UNIVERSITY OF CALIFORNIA

SANTA CRUZ

RNA SPLICING MECHANISMS IN RARE DISEASE

A dissertation submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MOLECULAR, CELL, AND DEVELOPMENTAL BIOLOGY

by

Victor Tse

June 2024

The Dissertation of Victor Tse is approved:

_____

Professor Jeremy Sanford, Chair

_____

Distinguished Professor Carol Greider

_____

Professor Melissa Jurica

_____

Professor Michael Stone

_____

Peter Biehl, Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Thesis Abstract

RNA Splicing Mechanisms in Rare Disease

By Victor Tse

The accurate expression of human genes requires precursor messenger RNA (pre-mRNA) splicing. Pre-mRNA splicing removes superfluous sequences (introns) from pre-mRNA and ligates protein-coding regions (exons) to yield mature messenger RNA (mRNA). To catalyze this reaction, the spliceosome, a large ribonucleoprotein complex, must assemble *de novo* on each intron. The spliceosome is composed of five core uracil-rich small nuclear RNAs and hundreds of associated proteins. A critical early step in the spliceosome assembly pathway is the recognition of exon-intron boundaries. These demarcations are defined by conserved consensus splicing sequences at the 5′ and 3′ ends of an intron. The efficient recruitment of early spliceosomal components to these landmarks is determined by *trans*-regulatory splicing factors that bind and activate *cis*-regulatory sequences that function as splicing enhancers or silencers. Together, the dynamic interplays between splicing factors and splicing regulatory sequences establish exon definition, defining and regulating an exon's identity for pre-mRNA splicing.

The exact mechanisms that control exon definition remain poorly understood due to *cis*- and *trans*-acting contextual and positional dependencies. The failure to

correctly define exons is a problem that can result in aberrant pre-mRNA splicing, a phenomenon that underlies numerous human diseases. Gaps in knowledge of the possible mechanisms that may regulate an exon's identity is a fundamental challenge that impedes the development of precision medicines for disease-linked exons.

Here, my thesis research addresses the fields' gap in knowledge surrounding mechanisms that control exon definition in a human disease context. Namely, my work adds to the growing understanding that certain exons are indeed fragile and highly vulnerable to pathogenic mutations that dysregulate splicing regulatory sequences. More intriguingly, my accomplished research suggests RNA structures may play a central role in exon identity than previously appreciated. Additionally, my cumulative research activity presents a novel hypothesis that RNA structures may have clinical relevance as a therapeutic target for precision RNA drugs. For these reasons, I believe my thesis work advances our understanding of the molecular mechanisms that control exon definition fidelity, and has potential broader impacts on developing precision medicine for patients with unmet needs.

Acknowledgements

To preface, I apologize if you are not mentioned in this section; there are simply too many people I have met and bonded with over the course of my individual and academic journey, so please do know that I am forever grateful for the coffee and conversations relating to the curiosities and wonders of life. To those who know how special you are to me, you are always on my mind; words cannot justify how much you and your support means to me.

I would like to first and foremost thank my parents, Jing Yao Tse and Qi Yi Ye, for their unconditional love. Without their sacrifices to persevere through socioeconomic and cultural barriers in an unfamiliar world to build futures for themselves and three boys—in spite of limited resources and guidance—I would not be in the position that I am in today, nor would I have learned how to remain resilient and disciplined in tackling the challenges of life. To my older brother and twin, Nathan and Vincent, thank you two for all of the support and laughs as we tackle this whacky game of life together.

Next, I would like to thank Jeremy Sanford for being the leading figure that shepherded me into scientific research, and the one who proactively gets me to think about what the next steps should be in respect to my science, and my life. Without his mentorship, encouragement, and constant drive to get me involved with as many

things as possible, I would not be the scientist and individual that I am today. Thank you, Jeremy. To the Sanford lab, current and past members, I am eternally grateful for the camaraderie and memories we all share, and I will miss these moments dearly. To Michael Stone, I am grateful for his mentorship and expertise, and to he and the Stone Research Group, I am thankful they welcomed me with a second lab home. To my thesis committee, and the UCSC biomedical research community, thank you for showing me what excellent science is as we work towards advancing humanity.

Chapter 1: Introduction to pre-mRNA splicing and its implications in human disease

The discovery of pre-mRNA splicing

DNA, the blueprint of life that cannot act on its own, must be transcribed into precursor messenger RNA (pre-mRNA) molecules to mediate the expression of genes (Cobb 2017; Crick 1970). Pre-mRNA contains protein-coding sequences (exons) and non-coding sequences (introns). Further processing of pre-mRNAs into mature messenger RNAs (mRNAs) is required for the transport and proper expression of genes to enable correct protein synthesis by ribosomes. In respect to the former aspect of what is required for their transport, processing of a pre-mRNA includes capping at the 5′ end (Both et al. 1975; Wei and Moss 1974), and polyadenylation at the 3′ end (Edmonds, Vaughan, and Nakazato 1971; S. Y. Lee, Mendecki, and Brawerman 1971; J. E. Darnell, Wall, and Tushinski 1971). In respect to the latter aspect of what is required to establish accurate gene expression, the removal of introns and joining of exons must take place to create a functional, interpretable mRNA. This process is known as pre-mRNA splicing and it was not discovered or understood until the late 1970s.

The first hint of pre-mRNA splicing was published by two research groups (the labs of Richard J. Roberts, and Phillip A. Sharp) that were contemporarily

studying where adenovirus mRNA mapped to within their respective DNA genome using electron microscopy (Berget, Moore, and Sharp 1977; Chow et al. 1977). Incidentally from their R-looping experiments, they observed that only specific segments of the genome hybridized to the adenovirus mRNA, and other parts did not hybridize at all, creating unpaired regions of "looped" DNA. The two groups essentially showed that only specific segments of genomic DNA are required for the synthesis of mature mRNA. In other words, these initial observations pointed to the notion that introns interspersing protein-coding sequences must be removed, followed by the joining of exon sequences to create functional mRNAs.

