UNIVERSITY OF CALIFORNIA

SANTA CRUZ

**Mechanistic studies using a humanized yeast strain sensitive to splicing inhibitors**

A dissertation submitted in partial satisfaction

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MOLECULAR, CELL, AND DEVELOPMENTAL BIOLOGY

by

Oarteze Hunter

December 2022

_____
Professor Manuel Ares Jr., Chair

_____
Associate Professor Angela Brooks

_____
Professor Melissa Jurica

_____
Professor Tracy Johnson

_____
Peter F. Biehl
Vice Provost and Dean of Graduate Studies

## Table of Contents

**List of Figures**

**Figure 4-2:** Plot of Intron Accumulation Index for blocked splicing caused by temperature-shifted *prp4-1* mutant yeast calculated by Clark et al. 2002

## List of Tables

## Dedication and Acknowledgements

I would like to take this time and space to thank all who have contributed to this work. First, I want to thank Dr. Manuel Ares Jr. for taking a chance on me. You have been my greatest inspiration and it has been an honor and a pleasure to work with you and be taught by you. Thank you, Manny. I will never forget our conversations, your lessons, and your funny metaphors. That time we burned my poster when I was feeling down will always stay with me. I promise I will make you proud as I grow as a scientist.

The Ares lab has been instrumental to this work. That includes current members Jason Talkish, PhD, Jen Quick-Clevelend, PhD, our lab mom Haller Igel, Research Associate Gena Lapointe, Graduate Student Michelle Seiwald, Junior Specialist Justin Sim, and my Undergrad Owen Jump. Thank you to John Paul Donohue and Sol Katzman for your help with the bioinformatics portions of this work. You all have been amazing and a family to me.

My PhD Thesis Committee has been an amazing team. I want to thank Dr. Angela Brooks, Dr. Melissa Jurica, and Dr. Tracy Johnson. I am so glad I met you and shared space with you all. You all have made me into the scientist I am today. Thank you and I can't wait for you all to see me grow as a scientist.

I would also like to thank my Qualifying Exam Committee: Dr. Needhi Bhalla, Dr. Alan Zahler, Dr. Seth Rubin, and Dr. Harry Noller. That exam was the most challenging in my life and I'm glad I saw it through with you all.

Thank you to Yuli Ortega, Xingci Situ, and the STEM Diversity Family! Much love.

My amazing (and wild) family. I don't know how I would have done it without yall. To my momma Brenda Brown. To Micco, Bre, Sean, Sikia, and Jacqueze. I love each one of you. I am so grateful to have yall's support and love surrounding me every day.

To the little ones (for now), Marty, Talia and Teyah, I love you. I can't wait to be there as you all grow up. This dissertation is dedicated to my niece Talia Porté. I love you so much.

Next, I want to thank and acknowledge three high school teachers. Without them I don't think I would be at UCSC. They were instrumental in getting into college from Hawthorne High School. I am forever grateful for the advocacy and dedication to see me through in my career in STEM. They are Arian Whitely, my first Black teacher. You taught me World History and AP Economics. Katie Kondo, who taught me Math (algebra and precalculus). I am forever grateful for when you advocated for me to get into Mr. T's summer geometry class so that I could be caught up for AP Calc senior year. You are truly an inspiration to me and I'm glad to know you. Lastly, Daniel Barnes who taught me Chemistry. Thank you for being one of the first to spark my interest in chemistry and science in general. As a biologist now, I still remember every chemistry lesson and the fun ways you taught it. Thank you all for being exceptional mentors to me and walking me through my

teenage years. So glad we are on a first name basis now lol.

I cannot end this section without thanking all my friends so I'm just gonna name drop a few: Reed and Daniela (Aquarius Support Group). Nestor Gutierrez and Noorafi Almajid. Anthony Smith (you're next!!!!), Dallas and Felicia Mickey (and family), Amir Williams. Alberto Reyes, Jon, Justin, Sari, Daniel, J, Nene, Steph, Jacques, Desi (you're next!!!). Lawrence Milan. Alison Hayes. Aitanna Parker and Lindsey. Cristobal, Angela, and Christina (ya'll next!!!). Randy. Asia Hilario and Omnia. Jon, Hasbun, and Jordan. Camila. Roberto. Richard and Dee. Andres McCree. Ana, Barah, and Adra (my rave family). Chulx. Jonathan Guzman, Sezy Reyes (and family), Alex Vargas (and family), Ray Cornejo, and Vivi Herrera. Brent and Enrique. Vanessa Iracheta and Alma Davis. Alex Jaramillo. Anise and Courtney Young. Kimberly and Destinee. Imani Bradford. Benji. Mariah. Dr. Jessie Suzuki, Dr. Londen Johnson. Gabby and Mercedes. My wife Caro and George. Steven Ruschman and Dallas Plumley. Adrian Davey. Andrew Olander (you're next!!!). Ruben, Miguel, Jose, and John. Marcy and Jeff. My former therapist Omar Geray. And this is not an exhaustive list. I'm so happy to be on this Earth living at the same time as you all.

## Abstract

### Mechanistic studies using a humanized yeast strain sensitive to splicing inhibitors

Oarteze Hunter

Gene expression is a critical process for cell viability and growth. One step in gene expression is RNA splicing, which removes non-coding introns from mRNA during gene expression. This work aims to answer the following outstanding questions in the field: 1) What are the consequences of blocking splicing for individual introns and 2) how does global gene expression change when splicing is blocked? Proper recognition of branchpoint sequence (BPS) and splice sites in the intron are required for splicing. U2 small nuclear RNA (snRNA) base pairing to the BPS is an early step in this recognition. The U2 snRNA-BPS helix contains a bulged branchpoint adenosine (BPA) that binds a pocket formed by HEAT repeats 15 and 16 of the budding yeast HSH155 protein. In human SF3B1 (homolog of HSH155), this pocket also binds splicing inhibitors such as Pladienolide-B (Plad-B). *HSH155* differs from SF3B1 in the amino acids lining the pocket, rendering yeast resistant to Plad-B. To study the consequences of splicing inhibition on individual introns and gene expression, we created a yeast strain sensitive to inhibitors. We replaced the *HSH155* BPA binding pocket with that of SF3B1, essentially “humanizing” splicing in budding yeast. We called this allele *hsh155- ds* for drug sensitive hsh155 mutant. *In vivo*, splicing inhibition occurs minutes after addition of Plad-B. *In vitro* splicing extracts from

*hsh155-ds* yeast treated with inhibitors show that spliceosomes are blocked at prespliceosomes (A-complex). RNA-sequencing of cells carrying WT *HSH155* treated with Plad-B produces no significant splicing changes compared to untreated cells carrying either *HSH155* or *hsh155-ds*, showing that (1) wild type cells are completely resistant to Plad-B and (2) the humanizing mutation does not affect splicing. In contrast, treatment of the *hsh155-ds* strain with Plad-B results in a dramatic accumulation of pre-mRNA for intron-containing genes. The effect of a block to splicing reverberates across the transcriptome, leading to gene expression changes transcriptome wide. Intronless ribosomal protein genes are downregulated, presumably in response to reduced expression of intron-containing ribosomal protein genes. The ability to chemically block splicing in a cell with few introns may reveal conserved regulatory connections between splicing and other steps of gene expression.

# PROLOGUE

Since the discovery of splicing in the 1970s, there has been a huge effort toward understanding its mechanism, its regulation, and its implications in disease. The work described here continues that tradition by analyzing the repercussions of splicing inhibition in yeast cells. An analysis of gene expression under splicing inhibition conditions will reveal what gene regulatory networks depend on splicing to function. We will be able to see more specifically how important RNA splicing is for the cell's overall gene expression program and how this program relates to disease.

To answer new questions about splicing, a solid understanding of the current state of splicing research is necessary. Thus, Chapter 1 lays the foundations for splicing and provides context about its role in gene expression. I will discuss the two-step transesterification mechanism of the spliceosome class of introns seen in all eukaryotes including yeast and human cells. I will focus on a key factor in the spliceosome, SF3B1, and its recognition of a sequence within the intron called the branchpoint sequence early during the splicing cycle. Next, I will discuss how recurrent cancer mutations in SF3B1 have been shown to alter the recognition of the branchpoint sequence. Recognition of an aberrant branchpoint sequence by mutant SF3B1 has downstream implications for the final spliced product and cancer.

Later in the chapter, I will pivot to yeast splicing. The yeast species *S. cerevisiae* has been and continues to be a good model system for

studying splicing. As such, we must understand how yeast introns are structured and how yeast splicing is regulated. Yeast splicing is comparatively simpler, but is still highly similar, at its core, to human splicing. It is this conservation that makes the yeast splicing system a powerful tool to understand the fundamental mechanisms of human splicing. Finally, Chapter 1 will continue with a brief history of the ways in which splicing has been inhibited. Splicing inhibition has been used to study many aspects of splicing. For example, discreet stages in the splicing cycle can be inhibited and therefore studied in more detail. A powerful class of small molecule inhibitors has recently drawn much attention. First shown to inhibit the growth of cancer cells in 2004, Pladienolide-B (Plad-B) emerged as a new candidate for cancer therapeutics. Researchers over the next decade focused on pinpointing where Plad-B binds in the cell, while also creating derivatives of Plad-B to maximize its cancer-reducing effects. We now know that these molecules bind the spliceosome, specifically SF3B1, to inhibit splicing. Studies on the chemical inhibition of splicing offer new insights into still unknown aspects of splicing. For example, the co-evolution of inhibitors and spliceosomes of various eukaryotic organisms has been vastly understudied. Another example is the consequences of blocked splicing on gene expression overall. Chapters 2-5 explores the consequences of blocked splicing on the gene expression landscape of *S. cerevisiae*. Chapter 1 concludes by introducing a new tool for studying

splicing inhibition in *S.* cerevisiae and why *S. cerevisiae* is the best model for this area of study.

Chapter 2 begins by describing how I built a new yeast strain to study splicing inhibition. To create inhibitor sensitive yeast, I used CRISPR/Cas9 genome-editing technology to “humanize” the yeast sequence that codes for the target of Plad-B and other inhibitors. This was feasible due to the high sequence similarity between human and yeast SF3B1. The region of SF3B1 that binds Plad-B differs from yeast by only a dozen amino acids. After confirming the humanizing mutations and checking that the protein is being expressed, I performed a time course experiment to analyze splicing immediately after addition of Plad-B to the humanized yeast. Not only was splicing in humanized yeast sensitive to Plad-B as shown by the presence of unspliced RNA, but inhibition was immediate. I concluded that splicing inhibition is fast and strong in the humanized strain.

After seeing how fast and strong splicing inhibition was *in vivo*, I wanted to see if I could measure splicing inhibition in a test tube. Thus, the next portion of Chapter 2 discusses my collaboration with the Ares lab postdoctoral researcher Jason Talkish. We performed two main *in vitro* experiments using the humanized strain. First, we tested splicing of the ACT1 intron using splicing extracts isolated from humanized yeast using denaturing gel electrophoresis. Next, we performed a similar reaction to observe spliceosome complex accumulation upon splicing inhibition by native gel

electrophoresis. As is the case *in vivo*, splicing is inhibited *in vitro*, and complexes are blocked at the step where SF3B1 is first introduced to the intron. Calculation of the concentration of inhibitor needed to inhibit splicing by 50% (IC50) in yeast is similar to that observed in human extracts. This was much needed evidence that humanized yeast are similarly sensitive to inhibitors as human cells, establishing these inhibitors as useful tools for our analysis.

Chapter 3 focuses on analyzing splicing inhibition by high throughput sequencing of total cellular RNA. RNA sequencing is a powerful technology that measures levels of RNA and thus gives an accurate picture of the amount of spliced and unspliced RNA in a sample. For this sequencing experiment, we collaborated with Scott Keursten and Illumina Inc. (Madison, WI) who performed the sequencing. We obtained raw reads from Scott and began processing them. Sol Katzmann aligned reads to the yeast genome and performed both differential gene expression analysis and identifying and quantifying splicing junctions.

Two main questions will be explored in Chapter 3. First, I will characterize the effect of the drug on wildtype yeast, ensuring that it is resistant. Then I will similarly characterize the effect of the drug on humanized yeast, ensuring that it is sensitive at different concentrations. I will explore whether all introns in yeast are equally sensitive to the drug, or whether some might be more resistant than others. Secondly, I will analyze

gene expression changes upon splicing inhibition. I found that intron-containing ribosomal protein genes (RPGs) were the most resistant to Plad-B inhibition compared to other genes. RPGs were also downregulated when splicing was blocked. Interestingly, I also saw the downregulation of intronless ribosomal protein genes, a regulatory connection between intron-containing and intronless gene expression in response to a splicing block. I also detected the downregulation of the purine synthesis pathway genes in my analysis which mostly comprises intronless genes, again, indicating that the effect of a block to splicing reverberates across the transcriptome. Next, we wanted to ask about co-transcriptional splicing. We now know that splicing happens at the same time the RNA is being made during transcription, in other words co-transcriptionally. We next sought out to measure co-transcriptional sensitivity to Plad-B using single molecule intron tracking (SMIT). This experiment was a collaboration with Ares lab postdoctoral researcher Jen Quick-Cleveland and the Karla Neugebauer lab at Yale University (New Haven, CT). I started by incubating wild type and humanized yeast with Plad-B for 15 minutes. I gave these samples to Dr. Quick Cleveland who performed a chromatin-associated RNA extraction following the protocol recommended by Karla Neugebauer's lab. The chromatin-associated nascent RNA was prepared for sequencing in a way that captures both the position of the RNA polymerase on the gene and the splicing state of the nascent RNA it is synthesizing as it is being transcribed

by RNA polymerase II. Tara Alpert, a graduate student in the Neugebauer lab at the time, performed the bioinformatic analysis. The results are plotted as the fraction of RNA spliced relative to the position of RNA polymerase II on the gene. We studied 62 intron-containing genes using this method. SMIT, along with total RNA sequencing, revealed just how important ribosomal protein gene expression is to the cell, especially when splicing is blocked. Finally, Chapter 5 explores important implications to the data collected and what could be done to strengthen the discoveries of this work.

# Chapter 1

# Introduction

## 1.1 Gene expression and RNA splicing

Gene expression is fundamental to maintaining the cell in a rapidly changing environment. Segments of DNA sequences called genes are converted to RNA that can either perform a certain function for the cell (non-coding RNA) or be translated from messenger RNA (mRNA) into protein. Gene expression covers a range of processes including identification of a gene by RNA polymerase and other proteins, generating an RNA molecule, RNA modification and processing, and translating the information into a functional protein product. These gene products then go on to perform many of the cell's functions. Alterations to any of these steps can lead to disease.

Gene expression involves the processing of RNA molecules that are created by RNA polymerase. RNA splicing is one critical process during gene expression. A gene may contain non-coding sequences called introns that interrupt the coding sequence, called exons (Gilbert 1978). The remaining exon sequences are joined together. There are different types of introns in nature, and they are removed by different mechanisms. Some introns in nature remove themselves through RNA catalysis reactions (self-splicing). Introns in nuclear protein-coding genes represent a separate class and are removed with the help of the

spliceosome. They are spliceosomal introns.

### 1.2 RNA Splicing Mechanism by the spliceosome

The biochemical removal of spliceosomal introns is like self-splicing introns and involves two chemical reactions (Sharp 1985). First, a specific RNA residue called the branchpoint A (BPA) within the intron uses its 2' hydroxyl group to perform a nucleophilic attack on the phosphate group at the 5' GU sequence splice site (5'SS) at the beginning of the intron (Figure 1-1). Second, the free 5' exon uses its 2' hydroxyl group on the last nucleotide to perform a similar nucleophilic attack on the phosphate group at the 3' AG sequence splice site (3' SS). The products from these two reactions are the spliced mRNA and the intron in the form of a lariat (Padgett et al. 1984). Splicing is performed by a large RNA-protein complex called the spliceosome. The spliceosome assembles on each intron from scratch. Specifically, it brings together the parts of the intron that will participate in the two chemical reactions, while coordinating magnesium ions (Wahl et al. 2009, Kastner et al. 2019). Spliceosomes are composed of 5 small nuclear RNAs (snRNAs), U1, U2, U4, U5, U6, and about 100 proteins (Wahl et al. 2009, Kastner et al. 2019) Each snRNA molecule is associated with a set of proteins to form small nuclear ribonucleoprotein particles (snRNPs) that participate in recognizing the intron and removing it. The RNA-RNA, RNA-protein, and protein-protein complexes that compose the

spliceosome require dynamic, stepwise arrangements for proper function throughout the splicing cycle. The spliceosome undergoes many ordered complexes starting at E to A, B, C, P, and ending with the intron-lariat (ILS) complex. Once both chemical reactions are complete, the spliceosome lets go of the spliced mRNA, the intron is degraded, and the spliceosome components are recycled for the next cycle of splicing of a new intron. Formation of the early complexes E and A are important for later complexes (B and C spliceosomes).

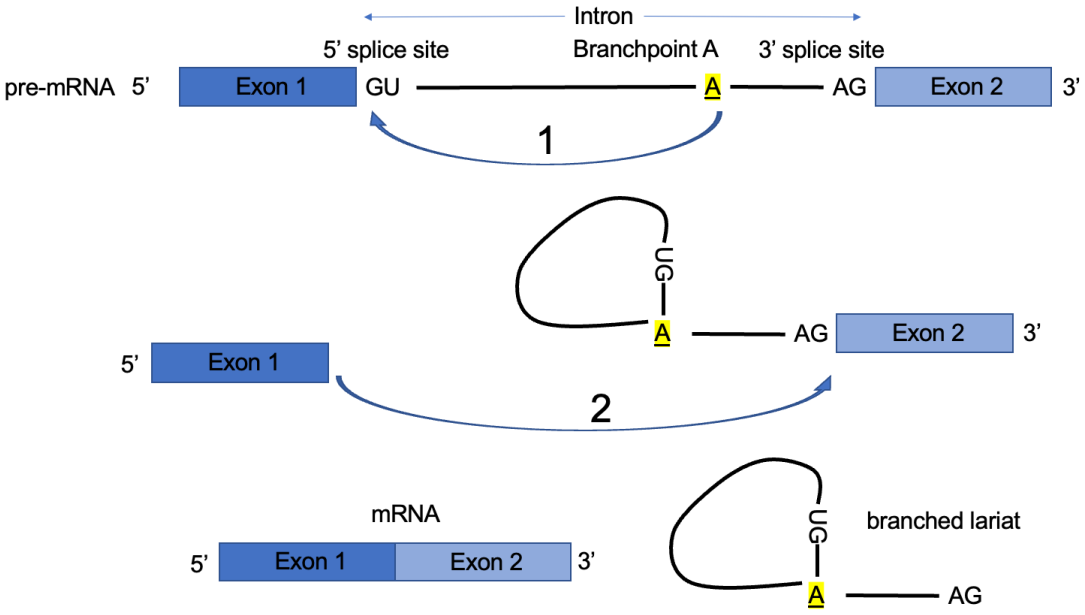



**Figure 1-1:** General two-step splicing mechanism for spliceosomal introns.

### 1.3 Commitment and pre-spliceosome complex formation

U1 snRNP uses its snRNA to base pair with a 5’ GU sequence (5’ splice site) that starts the intron (Plaschka et al. 2018, Zhuang and Weiner 1986). The branchpoint and 3'SS is also recognized. In multicellular organisms, SF1 protein and U2AF complex bind the conserved BPA and 3’SS AG sequence respectively (Berglund et al. 1998, Zamore et al. 1992). Budding yeast contains Msl5 and Mud2 that perform similar functions to recognize its introns (Wang et al. 2008). This is known as the commitment complex (also called the E complex) (Figure 1-2).

