mutant strains using splicing specific microarrays to characterize their splicing profile genome-wide at 16°C (Figure A.2 Panels C and D). Strikingly, all three *pht1* acetylation mutants phenocopy each other, suggesting that the ability to cycle acetylation is the primary function of the N-terminus of H2A.Z in splicing. To my surprise, the intron-specific splicing defect in *pht1Δ* was not phenocopied in the *pht1* acetylation mutants. Approximately half of the introns that are not efficiently spliced in the *pht1Δ* strain are unaffected or spliced more efficiently in the *pht1* acetylation mutant strains. Moreover, approximately half of the introns that are unaffected in the *pht1Δ* strain are less efficiently spliced in the *pht1* acetylation mutant strains, while approximately the other half of the *pht1Δ* unaffected intron are more efficiently spliced in the *pht1* acetylation mutant strains. The *pht1-ΔN* strain also had an increase in the number of intron that are more efficiently spliced, suggesting that the acetylation cycling in the N-terminus or other residues limit splicing efficiency. More work is required to dissect these residues from the acetylation sites in the N-terminus of Pht1. Additionally, there were not enough introns with splicing defects in the *pht1* acetylation mutant

strains to perform computational feature analysis, as in Chapter 1, so the characteristics of this class of introns is unknown. The fact that the splicing profile is different between the *pht1Δ* strain and in the *pht1* acetylation mutant strains suggests that the cycling of acetylation is required for splicing in a different mechanism than how deposition of H2A.Z by the Swr1 complex influences splicing.

Taken together, the cold-sensitive defect in *pht1Δ* may be linked to the role acetylation cycling plays in the protein. While the *pht1* acetylation mutant strains have a splicing defect, they also enhance splicing of some introns, which are not defective in *pht1Δ*, suggesting that their role in splicing is influenced by Pht1's acetylation status in way that is different from Swr1 complex's deposition and positioning of Pht1. More work is needed to dissect the dependency upon acetylation state from the role of Pht1 in splicing.



**Figure A.2: Acetylation cycling on Pht1 is important for splicing**

(A) Cartoon depicting mutant alleles of *pht1*: *pht1*-ΔN, *pht1*-4KQ, and *pht1*-4KR. Lysines mutated to glutamine or arginine are indicated. Amino acid length of each construct is labeled.



(B) Serial dilutions (1:5) of WT strain, *pht1* mutant strains, U1 splicing factor *usp107<sup>DAmP</sup>* strain, and their double mutant strains, grown at 16°C, 30°C, and 37°C.

(C) Cartoon depicting splicing-specific microarray probes.

(D) Splicing profile of *pht1*Δ, *pht1*-ΔN, *pht1*-4KQ, and *pht1*-4KR strains at 16°C. cDNA from each strain listed was competitively hybridized on the splicing-specific microarray against that from WT. A Splicing Index value was calculated for each intron by normalizing the log<sub>2</sub>-ratio of the intron change to junction change, multiplied by the exon change. The heat map shows the Splicing Index score for most introns in *S. pombe* of the indicated strain compared to WT. Gene order along the y-axis is the same for all arrays. Stacked bar plots below show the number of introns with a Splicing Index score greater than 0.3. Color scale shows distribution of the severity of splicing defect. n = 3 biological replicates.

## **Part C: Splicing factor chromatin immunoprecipitation (ChIP) protocol optimization in *S. pombe***

### **Background:**

The original first aim of my thesis project was to map and analyze splicing factor chromatin association from genome-wide in *S. pombe*. The rationale was that splicing factor association patterns across gene bodies provided information about spliceosome assembly order *in vivo* (Alpert et al. 2017). When I began this project, ChIP had been successfully used to monitor splicing factor association with chromatin in *S. cerevisiae* and *H. sapiens* (Kotovic et al. 2003; Gornemann et al. 2005; Lacadie and Rosbash 2005; Listerman et al. 2006; Tardiff and Rosbash 2006). These studies revealed that tagged splicing factors, representing snRNPs, associated co-transcriptionally across a gene body in the same order as they do *in vitro*, and their distribution across the gene body could be used as a proxy to describe “timing” of assembly or activation of the spliceosome (Alpert et al. 2017). While ChIP provides a static picture of cross-linked associations between protein and DNA, the unidirectional nature of transcription on highly expressed genes provides the opportunity to link site of splicing factor association with chromatin and spliceosome assembly progression.

The primary challenge of the splicing factor ChIP technique was generating a strong enough signal over background noise. Splicing factors are hypothesized by the field to be crosslinked to DNA through their interactions with RNA, as opposed to directly interacting with DNA (Bieberstein et al. 2014). This assumed-indirect crosslinking of splicing factors to DNA yields a much lower signal when compared to known DNA-interacting proteins such as RNAP II (Bieberstein et al. 2014). Therefore, this technique is challenging when the gene of interest is not

highly expressed and does not recruit the spliceosome frequently. I spent several years optimizing this technique for *S. pombe* in collaboration with Kristin Patrick.

### **Experimental Approach:**

In data not shown here, I optimized a variety of steps to a published *S. cerevisiae* splicing factor ChIP protocol that was commonly used in the Guthrie Lab (Kress et al. 2008). These steps included:

**Protein Tagging:** 1X FLAG, 3X FLAG, 6X FLAG, HA, Streptavidin were all tried on several different splicing factors, Usp103, Usp104, Usp105, Prp2, Bpb1, Lea1, Cdc28, Srp1, Prp45

**Culture:** OD<sub>600</sub> of culture, volume of culture

**Crosslinking:** Formaldehyde final concentration, duration of crosslinking

**Lysis:** Volume of crosslinked cell aliquots in each lysis tube, number of rounds of bead-beating, duration of bead-beating round, rest time between bead-beating rounds, protease concentrations in lysis buffer, additional centrifugation after lysis to separate chromatin-associated fraction

**Sonication:** volume of lysate in tubes during sonication, duration of sonication round, number of sonication rounds

**Immunoprecipitation:** duration of immunoprecipitation before and after bead addition, antibody final concentration, bead matrix material (sepharose vs. metallic DYNAbeads), bead concentration and bead preparation procedure (blocking), washing buffer stringency, temperature of immunoprecipitation and washing

**Elution and Clean-up:** microcentrifuge tube material (conventional vs. non-stick), SDS removal method (commercial kit vs. in-house method)

Several different highly and lowly expressed genes were queried using qPCR. To minimize the background signal for qPCR, I normalized to three amplicons from genomic regions with little predicted transcription, such as intergenic regions, the centromere and mating type locus.