Soon after these initial publications, the lab of Phillip A. Sharp and colleagues continued to unveil pre-mRNA splicing using a combination of molecular biology techniques and electron microscopy. Radiolabeled adenovirus DNA fragments were used to selectively hybridize and capture polyadenylated RNA produced from adenovirus-infected HeLa cells (Berk and Sharp 1977). Using an S1 endonuclease assay with these DNA:RNA hybrids, *Berk and Sharp* (1977) were then able to precisely size these RNA and map the exact positions where intronic sequences and exonic sequences originated from using the adenovirus genome. Through this assay, they validated that only specific sequences from the genome are protein-coding and are essential for a functional mRNA. This finding was further corroborated by contemporary research published by the lab of James E. Darnell, Jr. (Weber, Jelinek, and Darnell 1977). Additional work using electron microscopy examined steady-state

nuclear RNA of adenovirus-infected HeLa cells (Berget and Sharp 1979). Based on what was known at the time, *Berget and Sharp* (1979) detected nuclear mRNA intermediate products that were consistent with heterogeneous nuclear RNA (hnRNA) processing of pre-mRNA. These nuclear intermediates were determined to contain significantly less intronic sequences when mapped back to the adenovirus DNA genome. At this point in time, the collective data supported the split gene theory which hypothesized that the protein-coding sequences of genes are split up by interspersing sequences, and that these coding units must be spliced together to create an intact, functional gene (Gilbert 1978a; Periannan Senapathy 1995; P. Senapathy 1988, 1986; James E. Darnell Jr 1978).

The concept of pre-mRNA splicing began to take shape and demystify the long-standing question of how mammalian gene expression can be regulated. Through the proceeding widespread observation of this phenomenon across genes of other multicellular organisms (Krainer et al. 1984; Breathnach et al. 1978; Kedes and Steitz 1988, 1987), Richard J. Roberts and Phillip A. Sharp were both awarded the Nobel Prize in Physiology or Medicine in 1993 for discovering what is now understood as pre-mRNA splicing. However, as groundbreaking as these collective discoveries were, the biochemistry required for pre-mRNA splicing catalysis remained clouded and was just beginning to be unraveled during this same time period.

The spliceosome, a multi-megadalton ribonucleoprotein complex, catalyzes pre-mRNA splicing

The lab of Phillip A. Sharp and colleagues determined that a two-step transesterification phosphoryl transfer reaction enabled the removal of intronic sequences and the ligation of exonic sequences (Figure 1.1A) (Padgett et al. 1984; Ruskin et al. 1984). A relatively conserved adenosine within metazoan introns is critical for catalyzing these phosphoryl transfer reactions (Galej et al. 2016; Wan et al. 2016; Padgett et al. 1984; Zelin, Wang, and Silverman 2006). The catalysis of these phosphoryl transfer reactions are dependent on a multi-megadalton ribonucleoprotein complex known as the spliceosome (Wilkinson, Charenton, and Nagai 2020; Brody and Abelson 1985; Grabowski, Seiler, and Sharp 1985; Frendewey and Keller 1985; Wan et al. 2020). The spliceosome is composed of hundreds of proteins and involves five essential small nuclear ribonucleoproteins (snRNP) known as: U1, U2, U4, and U5 (Figure 1.1B) (Lerner and Steitz 1979; Bringmann and Lührmann 1986). The spliceosome assembles *de novo* on each intron (Konarska and Sharp 1987), and each individual snRNP contains a small nuclear RNA (snRNA) that recognizes key sequence elements within pre-mRNAs to catalyze intron removal and exon ligation (Kastner et al. 2019; Reddy and Busch 1988). Step-wise conformational changes to protein and snRNP associations, in addition to optimal ionic conditions, must occur for the spliceosome to become a functional catalytic core that executes the two phosphoryl transfer reactions that is pre-mRNA splicing (Frendewey and Keller 1985; Steitz and Steitz 1993).

A key initial conformation of the spliceosome during its assembly is known as the E complex (Plaschka et al. 2018; Michaud and Reed 1993). There are key sequence elements referred to as consensus splicing signals within a pre-mRNA that is required for E complex formation in metazoans: the branchpoint motif which contains the conserved adenosine, the polypyrimidine (poly-Y) tract in addition to the 3′ splice site (ss), and the 5′ss (Figure 1.2A) (Sheth et al. 2006; Zhuang and Weiner 1986; M. S. Wong, Kinney, and Krainer 2018; Michaud and Reed 1993; Berglund et al. 1997; Zorio and Blumenthal 1999; Mount et al. 1983). Together, SF1 binding to the branchpoint, heterodimers U2AF65 and U2AF35 binding to the poly-Y tract and 3′ss respectively, and U1 binding to the 5′ss, constitute the E complex. Intrinsically, these consensus splicing signals within pre-mRNA are fundamental for defining an exon from an intron; this is referred to as exon definition (De Conti, Baralle, and Buratti 2013; K. H. Lim et al. 2011; Robberson, Cote, and Berget 1990).

The interplays between auxiliary *cis*-acting sequence elements within a pre-mRNA and *trans*-acting RNA-binding proteins facilitate exon definition

Defining an exon's identity within a pre-mRNA is paramount to precisely splice protein-coding sequences together and remove introns. Yet, relative to lower eukaryotes such as *Saccharomyces cerevisiae*, a majority of these consensus splicing signals are typically degenerate and less conserved in higher eukaryotes like humans (Fu and Ares 2014; Izquierdo and Valcárcel 2006; Spingola et al. 1999; Qin et al.

2016). An additional layer of complexity is the size of exons in mammalian genes. Relative to introns that will exceed thousands of nucleotides in length (X. Hong, Scofield, and Lynch 2006; Shepard, McCreary, and Fedorov 2009), the average size of exons are typically less than 300 nucleotides in length (Lander et al. 2001). This stark contrast in exon size and the lack of sequence conservation raises the question of how an exon is properly defined from an intron by the spliceosome.

It has been established that auxiliary *cis*-acting sequence elements (Dirksen, Sun, and Rottman 1995; Reed and Maniatis 1986; Michaud and Reed 1993; Watakabe, Tanaka, and Shimura 1993; Z. Wang et al. 2004; Fairbrother et al. 2002; Ke et al. 2011), in addition to *trans*-acting splicing factors (Busch and Hertel 2012a; Hollander et al. 2016; Jurica and Moore 2003; R. Martinez-Contreras et al. 2007), act in concert to facilitate exon definition. *Cis*-regulatory elements are typically short sequence motifs residing within or adjacent to exonic sequences to enhance or silence pre-mRNA splicing (Figure 1.2B). These *cis*-regulatory elements are known as exonic and intronic splicing enhancers (ESEs, ISEs) and silencers (ESSs, ISSs) that generally affect the strength of proximal splice sites, influencing how an exon is recognized for splicing (Berget 1995). ESEs and ESSs are thought to play central roles in defining an exon's identity (Blencowe 2000). ESEs are predominantly purine-rich (Fairbrother et al. 2002; Eldridge et al. 1999; Lavigueur et al. 1993; Fu 1995; Manley and Tacke 1996; Graveley 2000; Zheng 2004), compared to ESSs which are usually pyrimidine-rich (Zheng, Huynen, and Baker 1998; Z. Wang et al.