Initial recognition of the 5’SS, BPA, and 3’SS is followed by addition of the U2 snRNP. SF1 is replaced by the U2 snRNP by Prp5 and Sub2 ATP helicases (Liang and Cheng 2015, Kistler and Guthrie 2001), forming the prespliceosome (also called A complex) (Figure 1-2). Accurate identification of the branchpoint sequence (BPS) during the formation of the prespliceosome early in the splicing cycle is important for gene expression. Branchpoint identification begins with the ATP-independent binding of the U1 snRNP to the pre-mRNA via the 5’ SS GU dinucleotide of the intron. Branchpoint binding protein Bbp (also called Msl5) and Mud2 binds the BPS and polypyrimidine tract (py-tract) respectively at the 3’ end of the mRNA (Wang et al. 2008, Krämer and Utans 1991, Kao et al. 2021). DEAD-box helicases Prp5 and Sub2 aids

the removal of Bbp-Mud2 and allows for extended pairing of U2 snRNA with the pre-mRNA intron (Perriman and Ares 2010, Kistler and Guthrie 2001, Liang and Cheng 2015). ATP hydrolysis by Prp5 causes extended base pairing of the U2 snRNA-pre-mRNA forming a bulged BPA that will act as the nucleophile in the first step of splicing.

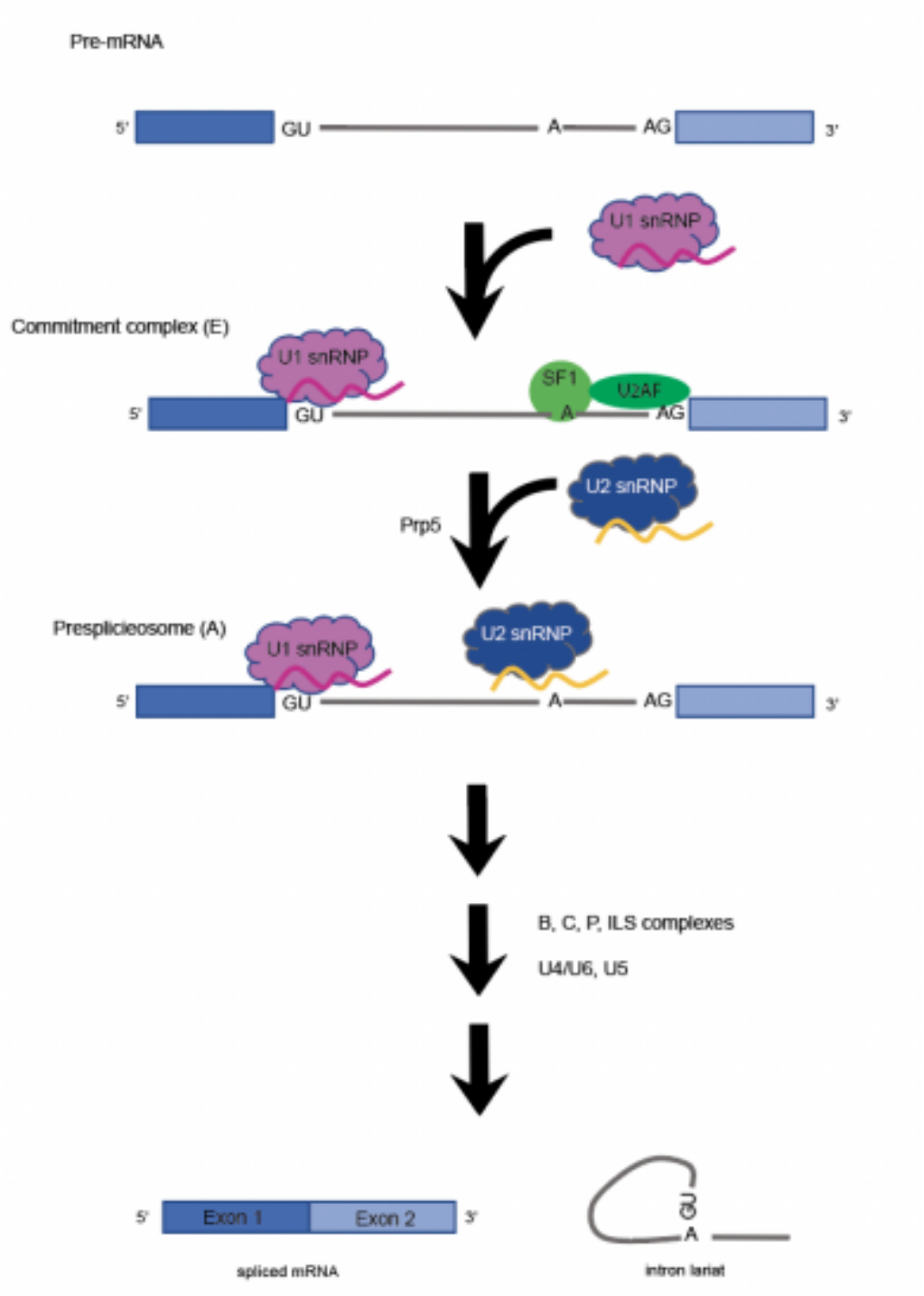



**Figure 1-2:** Early spliceosomal complexes recognize splice sites within the intron.

### 1.4 Recognition of the branch point adenosine during early splicing

Mammalian SF3B1 (also called SF3B155 or SAP155), and its budding yeast homolog Hsh155, of the SF3B complex are thought to stabilize the assembly of the U2 snRNP to the BPS (Gozani et al. 1998, Plaschka et al. 2019). Initial recruitment of the U2 snRNP leads to initial nucleotide contacts of U2 snRNP to the BPS that is mediated by a branch point stem-loop (BSL) structure within U2 snRNA (Perriman and Ares 2010). Removal of TAT-SF1 bound to SF3B1 (Cus2 bound to HSH155 in yeast) by the ATPase Prp5 allows the formation of an extended RNA helix formed by the U2 snRNA and pre-mRNA intron at the BPS bound by the SF3b1 HEAT repeat domain (Huntingtin, Elongation factor 3, protein phosphatase 2A, TOR1). *SF3B1*/Hsh155 HEAT repeats comprise approximately 22 antiparallel helix-turn-helix motifs that stabilize the U2 snRNA-BPS helix. The formation of this RNA helix complex bulges the nucleophilic BPA into a pocket formed by HEAT repeats 15 and 16 and PHF5A. The bulged BPA will participate in the first chemical step of splicing (Yan et al. 2016) (Figure 1-3).

### 1.5 Recurring cancer-associated mutations in SF3B1

Recurrent hotspot mutations in the HEAT repeats of *SF3B1* are associated with a variety of cancers such as myelodysplastic syndrome (MDS) and chronic lymphocytic leukemia (CLL), (Quesada et al. 2011, Yoshida et al. 2011, Wang et al. 2011), uveal melanoma (Harbour et al. 2013, Martin et al. 2013), breast (Stephens et al. 2012), and pancreatic cancer (Biankin et al. 2012, Bailey et al. 2016). K700E mutation (blue residue, Figure 1-3), in the sixth HEAT repeat (Hsh155 has 22 HEAT repeats), is the most recurring hematologic aberration and appears to result in recognition of a cryptic 3' splice site AG (3' SS) between 12 and 24 nucleotides upstream of the canonical 3'SS (DeBoever et al. 2015, Darman et a. 2015, Alsafadi et al. 2016) (Figure 1-4). The cryptic 3'SS coincides with a cryptic, yet functional, BPS-like sequence about 11-14 nucleotides upstream of the cryptic 3' SS (Darman et a. 2015, Alsafadi et al. 2016). The use of both a cryptic BPS and a 3'SS at least 7 nucleotides downstream results in the formation of a transcript with retained intron sequence that can be used to form an aberrant protein or be degraded by nonsense-mediated decay (NMD) (Figure 1-4).

Different explanations have been put forth concerning the alternative BPS/3'SS use including that hotspot mutations stabilize SF3B1 interactions with a cryptic BPS or destabilize SF3B1 interactions

with a canonical BPS (Darman et al. 2015; Alsafadi et al. 2016). MDS mutations placed in budding yeast have been shown to perturb splicing by increasing or decreasing accurate BPS selection, possibly by altered interaction with Prp5 (Carrocci et al. 2017; Tang et al. 2016). This could mean that BPS selection is driven by conformational changes in SF3B1 at or before the time the extended helix is formed. Thus, correct identification of the BPS during pre-spliceosome formation is important for generating functional mature mRNAs.

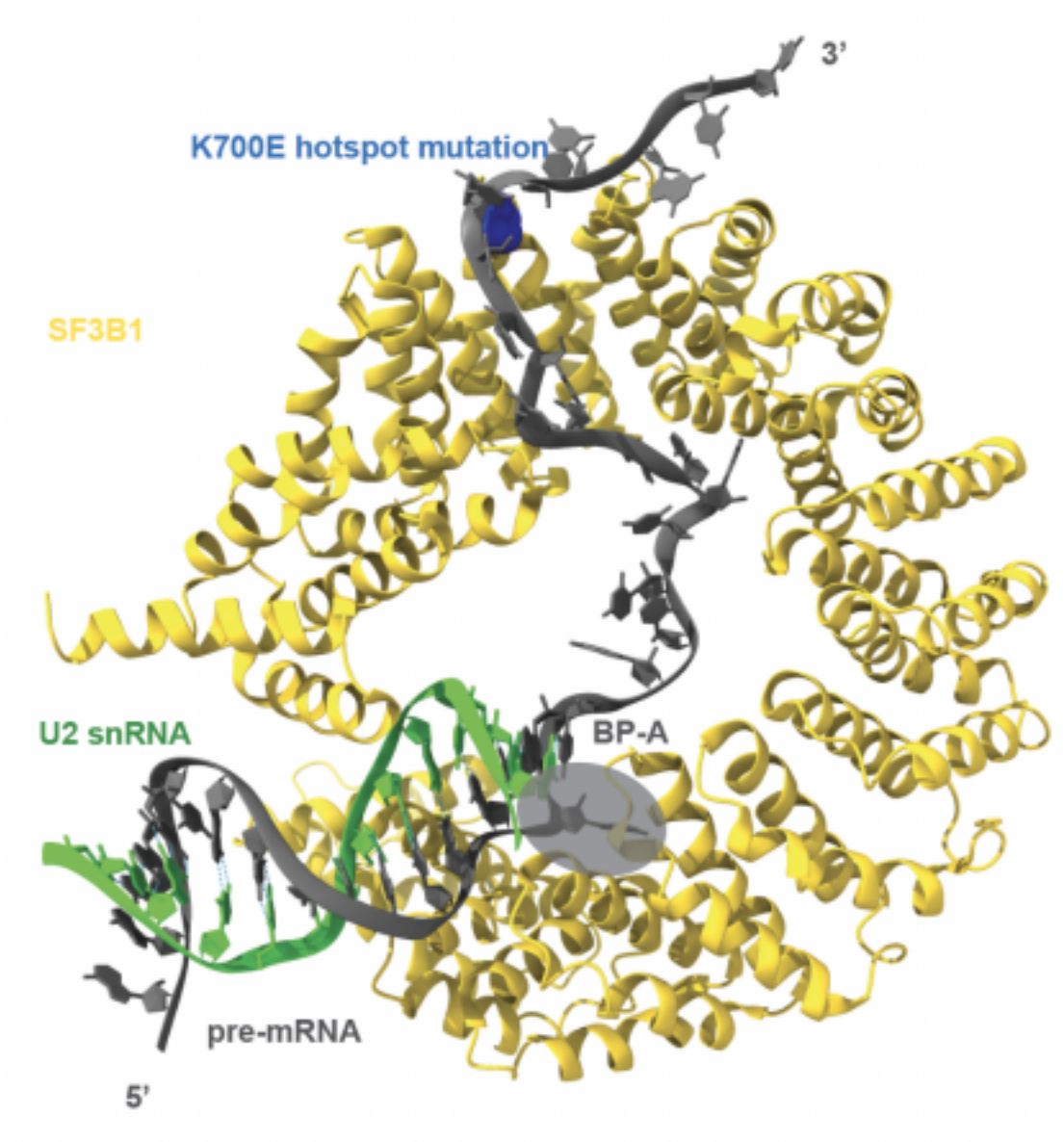



**Figure 1-3:** Cryo-EM model of SF3B1 bond to U2 snRNA-intron helix. SF3B1 shown in gold, U2 snRNA in green, and pre-mRNA intron in gray. The BPA (shaded) is shown building into a pocket partly formed by HEAT 15 and 16 of SF3B1. K700E hotspot mutation is shown as a blue sphere near the 3' end of the intron. Figure adapted from PDB: 5GM6.

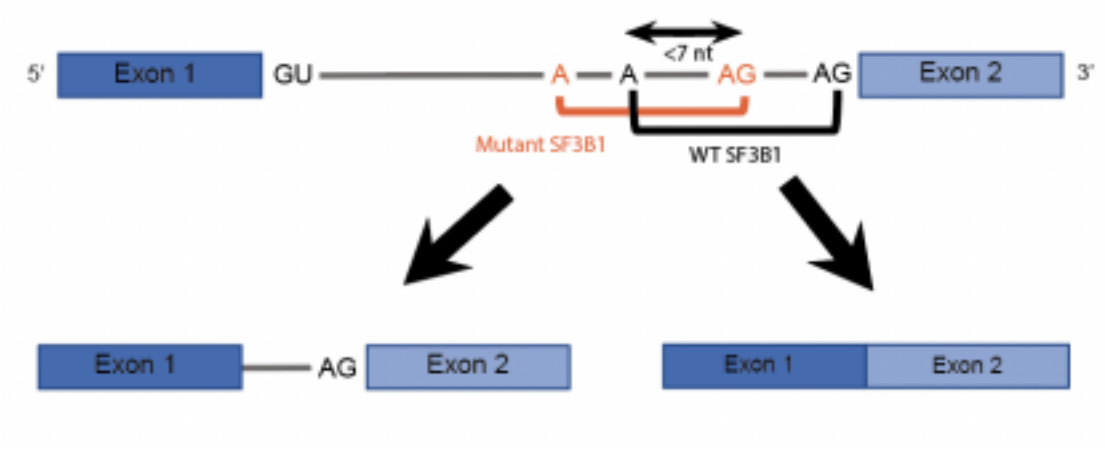



**Figure 1-4:** Mutant SF3B1 uses an alternative BPA/3' SS pair (orange) that leads to partial intron retention after splicing. The use of a BP-A/3' SS pair that are less than 7 nucleotides apart is not permitted due to steric hindrance of the spliceosome.

### 1.6 Yeast splicing regulation and introns

Introns are a staple of eukaryotic genes. Examination of the so-called “intronome” of *S. cerevisiae* might lead one to believe that introns are not a critical component of gene regulation. Present in only about 4-5% of protein coding genes, introns play a significant role in regulating gene expression in yeast (Ares et al. 1999). There is some evidence to suggest that the Sacharromycotica subphylum, which includes *S. cerevisiae*, experienced a widespread loss of introns from an intron-rich ancestor over evolutionary time (Irimia et al. 2007). This likely coincided with the loss of accessory factors of the spliceosome present in higher eukaryotes (Sales-Lee et al. 2021). About a third of the remaining genes with introns occupy the 5’ ends of ribosomal protein genes (Ares et al. 1999). Furthermore, why ribosomal protein genes seem resistant to the loss of introns in this scenario also remains a mystery.

Eukaryotes contain two ribosomal subunits, 60S and 40S. The 60S is made up of 3 rRNA molecules (28/25S, 5.8S, and 5S) and 46 ribosomal proteins, while the 40S is made up of one 18S rRNA molecule and 33 ribosomal proteins (Peña et al. 2017). Making ribosomes is a resource-intensive process that requires coordinated gene expression of these components. In addition, about 150 rDNA repeats and >100 small nucleolar RNAs (snoRNAs) and >200 assembly factors help generate ribosomes (Kressler et al. 2010). In *S. cerevisiae*, most ribosomal protein genes (RPGs) are made from duplicated paralogs (Wapinski et al 2009, Marcet-Houben and

Gabaldón 2015).

Deletions of individual paralogs show different phenotypes even though the protein sequence is similar, suggesting that they may perform distinct functions (Komili et al. 2007). Interestingly, it has been shown that paralogs with introns are differentially expressed and use splicing as a regulatory tool (Komili et al. 2007, Parenteau et al. 2011, Ghulam et al. 2020). For example, a highly expressed major paralog may bind to the intron of its minor counterpart to prevent its splicing, resulting in asymmetrical expression (Figure 1-5) (Komili et al. 2007, Parenteau et al. 2011, Ghulam et al. 2020). The crosstalk between ribosome biogenesis and splicing is vital.

In rapidly growing cells, 90% of splicing is dedicated to RPGs (Warner 1999). This observation raises the question of whether the splicing of ribosomal protein genes (RPGs) is regulated. Indeed, there is evidence that splicing plays a role in RPG regulation. For example, genetic experiments revealed the ribosomal protein L30 (eL30) regulates the splicing of its own transcript (Vilardell and Warner 1997). eL30 binds its own pre-mRNA (*RPL30*) which causes it to be exported to the cytoplasm. mRNA is then released from eL30 and is turned over by NMD (Vilardell et al. 2000). Furthermore, cap-binding protein Cbp80 was found to repress U2 snRNP recruitment to the RPL30 transcript (Bragulat et al. 2010). The regulation of these and other ribosomal protein genes in yeast underscores the importance of splicing in ribosome protein expression. One reason it might be beneficial for yeast to regulate the splicing of ribosomal protein genes

strongly is to control protein synthesis in response to a rapidly changing environment. For example, ribosomal pre-mRNAs “compete” for the attention of the spliceosome; shutting down the transcription of ribosomal mRNAs by rapamycin results in increased splicing of non-ribosomal mRNAs (Talkish et al. 2019b, Munding et al. 2013). Under such conditions, where ribosomal protein genes are repressed in budding yeast, splicing can be seen at previously unannotated splice sites which could also provide the raw material needed for new introns to evolve in the genome. These important lines of evidence provide clues into why splicing is so restricted to ribosomal protein genes in yeast.

Little is known about how the proposed intron-rich ancestor gave rise to intron-poor yeast species. One possible mechanism for this loss of introns is through the recombination of reverse transcribed RNA (cDNA) with genomic DNA (Zhu and Niu 2013, Fink 1987). However, other possible factors such as small molecules in the yeast environment have received limited attention. Work over the past 18 years has focused on chemical inhibitors of splicing in humans, worms, yeast, and plant organisms that could illuminate the role splicing plays in gene expression.

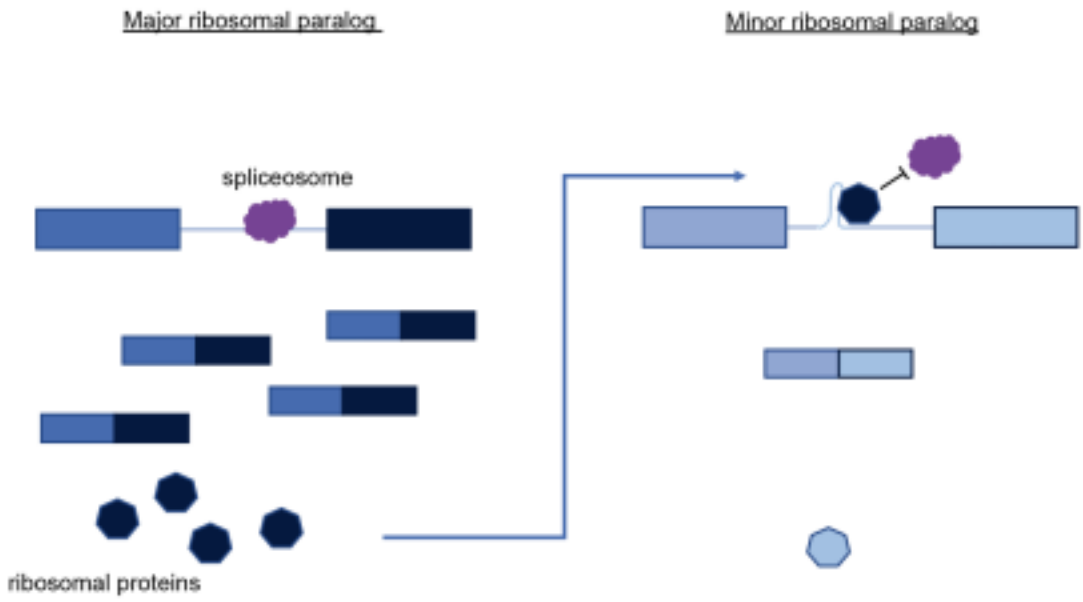



**Figure 1-5:** The major ribosomal protein paralog can repress the splicing of its minor paralog by binding to its intron and blocking access of the spliceosome.