### **ChIP Method:**

Day 1:

1. Grow 100ul culture to  $OD_{600}=0.5-0.8$ .
2. To crosslink, add 1.4mL 37% formaldehyde (Sigma) (final concentration 1%).
3. Incubate at room temperature for 15 minutes with occasional swirling.
4. To quench crosslinking reaction, add 2.5mL 2.5M glycine (Sigma).
5. Incubate at room temperature for 5 minutes with occasional swirling.
6. Transfer to vessels for centrifugation in 50  $OD_{600}$  aliquots.
7. Centrifuge to pellet in a tabletop 4°C centrifuge at 3000rpm for 5 minutes.
8. Wash pellet twice with 25mL 1XTBS, centrifuge as in step 7.
9. Resuspend pellet in 1mL 1XTBS and transfer to 2mL screw cap Eppendorf tube.
10. Remove supernatant and flash freeze in liquid nitrogen. Store at  $-80^{\circ}$

Day 2:

1. Aliquot Lysis Buffer and add fresh protease inhibitors (+ PIs) per recipe below.
2. Remove cell pellets from  $-80^{\circ}\text{C}$  freezer and thaw pellets on ice.

3. Add 1400 $\mu$ L of Lysis Buffer with PIs to the pellet, and resuspend the pellet by pipetting to avoid bubbles.
4. Add 200 $\mu$ L 4 $^{\circ}$ C 0.5mm silica glass beads.
5. Beadbeat in Mini-Beadbeater (Biospec Products) for 8 rounds of 1.5 minutes, with 2 minutes rest on ice in between rounds.
6. Use a 23G1 needle (BD) heated in a flame to puncture bottom of tubes and place inside a 5mL Falcon tube.
7. Centrifuge in a tabletop 4 $^{\circ}$ C centrifuge at 2000rpm for 3 minutes.
8. Discard supernatant and resuspend the heavy pellet in 200 $\mu$ L Lysis Buffer + PIs and transfer to an Eppendorf tube. Do not transfer silica beads.
9. Centrifuge in a tabletop 4 $^{\circ}$ C centrifuge at 8000rpm for 10 minutes.
10. Discard supernatant and resuspend pellet in 300 $\mu$ L Lysis Buffer + PIs.
11. Sonicate in a Bioruptor waterbath sonicator (Diagenode) at 4 $^{\circ}$ C for 6 rounds on HIGH, where each round is 10 minutes long with 30 seconds ON and 1 minute OFF.
12. Centrifuge in a tabletop 4 $^{\circ}$ C centrifuge at 14,000rpm for 10 minutes.
13. Transfer supernatant to new eppendorf tube. This supernatant is whole cell extract (WCE).
14. Add Lysis Buffer + PIs to bring WCE volume to 1000 $\mu$ L. Mix well.
15. Remove 100 $\mu$ L WCE to new Eppendorf. This is the INPUT control.
16. To IP, add antibody (for example M2 FLAG (F3165, Sigma)) to WCE and nutate 2 hours at 4 $^{\circ}$ C.
17. While the antibody is binding, prepare Protein G DYNAbead beads (Invitrogen) by aliquoting 60 $\mu$ L packed beads per IP to a new Eppendorf tube, counting one extra aliquot for easier pipetting. Use the 150magnetic rack to collect the beads for wash steps:

- A. Wash beads 3 times with 1 mL 1XTBS.
  - B. Wash beads 2 times with 1 mL Lysis Buffer + PIs.
  - C. Resuspend in 200 $\mu$ L with Lysis Buffer + PIs. Set aside on ice.
18. After 2 hour IP, add only 100 $\mu$ L washed DYNAbead suspension to the IP.
19. Nutate for 2 hours at 4°C.
20. Using a magnetic rack, discard the supernatant
21. Perform washes, keeping the beads using a magnetic rack:
- A. Twice for 5 minutes on nutator at 4°C with 1 mL Lysis Buffer.
  - B. Twice for 5 minutes on nutator at 4°C with 1 mL High Salt Lysis Buffer.
  - C. Twice for 5 minutes on nutator at 4°C with 1 mL Wash Buffer.
  - D. Move to non-stick Eppendorf tube.
  - E. Once for 5 minutes on nutator at room temperature with 1mL TE.
22. Elute ChIPed DNA from beads by adding 100 $\mu$ L Elution Buffer and incubating for 20 minutes at 70° (flick tube once or twice during incubation).
23. Flick tube again. Place on magnet.
24. Transfer eluate to new Eppendorf tube, set aside, but also keep the tube with beads for a second elution below.
25. Wash beads with 150 $\mu$ L TE + 0.67% SDS + 334  $\mu$ g/ml Proteinease K (made fresh)
26. Keep this eluate and add it to first eluate of ChIPed DNA that was set aside.
27. At the same time, add 200 $\mu$ L of TE + 1% SDS + 250 $\mu$ g/ml Proteinase K (made fresh) to the reserved 100 $\mu$ L INPUT aliquot.
28. Incubate both the combined ChIP eluate and INPUT at 65°C for at least 8 hours (overnight) to reverse crosslinking

Day 3:

1. Cool samples to room temperature.
2. Removed SDS from samples using NucleoSpin Gel Clean-up Columns with 1250 $\mu$ L NTB Buffer (Macherey-Nagel). Elute DNA from columns with 100 $\mu$ L water.
3. Combine separated DNA aliquots from original culture split into 50 OD<sub>600S</sub>.

### Solutions

10X TBS (500mL) – Store at 4°C

200mM Tris-Cl pH 7.6	100mL of 1M
1.5M NaCl	150mL of 5M

Lysis Buffer (50mL) – Store at 4°C

50mM HEPES-KOH pH 7.5	2.5mL of 1M
140mM NaCl	1.4mL of 5M
1mM EDTA	100 $\mu$ L of 0.5M
1% Triton X-100	2.5mL of 20%
0.1% Sodium Deoxycholate	0.5mL of 10%

Protease Inhibitors: 1 tablet/50mL EDTA-free Complete protease inhibitor (Roche)

Protease inhibitors (2mM PMSF, 20 $\mu$ g/mL leupeptin, 20 $\mu$ g/mL pepstatin, 20 $\mu$ g/mL aprotonin, 20 $\mu$ g/mL antipain).