2006; C. D. Chen, Kobayashi, and Helfman 1999; Z. Wang et al. 2004). All of these *cis*-regulatory sequences serve as binding sites for *trans*-acting splicing factors which comprise RNA-binding proteins (RBPs) that interact with spliceosomal components, promoting or inhibiting their recruitment to pre-mRNA.

There are generally two established classes of RBPs that can influence the stepwise assembly of the spliceosome: serine-arginine rich (SR) proteins (J. Y. Wu and Maniatis 1993; Z. Wang, Hoffmann, and Grabowski 1995; Stark et al. 1998; Graveley, Hertel, and Maniatis 2001a), and heterogeneous nuclear ribonucleoprotein particles (hnRNPs) (Mayeda, Helfman, and Krainer 1993a; Del Gatto-Konczak et al. 1999; Caputi et al. 1999a; Bilodeau et al. 2001). SR proteins conventionally bind splicing enhancers (Figure 1.3A), and hnRNPs conventionally bind splicing silencers (Figure 1.3B). Evidence shows that these two classes of RBPs can compete and counteract each other based on their relative abundance when bound to their respective *cis*-acting element on the same pre-mRNA (Zhu, Mayeda, and Krainer 2001; Mayeda and Krainer 1992a; Busch and Hertel 2012a; X. Yang et al. 1994; Pollard et al. 2002; Expert-Bezançon et al. 2004a; Guil et al. 2003; Disset et al. 2006). This antagonism between SR and hnRNP proteins modulates the strength of an exon's identity. Altogether, the collective strength and positions of these *cis*- and *trans*-acting regulators help define an exon's identity, and can promote alternative pre-mRNA splicing of an exon, a process where exons are differentially spliced into mRNA transcripts, diversifying the transcriptome and proteome (Nilsen and Graveley

2010; Gilbert 1978b; Barbosa-Morais et al. 2012; Q. Pan et al. 2008; E. T. Wang et al. 2008; Alt et al. 1980; Early et al. 1980).

Determinants of exon identity remain contextually and positionally dependent on local information

The *cis*- and *trans*- regulatory landscape that determines an exon's identity and its propensity to be spliced into an mRNA transcript are contextually and positionally dependent (Fu and Ares 2014). For example, purine-rich sequence motifs within introns have been shown to act as hnRNP binding sites to enhance splicing (Rebeca Martinez-Contreras et al. 2006). This is in spite of the fact that purine-rich motifs have been historically documented to be found within exons and are representative of splicing enhancers that are primarily recognized by SR proteins (H. X. Liu, Zhang, and Krainer 1998). Similarly, motifs representative of ESEs have been shown to repress splicing when they are found within introns, despite interactions detected between the motif and SR proteins (Kanopka, Mühlemann, and Akusjärvi 1996). These distinct findings demonstrate that *cis*- and *trans*- regulators may not necessarily follow conventional logic and data as expected. That is, perhaps how a splicing enhancer or silencer functions in one exon may not necessarily work the same for another exon due to unique features that underlie a specific exon's identity.

There remains a need to conclude if a deterministic rule or variable exists to ensure pre-mRNA splicing proceeds with high fidelity. It can be hypothesized that the

function of splicing regulatory sequences may be more strongly influenced by their position within a pre-mRNA, and likely by their adjacent *cis*- and *trans*- acting neighbors as well. As such, it is possible that the proximity of splicing regulatory sequences to each other is an important rule underlying optimal pre-mRNA splicing. However, relative to these canonical *cis*- and *trans*-regulatory mechanisms of pre-mRNA splicing, there are other features of a pre-mRNA whose role remains much more enigmatic, such as the impact of RNA structures in pre-mRNA splicing mechanisms. Perhaps the accessibility of a pre-mRNA, namely the availability of splicing regulatory sequences to RBPs and the spliceosome, plays a more universal determinant of pre-mRNA splicing efficiency.

RNA structures as an additional layer of pre-mRNA splicing regulation that remains underappreciated and understudied.

RNA structures are dynamic intramolecular base pairing interactions that occur between nucleotides that belong to the same RNA molecule (Figure 1.4A). Biochemical and thermodynamically favorable conditions will encourage an RNA molecule to fold and adopt the most stable structural conformation (Mathews and Turner 2006; Draper 2004). Phylogenetic comparative sequence analyses of transfer RNAs (tRNAs), telomerase RNA, and ribosomal RNAs have been extensively performed to identify conserved nucleotides that presumably participate in base pairing to form RNA structures (Gardner and Giegerich 2004; Rivas 2021; J. L. Chen, Opperman, and Greider 2002; J.-L. Chen and Greider 2004; Gutell 2014; Noller and

Woese 1981). Experimental validation indicates that it appears common for RNA structures to be evolutionarily conserved to confer key functions, such as in tRNA activity during protein synthesis (Rw et al. 1965; Sharp et al. 1985; Berg and Brandl 2021), and for telomerase function in telomere length maintenance at chromosome ends (J.-L. Chen, Blasco, and Greider 2000; Greider and Blackburn 1985, 1987, 1989; J. L. Chen, Opperman, and Greider 2002). Additionally, chemical probing and computational methods have been developed, refined, and widely adopted in recent years to ascertain the structure of RNA molecules (Figure 1.4B) (Spitale and Incarnato 2022; Marinus et al. 2021; Siegfried et al. 2014; Cordero and Das 2015). In spite of these established and emerging approaches to investigate RNA structure-function in the current era of RNA biology, there is a lack of research investigating the mechanistic role of RNA structures in regulating gene expression.