### 1.7 Chemical inhibition of splicing

In 2004, a class of small molecules called pladienolides were isolated from soil bacteria in Kanagawa, Japan. When human glioma cells (U251 cell line) were subjected an assay testing tumor activity in the presence of pladienolides A-G, Pladienolide-B (Plad-B) (Figure 1-6) was reported to have shown the most potent inhibition of growth (Sakai et al. 2004b). In 2007, Plad B was found to bind the spliceosomal SF3B complex using chemical and radioactive probes. The authors used a variety of Plad-B probes to determine the identity of its target. Plad-B was probed with the green fluorescent dye BODIPY-FL and $^{3}H$, and HeLa cells were treated with these probes. Fluorescence and radioactivity were found to be highest in the nucleus (Kotake et al. 2007). BODIPY-FL probe colocalized with nuclear speckle marker splicing factor SC-35 suggesting that Plad-B binds a protein or complex involved in splicing (Kotake et al. 2007). Indeed, the authors concluded that Plad-B crosslinks to SAP130 (SF3B3) (Kotake et al. 2007). They, however, could not rule out the possibility that Plad-B doesn't bind other SF3B proteins.

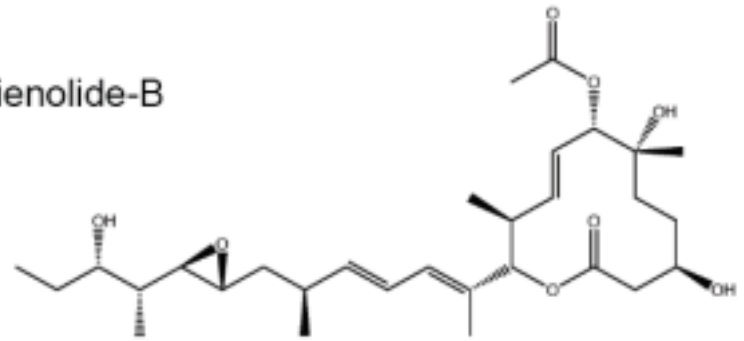
Biochemical studies using Plad-B and closely related compounds expanded the knowledge of the cellular location and effects of inhibitor binding. E7107, a Plad-B analog, was shown to prevent the binding of the U2 snRNP to pre mRNA (Folco et al. 2011). Stepwise selection with increasing concentration of inhibitors established Plad-B and E7107-

resistant colorectal cancer cell lines. The R1074H resistance mutation of SF3B1 implicated this residue as a significant binding platform for this class of splicing inhibitors (Yokoi et al. 2011). Additionally, colon cancer cells were exposed to low-dose E7107 and herboxidiene (HBD) (Figure 1-6), followed by whole exome sequencing. This generated more SF3B1 resistant mutations: V1087I, V1087A, and K1071E, and a resistance mutation in a second protein PHF5A, Y36C. Like the R1074H mutation, these sites all map to a pocket formed by SF3B1 and the interface between SF3B1 and PHF5A that normally binds the bulged BP-A residue early in splicing, providing key evidence of the exact binding location (Teng et al. 2017). Effenberger and colleagues showed certain modifications to the Plad-B structure are more critical than others for *in vitro* splicing and complex formation inhibition in HeLa cells (Effenberger et al. 2014). This indicates the importance of these parts for affecting *in vitro* splicing and inhibiting prespliceosome complex formation in HeLa cells. Cytological profile clustering comparing cells with and without drugs show that DNA synthesis, mitosis, and nuclear size and shape are the most affected by the drug, consistent with earlier studies (Mizui et al. 2004, Effenberger et al. 2014).

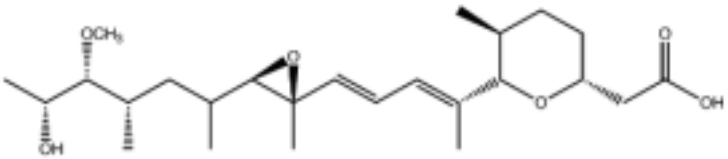
The explosion of solved spliceosome structures beginning in 2015 included early human spliceosome complexes bound to splicing inhibitors. The cryogenic electron microscopy (cryo-EM) structure of the Plad-B

analog E7107 bound to SF3B1 showed that inhibitors indeed use the same pocket as the BPA residue (Finci et al. 2018). It was proposed that inhibitors keep SF3B1 in an “open conformation” around a hinge-like region composed of HEAT repeats 15-17 that it adopts before fully binding the pre-mRNA, while it forms a “closed conformation” around the U2 snRNA-pre-mRNA helix in the absence of inhibitors (Finci et al. 2018). Crystal structure of U2 snRNP components in complex with Plad-B and spliceostatin A (SSA) confirmed this inhibitor mechanism of action, showing inhibited SF3B1 was locked in an open conformation relative to uninhibited SF3B1 around the hinge region (Cretu et al. 2018, Cretu et al. 2021). A derivative of SSA, thailanstatin A (Figure 1-6), was also found to show potent antitumor activity by preventing splicing *in vitro* in cancer cells (Ghosh et al. 2018, Nicolaou et al. 2021). SF3B1/PHF5A directly contacts the three major structural motifs of Plad-B, the aliphatic arm, the conjugated diene, and the macrolide ring (Figure 1-7). SSA, which is slightly different in structure, binds similar amino acid residues (Figure 1-7). A key difference in the case of the SSA complex, the C26 residue of PHF5A is important for covalent interaction with the drug (Figure 1- 7) (Cretu et al. 2021).

Pladienolide-B



Herboxidiene



Thailanstatin-A

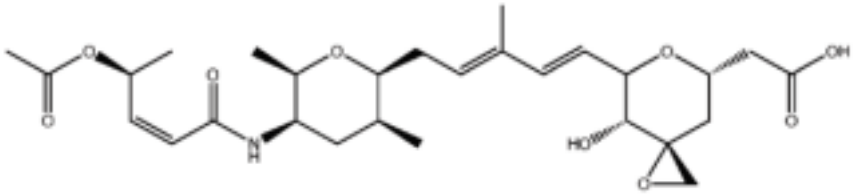
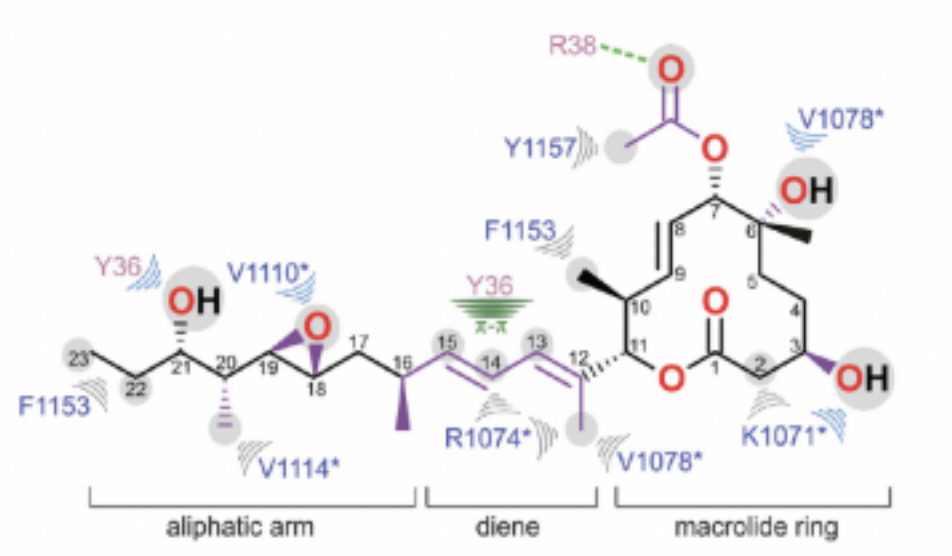
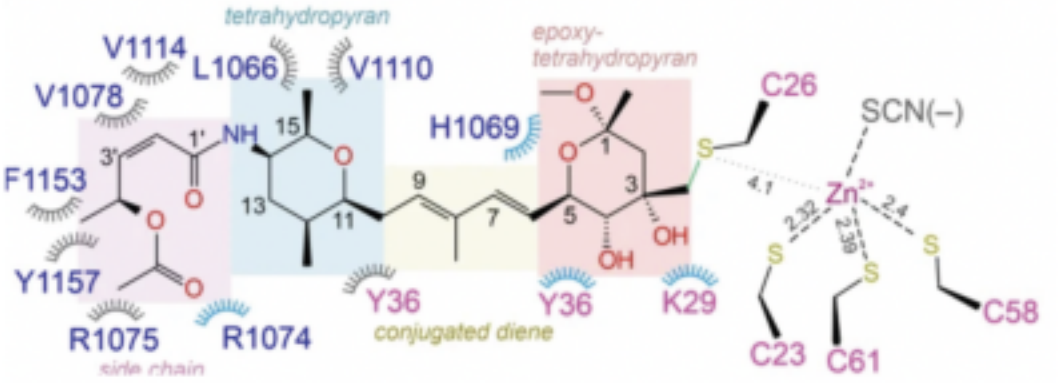


**Figure 1-6:** SF3B1-targeting splicing inhibitors.




SF3B1-PHF5A bound to Plad-B
(Cretu et al. 2018)




SF3B1-PHF5A bound to spliceostain-A
(Cretu et al. 2021)

**Figure 1-7:** SF3B1/PHF5A amino acid interactions with Plad-B and spliceostatin A. Blue residues are SF3B1 and pink residues are PHF5A. Figures adapted from Cretu et al. 2018 and Cretu et al. 2021.

Studies on the chemical inhibition of splicing in non-mammalian organisms have been keeping pace. Plad-B and herboxidiene (HBD, also called GEX1A) have been demonstrated to inhibit splicing in the plant Arabidopsis *thaliana* (Ling et al. 2017, AlShareef et al. 2017). Resistance mutations were created in the rice plant *Oryza sativa* SF3B1 homolog increasing its splicing efficiency in the presence of HBD (Butt et al. 2019). SSA also inhibits splicing in *Schizosaccharomyces pombe* (Lo et al. 2007). Other organisms such as *Caenorhabditis elegans* and *Saccharomyces cerevisiae* are naturally resistant to splicing inhibitors, suggesting an evolutionary change in the pocket environment that does not allow for inhibitor binding but still does allow for BPA binding.

Important amino acid residues in the HEAT 15-16 pocket may predict sensitivity or resistance to inhibitors. For an organism to be naturally resistant, SF3B1/PHF5A residues must bind BPA and not bind inhibitors. Some of the essential amino acids for binding inhibitors include SF3B1 V1078 and L1108. The conservation of many of these residues can be seen across a wide variety of organisms from *H. sapiens* to *S. cerevisiae* (Figure 1-8). These organisms have a wide variety of intron numbers per gene and splice site strengths within each intron.

Fascinatingly, inhibitor-sensitizing sf3b1/hsh155 mutations in *C.*

*elegans* and *S. cerevisiae* have been reported (Hansen et al. 2019, Serrat et al. 2019). Plasmid-based humanizing mutations of yeast strains lacking the drug transporters *PDR5, SNQ2,* and *YOR1* (strain JRY8012, Jeong et al. 2007) have shown great promise. However, the initial characterization of humanized hsh155 yeast in the presence of splicing inhibitors has been limited to growth assays and splicing assays on a few reporter genes.

The work presented here builds on recent studies humanizing yeast to study splicing inhibition. Humanized HSH155 gene on a plasmid has been used to study inhibition (Carrocci et al. 2018, Hansen et al. 2019). Humanized HEAT repeats 5-16 of yeast Hsh155 are viable and Plad-B can inhibit splicing (Carrocci et al. 2018, Hansen et al. 2019). Importantly, these studies showed that WT yeast is not affected. A single L777N substitution alone was not sufficient for inhibition, but when paired with N747V (NL/VN) rescues splicing inhibition (Carrocci et al. 2018, Hansen et al. 2019). The authors utilized an ACT1-CUP1 reporter gene to measure *in vivo* splicing and correlated the growth on copper to splicing.

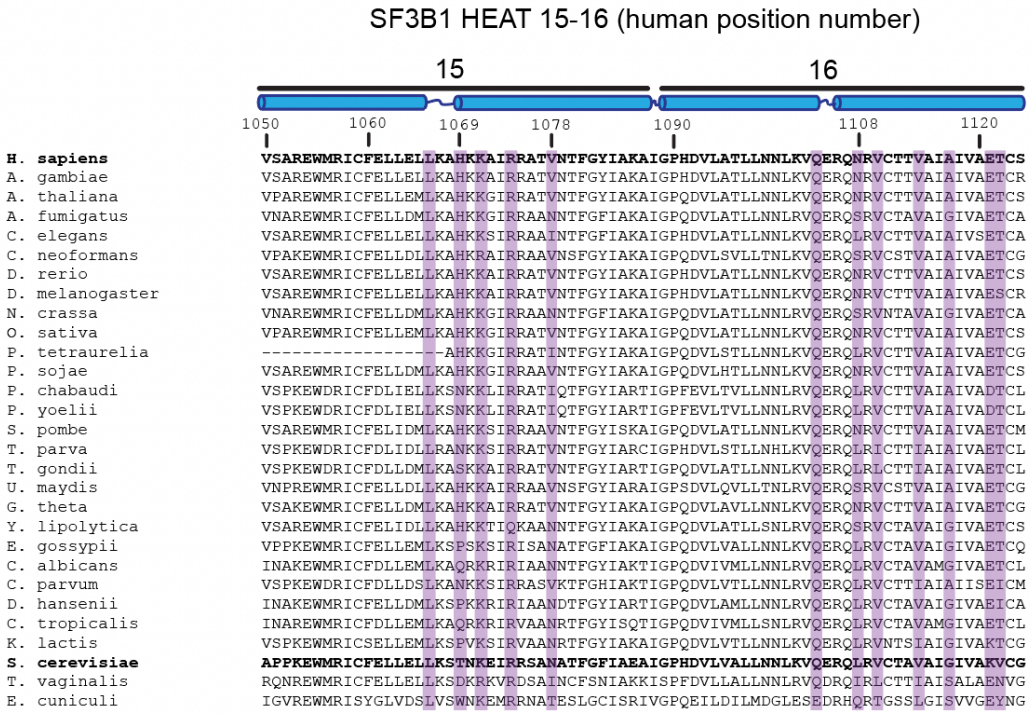



**Figure 1-8:** SF3B1 HEAT repeat 15-16 amino acid alignment from 29 organisms show high conservation. Human SF3B1 amino acid number is shown. Residues in purple are involved in BPA/Plad-B or SSA inhibitor binding as shown in Cretu et al. 2018 and Cretu et al. 2021. Blue cylinders represent alpha helices that correspond to amino acid position. *H. sapiens* and *S. cerevisiae* are bold to highlight amino acid comparison in naturally sensitive and resistant organisms respectively.

### 1.8 A new humanized yeast strain sensitive to splicing inhibitors

This work aims to answer the following outstanding questions in the field: 1) What are the consequences of blocking splicing for individual introns and 2) how does global gene expression change when splicing is blocked? To answer these questions, my work will 1) measure splicing inhibition of individual introns and 2) analyze transcriptome-wide gene expression changes that occur because of splicing inhibition, with special attention to genes without introns.

We began to address these questions by determining the best method to block splicing in yeast. Past methods of splicing inhibition mainly include temperature-sensitive splicing factor alleles (Clark et al. 2002). However, for the purposes of my studies, temperature shock had to be avoided as it could have a significant effect on gene expression (Herruer et al. 1988, Rodriguez Vargas et al. 2002). For this reason, we turned to the chemical inhibition of splicing. The use of small molecule splicing inhibitors avoids the need for temperature shocking yeast; however, yeast is naturally resistant to such inhibitors. Therefore, sensitizing yeast to chemical inhibitors at the amino acid level via mutation was determined to be the best route to block splicing in yeast.

Human spliceosomes have been shown to be sensitive to a wide class of SF3B1-targeting small molecules. Yeast and human spliceosomes are highly conserved. We, therefore, determined that

making “humanizing” mutations to the yeast homolog of SF3B1, called HSH155, directly in the genome would be the best route to creating yeast that is sensitive to splicing inhibitors. Additionally, yeast has few introns compared to the total number of human introns, which simplifies our efforts in understanding splicing. Blocking splicing in human cells may affect nearly every cellular process because almost all genes in human cells contain introns.

The successful creation of a humanized HSH155 yeast strain requires the following four conditions to be met; 1) Mutations must be stably integrated into the genome, 2) Mutations do not affect splicing in absence of drugs 3)  Mutations do not affect HSH155 protein level, and 4) Mutations do not create  temperature sensitivity. Once these conditions are met, the new humanized yeast will have a wide range of advantages and applications.

First, a more accurate dose-dependent response will be seen upon chemical inhibition. This is important because a higher dose of splicing inhibitors of Plad-B may be detrimental to cells. However, a smaller dose may reveal gene-specific information such as branchpoint choice without killing cells right away. On a similar note, humanized yeast will allow for greater control of inhibition. Humanized yeast can therefore be used as a tool to inhibit a very specific step of splicing. Additionally, a range of different inhibitor molecules that may have varying affinities for

humanized HSH155 can be exploited to study the mechanism of the spliceosome.

# Chapter 2

## Humanizing the Yeast Spliceosome

### 2.1 Design and creation of a humanized *hsh155* protein

Budding yeast is not sensitive to splicing inhibitors, so we opted to humanize the yeast homolog of SF3B1, called HSH155. We chose to replace the minimum number of amino acids at the part of the protein the drug binds. Humanizing was the best option when considering the need for stable expression of such a mutant. With this strain, we can block splicing and see how essential it is during yeast gene expression.

Humanizing amino acid substitutions in the HSH155 splicing protein was done directly in the genome using CRISPR/Cas9 (DiCarlo et al. 2013). Yeast also expresses several multi-drug transporters, which is an issue for inhibitor effectiveness due to drug efflux in the cell. We performed CRISPR editing in a background strain that was transporter deficient (JRY8012, Jeong et al. 2007). We designed a new humanized yeast strain that is a partially humanized HEAT 15 and a fully humanized HEAT 16 called *hsh155-ds* for “drug-sensitive hsh155 mutant” with a minimum set of 13 amino acid substitutions (Figure 2-1, Figure 2-2). Thus, we designed the change in such a way that the pocket-facing residues are the only ones being humanized for interaction with splicing inhibitors and to avoid humanizing other surfaces that might interact with yeast spliceosomal proteins (Figure 2-2). The next step was to ensure that

*hsh155-ds* is absent of any growth defects and stably expressed. We performed a growth assay comparing hsh*155-ds* to WT HSH155 yeast. We made 10X dilution spots of both strains and grew them up at 18°C, 30°C, and 37°C (Figure 2-3). For all temperatures tested, *hsh155-ds* yeast grew similarly to WT HSH155 yeast. We performed western blot analysis to detect WT Hsh155 and *hsh155-ds* protein. To do this, WT HSH155 and *hsh155-ds* GFP fusion proteins were created. Total protein was then extracted from log phase cells and run on an SDS PAGE gel and probed using anti-GFP antibody. Western blot analysis verified the expression of WT HSH155 and *hsh155-ds* fusion proteins at 137 kDa (Figure 2-4). We conclude that *hsh155-ds* are not temperature sensitive and are stably expressed in budding yeast.