High Salt Lysis Buffer (50mL) – Store at 4°C

50mM HEPES-KOH pH 7.5	2.5mL of 1M
500mM NaCl	5mL of 5M
1mM EDTA	100µL of 0.5M
1% Triton X-100	2.5mL of 20%
0.1% Sodium Deoxycholate	0.5mL of 10%

Wash Buffer (50mL) – Store at 4°C

10mM Tris-Cl pH 8.0	0.5mL of 1M
0.25M LiCl	3.125mL of 4M
0.5% NP-40	2.5mL of 10%
0.5% Sodium Deoxycholate	2.5mL of 10%
1mM EDTA	100µl of 0.5M

Elution Buffer (10mL) – Store at room temperature

50mM Tris-Cl pH 8.0	0.5 ml of 1M
10mM EDTA	200µL of 0.5M
1% SDS	1 ml of 10%

TE + 1% SDS (10mL) – Store at room temperature

10mM Tris-Cl pH 8.0	100µL of 1M
1mM EDTA	20µL of 0.5M
1% SDS	1 ml of 10%



Add 12.5 $\mu$ L of 20 mg/mL Proteinease K per 1mL TE+1% SDS (250 $\mu$ g/mL final  
Proteinase K concentration)

TE + 0.67% SDS (10 ml) – Store at room temperature

10mM Tris-Cl pH 8.0	100 $\mu$ L of 1M
1mM EDTA	20 $\mu$ L of 0.5M
0.67% SDS	670 $\mu$ L of 10%

Add 16.7 $\mu$ L of 20 mg/ml Proteinease K per 1mL TE + 0.67% SDS (334 $\mu$ g/mL final  
Proteinase K concentration)

### **Results and Discussion:**

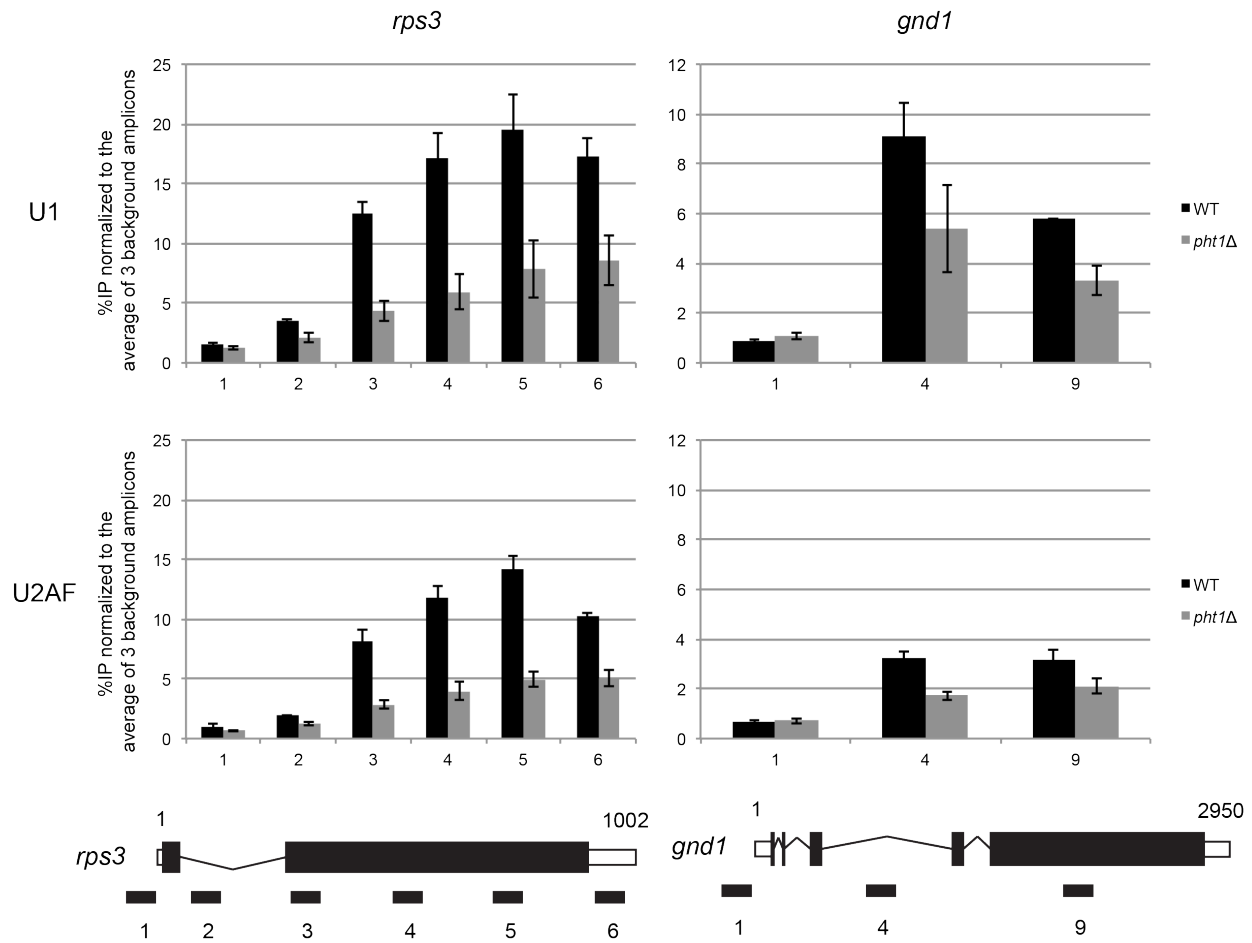
The resulting protocol was published as a supplemental method in Chapter 1 and used again in unpublished work in Chapter 2. It was also published in (Patrick et al. 2015), with additional modifications made by Kristin Patrick to optimize for her protein-tagging scheme. What follows in the base protocol for ChIP of splicing factors optimized for *S. pombe*.

In addition, I show successful examples of this technique for two splicing factors (Usp103 from U1 snRNP and Prp2 from U2AF) on two highly expressed genes (*rsp3* and *gnd1*) that were selected based on their high expression levels (Marguerat et al. 2012; Cvitkovic and Jurica 2013). In Figure A.3, U1 and U2AF are associated across the gene body in a 5' to 3' manner, with recruitment beginning downstream of the 5' SS, peaking in the gene body, and decreasing towards the transcription termination site, reminiscent of association patterns of U1

and U2 in *S. cerevisiae* (Lacadie and Rosbash 2005; Tardiff and Rosbash 2006). Strikingly, in the *pht1Δ* strain, the association of both splicing factors is lower on both genes, suggesting a defect in spliceosome recruitment in the absence of H2A.Z.

While the experiments presented here use qPCR as the read-out, this optimized *S. pombe* protocol was used to ChIP both splicing factors (Chapter 2) and chromatin factors (Chapter 1) with successful readout genome-wide by high-throughput sequencing. The challenge remains for generating reproducible signal on poorly expressed genes and the resolution limit of this technique needs to be taken into consideration when interpreting signals. Due to these challenges, I would caution careful interpretation of splicing factor ChIP experiments in *S. pombe* going forward and assert that more sensitive and higher resolution techniques are needed to monitor splicing factor assembly *in vivo* in the future.

A



**Figure A.3: U1 and U2AF splicing factor ChIP on two highly expressed genes in *S. pombe***

(A) ChIP-qPCR with anti-FLAG antibody in WT and *pht1*Δ strains expressing 3XFLAG-tagged Usp103 (Yhc1) and Prp2 (U2AF<sup>23</sup>). Bars represent %IP (IP/input) normalized to the average of three background amplicons. Error bars are ±SEM for each primer set, n = 3 biological replicates. Gene diagram below indicates approximate location of qPCR amplicons along *rps3* and *gnd1*.