In contrast to how well understood the biochemistry behind spliceosome assembly and catalysis is, the impact of RNA structures on *cis*- and *trans*-regulatory mechanisms of pre-mRNA splicing has largely been ignored, and has only started to become more scrutinized across genes (Jin, Yang, and Zhang 2011; Xing and Lee 2006; Warf and Berglund 2010). It is now established that RNA structures can form co-transcriptionally during pre-mRNA synthesis (Meyer and Miklós 2004; Tao Pan and Sosnick 2006; T. Saldi et al. 2021), and importantly, RNA structure's influence on pre-mRNA splicing efficiency was first realized in the late 20th century (Eperon et al. 1988; Muro et al. 1999; Goguel and Rosbash 1993). Using several intriguing

findings from recent decades, the proceeding text intends to illuminate and underscore the impactful roles RNA structures may contextually play in regulating pre-mRNA splicing.

Structural conformations adopted at the 5′ss of pre-mRNA can hinder spliceosome assembly. Stem-loop RNA structures can modulate 5′ss accessibility as observed in *Tau* exon 10 (Kar et al. 2011; Grover et al. 1999; Hutton et al. 1998; Jiang et al. 2000; Varani et al. 1999, 2000). The *Tau* gene encodes for a protein that is essential for the maintenance and function of neuronal microtubules (Avila et al. 2002; Iqbal et al. 2009). Studies indicate that the stem-loop structure occludes the 5′ss of *Tau* exon 10, weakening base pairing between the splice site and U1 snRNA to favor the exclusion of exon 10 from pre-mRNA splicing (Jiang et al. 2000; Kalbfuss, Mabon, and Misteli 2001; Varani et al. 2000, 1999). *Kar et al.* used RNA affinity pulldown assays coupled to mass spectrometry to demonstrate that the 5′ss stem-loop structure of *Tau* exon 10 interacts with p68 (Kar et al. 2011), an RNA helicase that plays key roles in spliceosome assembly and pre-mRNA splicing (Cheng et al. 2018). They validate that p68 can destabilize the stem-loop structure to enhance U1 snRNA base pairing to this 5′ss, restoring *Tau* exon 10 splicing. Together, these results indicate that RNA structures can suppress an exon's identity during early spliceosome assembly and prevent E complex formation, and that RNA helicases are likely key splicing regulators that modulate the accessibility of highly structured pre-mRNAs.

Not exclusive to the 5′ss, RNA structures that form and encompass the 3′ss can also inhibit pre-mRNA splicing by compromising spliceosome assembly. Research characterizing *SMN2* exon 7 has revealed key observations regarding contextual mechanisms that control the fidelity of its exon definition. *SMN2*, a gene duplication of *SMN1* resulting from primate evolution (Rochette, Gilbert, and Simard 2001), encodes for survival motor neuron (SMN) proteins, but its pre-mRNA is significantly spliced less efficiently relative to its paralog *SMN1* (Rochette, Gilbert, and Simard 2001; Lefebvre et al. 1995; Lorson et al. 1999; Monani et al. 1999). An innate C to T silent mutation in *SMN2* exon 7 disrupts a critical ESE that results in its inefficient pre-mRNA splicing (Cartegni and Krainer 2002a), and this mutation induces an extended inhibitory context that creates within this locus: multiple *cis*-elements that have been shown to be bound by repressive splicing factors that include hnRNP A1/A2, and an inhibitory terminal stem-loop structure that weakens the 3′ss (referred to as TSL1) (Natalia N. Singh, Androphy, and Singh 2004). The depletion of hnRNP A1/A2 restored exon inclusion (Kashima and Manley 2003a), but strikingly, destabilizing TSL1 alone is sufficient to restore *SMN2* exon 7 inclusion as well (Natalia N. Singh, Androphy, and Singh 2004). A recent report also showed that using a small molecule to target an inhibitory terminal stem-loop structure that occludes the 5′ss (referred to as TSL2) is also sufficient to enhance pre-mRNA splicing of *SMN2* exon 7 (Garcia-Lopez et al. 2018). These findings suggest that, in addition to canonical *cis*- and *trans*- inhibitory pre-mRNA splicing mechanisms, RNA structures by themselves can sufficiently inhibit exon definition.

RNA structures have also been reported to modulate RBP function in pre-mRNA splicing. A recent study on paracaspase *mucosa-associated lymphoid tissue protein 1* (*MALT1*) alternative pre-mRNA splicing regulation elegantly exemplifies this mechanism (Jones et al. 2022). When bound by the E3 ligase TRAF6, MALT1 promotes innate and adaptive immune responses by stimulating and regulating the activation of T cells (Noels et al. 2007; O'Neill et al. 2021; Sun et al. 2004). The inclusion of *MALT1* exon 7 creates an alternative *MALT1* mRNA isoform (*MATL1A*) that has an additional binding site for TRAF6, enhancing the activation of T cells (Meininger et al. 2016).

RBPs are critical regulators of spliceosome assembly and thereby pre-mRNA splicing fidelity. Prior work indicates that the inclusion or exclusion of *MALT1* exon 7 is controlled by the RBPs hnRNPU or hnRNPL (Meininger et al. 2016). *Jones et al.* demonstrated that the mutual exclusive binding of hnRNPL or hnRNPU to *MALT1* exon 7 and flanking intron sequences determines its inclusion or exclusion into mRNA, respectively (Jones et al. 2022). Unexpectedly, however, *Jones et al.* also discovered two distinct RNA structures that each sequester the 3′ and 5′ss, and additionally harbor hnRNPL and hnRNPU consensus binding sites. The authors experimentally show that hnRNPL destabilizes the RNA structures at the 3′ and 5′ss, increasing pre-mRNA splicing of *MALT1* exon 7. In contrast, hnRNPU stabilizes the RNA structures at the 3′ and 5′ss, inhibiting pre-mRNA splicing of *MALT1* exon 7.

Biochemical data also show that the presence of hnRNPL is required to enhance the accessibility of the 3′ss and facilitate U2AF association to the 3′ss. Together, these findings underscore the functional roles RNA structures can have on the accessibility of consensus splicing signals, as well as on the recruitment and influence of RBPs on the binding affinity of early spliceosomal components.