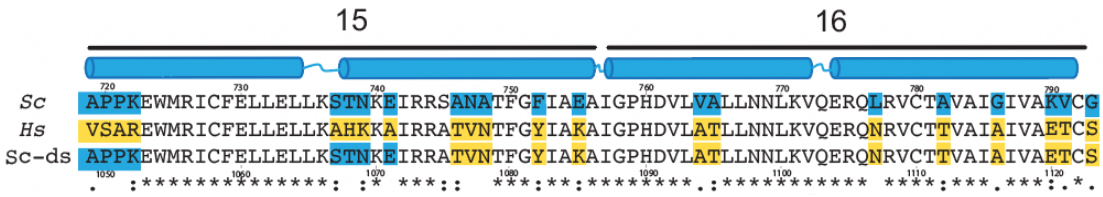



**Figure 2-1:** Amino acid alignment of the hinge region of SF3B1/HSH155 generated by Clustal Omega. ‘15’ and ‘16’ designate HEAT repeats. Blue cylinders depict alpha helical pairs (HEAT repeats). *“Sc,” “Hs,”* and *“Sc-ds' ‘indicate* the species *S. cerevisiae*, *H. sapiens*, and drug sensitive (humanized) yeast respectively. Blue boxes indicate the yeast amino acid sequence and gold boxes indicate the human amino acid sequence. Alignment consensus symbols follow Clustal Omega conventions: an * (asterisk) indicates a fully conserved amino acid, a : (colon) indicates similar amino acids, a . (period) indicates weakly similar amino acids, and no symbol indicates no similarity. Respective amino acid numbers are shown. *Sc-ds* amino acid number shows SF3B1 number relative to yeast.

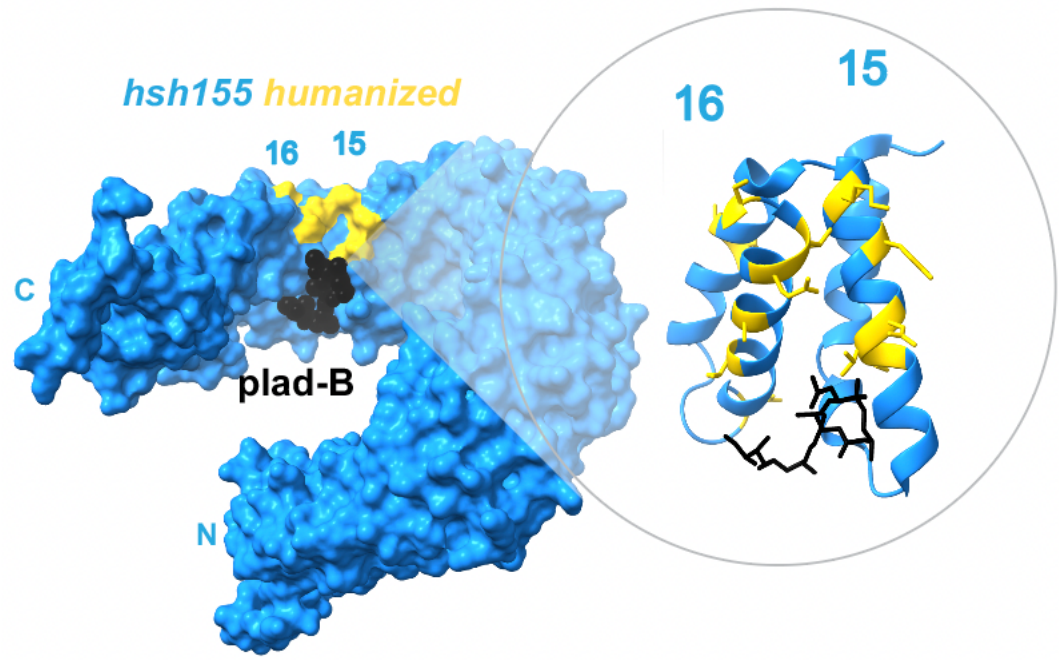



**Figure 2-2:** Crystal structure of Plad-B (black) bound to humanized *hsh155-ds* protein (blue with humanizing mutations in gold). Figure adapted from Cretu et al., 2018 (PDB: 6EN4). '15' and '16' designate HEAT repeats. N and C terminus are shown. Right circle shows a zoomed-in perspective of the Plad-B pocket in ribbon diagram (*hsh155-ds* alpha helices), sphere model (Plad-B) and stick model (amino acid residues).

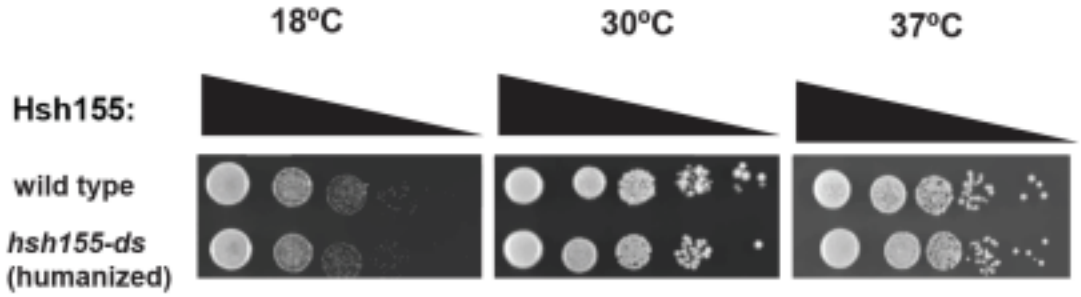



**Figure 2-3:** Growth assay comparing WT Hsh155 and *hsh155-ds* yeast on YPD plates to test for temperature sensitivity. Strains were diluted 10X down (right triangles) a total of four times. Temperatures tested are shown in °C.

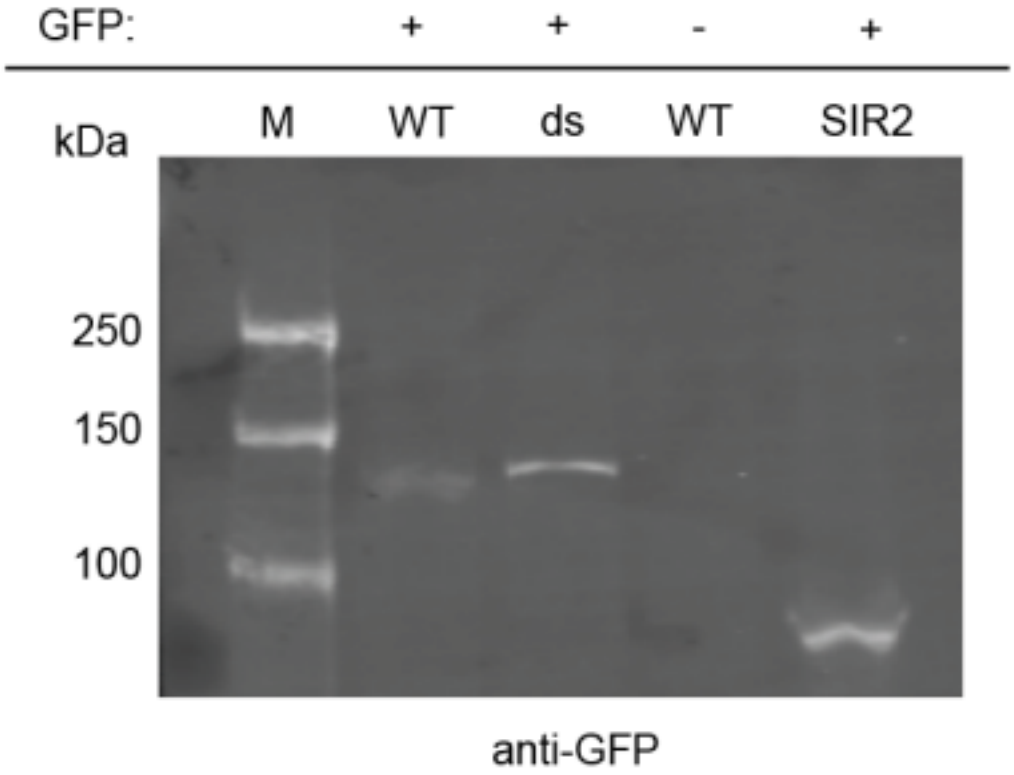



**Figure 2-4:** Western blot analysis of WT Hsh155-GFP+ and *hsh155-ds*-GFP+ fusion proteins. WT Hsh155-GFP+ and *hsh155-ds*-GFP+, WT Hsh155-GFP-, and Sir2-GFP+ control fusion proteins run on 4-12% gradient SDS-PAGE gel. Hsh155-GFP+ fusion proteins each run at ~137 kDa while Sir2-GFP+ antibody control fusion runs at ~90 kDa. Blot was probed with mouse anti GFP antibody.

### 2.2 Plad-B inhibits splicing within minutes *in vivo*

Having designed a minimum set of mutations predicted to enable Plad B to bind humanized yeast *hsh155-ds* in a yeast strain lacking drug exporter proteins, we next wanted to determine whether and how rapidly Plad-B might inhibit splicing. We performed a time course of inhibition by adding Plad-B or as a control, an equal amount of the carrier DMSO, for 0, 1, 4, or 8 minutes, followed by isolating RNA immediately afterward using a rapid freezing method similar to (Kos and Tollervey 2010). Isolated RNA was used for RT PCR using primers to the first intron of the *MATa1* gene (Figure 2-5). *MATa1* mRNA has a high turnover rate (Miller 1984) which means that a block in splicing that prevents the synthesis of new mRNA would be immediately visible in our experiment as rapid loss of mRNA, even if pre-mRNA was unstable and did not accumulate. The addition of 5 μM Plad-B was enough to inhibit splicing within the 4-minute mark (top, 'US' band, Figure 2-6). Inhibition continued to increase into the 8-minute mark (Figure 2-6, Figure 2-7). DMSO incubated (0 μM Plad-B) cells showed no inhibition (bottom, 'S' band, Figure 2-6, Figure 2-7). This experiment enabled us to establish that splicing inhibition occurs within minutes and is strong *in vivo*, at least for the *MATa1* first intron.

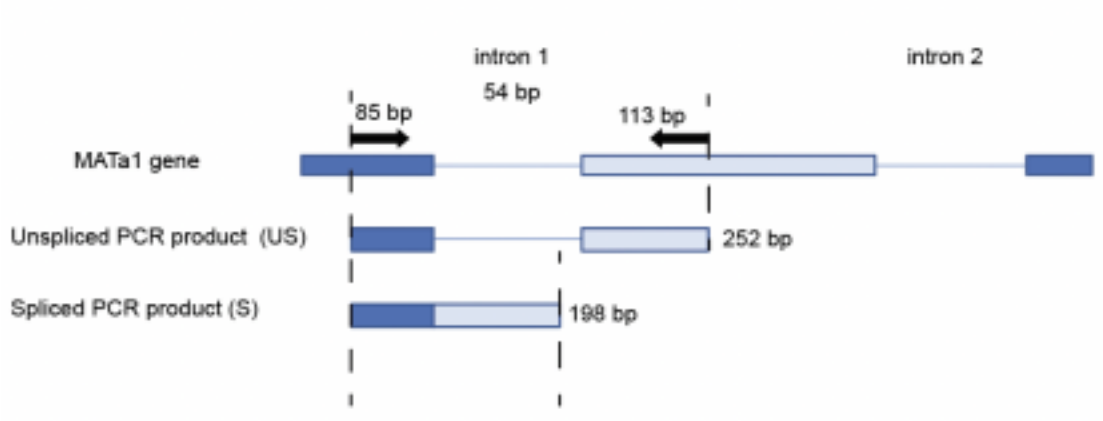



**Figure 2-5:** Schematic of the RT-PCR experiment and the expected results. MATa1 gene shown as blue boxes (exons) connected by thin lines (introns). Introns 1 and 2 are labeled. The PCR product expected in this experiment is shown flanked by two broken lines using primers that sit in exon 1 and exon 2 (arrows). The PCR product for exon 1 is 85 bp, the intron is 54 bp, and exon 2 is 113 bp. This creates an unspliced (US) PCR product that is 252 bp and a spliced (S) product that is 198 bp.

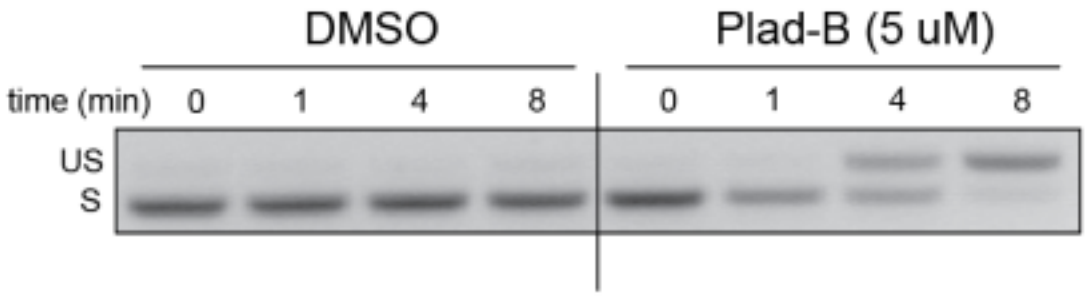



**Figure 2-6:** Gel image of the RT-PCR. 0, 1,4 and 8-minute timepoints were measured for both DMSO and Plad-B (5 μM) conditions in *hsh155-ds* yeast. 'US' and 'S' indicate unspliced and spliced bands respectively.

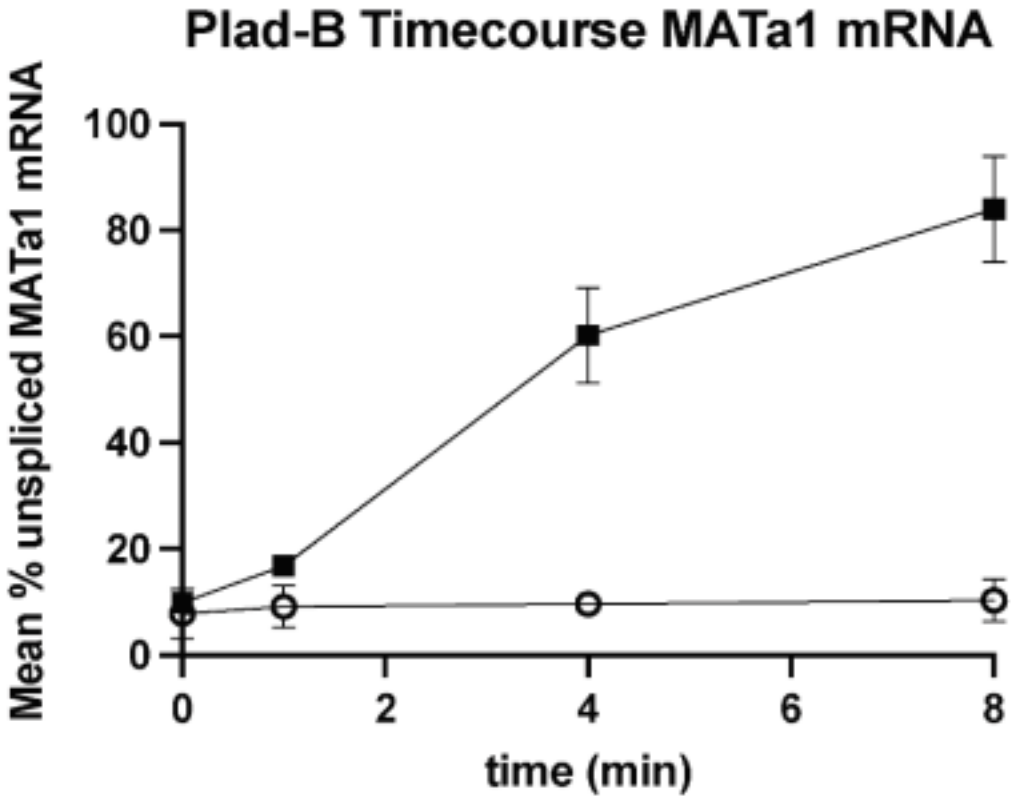



**Figure 2-7:** Quantification of gel image showing the mean % unspliced MATa1 mRNA from Figure 2-6 generated from ImageJ. Standard deviation was calculated and is represented by error bars. Square points represent hsh*155-ds* incubated 5 μM Plad-B. Open circles represent *hsh155-ds* incubated 0 μM Plad-B (DMSO).

## 2.3 Splicing inhibitors block splicing in yeast extracts with similar potency as in human extracts

Studies in human nuclear extracts indicate that Plad-B and other inhibitors block splicing at or before the ATP-dependent step of pre spliceosome formation when the BPA of the U2-branchpoint duplex is bound in the pocket formed by SF3B1 and its interface with PHF5A (Roybal and Jurica 2010, Corrionero et al. 2011, Cretu et al. 2018). We hypothesized that Plad-B would inhibit similar steps in the same way in splicing extracts from the *hsh155-ds* yeast strain. To measure this inhibition, we performed *in vitro* splicing in extracts from *hsh155-ds* at different concentrations of Plad-B using a radioactive actin pre-mRNA substrate. After splicing we separated the reaction products on a denaturing gel and exposed the gel to a phosphorimager screen for visualization of the precursor, the splicing intermediates (free exon and the lariat-2nd exon RNA), and splicing products (ligated exons and free lariat intron). With increasing concentration of Plad-B, we observed the decreasing presence of both splicing intermediates and products, Figure 2-8A). We concluded splicing could be blocked *in vitro* using the ACT1 substrate.

We carried out native (nondenaturing) gel electrophoresis testing spliceosome complex formation in the presence of increasing

concentrations of Plad-B. Like our splicing assay, native gel analysis revealed the decreasing presence of pre-spliceosomes with increasing concentrations of the drug (Figure 2-8B). The precursor complex to the prespliceosome is the commitment complex. We did not observe a decreasing commitment complex indicating they are not inhibited by Plad-B. We concluded that splicing inhibition takes hold at a very specific step, after the formation of the commitment complex but before the formation of the prespliceosome.

Quantification of the *in vitro* splicing assay and native gel analysis allowed us to calculate the half maximal inhibitory concentration (IC50) to be ~25 nM and ~27 nM respectively (Figure 2-8, C and D). We repeated the *in vitro* splicing reaction and native gel analysis using HBD and generated similar results (Figure 2-9A, Figure 2-9B). Interestingly, these values closely compare to IC50 values calculated from studies that inhibited human and yeast cells (Sakai et al. 2004a, Yokoi et al. 2011, Effenberger et al. 2017, Carrocci et al. 2018). This was, therefore, compelling evidence that *hsh155- ds* inhibition is comparable to human splicing inhibition.