## REFERENCES

- Adam M, Robert F, Laroche M, Gaudreau L. 2001. H2A.Z is required for global chromatin integrity and for recruitment of RNA polymerase II under specific conditions. *Molecular and cellular biology* **21**: 6270-6279.
- Ahmed S, Dul B, Qiu X, Walworth NC. 2007. Msc1 acts through histone H2A.Z to promote chromosome stability in *Schizosaccharomyces pombe*. *Genetics* **177**: 1487-1497.
- Alexander RD, Innocente SA, Barrass JD, Beggs JD. 2010. Splicing-dependent RNA polymerase pausing in yeast. *Molecular cell* **40**: 582-593.
- Allen JJ. 2008. Development and application of technologies to study individual kinase substrate relationships. in (Order No 3328096) Available from Dissertations & Theses @ University of California; ProQuest Dissertations & Theses A&I (304656821), University of California, San Francisco.
- Alpert T, Herzog L, Neugebauer KM. 2017. Perfect timing: splicing and transcription rates in living cells. *Wiley interdisciplinary reviews RNA* **8**.
- Ares M, Jr., Grate L, Pauling MH. 1999. A handful of intron-containing genes produces the lion's share of yeast mRNA. *Rna* **5**: 1138-1139.
- Ast G. 2004. How did alternative splicing evolve? *Nature reviews Genetics* **5**: 773-782.
- Awan AR, Manfredo A, Pleiss JA. 2013. Lariat sequencing in a unicellular yeast identifies regulated alternative splicing of exons that are evolutionarily conserved with humans. *Proceedings of the National Academy of Sciences of the United States of America* **110**: 12762-12767.

- Babiarz JE, Halley JE, Rine J. 2006. Telomeric heterochromatin boundaries require NuA4-dependent acetylation of histone variant H2A.Z in *Saccharomyces cerevisiae*. *Genes & development* **20**: 700-710.
- Basnet H, Su XB, Tan Y, Meisenhelder J, Merkurjev D, Ohgi KA, Hunter T, Pillus L, Rosenfeld MG. 2014. Tyrosine phosphorylation of histone H2A by CK2 regulates transcriptional elongation. *Nature* **516**: 267-271.
- Baxevanis AD, Arents G, Moudrianakis EN, Landsman D. 1995. A variety of DNA-binding and multimeric proteins contain the histone fold motif. *Nucleic acids research* **23**: 2685-2691.
- Bergkessel M. 2010. The regulation of pre-mRNA splicing in response to environmental stress in budding yeast in (*Order No 3438937*) Available from *Dissertations & Theses @ University of California; ProQuest Dissertations & Theses A&I (849714337)*. University of California, San Francisco.
- Bergkessel M, Whitworth GB, Guthrie C. 2011. Diverse environmental stresses elicit distinct responses at the level of pre-mRNA processing in yeast. *Rna* **17**: 1461-1478.
- Bertram K, Agafonov DE, Dybkov O, Haselbach D, Leelaram MN, Will CL, Urlaub H, Kastner B, Luhrmann R, Stark H. 2017a. Cryo-EM Structure of a Pre-catalytic Human Spliceosome Primed for Activation. *Cell* **170**: 701-713 e711.
- Bertram K, Agafonov DE, Liu WT, Dybkov O, Will CL, Hartmuth K, Urlaub H, Kastner B, Stark H, Luhrmann R. 2017b. Cryo-EM structure of a human spliceosome activated for step 2 of splicing. *Nature* **542**: 318-323.
- Bieberstein NI, Carrillo Oesterreich F, Straube K, Neugebauer KM. 2012. First exon length controls active chromatin signatures and transcription. *Cell reports* **2**: 62-68.

- Bieberstein NI, Straube K, Neugebauer KM. 2014. Chromatin immunoprecipitation approaches to determine co-transcriptional nature of splicing. *Methods in molecular biology (Clifton, NJ)* **1126**: 315-323.
- Bitton DA, Atkinson SR, Rallis C, Smith GC, Ellis DA, Chen YY, Malecki M, Codlin S, Lemay JF, Cotobal C et al. 2015. Widespread exon skipping triggers degradation by nuclear RNA surveillance in fission yeast. *Genome research* **25**: 884-896.
- Blencowe BJ. 2006. Alternative splicing: new insights from global analyses. *Cell* **126**: 37-47.
- Bohnsack MT, Martin R, Granneman S, Ruprecht M, Schleiff E, Tollervey D. 2009. Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. *Molecular cell* **36**: 583-592.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114-2120.
- Bon E, Casaregola S, Blandin G, Llorente B, Neuveglise C, Munsterkötter M, Guldener U, Mewes HW, Van Helden J, Dujon B et al. 2003. Molecular evolution of eukaryotic genomes: hemiascomycetous yeast spliceosomal introns. *Nucleic acids research* **31**: 1121-1135.
- Braberg H, Jin H, Moehle EA, Chan YA, Wang S, Shales M, Benschop JJ, Morris JH, Qiu C, Hu F et al. 2013. From structure to systems: high-resolution, quantitative genetic analysis of RNA polymerase II. *Cell* **154**: 775-788.
- Bresson S, Tuck A, Staneva D, Tollervey D. 2017. Nuclear RNA Decay Pathways Aid Rapid Remodeling of Gene Expression in Yeast. *Molecular cell* **65**: 787-800 e785.
- Campos EI, Reinberg D. 2009. Histones: annotating chromatin. *Annual review of genetics* **43**: 559-599.

- Carnahan RH, Feoktistova A, Ren L, Niessen S, Yates JR, 3rd, Gould KL. 2005. Dim1p is required for efficient splicing and export of mRNA encoding lid1p, a component of the fission yeast anaphase-promoting complex. *Eukaryotic cell* **4**: 577-587.
- Carrillo Oesterreich F, Preibisch S, Neugebauer KM. 2010. Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. *Molecular cell* **40**: 571-581.
- Chathoth KT, Barrass JD, Webb S, Beggs JD. 2014. A splicing-dependent transcriptional checkpoint associated with prespliceosome formation. *Molecular cell* **53**: 779-790.
- Chen M, Manley JL. 2009. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nature reviews Molecular cell biology* **10**: 741-754.
- Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G, Chan ET, Christie KR, Costanzo MC, Dwight SS, Engel SR et al. 2012. Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic acids research* **40**: D700-705.
- Clement-Ziza M, Marsellach FX, Codlin S, Papadakis MA, Reinhardt S, Rodriguez-Lopez M, Martin S, Marguerat S, Schmidt A, Lee E et al. 2014. Natural genetic variation impacts expression levels of coding, non-coding, and antisense transcripts in fission yeast. *Molecular systems biology* **10**: 764.
- Cordin O, Hahn D, Beggs JD. 2012. Structure, function and regulation of spliceosomal RNA helicases. *Current opinion in cell biology* **24**: 431-438.
- Cramer P, Pesce CG, Baralle FE, Kornblihtt AR. 1997. Functional association between promoter structure and transcript alternative splicing. *Proceedings of the National Academy of Sciences of the United States of America* **94**: 11456-11460.
- Cvitkovic I, Jurica MS. 2013. Spliceosome database: a tool for tracking components of the spliceosome. *Nucleic acids research* **41**: D132-141.