As demonstrated by the case of the *Drosophila melanogaster Down Syndrome Cell Adhesion Molecule* gene (*DSCAM*), RNA structures can also control alternative pre-mRNA splicing to generate an array of mRNA transcripts in eukaryotes (May et al. 2011; Graveley 2005; X. Wang et al. 2012; Dong et al. 2021; Xu et al. 2019; W. Hong et al. 2021; Anastassiou, Liu, and Varadan 2006). The *DSCAM* gene contains four regional clusters of protein-coding sequences that comprise 95 alternative exons; exons 4, 6, 9, and 17 make up each of these clusters. Each alternative exon from each cluster is mutually exclusively spliced to generate a broad array of mRNA transcripts that potentially synthesizes 38,016 unique axon receptors. Among these clusters, the exon 6 cluster has been extensively studied due to it having been initially identified to contain a large portion of conserved intronic sequences in the *DSCAM* gene across a wide range of insect species (Anastassiou, Liu, and Varadan 2006; Graveley 2005). These intronic sequences, referred to as “docking” and “selector” sites within the *DSCAM* exon 6 cluster, were predicted to base pair together to form duplex RNA structures based on a comparative genomics approach. Proceeding work from other

researchers using a similar approach also revealed docking and selector sites in the *DSCAM* exon 4 and 9 clusters (Y. Yang et al. 2011).

The functional relevance of the docking and selector elements were experimentally tested and validated using the *DSCAM* exon 6 cluster (May et al. 2011; X. Wang et al. 2012), and other alternative *DSCAM* exons from the exon 4 and 9 clusters (W. Hong et al. 2020, 2021). Generally, the docking element exists upstream or downstream of the first or last alternative exon of each cluster, and a selector element flanks upstream of each alternative exon. These studies collectively show that the relative strength in base pairing potential between the docking site and a selector site is essential for mutually exclusive pre-mRNA splicing of a distinct alternative exon. Although this "structural code" within pre-mRNA appears unique to the *DSCAM* gene of eukaryotic insects, these data do support the notion that perhaps other eukaryotic gene architecture may have conserved sequence elements that facilitate structural conformations that optimally position splicing regulatory sequences, and therefore exons, for efficient pre-mRNA splicing. Together, these data form the hypothesis that favorable RNA structure conformations can juxtapose a preceding exon with a specific proceeding exon to guide their precise ligation.

Concluding remarks - current gaps in knowledge surrounding pre-mRNA splicing mechanisms

The interplay between auxiliary *cis*-acting elements and *trans*-acting factors serve to aid or antagonize exon definition, enhancing or silencing pre-mRNA splicing fidelity. These mechanisms controlling exon definition can also expand the transcriptome and proteome. However, it remains unclear what rules may broadly govern these pre-mRNA splicing mechanisms across the human genome. It remains poorly understood what characteristics ideally define an exon; are canonical *cis*- and *trans*- regulators stronger determinants of exon identity, or can RNA structural features play a more deterministic effect?

The positional context of *cis*- and *trans*- regulatory elements involved in an exon's identity make it difficult to extrapolate the functional relevance of RNA structures discovered in specific exons to less characterized exons. However, research over the last decades has produced a vignette that allows us to better understand the potential biological significance RNA structures may generally have on pre-mRNA splicing mechanisms. Evidence clearly shows that RNA structures across different gene contexts can affect various biological processes, including pre-mRNA splicing (Figure 1.4C), as showcased by the examples of *Tau* exon 10, *SMN2* exon 7, *MALT1* exon 7, and the *DSCAM* gene. An intriguing question these findings raise is whether exonic and/or intronic RNA structures play a larger antagonistic or supporting role in exon definition by modulating RBP activity. Additionally, the question remains whether the sequences comprising the structure, or the physical structure itself, is more important for their function in pre-mRNA splicing regulation.

Connecting RNA structures' function to pre-mRNA splicing mechanisms in the context of co-transcription is another aspect to consider. The exclusive transcription of nascent pre-mRNAs by RNA polymerase II (Pol II) are shown to be spliced more efficiently (Hicks et al. 2006), and that Pol II specific transcription also appears to regulate pre-mRNA splicing by influencing the recruitment of key splicing factors such as SR proteins (Cramer et al. 1997, 1999; Kornblihtt et al. 2004). It has also been reported that pre-mRNA structure formation correlates with rapid or slow transcription of nascent pre-mRNAs (Spitale and Incarnato 2022; Tao Pan and Sosnick 2006; T. Pan et al. 1999; T. N. Wong, Sosnick, and Pan 2007), and that the folding and adoption of RNA structures can form in sub-milliseconds (Pabit et al. 2013). Evidence indicates that elongation rate coupled to RNA folding dynamics can impact pre-mRNA splicing by modulating the proximity and availability of splice sites and splicing regulatory sequences (Goguel and Rosbash 1993; Kalinina et al. 2021). Together, there is an emerging model where the rate of transcription elongation and the kinetics of pre-mRNA folding can regulate pre-mRNA splicing efficiency (T. Saldi et al. 2021; A. M. Yu et al. 2021; Tassa Saldi, Fong, and Bentley 2018; Fong et al. 2014; T. Saldi et al. 2016). However, a question that arises from these observations is whether this co-transcriptional mechanism mediated by RNA structures is a conserved phenomenon that bypasses the need for exon definition, or contributes to establishing exon definition. That is, it is unclear how widespread this level of

regulation may be across metazoans, and whether or not such a mechanism is context-dependent and required for specific exons within the genome.

Deciphering if there are general rules and patterns behind pre-mRNA splicing regulation by RNA structures remains obscure. In spite of this, the selected examples emphasize an emerging and consistent observation that RNA structures are much more than mere physical properties of a pre-mRNA transcript. As a whole, there indeed remain gaps in knowledge researchers must address to refine our understanding of the mechanisms that regulate or dysregulate pre-mRNA splicing fidelity for any given exon.

Human disease-causing mutations' impact on pre-mRNA splicing

Aberrant pre-mRNA splicing is an established hallmark of numerous human diseases, ranging from rare inherited disorders to cancers (H. K. Kim et al. 2018; Anna and Monika 2018a; R. K. Singh and Cooper 2012a; G.-S. Wang and Cooper 2007a; Garcia-Blanco, Baraniak, and Lasda 2004a; Faustino and Cooper 2003a; Cáceres and Kornblihtt 2002; Lord and Baralle 2021a; E. Wang and Aifantis 2020a; Fredericks et al. 2015a; Sterne-Weiler et al. 2011a; Sterne-Weiler and Sanford 2014a; Kitamura and Nimura 2021). Aberrant pre-mRNA splicing can generate mRNAs that produce truncated non-functional proteins, or proteins with unexpected and antagonistic functions (Fackenthal and Godley 2008). Furthermore, aberrant pre-mRNA splicing can result in the production of faulty mRNAs that lack important

protein-coding sequences or retain non-coding intronic sequences, destabilizing the mRNA for degradation by inducing nonsense-mediated decay (Puisac et al. 2013; S. H. Kim et al. 2009). These effects induced by aberrant pre-mRNA splicing can debilitate cell function and promote disease phenotypes, and often, the underlying cause of this can stem from how an exon's identity is being dysregulated.