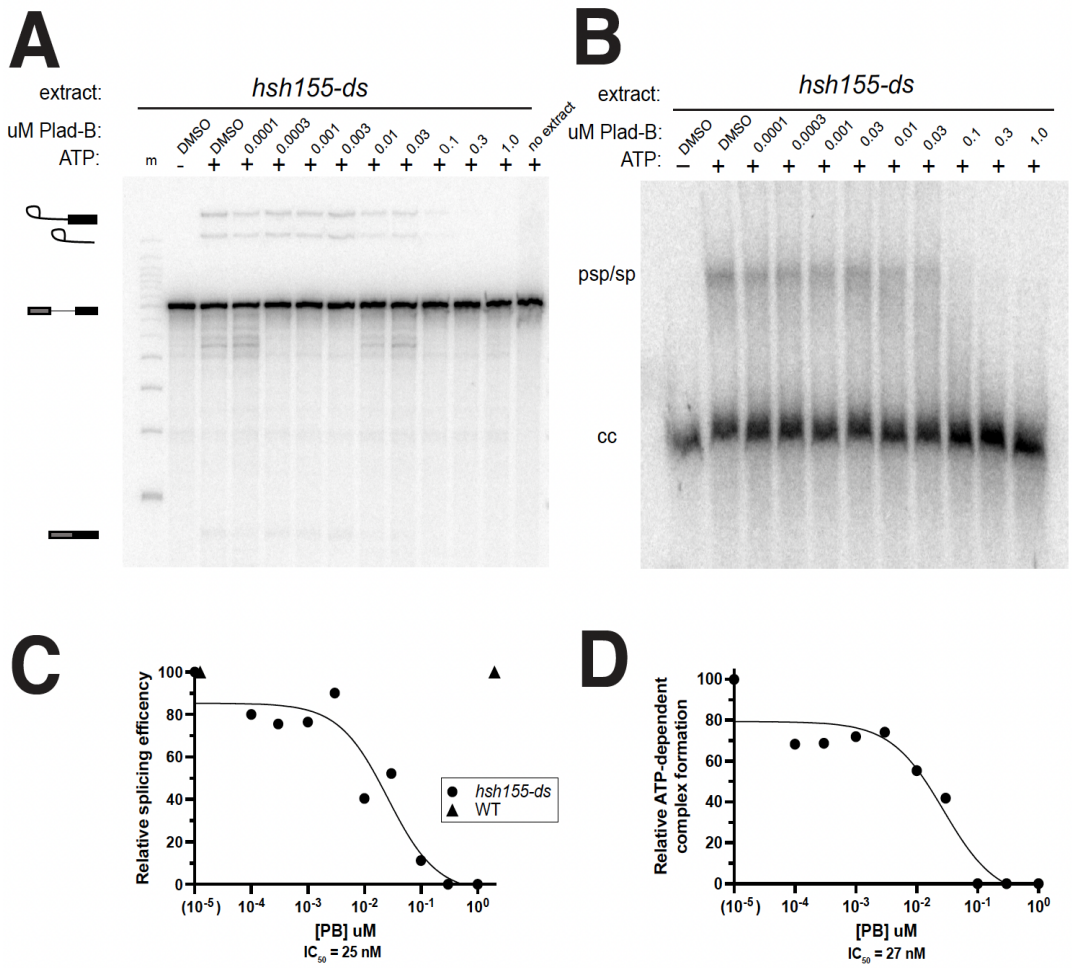



**Figure 2-8: In vitro analysis of Plad-B-treated *hsh155-ds* yeast. A)** *In vitro* splicing reaction assay using a $^{32}$P-radiolabeled actin pre-mRNA run on a 6% acrylamide/8M urea denaturing gel using *hsh155-ds* extract. Lariat-3' exon, lariat, pre-mRNA, and mature mRNA are shown. ATP was used at 20 mM (+) or 0 mM (-). Lane m is the marker. DMSO (0 μM Plad-B) lanes are shown in the presence and and absence of Plad-B. Lanes show increasing concentrations of Plad-B. 'no extract' control is shown in the last lane **B)** Non denaturing agarose gel analysis of *hsh155-ds* extract complexes. psp/sp and cc indicate "pre-spliceosome and spliceosomes' and "commitment complex" respectively. **C)** Quantification of splicing assay measuring splicing efficiency of Plad-B-treated *hsh155-ds* extract relative to DMSO-treated extract. **D)** Quantification of splicing assay measuring ATP-dependent complex formation for *hsh155-ds* extract in the

presence of increasing concentration of Plad-B relative to DMSO-treated extract.

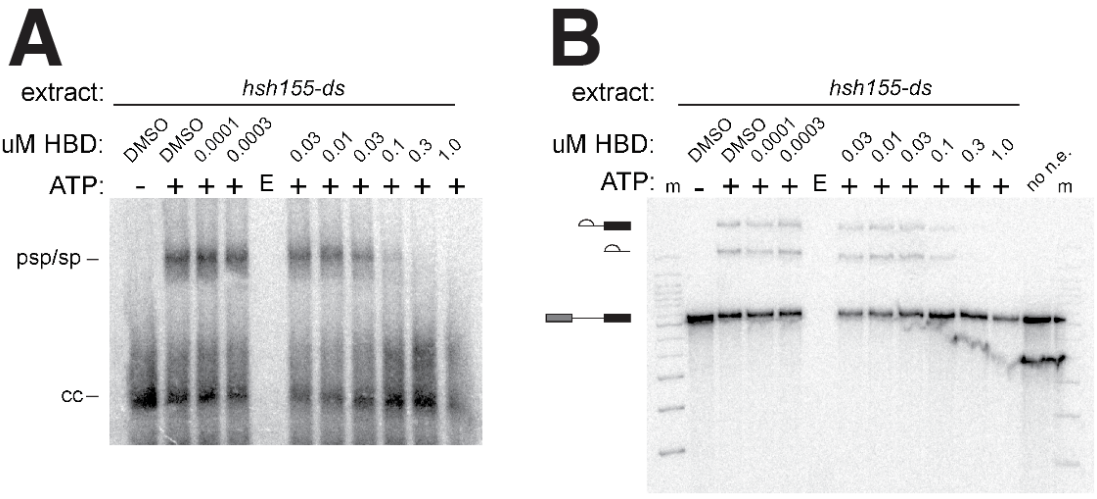



**Figure 2-9: *In vitro* analysis of Herboxidiene-treated (HBD) *hsh155-ds* yeast. A)** Non-denaturing agarose gel analysis of *hsh155-ds* extract complexes. psp/sp and cc indicate “pre-spliceosome and spliceosomes” and “commitment complex” respectively. Lane labeled ‘E’ is empty. **B)** *In vitro* splicing reaction assay using a $^{32}$P-radiolabeled actin pre-mRNA run on a 6% acrylamide/8M urea denaturing gel using *hsh155-ds* extract. Lariat-3’ exon, lariat, pre-mRNA, and mature mRNA are shown. ATP was used at 20 mM (+) or 0 mM (-). Lane m is the marker. DMSO (0 µM Plad-B) lanes are shown in the presence and absence of Plad-B. Lanes show increasing concentrations of Plad-B. ‘no extract’ control is shown in the last lane. Lane labeled ‘E’ is empty.

# Chapter 3

# Splicing and Gene expression in response to Pladienolide-B and Thailanstatin A in *hsh155-ds* yeast

## 3.1 RNA-sequencing of *hsh155-ds* yeast total mRNA treated with Pladienolide-B and Thailanstatin A

We performed total RNA-sequencing (RNA-seq) to measure splicing and mRNA levels for every gene across the genome. We treated WT HSH155 and *hsh155-ds* yeast with DMSO, Plad-B, and Thail-A. Six duplicate treatments (12 samples total) were performed (Figure 3-1). *hsh155-ds* were treated with either DMSO, 0.5 μM Plad-B, 5 uM Plad-B, or 5 μM Thail-A. WT HSH155 yeast was treated with either DMSO or 5 μM Plad-B. After 1 hour of treatment, total RNA was extracted, and libraries were prepared for Illumina sequencing. We obtained a total of more than 1 billion sequencing reads for all samples in both duplicates. Raw sequencing reads were trimmed, normalized, and aligned to the annotated *S. cerevisiae* genome, SacCer3, to obtain coverage of various regions.

Because splicing inhibition is dose-dependent, we opted to use high (5 μM) and low (0.5 μM) concentrations of Plad-B during treatment. We hypothesized that the amount of intron reads would be higher in 5 μM Plad-B than at 0.5 μM Plad-B. Our rationale for treating *hsh155-ds* cells with two different doses of Plad-B is that a higher dose such as 5 μM of

splicing inhibitors may be harmful overall for the cell, but a smaller dose such as 0.5 μM may uncover differential gene expression patterns or alternative splice site choices without killing cells immediately. Thail-A is suspected to have a slightly different binding mechanism, as it is an analog of spliceostatin A (Figure 1-7). We included Thail-A at 5 μM to observe any difference in splicing inhibition when compared to 5 μM Plad-B.

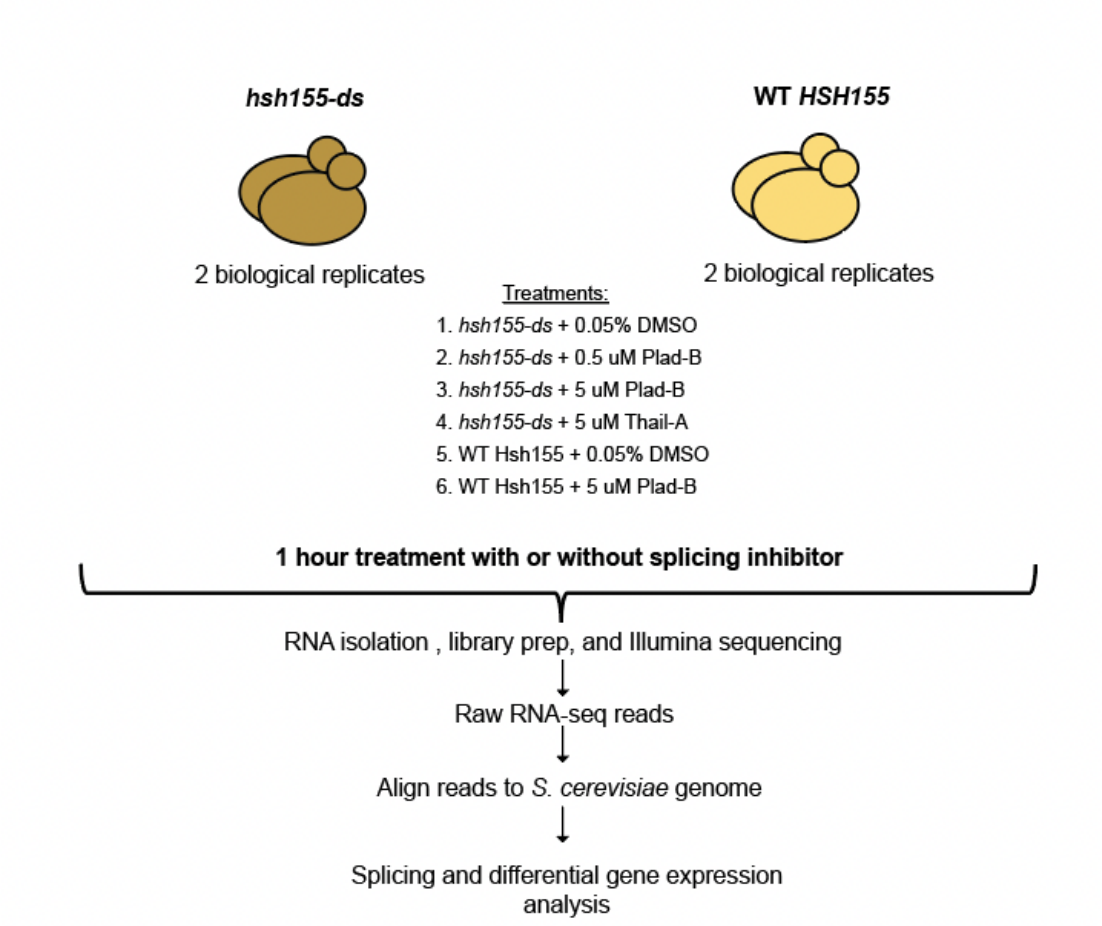



**Figure 3-1:** RNA-sequencing library pipelie.

### 3.2 Measuring the splicing efficiency of individual introns after drug treatment

Addition of splicing inhibitors to human cells or yeast containing a humanized hsh155 allele causes an accumulation of introns. Whether or not different introns are equally inhibited by Plad-B and other splicing inhibitors has largely remained unexplored. The need to compare introns individually is vital to understanding inhibition across the transcriptome. To that end, we analyzed individual introns from all genes in yeast after drug treatment.

We identified and quantified retained intron (RI) events from intron containing genes in our RNA-seq dataset using the software JunctionCounts (https://github.com/ajw2329/junctionCounts). RI events were quantified using the percent spliced value (% spliced). The % spliced is the number of reads that span the exon-exon junction (spliced counts) divided by the number of reads that span the exon-exon junction (spliced counts) plus the number of reads that span the intron-exon junction (unspliced counts) for a given gene (Figure 3-2). A low % spliced value indicates a lower splicing efficiency, while a higher % spliced value indicates a higher splicing efficiency.

The filtering process is as follows. The average unspliced and spliced counts for each gene was taken across the 12 samples and summed. Genes with 100 average counts or less were removed as were

2-intron skipping events. Non-RI splicing events were also removed. This resulted in 247 splicing events for this analysis.

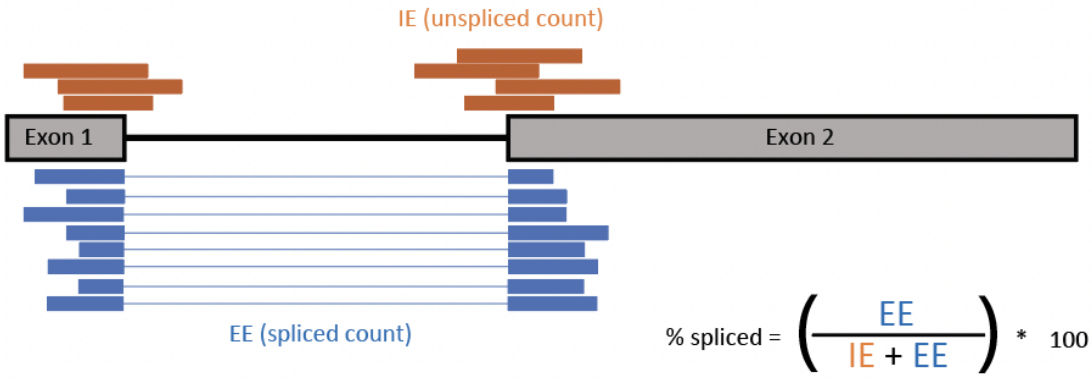



**Figure 3-2:** JunctionCounts schematic for calculation of splicing efficiency (% spliced). EE indicates exon-exon junctions and IE indicates intron-exon junctions.

### 3.3 Measuring the splicing efficiency between replicate samples

An important quality control step in RNA sequencing analysis is consistent measurements across replicates. For this reason, we plotted the % spliced values of each replicate for all six treatments and examined its correlation relative to each other. All six treatments resulted in a correlation rank coefficient ($r^2$) of at least 0.9800 (Figure 3-3), indicating good agreement between replicates. We conclude that each treatment has consistent measurements across replicates and can be used to analyze differences in splicing efficiency comparing various treatments.

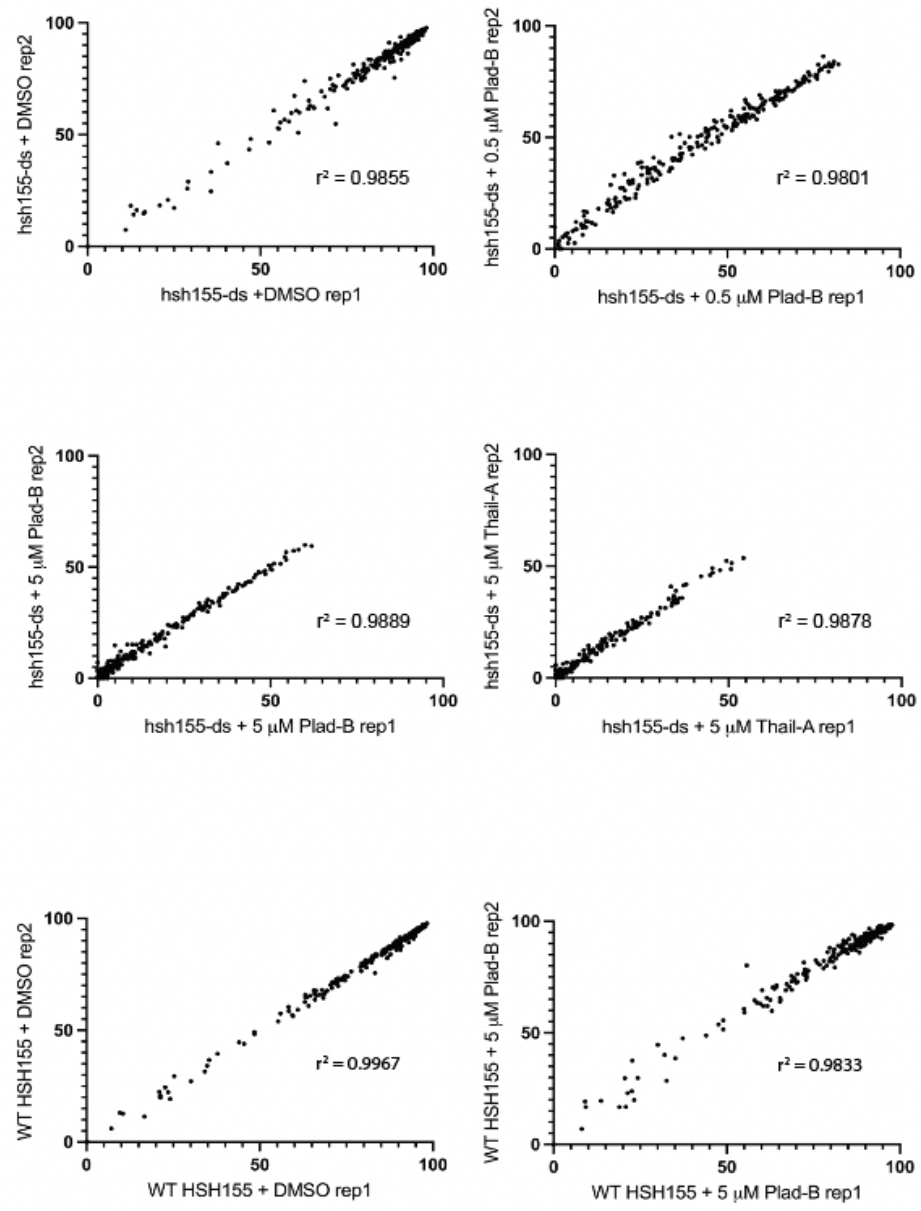



**Figure 3-3:** Correlation plot of replicate 1 (rep1) versus replicate 2 (rep2) for all 6 treatments of the RNA-sequencing experiment. Rank correlation coefficient is shown as an $r^2$ value for each plot.

### 3.4 Are introns in wild type HSH155 yeast sensitive to Plad-B?

Next, we wanted to confirm that yeast harboring the WT HSH155 allele were insensitive to Plad-B in our RNA-seq dataset. To validate this, the average % spliced values of WT HSH155 + 0 μM Plad-B (DMSO) across both replicates was plotted against the average % spliced values of WT HSH155 + 5 μM Plad-B, the highest concentration used. We found that these two treatments were very correlated with an $r^2$ of 0.9942 (Figure 3-4). This demonstrates WT HSH155 yeast are not sensitive to Plad-B, even at concentrations as high as 5 μM.

We next wanted to rule out splicing differences between WT HSH155 and *hsh155-ds* in the absence of splicing inhibitors We observed a high degree of agreement in the splicing efficiencies of each gene between these two different strains with an $r^2$ of 0.9977 (Figure 3-4). Although we could not exclude individual genes that might be spliced significantly different in the absence of any drug, we can conclude that the *hsh155-ds* mutations do not greatly alter splicing between WT HSH155 and *hsh155-ds* in the absence of drug.