- Damgaard CK, Kahns S, Lykke-Andersen S, Nielsen AL, Jensen TH, Kjems J. 2008. A 5' splice site enhances the recruitment of basal transcription initiation factors in vivo. *Molecular cell* **29**: 271-278.
- David L, Huber W, Granovskaia M, Toedling J, Palm CJ, Bofkin L, Jones T, Davis RW, Steinmetz LM. 2006. A high-resolution map of transcription in the yeast genome. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 5320-5325.
- Davis CA, Grate L, Spingola M, Ares M, Jr. 2000. Test of intron predictions reveals novel splice sites, alternatively spliced mRNAs and new introns in meiotically regulated genes of yeast. *Nucleic acids research* **28**: 1700-1706.
- de Almeida SF, Carmo-Fonseca M. 2014. Reciprocal regulatory links between cotranscriptional splicing and chromatin. *Seminars in cell & developmental biology* **32**: 2-10.
- de Almeida SF, Grosso AR, Koch F, Fenouil R, Carvalho S, Andrade J, Levezinho H, Gut M, Eick D, Gut I et al. 2011. Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. *Nature structural & molecular biology* **18**: 977-983.
- de la Mata M, Alonso CR, Kadener S, Fededa JP, Blaustein M, Pelisch F, Cramer P, Bentley D, Kornblihtt AR. 2003. A slow RNA polymerase II affects alternative splicing in vivo. *Molecular cell* **12**: 525-532.
- Delhomme N, Padioleau I, Furlong EE, Steinmetz LM. 2012. easyRNASeq: a bioconductor package for processing RNA-Seq data. *Bioinformatics* **28**: 2532-2533.



- Dumesic PA, Homer CM, Moresco JJ, Pack LR, Shanle EK, Coyle SM, Strahl BD, Fujimori DG, Yates JR, 3rd, Madhani HD. 2015. Product binding enforces the genomic specificity of a yeast polycomb repressive complex. *Cell* **160**: 204-218.
- Engel SR, Dietrich FS, Fisk DG, Binkley G, Balakrishnan R, Costanzo MC, Dwight SS, Hitz BC, Karra K, Nash RS et al. 2014. The reference genome sequence of *Saccharomyces cerevisiae*: then and now. *G3 (Bethesda, Md)* **4**: 389-398.
- Fabrizio P, Dannenberg J, Dube P, Kastner B, Stark H, Urlaub H, Luhrmann R. 2009. The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. *Molecular cell* **36**: 593-608.
- Fackenthal JD, Godley LA. 2008. Aberrant RNA splicing and its functional consequences in cancer cells. *Disease models & mechanisms* **1**: 37-42.
- Fiedler D, Braberg H, Mehta M, Chechik G, Cagney G, Mukherjee P, Silva AC, Shales M, Collins SR, van Wageningen S et al. 2009. Functional organization of the *S. cerevisiae* phosphorylation network. *Cell* **136**: 952-963.
- Fong N, Kim H, Zhou Y, Ji X, Qiu J, Saldi T, Diener K, Jones K, Fu XD, Bentley DL. 2014. Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes & development* **28**: 2663-2676.
- Fong YW, Zhou Q. 2001. Stimulatory effect of splicing factors on transcriptional elongation. *Nature* **414**: 929-933.
- Fourmann JB, Dybkov O, Agafonov DE, Tauchert MJ, Urlaub H, Ficner R, Fabrizio P, Luhrmann R. 2016. The target of the DEAH-box NTP triphosphatase Prp43 in *Saccharomyces cerevisiae* spliceosomes is the U2 snRNP-intron interaction. *eLife* **5**.

- Furger A, O'Sullivan JM, Binnie A, Lee BA, Proudfoot NJ. 2002. Promoter proximal splice sites enhance transcription. *Genes & development* **16**: 2792-2799.
- Galej WP, Wilkinson ME, Fica SM, Oubridge C, Newman AJ, Nagai K. 2016. Cryo-EM structure of the spliceosome immediately after branching. *Nature* **537**: 197-201.
- Gomez Acuna LI, Fiszbein A, Allo M, Schor IE, Kornbliht AR. 2013. Connections between chromatin signatures and splicing. *Wiley interdisciplinary reviews RNA* **4**: 77-91.
- Gornemann J, Kotovic KM, Hujer K, Neugebauer KM. 2005. Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Molecular cell* **19**: 53-63.
- Halley JE, Kaplan T, Wang AY, Kobor MS, Rine J. 2010. Roles for H2A.Z and its acetylation in GAL1 transcription and gene induction, but not GAL1-transcriptional memory. *PLoS biology* **8**: e1000401.
- Han M, Grunstein M. 1988. Nucleosome loss activates yeast downstream promoters in vivo. *Cell* **55**: 1137-1145.
- Heckman DS, Geiser DM, Eidell BR, Stauffer RL, Kardos NL, Hedges SB. 2001. Molecular evidence for the early colonization of land by fungi and plants. *Science* **293**: 1129-1133.
- Herzel L, Ottoz DSM, Alpert T, Neugebauer KM. 2017. Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nature reviews Molecular cell biology*.
- Hodges C, Bintu L, Lubkowska L, Kashlev M, Bustamante C. 2009. Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. *Science* **325**: 626-628.
- Hoffman CS, Wood V, Fantes PA. 2015. An Ancient Yeast for Young Geneticists: A Primer on the *Schizosaccharomyces pombe* Model System. *Genetics* **201**: 403-423.