Single-nucleotide mutations in splicing regulatory sequences (Figure 1.5A), especially within genes that produce critical *trans*-acting RBPs including spliceosomal components (Figure 1.5B), can cause aberrant pre-mRNA splicing. It is accepted that at least 10% of disease-causing mutations disrupt canonical splice sites, resulting in aberrant pre-mRNA splicing (Krawczak et al. 2007a). Additionally, at least one-third of known disease-causing mutations are predicted to disrupt auxiliary *cis*-acting elements and induce aberrant pre-mRNA splicing (K. H. Lim et al. 2011; Krawczak et al. 2007a; Sterne-Weiler et al. 2011a; Fredericks et al. 2015a). Furthermore, disease-causing mutations that result in pathogenic phenotypes have been shown to cause aberrant pre-mRNA splicing by altering *cis*- or *trans*- regulatory mechanisms (Cartegni et al. 2006a; Cartegni and Krainer 2002a; Kashima and Manley 2003a; N. K. Singh et al. 2006; Yimin Hua et al. 2008; Natalia N. Singh and Singh 2011; Natalia N. Singh, Singh, and Androphy 2007; Zatkova et al. 2004). These mutation-driven effects can therefore have large-scale effects on exon definition and pre-mRNA splicing fidelity, perturbing the strength of splice sites, disrupting an ESE, creating an ESS, or possibly altering the accessibility of splicing

regulatory sequences. The impact of disease-causing mutations will vary based on the contextual regulatory landscape of an exon that is being impacted, but previous work clearly indicates such mutations are sufficient to induce aberrant pre-mRNA splicing and cause disease phenotypes. Thus, more work is required to better connect how dysregulation of exon identity can promote molecular phenotypes that cause disease.

Prior research predicting and investigating the potential impact mutations have on exon identity and pre-mRNA splicing mechanisms

Our previous work determined that a large dataset of disease-causing mutations from the Human Gene Mutation Database (HGMD) dominated in creating ESSs and in disrupting ESEs (Sterne-Weiler et al. 2011a; Sterne-Weiler and Sanford 2014a). It was discovered that a strong ESS motif (ACUAGG) was predicted to be created within 67 genes by 83 different point mutations. To functionally validate the power of this analysis pipeline through *in vivo* splicing assays, using this ESS motif as an example, three of the 67 genes—OPA1, PYGM, and TFR2—were selected based on the respective exons' size, and the strengths of their splice sites. Relative to their wild-type controls, reporters containing this predicted ESS motif resulted in significant aberrant pre-mRNA splicing for all exons tested. RNA affinity chromatography, mass spectrometry, and RNA interference experiments indicate that members from the hnRNP family, SRSF3, and PTBP1 likely form a silencing complex that activates this ESS motif to induce aberrant pre-mRNA splicing. The predictions from our computational approach are consistent with data from other labs

(Fairbrother et al. 2002; K. H. Lim et al. 2011), and our experimental results are consistent with previous reports (Zhu, Mayeda, and Krainer 2001; Y. Yu et al. 2008). Our research substantiates the impact disease-causing mutations have in perturbing exon identity, and how the loss of exon definition serves as a common mechanism for aberrant pre-mRNA splicing in human disease.

Although it is clear disease-causing mutations can impact determinants of exon definition, it remains unclear how these mutations alter spliceosome assembly or its catalytic function. The massively parallel splicing reporter assay (MaPSy) designed by *Somedi et al.* tested 4,964 HGMD disease-causing mutations' effect on splicing efficiency *in vivo* and *in vitro* (Soemedi et al. 2017). Of note, using a density-gradient centrifugation approach, it was revealed that about 10% of disease-causing mutations from this study disrupted multiple stages of spliceosome assembly. Different mutations found in the same exon prevented the spliceosome from transitioning into the proceeding complex, causing it to stall at the preceding stage of spliceosome assembly. The authors hypothesized that this stalling phenomenon is likely due to induced RNA structures, depletion of strong splice sites, or the loss or gain of cis-elements as a result of disease-causing mutations. Though the authors did not perform functional assays to test these hypotheses, these predictions postulate a mechanism for aberrant pre-mRNA splicing where disease-causing mutations are sufficient by themselves to obstruct the assembly of the catalytic spliceosome, leading to aberrant pre-mRNA splicing. *Somedi et al.* also

suggest that distinct mutations can also have pleiotropic effects, dysregulating functional RNA elements or spliceosome assembly directly. Thus, it can also be hypothesized that aberrant pre-mRNA splicing of some exons may primarily be due to problems with the physical assembly of the spliceosome, as opposed to just the loss of *cis*-elements that are involved in the fidelity of exon definition. What remains unclear is if this is indeed the case, or if both of these mechanisms are simultaneously being dysregulated to cause aberrant pre-mRNA splicing.

Disease-causing mutations may also alter RNA structures native to the wild-type ensemble, possibly promoting aberrant pre-mRNA splicing (Halvorsen et al. 2010; Salari et al. 2013). It is compelling to hypothesize that perhaps mutations can induce aberrant pre-mRNA splicing by altering the accessibility of splicing regulatory sequences, especially splice sites for initiating E complex formation. It is also intriguing to consider that perhaps a mutation may enhance the accessibility of a splicing silencer that was previously sequestered, or that a splicing enhancer is now sequestered due to inaccessibility resulting from new structural constraints. However, more experimentation is needed to explore these computational predictions and build models that validate or reject hypotheses on pre-mRNA structure and function.

The identity of most exons likely have unique central determinants of pre-mRNA splicing that it cannot survive without. That is, mutations that cripple key regulatory features that define a specific exon's identity will increase the likelihood of

its aberrant pre-mRNA splicing. Thus, there are probably context-dependent mechanisms that are required for exon definition fidelity, so as to prevent aberrant pre-mRNA splicing. Uncovering what these molecular mechanisms are for each exon and how they are affected by mutations is tantamount to understanding the etiology of a disease, as well as identifying potential therapeutic targets for precision medicine.