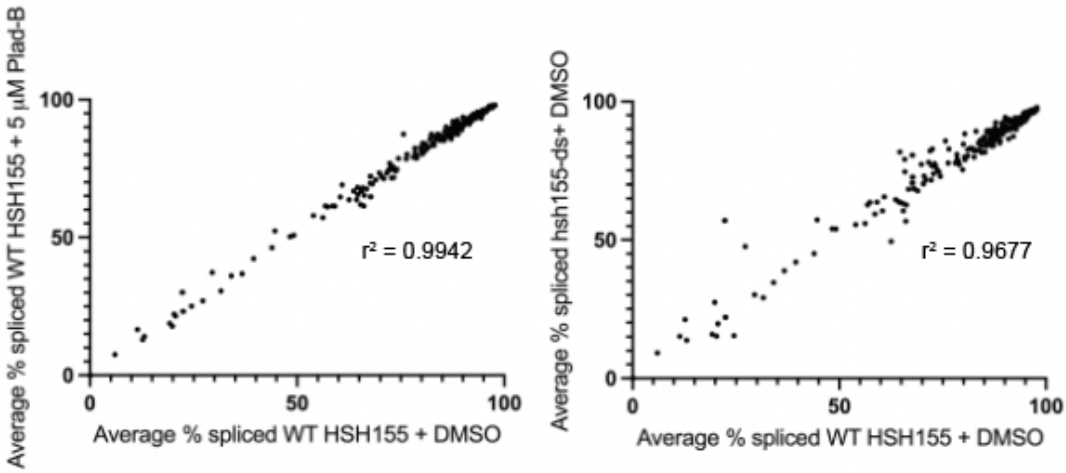



**Figure 3-4:** Correlation plot of average % spliced of WT HSH155 with DMSO versus WT HSH155 with 5 µM Plad-B (left). Correlation plot of average % spliced of WT HSH155 with DMSO versus *hsh155-ds* with DMSO (right). Rank correlation coefficient is shown as an $r^2$ value for each plot.

### 3.5 Are all introns in hsh155-ds sensitive to Plad-B?

The RNA-seq data set we obtained enabled us to view the splicing efficiency for almost all hsh155-ds introns. We compared the % spliced values for hsh155-ds at 0.5 μM and 5 μM to hsh155-ds with no drug. We observe an apparent range of sensitivities to both 0.5 μM Plad-B and 5 μM Plad-B, with 5 μM showing a stronger effect. respectively. Noticeably, no genes reached higher than 80% spliced in the 0.5 μM Plad-B condition and 60% spliced in the 5 μM Plad-B condition (Figure 3-5). We conclude that most introns are sensitive to Plad-B, although a minority continue to be apparently resistant, even in the presence of high concentrations of Plad-B.

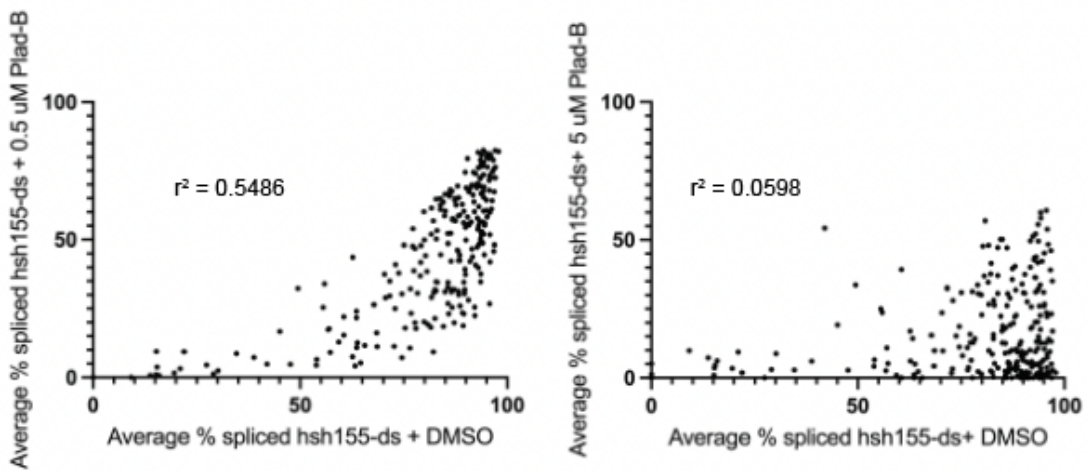



**Figure 3-5:** Correlation plot of average % spliced of *hsh155-ds* with DMSO versus *hsh155-ds* with 0.5 µM Plad-B (left). Correlation plot of average % spliced of *hsh155-ds* with DMSO versus hsh155-ds with 5 µM Plad-B (right). Rank correlation coefficient is shown as an $r^2$ value for each plot.

### 3.6 Gene expression patterns in response to splicing inhibition

To begin to understand gene expression changes in response to splicing inhibition, we ran differential gene expression analysis on the total RNA-seq dataset using DESeq2 (Love et al. 2014). We first noticed that most ribosomal protein genes were significantly down regulated when we compared *hsh155-ds* samples with DMSO to samples containing 5 μM Plad-B (Figure 3-6, blue and green points). This suggests ribosomal protein gene expression is sensitive to splicing inhibition in yeast. This effect was also in samples with 0.5 μM Plad-B (low concentration) and 5 μM Thailanstain-A (data not shown). Interestingly, the decrease in gene expression of intronless ribosomal protein genes (blue points) mirrored the decrease in intron containing ribosomal ribosomal protein genes (green points). We conclude that intron-containing and intronless ribosomal genes are downregulated upon splicing inhibition.

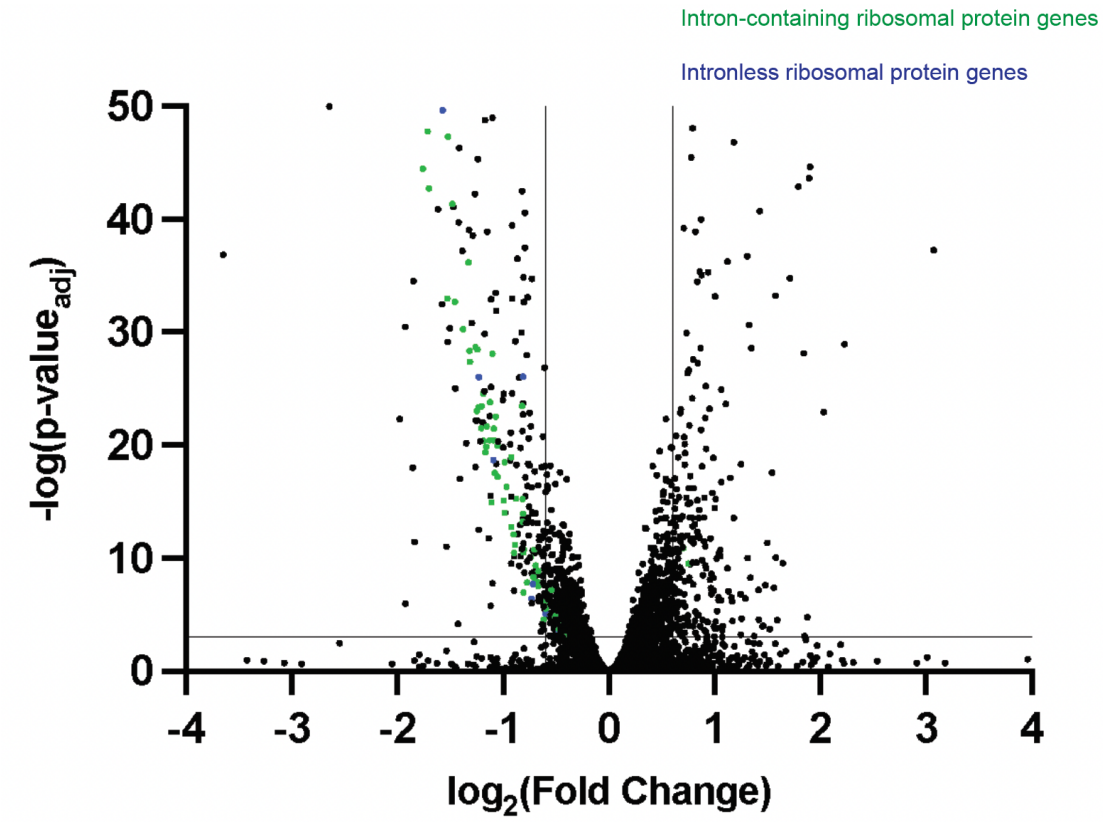



**Figure 3-6: Volcano plot showing differential gene expression analysis of hsh155-ds in the presence of 5 µM Plad-B compared to *hsh155-ds* in the absence of Plad-B.** Points to the left of $\log_2$(Fold Change) = 0 are downregulated and points to the right of $\log_2$(Fold Change) = 0 are upregulated. Y-axis shows log scale of the adjusted p-value. X-axis shows log scale of expression fold-change of *hsh155-ds* with 5 µM Plad-B relative to hsh*155-ds* with DMSO. Horizontal line -log (p-value$_{adj}$) = 3 indicates an adjusted p-value of 0.001. Points above -log (p-value$_{adj}$) = 3 are considered significant. Vertical lines at x = -0.6 and x = 0.6 indicate a $\log_2$(Fold Change) = 0.6 (Fold Change = 1.5). Intron-containing ribosomal protein genes shown in green. Intronless ribosomal protein genes shown in blue.

### 3.7 De novo purine synthesis pathway is downregulated after a block in splicing

To identify non-ribosomal protein genes that were differentially expressed, we performed gene ontology (GO) analysis of all downregulated non-RPG intronless genes for the *hsh155-ds* in 5 μM Plad-B treatment. GO analysis takes a list of genes and finds what cellular processes (GO terms) they have in common. A p-value is calculated to determine the likelihood that the observed annotation of a particular GO term to our differentially expressed group of genes occurs by chance. For downregulated genes in *hsh155-ds* plus 5 μM Plad-B, we observed an enrichment of genes involved in de novo purine synthesis including ADE1, ADE2, ADE6, ADE12, ADE17, URA1, URA2, URA4, and HPT1 (Table 3-1). We conclude that purine synthesis is negatively impacted by a block in splicing.

| GOID | GO TERM | PVALUE | DOWNREGULATED GENE |
|---|---|---|---|
| GO:0046040 | IMP metabolic process | 4.12565E-10 | ADE2, ADE12, ADE6, ADE57, ADE17, HPT1, ADE1, ADE13 |
| GO:0009156 | ribonucleoside monophosphate biosynthetic process | 5.10706E-10 | ADE13, HPT1, URA4, ADE17, URA2, ADE57, ADE12, ADE1, ADE6, ADE2, URA1 |

**Table 3-1:** Gene ontology analysis of differentially expressed downregulated non-RPGs in response to 5 μM Plad-B.

# Chapter 4

## Inhibition of Co-transcriptional splicing

### 4.1 How is splicing co-transcriptionally inhibited?

It has been appreciated that pre-mRNA splicing occurs as soon as RNA polymerase transcribes past the ends of introns (Oesterreich et al. 2016). Research groups since this realization have come up with various methods of measuring connections between transcription elongation rate and splicing (Osheim et al. 1985, Howe et al. 2003, Braberg et al. 2013, Moehle et al. 2014, Aslanzadeh et al. 2018). Moreover, further acknowledgment of splicing differences on a for individual introns is important to fully understand this crosstalk. For example, splicing kinetics have been shown to be regulated by splice site strength from gene-to-gene (Oesterreich et al. 2016). The clear interplay between splicing and transcription rate motivated us to consider the effects co-transcriptional splicing inhibition may have gene-to gene. We employed single-molecule intron tracking (SMIT) to study these effects (Oesterreich et al. 2016).

SMIT is a single-molecule RNA-seq method that measures spliced mRNA reads at a given base pair position of RNA polymerase II (Pol II). WT and humanized yeast were incubated with 5 μM Plad-B for 15 minutes. Nascent RNA was then purified from chromatin, followed by a 3' end linker added to target the position Pol II for each nascent RNA, followed by PCR using gene-specific primers hybridized to the first

exons of 62 intron containing genes. After paired-end sequencing, splicing state was determined.

### 4.2 Ribosomal protein introns are more resistant to splicing inhibitors than other introns during co-transcriptional splicing

We detected variations in co-transcriptional splicing from gene-to gene. The most striking observation in out SMIT analysis was that in most ribosomal protein genes such as *RPS13, RPL35B, and RPS27B* seemed to be uninhibited by the presence of 5 μM Plad-B in the *hsh155-ds* strain when compared to WT HSH155 cells (Figure 4-1, red line versus blue line). Both WT HSH155 and *hsh155-ds* spliced at approximately equal rates for ribosomal protein genes. This contrasted to other genes such as *ARP9, IST1,* and *ERD2* which showed dramatically inhibited *hsh155-ds* co-transcriptional splicing in *hsh155-ds* with Plad-B compared to WT HSH155 yeast (Figure 4-1). In the case of *ARP9*, *hsh155-ds* yeast did not reach WT HSH155 co-transcriptional splicing rate, staying under 25% spliced for at least 1000 bp after the 3'SS. We conclude that splicing of nascent transcripts from ribosomal protein genes are less inhibited by Plad-B, whereas other genes are clearly sensitive to splicing inhibitors in a co transcriptional context.

Splicing inhibition has been used to study gene expression using a temperature-sensitive prp4-1 allele (Clark et al. 2002). Prp4 is essential

for splicing and shifting a temperature-sensitive prp4-1 strain can render splicing blocked. Using microarray probes that hybridize to the introns of various genes, the amount of unspliced RNA in this mutant was measured when cells were heat shifted to 37°C (Clark et al. 2002). The amount of intron present was measured as a calculated intron accumulation index. We therefore wanted to compare the intron microarray data from Clark et al. to an intron accumulation index of Plad-B inhibited *hsh155-ds*, using read counts that span the intron-exon junction.

We found a clear clustering on ribosomal protein genes showing a lower intron accumulation index relative to non-ribosomal protein genes (Figure 4-2, blue points). This was the case comparing microarray data to 5 uM Plad-B-inhibited. Genes from the SMIT analysis including *RPS13, RPL35B, and RPS27B* clustered on the lower end of the intron accumulation index while *ARP9, IST1,* and *ERD2* clustered into the non-ribosomal on the higher end of the intron accumulation index (green points). We concluded that certain introns within ribosomal protein genes are more resistant to Plad-B than compared to non-ribosomal protein genes.

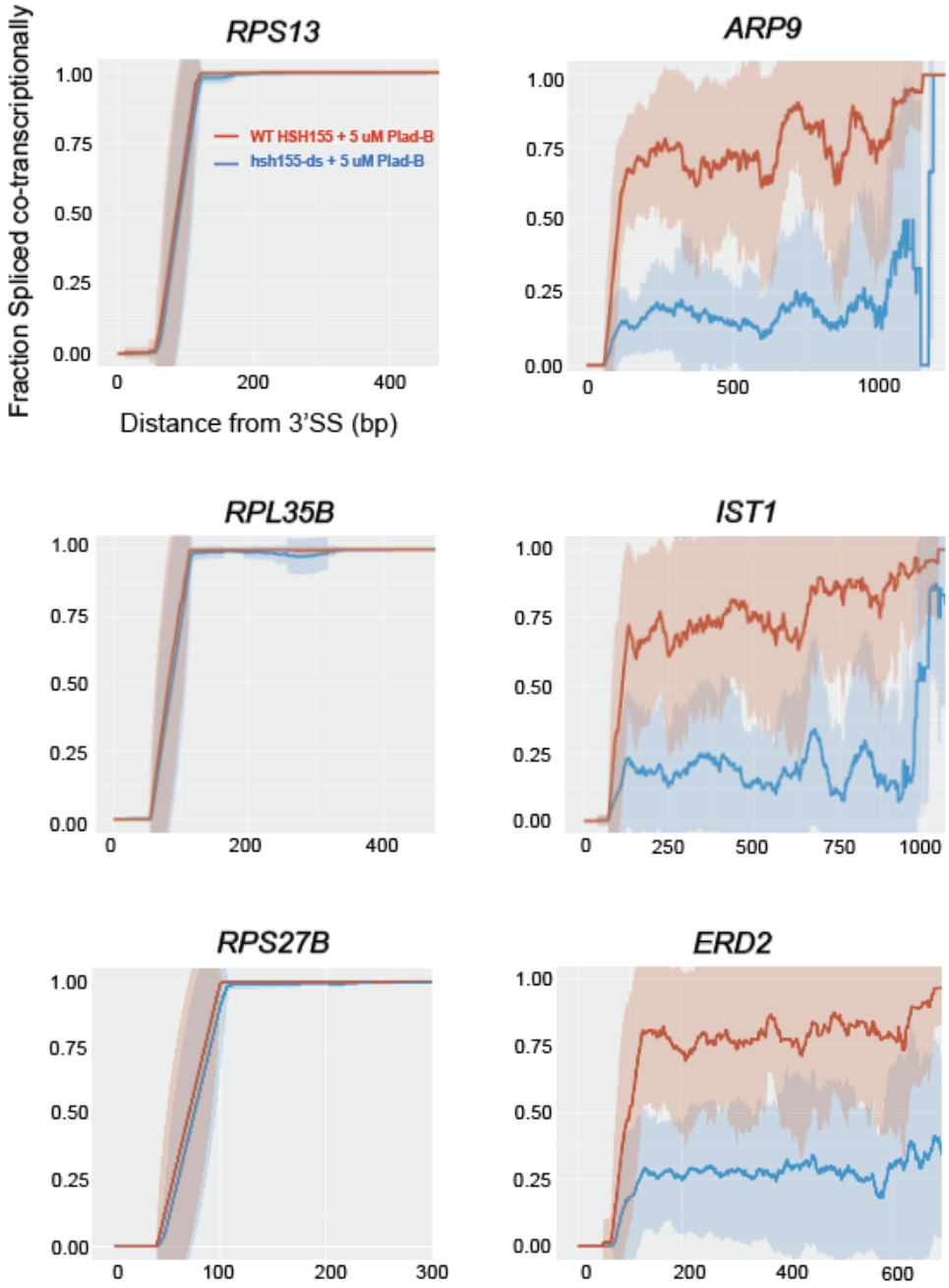



**Figure 4-1:** SMIT overlay profiles of 6 endogenous genes plotting fraction spliced co-transcriptionally versus distance from the 3' splice site (3'SS). Red line is WT HSH155 with 5 μM Plad-B. Blue line is *hsh155-ds* with 5 μM Plad B.

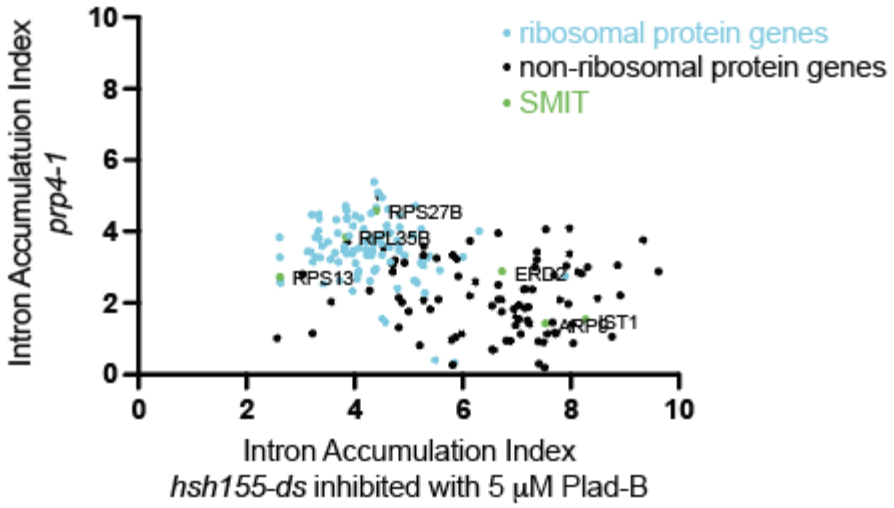



**Figure 4-2:** Plot of Intron Accumulation Index for blocked splicing caused by temperature-shifted *prp4-1* mutant yeast calculated by Clark et al. 2002. versus Intron Accumulation Index calculated in this work using retained intron junction counts from total RNA-sequencing of *hsh155-ds* inhibited by 5 µM Plad-B.