- Homann OR, Johnson AD. 2010. MochiView: versatile software for genome browsing and DNA motif analysis. *BMC biology* **8**: 49.
- Horowitz DS. 2011. The splice is right: guarantors of fidelity in pre-mRNA splicing. *Rna* **17**: 551-554.
- Hossain MA, Claggett JM, Edwards SR, Shi A, Pennebaker SL, Cheng MY, Hasty J, Johnson TL. 2016. Posttranscriptional Regulation of Gcr1 Expression and Activity Is Crucial for Metabolic Adjustment in Response to Glucose Availability. *Molecular cell* **62**: 346-358.
- Huang T, Vilardell J, Query CC. 2002. Pre-spliceosome formation in *S.pombe* requires a stable complex of SF1-U2AF(59)-U2AF(23). *The EMBO journal* **21**: 5516-5526.
- Huang Y, Li W, Yao X, Lin QJ, Yin JW, Liang Y, Heiner M, Tian B, Hui J, Wang G. 2012. Mediator complex regulates alternative mRNA processing via the MED23 subunit. *Molecular cell* **45**: 459-469.
- Huff JT, Plocik AM, Guthrie C, Yamamoto KR. 2010. Reciprocal intronic and exonic histone modification regions in humans. *Nature structural & molecular biology* **17**: 1495-1499.
- Kalsotra A, Cooper TA. 2011. Functional consequences of developmentally regulated alternative splicing. *Nature reviews Genetics* **12**: 715-729.
- Kaufner NF, Potashkin J. 2000. Analysis of the splicing machinery in fission yeast: a comparison with budding yeast and mammals. *Nucleic acids research* **28**: 3003-3010.
- Keogh MC, Mennella TA, Sawa C, Berthelet S, Krogan NJ, Wolek A, Podolny V, Carpenter LR, Greenblatt JF, Baetz K et al. 2006. The *Saccharomyces cerevisiae* histone H2A variant Htz1 is acetylated by NuA4. *Genes & development* **20**: 660-665.
- Keren H, Lev-Maor G, Ast G. 2010. Alternative splicing and evolution: diversification, exon definition and function. *Nature reviews Genetics* **11**: 345-355.

- Kim DU, Hayles J, Kim D, Wood V, Park HO, Won M, Yoo HS, Duhig T, Nam M, Palmer G et al. 2010. Analysis of a genome-wide set of gene deletions in the fission yeast *Schizosaccharomyces pombe*. *Nature biotechnology* **28**: 617-623.
- Kim Guisbert KS, Li H, Guthrie C. 2007. Alternative 3' pre-mRNA processing in *Saccharomyces cerevisiae* is modulated by Nab4/Hrp1 in vivo. *PLoS biology* **5**: e6.
- Kim HS, Vanoosthuyse V, Fillingham J, Roguev A, Watt S, Kislinger T, Treyer A, Carpenter LR, Bennett CS, Emili A et al. 2009. An acetylated form of histone H2A.Z regulates chromosome architecture in *Schizosaccharomyces pombe*. *Nature structural & molecular biology* **16**: 1286-1293.
- Kim S, Kim H, Fong N, Erickson B, Bentley DL. 2011. Pre-mRNA splicing is a determinant of histone H3K36 methylation. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 13564-13569.
- Kobor MS, Venkatasubrahmanyam S, Meneghini MD, Gin JW, Jennings JL, Link AJ, Madhani HD, Rine J. 2004. A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS biology* **2**: E131.
- Koodathingal P, Novak T, Piccirilli JA, Staley JP. 2010. The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA splicing. *Molecular cell* **39**: 385-395.
- Koodathingal P, Staley JP. 2013. Splicing fidelity: DEAD/H-box ATPases as molecular clocks. *RNA biology* **10**: 1073-1079.
- Kornblihtt AR. 2005. Promoter usage and alternative splicing. *Current opinion in cell biology* **17**: 262-268.

- Kotovic KM, Lockshon D, Boric L, Neugebauer KM. 2003. Cotranscriptional recruitment of the U1 snRNP to intron-containing genes in yeast. *Molecular and cellular biology* **23**: 5768-5779.
- Kress TL, Krogan NJ, Guthrie C. 2008. A single SR-like protein, Npl3, promotes pre-mRNA splicing in budding yeast. *Molecular cell* **32**: 727-734.
- Krogan NJ, Keogh MC, Datta N, Sawa C, Ryan OW, Ding H, Haw RA, Pootoolal J, Tong A, Canadien V et al. 2003. A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Molecular cell* **12**: 1565-1576.
- Kuang Z, Cai L, Zhang X, Ji H, Tu BP, Boeke JD. 2014. High-temporal-resolution view of transcription and chromatin states across distinct metabolic states in budding yeast. *Nature structural & molecular biology* **21**: 854-863.
- Kuhn AN, Kaufer NF. 2003. Pre-mRNA splicing in *Schizosaccharomyces pombe*: regulatory role of a kinase conserved from fission yeast to mammals. *Current genetics* **42**: 241-251.
- Kupfer DM, Drabenstot SD, Buchanan KL, Lai H, Zhu H, Dyer DW, Roe BA, Murphy JW. 2004. Introns and splicing elements of five diverse fungi. *Eukaryotic cell* **3**: 1088-1100.
- Lacadie SA, Rosbash M. 2005. Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5'ss base pairing in yeast. *Molecular cell* **19**: 65-75.
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**: R25.
- Larson A, Fair BJ, Pleiss JA. 2016. Interconnections Between RNA-Processing Pathways Revealed by a Sequencing-Based Genetic Screen for Pre-mRNA Splicing Mutants in Fission Yeast. *G3 (Bethesda, Md)* **6**: 1513-1523.

- Li B, Carey M, Workman JL. 2007. The role of chromatin during transcription. *Cell* **128**: 707-719.
- Lipp JJ, Marvin MC, Shokat KM, Guthrie C. 2015. SR protein kinases promote splicing of nonconsensus introns. *Nature structural & molecular biology* **22**: 611-617.
- Listerman I, Sapra AK, Neugebauer KM. 2006. Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nature structural & molecular biology* **13**: 815-822.
- Luco RF, Allo M, Schor IE, Kornblihtt AR, Misteli T. 2011. Epigenetics in alternative pre-mRNA splicing. *Cell* **144**: 16-26.
- Lund MK, Kress TL, Guthrie C. 2008. Autoregulation of Npl3, a yeast SR protein, requires a novel downstream region and serine phosphorylation. *Molecular and cellular biology* **28**: 3873-3881.
- Marguerat S, Schmidt A, Codlin S, Chen W, Aebersold R, Bahler J. 2012. Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. *Cell* **151**: 671-683.
- Mayas RM, Maita H, Semlow DR, Staley JP. 2010. Spliceosome discards intermediates via the DEAH box ATPase Prp43p. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 10020-10025.
- Merkhofer EC, Hu P, Johnson TL. 2014. Introduction to cotranscriptional RNA splicing. *Methods in molecular biology (Clifton, NJ)* **1126**: 83-96.
- Millar CB, Xu F, Zhang K, Grunstein M. 2006. Acetylation of H2AZ Lys 14 is associated with genome-wide gene activity in yeast. *Genes & development* **20**: 711-722.