Antisense oligonucleotides (ASOs) as precision RNA drugs to modulate gene expression

ASOs are multifaceted nucleic acid oligomers that exploit the simple chemical logic of base pairing to regulate gene expression (Figure 1.6). Beyond their use in: gene silencing through RNA interference (RNAi) (Hu, Liu, and Corey 2010; J. Liu, Hu, and Corey 2012; Corey 2007; Fire et al. 1998), as ligands for biomolecules (Ellington and Szostak 1990; Tuerk and Gold 1990; Berezovski et al. 2008; Sullenger et al. 1990), and pre-mRNA splicing modulation (Takeshima et al. 1995; van Deutekom et al. 2001; Dunckley et al. 1998; Natalia N. Singh, Luo, and Singh 2018; Kole, Shukla, and Akhtar 1991), ASOs are seeing more novel applications in the clinic (Egli and Manoharan 2023; Crooke et al. 2018; Jadhav et al. 2024; Ni et al. 2011). Within the last decades, splice-modulating ASOs in particular have become an established therapeutic modality to treat human diseases caused by aberrant pre-mRNA splicing (Rigo, Seth, and Bennett 2014; M. A. Havens, Duelli, and Hastings 2013a; Y. Hua et al. 2010a; J. Kim et al. 2019; Dominski and Kole 1993; Disterer et al. 2014; Echigoya and Yokota 2014; Mendell et al. 2013a; Corey 2017a).

ASOs can be designed with high specificity to their DNA and RNA targets; the phosphate backbone, in addition to the utility of the 2′-OH of a ribose sugar, can support chemical modifications that infer significant stability and nuclease resistance *in vivo* (Mallory A. Havens and Hastings 2016; Egli and Manoharan 2023; Summerton 1999; Rigo, Seth, and Bennett 2014). Although there are challenges with the toxicity and delivery of splice-modulating ASOs (Sharma and Watts 2015; Geary 2009; Gagliardi and Ashizawa 2021), which continue to be addressed and advanced using model systems (Byrnes et al. 2024; Juliano 2016), their impact as precision medicine to treat human diseases is becoming clear.

Splice-modulating ASOs achieve their potential therapeutic effect by targeting aberrant pre-mRNA splicing mechanisms. The first proof-of-concept experiment underscoring ASOs' capacity as splice modulators was demonstrated using the human beta-globin pre-mRNA, where a bound ASO sterically blocked the use of cryptic splice sites to redirect correct splice site usage in a thalassemia context (Dominski and Kole 1993). The application of ASOs in aberrant pre-mRNA splicing is to promote the production of functional or desired mRNA transcripts. This, in turn, should attenuate respective protein levels in diseased cells to ameliorate mutation-induced phenotypes. However, although *Dominski and Kole* did not study the effects of their splice-modulating ASOs on rescuing protein levels, proceeding work from other groups studying other genes cemented ASOs as viable RNA drugs that can enhance pre-mRNA splicing and rescue protein function. An exemplary case

pertains to the research and eventual conceptualization of an FDA-approved splice-modulating ASO that effectively cures spinal muscular atrophy (SMA).

SMA - a quintessential model exemplifying pre-mRNA splicing mechanisms as a therapeutic target for splice-modulating RNA drugs

SMA is a leading cause of infant mortality (Prior et al. 2010; Sugarman et al. 2012; Larson et al. 2015). It is a well-characterized example of a rare disease caused by aberrant pre-mRNA splicing. The gene paralogs *SMN1* and *SMN2* are implicated in causing SMA (Kolb and Kissel 2015). *SMN1* encodes for survival motor neuron (SMN) proteins that play critical roles in neuronal pathways (Zhang et al. 2006), and gene regulatory processes including pre-mRNA splicing (R. N. Singh et al. 2017; Pellizzoni, Yong, and Dreyfuss 2002; Kolb, Battle, and Dreyfuss 2007). As mentioned in **Chapter 1: Introduction to pre-mRNA splicing**, *SMN2*, a paralog of *SMN1*, encodes for SMN proteins but is unable to produce enough full-length functional SMN proteins as *SMN1* can (Figure 1.7A). This is due to a single nucleotide difference, where a C to T silent mutation within *SMN2* exon 7 causes it to be spliced into *SMN2* mRNA ~20% of the time, relative to *SMN1* exon 7 which is spliced efficiently >95% of the time into *SMN1* mRNA. Thus, the common cause of SMA is the result of mutations that render both copies of the *SMN1* gene to be non-functional (Lefebvre et al. 1995; Sumner 2007; Sugarman et al. 2012).

It is key to understand that the C to T mutation in *SMN2* exon 7 disrupts an experimentally validated SRSF1-dependent ESE that is otherwise intact in *SMN1* exon 7 (Cartegni and Krainer 2002a), and creates in *SMN2* exon 7 binding sites for the splicing silencing factor hnRNPA1 (Natalia N. Singh, Androphy, and Singh 2004; Kashima and Manley 2003a). Additionally, *SMN2* exon 7 adopts an unfavorable structural conformation that contains a splicing inhibitory RNA structure that weakens the 3′ss (Natalia N. Singh, Androphy, and Singh 2004), as well as the 5′ss (Garcia-Lopez et al. 2018). Collectively, these mechanisms debilitate the identity of *SMN2* exon 7 and induce aberrant pre-mRNA splicing relative to its paralog. Thus, in the absence of functional *SMN1* alleles, this C to T change in *SMN2* exon 7 is therefore widely accepted as an etiology of SMA (Rochette, Gilbert, and Simard 2001; Sumner 2007; Lefebvre et al. 1995; Sugarman et al. 2012). Although *SMN2* does not efficiently produce enough SMN proteins to prevent SMA, SMA patients still typically have at least one functioning copy of *SMN2* present, which data shows can lessen SMA phenotypes and symptoms (Wirth et al. 2006; Prior et al. 2009). As such, developing therapeutic strategies to enhance *SMN2* exon 7 pre-mRNA splicing has become a focal point to counteract SMA phenotypes induced by mutation-induced aberrant pre-mRNA splicing.