# Chapter 5

# Discussion and Future Directions

## 5.1 Summary

In this study we created and used a yeast strain to study the consequences of splicing inhibition. The strain contains 13 humanizing amino acid substitutions in the HSH155 protein, a critical component of the spliceosome. We focused on humanizing the hinge-like region upon which HSH155 protein closes around the U2 snRNA-pre-mRNA helix. In the presence of a class of spliceostatin-like inhibitors that bind this region however, *hsh155-ds* remains in an open conformation. We failed to detect significant growth defects at 18°C, 30°C, and 37°C compared to WT HSH155 yeast (Figure 2-3). This mutant strain, termed *hsh155-ds*, was then used to further study splicing inhibition.

To evaluate splicing inhibition, we incubated *hsh155-ds* yeast with 5 μM Plad-B or DMSO for various time points and performed RT-PCR on the unstable MATa1 transcript. We found that splicing inhibition is rapid and strong in the presence of Plad-B, accumulating upwards of 80% unspliced MATa*1* mRNA by 8 minutes (Figure 2-5, Figure 2-6, Figure 2-7). We then performed *in vitro* analysis on *hsh155-ds* yeast at various Plad-B concentrations, analyzing splicing reactions and spliceosome complex formation. As expected, splicing was inhibited *in vitro* as shown by the decreasing amounts of splicing products and pre-spliceosome complexes

(Figure 2-8A, Figure 2-8). We calculated the IC50 for *hsh155-ds* in both denaturing and nondenaturing conditions to be 25 nM and 27 nM respectively, the same order of magnitude as human splicing inhibition by the same method. Splicing in this strain can, therefore, be chemically inhibited *in vivo* and *in vitro* (Figure 2-8C, Figure 2-8D).

Our RNA-seq experiments captured a wider range of genes for analysis. Replicate splicing efficiencies were highly correlated for all 6 treatments (Figure 3-3). Comparing the splicing efficiencies of WT HSH155 with DMSO to WT HSH155 with 5 μM Plad-B indicating WT HSH155 yeast is resistant to Plad-B (Figure 3-4). When *hsh155-ds* with DMSO was compared to WT HSH155 with DMSO, we found that humanizing mutations have little to no effect on splicing (Figure 3-4). We then compared hsh155-ds incubated with 0.5 μM and 5 μM Plad-B to hsh155-ds in the DMSO carrier. These comparisons confirmed the dose-dependency of inhibitor concentration (Figure 3-5). Additionally, we observed a range of splicing efficiencies in response to Plad-B.

Differential gene expression analysis was necessary to capture individual gene expression changes more fully in response to splicing inhibition. Interestingly, splicing components including HSH155 did not significantly change in expression with any treatment. Instead, we found that yeast splicing is closely tied to ribosome protein gene expression and ribosome biogenesis. Blocking splicing in yeast has a strong negative effect

on RPG expression (Figure 3-6). Intronless RPGs were not spared, presumably by coordinated expression between intronless and intron-containing RPGs during ribosome biogenesis and assembly.

GO analysis revealed the downregulation of the important purine synthesis pathway genes in response to splicing inhibition. We observed gene expression changes for genes without introns upon inhibiting genes with introns. ADE1, ADE2, ADE6, ADE12, ADE17, URA1, URA2, URA4, and HPT1 stood out by consistently being downregulated in *hsh155-ds* for all drug treatments compared to DMSO treatment, suggesting a connection between splicing and purine biosynthesis (Table 3-1).

Inhibition of co-transcriptional splicing was analyzed. Using single molecule intron tracking (SMIT), we found that most ribosomal protein genes (including *RPS13, RPL35B, and RPS27B*) were less inhibited during the 15-minute incubation of 5 μM Plad-B (Figure 4-1). This contrasted with non-ribosomal protein genes (including *ARP9, IST1,* and *ERD2*) which were dramatically inhibited by 5 μM Plad-B. To confirm the apparent resistance of some RPGs as seen by SMIT, we compared total RNA-seq data to *prp4-1* splicing inhibition data by Clark et al. 2002. We see clear clustering of most, but not all, RPGs showing a lower intron accumulation index (Figure 4-2).

### 5.2 Discussion and Future Directions

We introduced an improved humanized yeast strain with a minimal set of 13 amino acid substitutions. They are stably integrated directly into the genome of yeast without the need for plasmids and genomic HSH155 gene replacement. This improved method for humanizing yeast, therefore, provides a feasible way to fine-tune chemical inhibition to measure a range of responses, rather than just primer extension and survivability on copper. Thus, the creation of such a strain is impactful for studying other steps in the gene expression pathway.

Our study was able to gather splicing efficacies on a per intron basis after splicing inhibition. However, the apparent splicing inhibition observed using % spliced as a measure is likely going to be influenced by mRNA transcription and decay. To make this data stronger, an accurate calculation of how much RNA for each gene was present at the time inhibitor was added to cell culture and how much mRNA was produced in the presence of inhibitor after addition remains crucial. These corrected calculations will include transcription, decay, and splicing rates for each gene.

Blocking splicing in *S. cerevisiae* has many potential consequences for gene expression. The greatest effect was the expression of ribosomal protein genes. These observations imply a vulnerability of RPGs to splicing inhibition. Interestingly, the decrease in gene expression of intron-containing ribosomal protein genes was mirrored by the decrease in intronless ribosomal protein

genes. I propose a model in which splicing inhibition leads to the accumulation of intron-containing transcripts that either cannot be translated or are translated into aberrant protein. In either case, the loss of wild type protein expression signals related intronless genes (intronless RPGs in this case) to be downregulated. The exact mechanism of this signal is still unclear.

Another example of this phenomena was observed in the downregulation of genes in the purine biosynthetic pathway. Blocking splicing led to the downregulation of several intronless genes in this pathway. One possible explanation for this observed downregulation is that the IMD4 intron-containing gene, which is a key player in this pathway, is also being downregulated due to its splicing being inhibited by Plad-B. I propose that the inhibition of IMD4 leads to the loss of IMD4 protein signaling intronless genes in this pathway to be downregulated. Interestingly, HIS1 was also observed to be downregulated. This is consistent with ADE and HIS1 being involved in the same pathway (Rébora et al. 2005). Future experiments will validate this hypothesis by rescuing IMD4 splicing inhibition in Plad-B-treated cells using an intronless IMD4 plasmid vector and analyzing expression of intronless genes in the purine biosynthetic pathway.

Our strain is useful for studying a wide variety of gene expression processes. Recent work, for example, has revealed intron secondary structures by the enrichment of intron sequence after splicing inhibition using

our strain (Rangan et al. 2022). It is also practical to simultaneously block transcription and splicing followed by measurement of RNA decay. Moreover, the kinetics of individual inhibitor molecules can be studied in our simple humanized model. How long splicing inhibitor binding takes, and affinity properties are of great importance for using these molecules as cancer therapies.

## References


Alsafadi S, Houy A, Battistella A, Popova T, Wassef M, Henry E, Tirode F, Constantinou A, Piperno-Neumann S, Roman-Roman S, et al. 2016. Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage. *Nat Commun* **7**: 10615.

AlShareef S, Ling Y, Butt H, Mariappan KG, Benhamed M, Mahfouz MM. 2017. Herboxidiene triggers splicing repression and abiotic stress responses in plants. *BMC Genomics* **18**. http://dx.doi.org/10.1186/s12864-017-3656-z.

Ares M, Grate L, Pauling MH. 1999. A handful of intron-containing genes produces the lion's share of yeast mRNA. *RNA* **5**: 1138–1139. http://dx.doi.org/10.1017/s1355838299991379.

Ares M Jr. 2013. Analysis of splicing in vitro using extracts of Saccharomyces cerevisiae. *Cold Spring Harb Protoc* **2013**: 982–985.

Aslanzadeh V, Huang Y, Sanguinetti G, Beggs JD. 2018. Corrigendum: Transcription rate strongly affects splicing fidelity and cotranscriptionality in budding yeast. *Genome Res* **28**: 606.2.

Bailey P, Chang DK, Nones K, Johns AL, Patch A-M, Gingras M-C, Miller DK, Christ AN, Bruxner TJC, Quinn MC, et al. 2016. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* **531**: 47–52.

Berglund JA, Abovich N, Rosbash M. 1998. A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes Dev* **12**: 858–867.

Biankin AV, Waddell N, Kassahn KS, Gingras M-C, Muthuswamy LB, Johns AL, Miller DK, Wilson PJ, Patch A-M, Wu J, et al. 2012. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature* **491**: 399–405.

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.

Bragulat M, Meyer M, Macías S, Camats M, Labrador M, Vilardell J. 2010. *RPL30* regulation of splicing reveals distinct roles for Cbp80 in U1 and U2 snRNP cotranscriptional recruitment. *RNA* **16**: 2033–2041. http://dx.doi.org/10.1261/rna.2366310.

Butt H, Eid A, Momin AA, Bazin J, Crespi M, Arold ST, Mahfouz MM. 2019. CRISPR directed evolution of the spliceosome for resistance to splicing inhibitors. *Genome Biol* **20**: 73.

Carrocci TJ, Paulson JC, Hoskins AA. 2018. Functional analysis of Hsh155/SF3b1 interactions with the U2 snRNA/branch site duplex. *RNA*

**24**: 1028–1040.

Carrocci TJ, Zoerner DM, Paulson JC, Hoskins AA. 2017. SF3b1 mutations associated with myelodysplastic syndromes alter the fidelity of branchsite selection in yeast. *Nucleic Acids Res* **45**: 4837–4852.

Clark TA, Sugnet CW, Ares M Jr. 2002. Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. *Science* **296**: 907–910.

Corrionero A, Miñana B, Valcárcel J. 2011. Reduced fidelity of branch point recognition and alternative splicing induced by the anti-tumor drug spliceostatin A. *Genes Dev* **25**: 445–459.

Cretu C, Agrawal AA, Cook A, Will CL, Fekkes P, Smith PG, Lührmann R, Larsen N, Buonamici S, Pena V. 2018. Structural Basis of Splicing Modulation by Antitumor Macrolide Compounds. *Mol Cell* **70**: 265–273.e8.

Cretu C, Gee P, Liu X, Agrawal A, Nguyen T-V, Ghosh AK, Cook A, Jurica M, Larsen NA, Pena V. 2021. Structural basis of intron selection by U2 snRNP in the presence of covalent inhibitors. *Nature Communications* **12**. http://dx.doi.org/10.1038/s41467-021-24741-1.

Darman RB, Seiler M, Agrawal AA, Lim KH, Peng S, Aird D, Bailey SL, Bhavsar EB, Chan B, Colla S, et al. 2015. Cancer-Associated SF3B1

Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point. *Cell Rep* **13**: 1033–1045.

DeBoever C, Ghia EM, Shepard PJ, Rassenti L, Barrett CL, Jepsen K, Jamieson CHM, Carson D, Kipps TJ, Frazer KA. 2015. Transcriptome sequencing reveals potential mechanism of cryptic 3' splice site selection in SF3B1-mutated cancers. *PLoS Comput Biol* **11**: e1004105.

DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. 2013. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Res* **41**: 4336–4343.

Effenberger KA, Anderson DD, Bray WM, Prichard BE, Ma N, Adams MS, Ghosh AK, Jurica MS. 2014. Coherence between cellular responses and in vitro splicing inhibition for the anti-tumor drug pladienolide B and its analogs. *J Biol Chem* **289**: 1938–1947.

Effenberger KA, Perriman RJ, Bray WM, Lokey RS, Ares M Jr, Jurica MS. 2013. A high-throughput splicing assay identifies new classes of inhibitors of human and yeast spliceosomes. *J Biomol Screen* **18**: 1110–1120.

Effenberger KA, Urabe VK, Jurica MS. 2017. Modulating splicing with small molecular inhibitors of the spliceosome. *Wiley Interdiscip Rev RNA* **8**. http://dx.doi.org/10.1002/wrna.1381.

Finci LI, Zhang X, Huang X, Zhou Q, Tsai J, Teng T, Agrawal A, Chan B, Irwin S, Karr C, et al. 2018. The cryo-EM structure of the SF3b spliceosome complex bound to a splicing modulator reveals a pre-mRNA substrate competitive mechanism of action. *Genes Dev* **32**: 309–320.

Fink GR. 1987. Pseudogenes in yeast? *Cell* **49**: 5–6.

Folco EG, Coil KE, Reed R. 2011. The anti-tumor drug E7107 reveals an essential role for SF3b in remodeling U2 snRNP to expose the branch point-binding region. *Genes Dev* **25**: 440–444.

Gallina I, Colding C, Henriksen P, Beli P, Nakamura K, Offman J, Mathiasen DP, Silva S, Hoffmann E, Groth A, et al. 2015. Cmr1/WDR76 defines a nuclear genotoxic stress body linking genome integrity and protein quality control. *Nat Commun* **6**: 6533.

Ghosh AK, Veitschegger AM, Nie S, Relitti N, MacRae AJ, Jurica MS. 2018. Enantioselective Synthesis of Thailanstatin A Methyl Ester and Evaluation of in Vitro Splicing Inhibition. *J Org Chem* **83**: 5187–5198.

Ghulam MM, Catala M, Abou Elela S. 2020. Differential expression of duplicated ribosomal protein genes modifies ribosome composition in response to stress. *Nucleic Acids Res* **48**: 1954–1968.

Gilbert W. 1978. Why genes in pieces? *Nature* **271**: 501.

Gozani O, Potashkin J, Reed R. 1998. A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. *Mol Cell Biol* **18**: 4752–4760.

Hansen SR, Nikolai BJ, Spreacker PJ, Carrocci TJ, Hoskins AA. 2019. Chemical Inhibition of Pre-mRNA Splicing in Living Saccharomyces cerevisiae. *Cell Chem Biol* **26**: 443–448.e3.

Harbour JW, Roberson EDO, Anbunathan H, Onken MD, Worley LA, Bowcock AM. 2013. Recurrent mutations at codon 625 of the splicing factor SF3B1 in uveal melanoma. *Nat Genet* **45**: 133–135.

Herruer MH, Mager WH, Raué HA, Vreken P, Wilms E, Planta RJ. 1988. Mild temperature shock affects transcription of yeast ribosomal protein genes as well as the stability of their mRNAs. *Nucleic Acids Res* **16**: 7917–7929.

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

Irimia M, Penny D, Roy SW. 2007. Coevolution of genomic intron number and splice sites. *Trends Genet* **23**: 321–325.

Janke C, Magiera MM, Rathfelder N, Taxis C, Reber S, Maekawa H, Moreno-Borchart A, Doenges G, Schwob E, Schiebel E, et al. 2004. A versatile

toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**: 947–962.

Jeong H, Herskowitz I, Kroetz DL, Rine J. 2007. Function-altering SNPs in the human multidrug transporter gene ABCB1 identified using a Saccharomyces-based assay. *PLoS Genet* **3**: e39.

Kao C-Y, Cao E-C, Wai HL, Cheng S-C. 2021. Evidence for complex dynamics during U2 snRNP selection of the intron branchpoint. *Nucleic Acids Res* **49**: 9965–9977.

Kastner B, Will CL, Stark H, Lührmann R. 2019. Structural Insights into Nuclear pre-mRNA Splicing in Higher Eukaryotes. *Cold Spring Harb Perspect Biol* **11**. http://dx.doi.org/10.1101/cshperspect.a032417.

Kistler AL, Guthrie C. 2001. Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for sub2, an essential spliceosomal ATPase. *Genes Dev* **15**: 42–49.

Komili S, Farny NG, Roth FP, Silver PA. 2007. Functional specificity among ribosomal proteins regulates gene expression. *Cell* **131**: 557–571.

Kos M, Tollervey D. 2010. Yeast pre-rRNA processing and modification occur cotranscriptionally. *Mol Cell* **37**: 809–820.

Kotake Y, Sagane K, Owa T, Mimori-Kiyosue Y, Shimizu H, Uesugi M,

Ishihama Y, Iwata M, Mizui Y. 2007. Splicing factor SF3b as a target of the antitumor natural product pladienolide. *Nat Chem Biol* **3**: 570–575.

Krämer A, Utans U. 1991. Three protein factors (SF1, SF3 and U2AF) function in pre-splicing complex formation in addition to snRNPs. *The EMBO Journal* **10**: 1503–1509. http://dx.doi.org/10.1002/j.1460-2075.1991.tb07670.x.

Kressler D, Hurt E, Bassler J. 2010. Driving ribosome assembly. *Biochim Biophys Acta* **1803**: 673–683.

Liang W-W, Cheng S-C. 2015. A novel mechanism for Prp5 function in prespliceosome formation and proofreading the branch site sequence. *Genes Dev* **29**: 81–93.

Ling Y, Alshareef S, Butt H, Lozano-Juste J, Li L, Galal AA, Moustafa A, Momin AA, Tashkandi M, Richardson DN, et al. 2017. Pre-mRNA splicing repression triggers abiotic stress signaling in plants. *Plant J* **89**: 291–309.

Lo C-W, Kaida D, Nishimura S, Matsuyama A, Yashiroda Y, Taoka H, Ishigami K, Watanabe H, Nakajima H, Tani T, et al. 2007. Inhibition of splicing and nuclear retention of pre-mRNA by spliceostatin A in fission yeast. *Biochem Biophys Res Commun* **364**: 573–577.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Marcet-Houben M, Gabaldón T. 2015. Beyond the Whole-Genome Duplication: Phylogenetic Evidence for an Ancient Interspecies Hybridization in the Baker's Yeast Lineage. *PLoS Biol* **13**: e1002220.

Martin M, Maßhöfer L, Temming P, Rahmann S, Metz C, Bornfeld N, van de Nes J, Klein-Hitpass L, Hinnebusch AG, Horsthemke B, et al. 2013. Exome sequencing identifies recurrent somatic mutations in EIF1AX and SF3B1 in uveal melanoma with disomy 3. *Nat Genet* **45**: 933–936.

Miller AM. 1984. The yeast MATa1 gene contains two introns. *EMBO J* **3**: 1061–1065.

Mizui Y, Sakai T, Iwata M, Uenaka T, Okamoto K, Shimizu H, Yamori T, Yoshimatsu K, Asada M. 2004. Pladienolides, new substances from culture of Streptomyces platensis Mer-11107. III. In vitro and in vivo antitumor activities. *J Antibiot* **57**: 188–196.

Moehle EA, Braberg H, Krogan NJ, Guthrie C. 2014. Adventures in time and space: splicing efficiency and RNA polymerase II elongation rate. *RNA Biol* **11**: 313–319.

Munding EM, Shiue L, Katzman S, Donohue JP, Ares M Jr. 2013. Competition between pre-mRNAs for the splicing machinery drives global regulation of splicing. *Mol Cell* **51**: 338–348.

Nicolaou KC, Rekula SR, Kumar SM, Podilapu AR, Matuszak RP, Jung PM,

Lam LT, Phillips AC, Lyssikatos J, Munneke S, et al. 2021. Design, Synthesis, and Biological Investigation of Thailanstatin A and Spliceostatin D Analogues Containing Tetrahydropyran, Tetrahydrooxazine, and Fluorinated Structural Motifs. *J Org Chem* **86**: 2499–2521.

Oesterreich FC, Herzel L, Straube K, Hujer K, Howard J, Neugebauer KM. 2016. Splicing of Nascent RNA Coincides with Intron Exit from RNA Polymerase II. *Cell* **165**: 372–381.

Osheim YN, Miller OL Jr, Beyer AL. 1985. RNP particles at splice junction sequences on Drosophila chorion transcripts. *Cell* **43**: 143–151.

Padgett RA, Konarska MM, Grabowski PJ, Hardy SF, Sharp PA. 1984. Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. *Science* **225**: 898–903.

Parenteau J, Durand M, Morin G, Gagnon J, Lucier J-F, Wellinger RJ, Chabot B, Elela SA. 2011. Introns within ribosomal protein genes regulate the production and function of yeast ribosomes. *Cell* **147**: 320–331.