- Mizuguchi G, Shen X, Landry J, Wu WH, Sen S, Wu C. 2004. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* **303**: 343-348.
- Molin C, Jauhiainen A, Warringer J, Nerman O, Sunnerhagen P. 2009. mRNA stability changes precede changes in steady-state mRNA amounts during hyperosmotic stress. *Rna* **15**: 600-614.
- Monsalve M, Wu Z, Adelmant G, Puigserver P, Fan M, Spiegelman BM. 2000. Direct coupling of transcription and mRNA processing through the thermogenic coactivator PGC-1. *Molecular cell* **6**: 307-316.
- Morillo-Huesca M, Clemente-Ruiz M, Andujar E, Prado F. 2010. The SWR1 histone replacement complex causes genetic instability and genome-wide transcription misregulation in the absence of H2A.Z. *PloS one* **5**: e12143.
- Morris DP, Greenleaf AL. 2000. The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. *The Journal of biological chemistry* **275**: 39935-39943.
- Munding EM, Igel AH, Shiue L, Dorighi KM, Trevino LR, Ares M, Jr. 2010. Integration of a splicing regulatory network within the meiotic gene expression program of *Saccharomyces cerevisiae*. *Genes & development* **24**: 2693-2704.
- Naftelberg S, Schor IE, Ast G, Kornblihtt AR. 2015. Regulation of alternative splicing through coupling with transcription and chromatin structure. *Annual review of biochemistry* **84**: 165-198.
- Neves LT, Douglass S, Spreafico R, Venkataramanan S, Kress TL, Johnson TL. 2017. The histone variant H2A.Z promotes efficient cotranscriptional splicing in *S. cerevisiae*. *Genes & development* **31**: 702-717.

- Newo AN, Lutzberger M, Bottner CA, Wehland J, Wissing J, Jansch L, Kaufer NF. 2007. Proteomic analysis of the U1 snRNP of *Schizosaccharomyces pombe* reveals three essential organism-specific proteins. *Nucleic acids research* **35**: 1391-1401.
- Nguyen THD, Galej WP, Bai XC, Oubridge C, Newman AJ, Scheres SHW, Nagai K. 2016. Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. *Nature* **530**: 298-302.
- Noble SM, Guthrie C. 1996. Identification of novel genes required for yeast pre-mRNA splicing by means of cold-sensitive mutations. *Genetics* **143**: 67-80.
- Ohi MD, Link AJ, Ren L, Jennings JL, McDonald WH, Gould KL. 2002. Proteomics analysis reveals stable multiprotein complexes in both fission and budding yeasts containing Myb-related Cdc5p/Cef1p, novel pre-mRNA splicing factors, and snRNAs. *Molecular and cellular biology* **22**: 2011-2024.
- Papamichos-Chronakis M, Watanabe S, Rando OJ, Peterson CL. 2011. Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell* **144**: 200-213.
- Patrick KL, Ryan CJ, Xu J, Lipp JJ, Nissen KE, Roguev A, Shales M, Krogan NJ, Guthrie C. 2015. Genetic interaction mapping reveals a role for the SWI/SNF nucleosome remodeler in spliceosome activation in fission yeast. *PLoS genetics* **11**: e1005074.
- Pleiss JA, Whitworth GB, Bergkessel M, Guthrie C. 2007a. Rapid, transcript-specific changes in splicing in response to environmental stress. *Molecular cell* **27**: 928-937.
- . 2007b. Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. *PLoS biology* **5**: e90.



- Price AM, Gornemann J, Guthrie C, Brow DA. 2014. An unanticipated early function of DEAD-box ATPase Prp28 during commitment to splicing is modulated by U5 snRNP protein Prp8. *Rna* **20**: 46-60.
- Rauhut R, Fabrizio P, Dybkov O, Hartmuth K, Pena V, Chari A, Kumar V, Lee CT, Urlaub H, Kastner B et al. 2016. Molecular architecture of the *Saccharomyces cerevisiae* activated spliceosome. *Science* **353**: 1399-1405.
- Roguev A, Bandyopadhyay S, Zofall M, Zhang K, Fischer T, Collins SR, Qu H, Shales M, Park HO, Hayles J et al. 2008. Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. *Science* **322**: 405-410.
- Romero-Santacreu L, Moreno J, Perez-Ortin JE, Alepuz P. 2009. Specific and global regulation of mRNA stability during osmotic stress in *Saccharomyces cerevisiae*. *Rna* **15**: 1110-1120.
- Rosbash M, Harris PK, Woolford JL, Jr., Teem JL. 1981. The effect of temperature-sensitive RNA mutants on the transcription products from cloned ribosomal protein genes of yeast. *Cell* **24**: 679-686.
- Rosonina E, Ip JY, Calarco JA, Bakowski MA, Emili A, McCracken S, Tucker P, Ingles CJ, Blencowe BJ. 2005. Role for PSF in mediating transcriptional activator-dependent stimulation of pre-mRNA processing in vivo. *Molecular and cellular biology* **25**: 6734-6746.
- Rudra D, Mallick J, Zhao Y, Warner JR. 2007. Potential interface between ribosomal protein production and pre-rRNA processing. *Molecular and cellular biology* **27**: 4815-4824.

- Ryan CJ, Roguev A, Patrick K, Xu J, Jahari H, Tong Z, Beltrao P, Shales M, Qu H, Collins SR et al. 2012. Hierarchical modularity and the evolution of genetic interactomes across species. *Molecular cell* **46**: 691-704.
- Sadeghi L, Bonilla C, Stralfors A, Ekwall K, Svensson JP. 2011. Podbat: a novel genomic tool reveals Swr1-independent H2A.Z incorporation at gene coding sequences through epigenetic meta-analysis. *PLoS computational biology* **7**: e1002163.
- Saint-Andre V, Batsche E, Rachez C, Muchardt C. 2011. Histone H3 lysine 9 trimethylation and HP1gamma favor inclusion of alternative exons. *Nature structural & molecular biology* **18**: 337-344.
- Schuldiner M, Collins SR, Thompson NJ, Denic V, Bhamidipati A, Punna T, Ihmels J, Andrews B, Boone C, Greenblatt JF et al. 2005. Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. *Cell* **123**: 507-519.
- Schwartz S, Ast G. 2010. Chromatin density and splicing destiny: on the cross-talk between chromatin structure and splicing. *The EMBO journal* **29**: 1629-1636.
- Semlow DR, Blanco MR, Walter NG, Staley JP. 2016. Spliceosomal DEAH-Box ATPases Remodel Pre-mRNA to Activate Alternative Splice Sites. *Cell* **164**: 985-998.
- Shao W, Kim HS, Cao Y, Xu YZ, Query CC. 2012. A U1-U2 snRNP interaction network during intron definition. *Molecular and cellular biology* **32**: 470-478.
- Shevchenko A, Roguev A, Schaft D, Buchanan L, Habermann B, Sakalar C, Thomas H, Krogan NJ, Shevchenko A, Stewart AF. 2008. Chromatin Central: towards the comparative proteome by accurate mapping of the yeast proteomic environment. *Genome biology* **9**: R167.