Peña C, Hurt E, Panse VG. 2017. Eukaryotic ribosome assembly, transport and quality control. *Nat Struct Mol Biol* **24**: 689–699.

Perriman R, Ares M. 2010. Invariant U2 snRNA Nucleotides Form a Stem Loop to Recognize the Intron Early in Splicing. *Molecular Cell* **38**: 416–

427. http://dx.doi.org/10.1016/j.molcel.2010.02.036.

Plaschka C, Lin P-C, Charenton C, Nagai K. 2018. Prespliceosome structure provides insights into spliceosome assembly and regulation. *Nature* **559**: 419–422.

Plaschka C, Newman AJ, Nagai K. 2019. Structural Basis of Nuclear pre-mRNA Splicing: Lessons from Yeast. *Cold Spring Harbor Perspectives in Biology* **11**: a032391. http://dx.doi.org/10.1101/cshperspect.a032391.

Quesada V, Conde L, Villamor N, Ordóñez GR, Jares P, Bassaganyas L, Ramsay AJ, Beà S, Pinyol M, Martínez-Trillos A, et al. 2011. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet* **44**: 47–52.

Rangan R, Hunter O, Pham P, Ares M, Das R. RNA structure landscape of *S. cerevisiae* introns. http://dx.doi.org/10.1101/2022.07.22.501175.

Rébora K, Laloo B, Daignan-Fornier B. 2005. Revisiting purine-histidine cross-pathway regulation in Saccharomyces cerevisiae: a central role for a small molecule. *Genetics* **170**: 61–70.

Rodriguez-Vargas S, Estruch F, Randez-Gil F. 2002. Gene expression analysis of cold and freeze stress in Baker's yeast. *Appl Environ Microbiol* **68**: 3024–3030.

Roybal GA, Jurica MS. 2010. Spliceostatin A inhibits spliceosome assembly subsequent to prespliceosome formation. *Nucleic Acids Res* **38**: 6664–6672.

Sakai T, Asai N, Okuda A, Kawamura N, Mizui Y. 2004a. Pladienolides, new substances from culture of Streptomyces platensis Mer-11107. II. Physico-chemical properties and structure elucidation. *J Antibiot* **57**: 180–187.

Sakai T, Sameshima T, Matsufuji M, Kawamura N, Dobashi K, Mizui Y. 2004b. Pladienolides, new substances from culture of Streptomyces platensis Mer-11107. I. Taxonomy, fermentation, isolation and screening. *J Antibiot* **57**: 173–179.

Sales-Lee J, Perry DS, Bowser BA, Diedrich JK, Rao B, Beusch I, Yates JR 3rd, Roy SW, Madhani HD. 2021. Coupling of spliceosome complexity to intron diversity. *Curr Biol* **31**: 4898–4910.e4.

Serrat X, Kukhtar D, Cornes E, Esteve-Codina A, Benlloch H, Cecere G, Cerón J. 2019. CRISPR editing of sftb-1/SF3B1 in Caenorhabditis elegans allows the identification of synthetic interactions with cancer-related mutations and the chemical inhibition of splicing. *PLoS Genet* **15**: e1008464.

Sharp PA. 1985. On the origin of RNA splicing and introns. *Cell* **42**: 397–400.

Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, Nik-Zainal S, Martin S, Varela I, Bignell GR, et al. 2012. The landscape of cancer genes and mutational processes in breast cancer. *Nature* **486**: 400–404.

Stevens SW, Abelson J. 2002. Yeast pre-mRNA splicing: methods, mechanisms, and machinery. *Methods Enzymol* **351**: 200–220.

Talkish J, Igel H, Hunter O, Horner SW, Jeffery NN, Leach JR, Jenkins JL, Kielkopf CL, Ares M. 2019a. Cus2 enforces the first ATP-dependent step of splicing by binding to yeast SF3b1 through a UHM-ULM interaction. *RNA* **25**: 1020–1037.

Talkish J, Igel H, Perriman RJ, Shiue L, Katzman S, Munding EM, Shelansky R, Donohue JP, Ares M Jr. 2019b. Rapidly evolving protointrons in Saccharomyces genomes revealed by a hungry spliceosome. *PLoS Genet* **15**: e1008249.

Tang Q, Rodriguez-Santiago S, Wang J, Pu J, Yuste A, Gupta V, Moldón A, Xu Y-Z, Query CC. 2016. SF3B1/Hsh155 HEAT motif mutations affect interaction with the spliceosomal ATPase Prp5, resulting in altered branch site selectivity in pre-mRNA splicing. *Genes Dev* **30**: 2710–2723.

Vilardell J, Chartrand P, Singer RH, Warner JR. 2000. The odyssey of a regulated transcript. *RNA* **6**: 1773–1780.

Vilardell J, Warner JR. 1997. Ribosomal protein L32 of Saccharomyces cerevisiae influences both the splicing of its own transcript and the processing of rRNA. *Molecular and Cellular Biology* **17**: 1959–1965. http://dx.doi.org/10.1128/mcb.17.4.1959.

Wahl MC, Will CL, Lührmann R. 2009. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell* **136**: 701–718. http://dx.doi.org/10.1016/j.cell.2009.02.009.

Wang L, Lawrence MS, Wan Y, Stojanov P, Sougnez C, Stevenson K, Werner L, Sivachenko A, DeLuca DS, Zhang L, et al. 2011. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N Engl J Med* **365**: 2497–2506.

Wang Q, Zhang L, Lynn B, Rymond BC. 2008. A BBP-Mud2p heterodimer mediates branchpoint recognition and influences splicing substrate abundance in budding yeast. *Nucleic Acids Res* **36**: 2787–2798.

Wapinski I, Pfiffner J, French C, Socha A, Thompson DA, Regev A. 2010. Gene duplication and the evolution of ribosomal protein gene regulation in yeast. *Proc Natl Acad Sci U S A* **107**: 5505–5510.

Warner JR. 1999. The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci* **24**: 437–440.

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.

Yokoi A, Kotake Y, Takahashi K, Kadowaki T, Matsumoto Y, Minoshima Y, Sugi NH, Sagane K, Hamaguchi M, Iwata M, et al. 2011. Biological validation that SF3b is a target of the antitumor macrolide pladienolide. *FEBS J* **278**: 4870–4880.

Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, Sato Y, Sato-Otsubo A, Kon A, Nagasaki M, et al. 2011. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**: 64–69.

Zamore PD, Patton JG, Green MR. 1992. Cloning and domain structure of the mammalian splicing factor U2AF. *Nature* **355**: 609–614.

Zhuang Y, Weiner AM. 1986. A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. *Cell* **46**: 827–835. http://dx.doi.org/10.1016/0092-8674(86)90064-4.

Zhu T, Niu D-K. 2013. Mechanisms of intron loss and gain in the fission yeast Schizosaccharomyces. *PLoS One* **8**: e61683.

## Appendix

## MATERIALS AND METHODS

### Strains:

*YOH001*

YOH001 was constructed from a JRY8012 background (BY4741 MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 with Δpdr5 Δsnq2 Δyor1, (Jeong et al. 2007) using the CRISPR/Cas9 yeast system. First, co-transformation of pET416-TEF1p-Cas9 plasmid (DiCarlo et al. 2013) with a URA3 selectable marker and a double-stranded rescue 1021 bp oligonucleotide carrying humanizing mutations was performed. CRISPR/Cas9 system uses a designed guide RNA (gRNA) in complex with Cas9 endonuclease, which generates a double-stranded DNA break (DSB). We designed the gRNA to have homology near the hinge region of Hsh155. Transformant colonies were selected from a plate lacking uracil and restreaked on another plate lacking uracil. Next, colonies were restreaked on plates containing 5-fluoroorotic acid to select for colonies that lost the plasmid. These colonies were then grown on YDP media plates (yeast extract, peptone, dextrose) and subjected to colony PCR to amplify the genomic DNA coding region of HEATs 15 and 16. Humanizing mutations in this region were confirmed by Sanger sequencing of the PCR product (UC Berkeley DNA sequencing facility). This strain was used in all experiments unless otherwise indicated.

*YOH001-GFP*

YOH001 (and WT Hsh155) strains were transformed with a 2493 bp PCR product bearing GFP and HIS3 genes with homologous arms to the C-terminus of the DNA coding region of the *HSH155* gene. The PCR product was generated from a pYM28-HIS3MX6 plasmid (Janke et al 2004) using S2/S3 primers (DiCarlo et al. 2013; Janke et al. 2004) adapted for C-terminal *HSH155* gene-tagging. Transformants were selected from a plate lacking histidine. Colony PCR was performed on WT HSH155 and *hsh155-ds* strains for confirm integration of GFP and the PCR product further confirmed by Sanger sequencing (UC Berkeley Sequencing Facility).

*Growth Assay*

WT and *hsh155-ds* cells were cultured overnight in 5mL YPD media at 30℃ shaking at 220 rpm. Cells were then spun down for 5 mins and washed with 5mL sterile water twice. Cells were then diluted down and allowed to grow up to OD600= 0.5. 100 ul of this culture was loaded in 96-well plates and diluted by 10X a total of four times. 5 µL of cells from each dilution were dropped onto YPD plates and incubated at 18°C, 30°C, and 37°C for two days. Experiment was done in triplicate.

*Western Blot*

WT HSH155-GFP, *hsh155-ds*-GFP, and Sir2-GFP cell were grown to log phase in 30°C overnight with shaking at 220 rpm. The next day, cells were harvested at OD600=0.5. Whole-cell extracts were prepared by trichloroacetic acid extraction as previously described by Gallina et al. 2015. Extracts were then run on a NuPAGE 4-12% Bis-Tris polyacrylamide gel (Invitrogen) for 1 hour at 100V. Next, proteins were transferred to a methanol treated PVDF membrane (Bio-Rad) for 1 hour at 350 mAmps on ice. GFP fusion protein were probed using primary mouse anti-GFP antibody (Santa Cruz Biotechnology) overnight in 4°C. The next day, secondary goat anti-mouse (680RD, LiCor) antibody was used to probe the membrane. The membrane was imaged on the LiCor infrared system.

*Time Course*

YOH001 cells were cultured overnight and diluted down to OD600=0.5 the next day. Plad-B was diluted in culture to 5 μM (0.05% DMSO) and incubated for either 0, 1, 4, or 8 minutes. After time, 8-9 mL cells were placed in 10 mL freezing 100% ethanol (pre-chilled on dry ice) to stop metabolism instantaneously (Kos and Tollervey 2010). Total RNA was extracted from all samples by hot phenol/chloroform extraction. Reverse transcription polymerase chain reaction (RT-PCR) was performed on the RNA extracted using primers annealing to the first intron of the *MATA1* gene. Bnads were

quantified using ImageJ (NIH) and data was visualized using GraphPad Prism (v9.4.0 for macOS).

Single Molecule Intron Tracking (SMIT)

Wild type and humanized yeast cells were grown in YPD media overnight. The next day, cells were incubated with 5 μM Plad-B at OD600 = 0.5 for 15 mins. From here, chromatin preparation, gene targeting, and sequencing were done as described in Oesterreich et al. 2016.

*Splicing extract preparation and in vitro splicing assays*

$^{32}$P-radiolabeled actin pre-mRNA was transcribed in vitro using the MEGAscript T7 transcription kit (Invitrogen). *S. cerevisiae* splicing extracts were prepared from the *hsh155 hum 15.7-16* strain using the liquid nitrogen method described in (Stevens and Abelson 2002), except frozen cells were disrupted using a Retsch MM301 ball mill for 3 minutes at 10 Hz for 5 cycles. ATP was depleted from extracts for 20 minutes at 23 °C, using 1U of *S. cerevisiae* hexokinase (Sigma-Aldrich) and 16 mM D-glucose, in the presence of DMSO or the indicated concentration of pladienolide-B (Santa Cruz Biotechnology). After depletion of ATP, $^{32}$P-radiolabled actin pre-mRNA was added to a final concentration of 0.4 nM and standard splicing reactions were carried out in the presence of water or 20 mM ATP at 23 °C for 20 minutes as described in (Ares 2013). To visualize the precursors and products of the

splicing reaction, reactions were quenched in 200 µL of RNA extraction buffer (0.3 M NaOAc, 0.2% SDS, 1 mM EDTA, 10 µg/mL proteinase K) and incubated at 65°C for 10 min. RNA was extracted from the reactions using 200 µl of acid phenol (VWR), ethanol precipitated, resolved by electrophoresis on 6% acrylamide/8M urea gels, and detected by phosphorimaging. Splicing complexes were visualized by mixing splicing reactions with 2X native loading dye (20 mM Tris/glycine, 25% glycerol, 0.1% bromophenol blue, and 1 mg/mL heparin) loaded directly on 2.1% agarose gels as described in (Effenberger et al. 2013) and visualized by phosphorimaging. Splicing efficiency and ATP-dependent complex formation were quantified using ImageJ (NIH) and data was visualized using GraphPad Prism (v9.4.0 for macOS).

*Reverse transcription Polymerase Chain Reaction (RT-PCR) and Quantification*

After extraction, total RNA was subjected to DNAse treatment (TURBO DNAse, Invitrogen). First strand (FS) synthesis was performed using a 5X FS master mix composed of 0.5 µL 1X FS buffer (Invitrogen), 0.5 µL random primer mix, 5 µg RNA, and up to 7 µL of water. This mixture was heated to 95℃ for 1 minute, followed by 65℃ for 1 minute, then room temperature for 1-2 minutes. All 7 µL were added to 5 µL of 1X reverse transcrptitase (RTase) master mix. 1X RTase master mix included 1.5 µL 5X FS buffer, 1 µL 0.1 M

dithiothreitol (DTT, Invitrogen), 1 µL 10 mM dNTPs (Thermo Scientific), 0.5 µL 40 U/µL RNAse inhibitor (RNAseIN, Promega), and 0.5 µL 200 U/uL SuperScript III RTase (Invitrogen) or water for no RTase control. The final 12 uL mixture was incubvated at room temperature for 5 minutes followed by 48℃ for 25 minutes. Zymo Reasech DNA Clean & Concentrator-5 kit was used to clean single stranded DNA (ssDNA). PCR was performed on ssDNA and products were run on an 2% agarose gel. Bands were quantified using ImageJ (NIH) and data was visualized using GraphPad Prism.

*RNA-Sequencing of Total RNA: Library Prep*

Six cDNA libraries were created from the following treatments: WT Hsh55 treated with 0.05% DMSO (0 µM inhibitor), WT Hsh155 treated with 5 µM Plad-B, *hsh155-ds* cells treated with 0.05% DMSO (0 µM inhibitor), *hsh155-ds* treated with 0.5 µM Plad-B, *hsh155-ds* treated with 5 µM Plad-B, and *hsh155-ds* treated with 5 µM Thail-A. Total RNA was extracted from cells. The quality of extracted RNA was assessed using High Sensitivity RNA TapeStation (Agilent). rRNA depletion was done using the Illumina Stranded Total RNA Prep kit, Ligation with Ribo-Zero Plus kit. After barcoding and amplification, libraries were then sequenced using the NovaSeq 6000 system (Illumina Inc).

*Read alignment, splicing and gene expression analysis*

Raw sequencing reads were aligned to the annotated SacCer3 genome using STAR. To calculates splicing efficiency for intron-containing genes, JunctionCounts was used (https://github.com/ajw2329/junctionCounts). The % spliced was calculated by dividing the exon-exon junction count by the sum the exon-exon and intron-exon count. Gene expression analysis was performed using DESeq2 (Love et al. 2014). Data was visualized using GraphPad Prism.

*Gene Ontology (GO) Analysis*

Introness genes found to be significantly downregulated according DESeq2 differential gene expression analysis were fed into the Gene Ontology Term Finder tool (Version 0.86) from the Saccharomyces Genome Database (https://www.yeastgenome.org). Gene Ontology aspect “Process” was chosen using a p-value of 0.01.

**PLASMIDS**

| **Plasmid** | **Description** | **Source** |
|---|---|---|
| pET416-TEF1p-Cas9 | Cas9, guide RNA guiding Cas9 to genomic region coding for HEATs 15 and 16 of HSH155 | (DiCarlo et al. 2013) |

| | | |
|---|---|---|
| pYM28-GFP-HIS3MX6 | EGFP, HIS3MX6 marker | (DiCarlo et al. 2013; Janke et al. 2004) |

**OLIGONUCLEOTIDES USED IN THIS STUDY**

| **Name** | **Sequence of top strand 5' to 3'** | **Purpose** |
|---|---|---|
| CRISPR guide RNA | TCCGCAGTGAAAGATAAATG ATCTTCAGCAATGAATCCAA ACGGTTTTAGAGCTAGAAAT AGCAAG | Guides Cas9 endonuclease to genomic region coding for HEATs 15 and 16 of HSH155 |
| CRISPR humanized rescue oligo | 1021 bp sequence included as separate text file | DNA oligo co-transformed with CRISPR plasmid to rescue double stranded break. Includes humanizing mutations for all yeast-human differences in HEATs 15 and 16 of HSH155 |
| GFP tag PCR product | 2492 bp sequence included as separate text file | PCR product generated from pYM28-GFP-HIS3MX6 using S2/S3 primers. This PCR was transformed and |

| | | |
|---|---|---|
| | | integrated into the genome to tag HSH155 with GFP |
| Primer S2 | GTGATGATATAGGTGTGTCA AGTAAAATATTCTTACAAGT TGTGGTTATTTATATGCTCT ATATATATTCAATCGATGAA TTCGAGCTCG | Primer 1 to generate GFP tag PCR product last 19 nucleotides anneal to pYM28-GFP-HIS3MX6 plasmid and creates a PCR product with GFP and HIS3MX6 marke with homology C-terminus of HSH155 |
| Primer S3 | GCCATGGTACCTTTTTACCC CGTTACACCAGACAACAATG AAGAATATATAGAAGAACTG GATTTAGTTCTGCGTACGCT GCAGGTCGAC | Primer 2 to generate GFP tag PCR product last 18 nucleotides anneal to pYM28-GFP-HIS3MX6 plasmid and creates a PCR product with GFP and HIS3MX6 marke with homology C-terminus of HSH155 |

| | | |
|---|---|---|
| matF | GGCGGAAAACATAAACAGAACTC | Forward primer of MATa1 RT-PCR time course. Sits in first exon. |
| matR | AAAGAGAAGAGCCCAAAGGG | Reverse primer of MATa1 RT-PCR time course. Sits in second exon. |
| rRNA depletion oligos (RNAse H) | Included as a separtae .xlsx file | DNA oligos that anneal to rRNA molecules for RNAse H depletion |
| SMIT oligos | Oligo list can be found in Oesterreich et al. 2016 | Targeting SMIT genes for sequencing |

## STRAINS USED IN THIS STUDY

| **Name** | **Genotype** | **Description** | **Source** |
|---|---|---|---|
| JRY8012 | BY4741 MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 Δpdr5 Δs*nq2 Δyor1;* WT *HSH155* | BY4741 with deletion of three drug transporters *PDR5* *SNQ2* and *YOR1* | (Jeong et al. 2007) |

| | | | |
|---|---|---|---|
| JRY8012-GFP | BY4741 *Δpdr5 Δsnq2 Δyor1; HSH155-GFP*+ | C-terminal green fluorescent protein tag on genomic HSH155 | This study |
| YOH001 | BY4741 *Δpdr5 Δsnq2 Δyor1; hsh155-ds* | Adapted from JRY8012 strain. Includes humanizing amino acid changes in HSH155 gene from amino acids 746 to 793 | This study |
| YOH001-GFP | BY4741 *Δpdr5 Δsnq2 Δyor1; hsh155-ds-GFP* | YOH001 strain with GFP following the C-terminal end of *HSH155* gene | This study |
| SIR2-GFP | Sir2-GFP | SIR2-GFP fusion protein | Kind gift from Dr. Rohinton Kamakaka |