- Sims RJ, 3rd, Millhouse S, Chen CF, Lewis BA, Erdjument-Bromage H, Tempst P, Manley JL, Reinberg D. 2007. Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Molecular cell* **28**: 665-676.
- Spiczki M. 2000. Where does fission yeast sit on the tree of life? *Genome biology* **1**: REVIEWS1011.
- Sisodia SS, Sollner-Webb B, Cleveland DW. 1987. Specificity of RNA maturation pathways: RNAs transcribed by RNA polymerase III are not substrates for splicing or polyadenylation. *Molecular and cellular biology* **7**: 3602-3612.
- Sorenson MR, Stevens SW. 2014. Rapid identification of mRNA processing defects with a novel single-cell yeast reporter. *Rna* **20**: 732-745.
- Soucek S, Zeng Y, Bellur DL, Bergkessel M, Morris KJ, Deng Q, Duong D, Seyfried NT, Guthrie C, Staley JP et al. 2016. The Evolutionarily-conserved Polyadenosine RNA Binding Protein, Nab2, Cooperates with Splicing Machinery to Regulate the Fate of pre-mRNA. *Molecular and cellular biology*.
- Spingola M, Grate L, Haussler D, Ares M, Jr. 1999. Genome-wide bioinformatic and molecular analysis of introns in *Saccharomyces cerevisiae*. *Rna* **5**: 221-234.
- Sridharan V, Heimiller J, Singh R. 2011. Genomic mRNA profiling reveals compensatory mechanisms for the requirement of the essential splicing factor U2AF. *Molecular and cellular biology* **31**: 652-661.
- Staley JP, Guthrie C. 1998. Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **92**: 315-326.

- Stepankiw N, Raghavan M, Fogarty EA, Grimson A, Pleiss JA. 2015. Widespread alternative and aberrant splicing revealed by lariat sequencing. *Nucleic acids research* **43**: 8488-8501.
- Subramanian V, Fields PA, Boyer LA. 2015. H2A.Z: a molecular rheostat for transcriptional control. *F1000prime reports* **7**: 01.
- Suto RK, Clarkson MJ, Tremethick DJ, Luger K. 2000. Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nature structural biology* **7**: 1121-1124.
- Tardiff DF, Lacadie SA, Rosbash M. 2006. A genome-wide analysis indicates that yeast pre-mRNA splicing is predominantly posttranscriptional. *Molecular cell* **24**: 917-929.
- Tardiff DF, Rosbash M. 2006. Arrested yeast splicing complexes indicate stepwise snRNP recruitment during in vivo spliceosome assembly. *Rna* **12**: 968-979.
- Tolstorukov MY, Goldman JA, Gilbert C, Ogryzko V, Kingston RE, Park PJ. 2012. Histone variant H2A.Bbd is associated with active transcription and mRNA processing in human cells. *Molecular cell* **47**: 596-607.
- Tseng CK, Liu HL, Cheng SC. 2011. DEAH-box ATPase Prp16 has dual roles in remodeling of the spliceosome in catalytic steps. *Rna* **17**: 145-154.
- Van Alst NE, Wellington M, Clark VL, Haidaris CG, Iglewski BH. 2009. Nitrite reductase NirS is required for type III secretion system expression and virulence in the human monocyte cell line THP-1 by *Pseudomonas aeruginosa*. *Infect Immun* **77**: 4446-4454.
- Vanichkina DP, Schmitz U, Wong JJ, Rasko JEJ. 2017. Challenges in defining the role of intron retention in normal biology and disease. *Seminars in cell & developmental biology*.
- Vijayraghavan U, Company M, Abelson J. 1989. Isolation and characterization of pre-mRNA splicing mutants of *Saccharomyces cerevisiae*. *Genes & development* **3**: 1206-1216.

- Wahl MC, Will CL, Luhrmann R. 2009. The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**: 701-718.
- Wan R, Yan C, Bai R, Huang G, Shi Y. 2016. Structure of a yeast catalytic step I spliceosome at 3.4 Å resolution. *Science* **353**: 895-904.
- Wang J, Zhang J, Li K, Zhao W, Cui Q. 2012. SpliceDisease database: linking RNA splicing and disease. *Nucleic acids research* **40**: D1055-1059.
- Warner JR. 1999. The economics of ribosome biosynthesis in yeast. *Trends in biochemical sciences* **24**: 437-440.
- Wassarman DA, Steitz JA. 1992. Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing. *Science* **257**: 1918-1925.
- Weber CM, Henikoff JG, Henikoff S. 2010. H2A.Z nucleosomes enriched over active genes are homotypic. *Nature structural & molecular biology* **17**: 1500-1507.
- Wilmes GM, Bergkessel M, Bandyopadhyay S, Shales M, Braberg H, Cagney G, Collins SR, Whitworth GB, Kress TL, Weissman JS et al. 2008. A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. *Molecular cell* **32**: 735-746.
- Wood V, Gwilliam R, Rajandream MA, Lyne M, Lyne R, Stewart A, Sgouros J, Peat N, Hayles J, Baker S et al. 2002. The genome sequence of *Schizosaccharomyces pombe*. *Nature* **415**: 871-880.
- Wood V, Harris MA, McDowall MD, Rutherford K, Vaughan BW, Staines DM, Aslett M, Lock A, Bahler J, Kersey PJ et al. 2012. PomBase: a comprehensive online resource for fission yeast. *Nucleic acids research* **40**: D695-699.

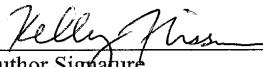
- Workman JL, Kingston RE. 1998. Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annual review of biochemistry* **67**: 545-579.
- Wu WH, Alami S, Luk E, Wu CH, Sen S, Mizuguchi G, Wei D, Wu C. 2005. Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. *Nature structural & molecular biology* **12**: 1064-1071.
- Yan C, Hang J, Wan R, Huang M, Wong CC, Shi Y. 2015. Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science* **349**: 1182-1191.
- Yan C, Wan R, Bai R, Huang G, Shi Y. 2016. Structure of a yeast activated spliceosome at 3.5 Å resolution. *Science* **353**: 904-911.
- . 2017. Structure of a yeast step II catalytically activated spliceosome. *Science* **355**: 149-155.
- Yofe I, Zafir Z, Blau R, Schuldiner M, Tuller T, Shapiro E, Ben-Yehezkel T. 2014. Accurate, model-based tuning of synthetic gene expression using introns in *S. cerevisiae*. *PLoS genetics* **10**: e1004407.
- Zhou J, Fan JY, Rangasamy D, Tremethick DJ. 2007. The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. *Nature structural & molecular biology* **14**: 1070-1076.
- Zlatanova J, Thakar A. 2008. H2A.Z: view from the top. *Structure (London, England : 1993)* **16**: 166-179.
- Zofall M, Fischer T, Zhang K, Zhou M, Cui B, Veenstra TD, Grewal SI. 2009. Histone H2A.Z cooperates with RNAi and heterochromatin factors to suppress antisense RNAs. *Nature* **461**: 419-422.

**Publishing Agreement**

*It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.*

***Please sign the following statement:***

*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.*

  
\_\_\_\_\_  
Author Signature

09/01/17  
Date