Spanning over a decade, research from the lab of Ravindra N. Singh and Adrian R. Krainer conceptualized the development of a splice-modulating ASO that rescues *SMN2* exon 7 pre-mRNA splicing. Made possible by an important research

breakthrough from the Singh lab, they discovered that targeting an ISS termed ISS-N1 (intronic splicing silencer N1) blocked the binding of hnRNPA1 silencing factor proximal to the 5′ss of *SMN2* exon 7 (N. K. Singh et al. 2006; N. N. Singh, Androphy, and Singh 2004). Following studies revealed that targeting ISS-N1 also perturbed associated regulatory mechanisms which ultimately enhance *SMN2* exon 7 identity (Natalia N. Singh et al. 2017; Natalia N. Singh and Singh 2011; N. N. Singh et al. 2011, 2013; Natalia N. Singh, Singh, and Androphy 2007). Contemporarily to the discovery of ISS-N1 by Singh and colleagues, an initial ASO screen was performed on *SMN2* exon 7 using a cell-based splicing reporter system by the lab of Adrian Krainer in collaboration with Ionis Pharmaceuticals (formerly Isis Pharmaceuticals), which initially did not identify a lead ASO (Yimin Hua et al. 2007). Performing additional ASO screens that this time targets intronic sequences flanking *SMN2* exon 7 in preclinical mice models, the Krainer Lab and Ionis Pharmaceuticals subsequently identified various ASOs that can rescue *SMN2* exon 7 inclusion (Yimin Hua et al. 2008), revealing an intriguing avenue of further investigation.

As a major breakthrough finding from this initial preclinical study by the Krainer lab (Yimin Hua et al. 2008), the lead ASO compound labeled as ASO 10-27 showed the highest efficacy in rescuing *SMN2* exon 7 pre-mRNA splicing by targeting ISS-N1 (Figure 1.7B). Further work by the Krainer lab and Ionis Pharmaceuticals demonstrated that ASO 10-27 is capable of restoring the production of functional SMN proteins and ameliorating SMA phenotypes in preclinical mice

studies (Y. Hua et al. 2010a; Passini et al. 2011; Y. Hua et al. 2011). ASO 10-27, now identified as Nusinersen, reported promising results in clinical phase trials (Chiriboga et al. 2016; Haché et al. 2016; Garber 2016; Finkel et al. 2016), and Nusinersen received FDA approval soon after the successful completion and termination of all clinical trials. Nusinersen (commercially distributed as Spinraza) is the first-ever FDA-approved drug to treat and effectively cure SMA by targeting a pre-mRNA splicing mechanism (Ottesen 2017). Collectively, this success story of SMA is a model case that underscores the need to better understand the underlying molecular mechanisms behind splicing-linked diseases to develop precision therapeutics.

Concluding remarks - understanding pathogenic mutations' impact on pre-mRNA splicing mechanisms can enable drug discovery for human disease

Mutation-induced aberrant pre-mRNA splicing is an established etiology of rare disease and cancers, but determining if there are mechanistic rules that promote or prevent the predisposition of an exon to aberrant pre-mRNA splicing remains less clear. ASOs can be a viable approach to ameliorate molecular disease phenotypes, but despite their power, there are currently only ten FDA-approved ASO therapies (Aartsma-Rus and Corey 2020). More pressingly, three of the ten are FDA-approved ASOs that work to correct aberrant pre-mRNA splicing. Understanding how mutations may sensitize an exon to aberrant pre-mRNA splicing is a key step forward in refining our knowledge of rules that may regulate exon identity, and in developing precision medicine that can help patients with unmet needs. As such, in the context of

pre-mRNA splicing in rare disease, more research is needed to identify: (1) potential therapeutic targets, (2) the mechanism(s) affected by mutations, and (3) the mechanism of action by which splice-modulating ASO rescue pre-mRNA splicing.

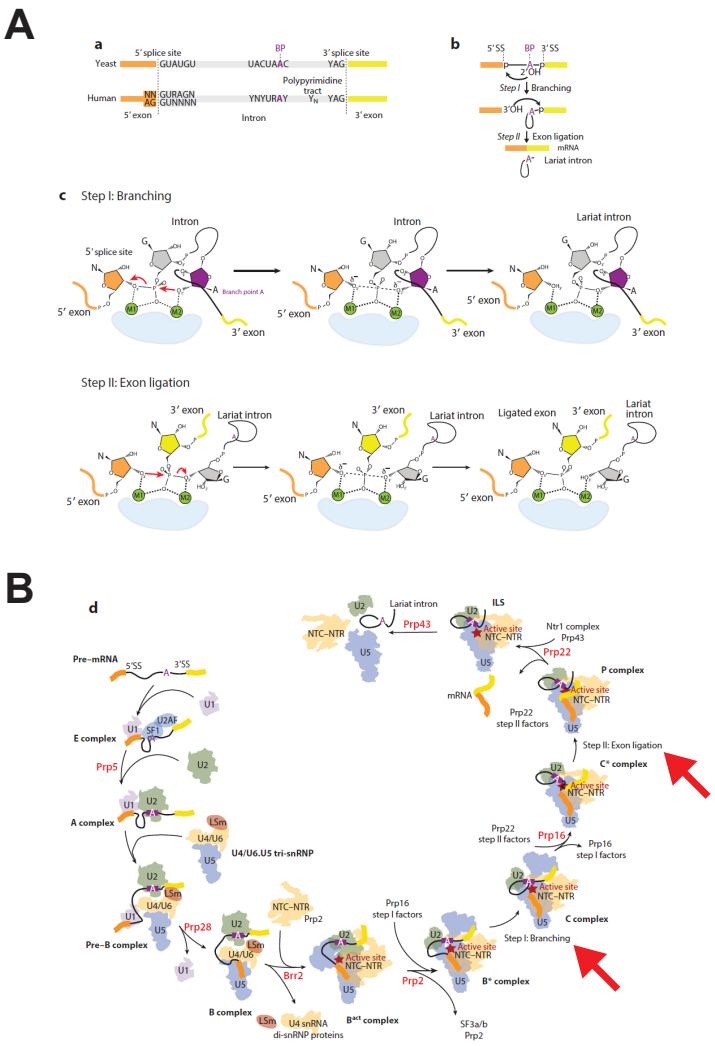


Figure 1.1 - Biochemistry of pre-mRNA splicing
(A) A schematic depicting consensus splicing signals required for catalyzing the two transesterification
phophoryl transfer reactions (branching and exon ligation) required to join exons together, and remove introns. Consensus splicing signals are shown for both the Yeast and Human genome to emphasize degeneracy in sequence conservation.
(B) A simplified depiction of the spliceosome assembly pathway. The transesterification phosphoryl transfer reactions that enable branching and exon ligation are annotated (red arrows) to emphasize the step-wise assembly and conformational changes required to catalyze pre-mRNA splicing.
Both panels are adapted from Wilkinson, M. E., Charenton, C., and Nagai., K., 2020, Annual Review of Biochemistry.

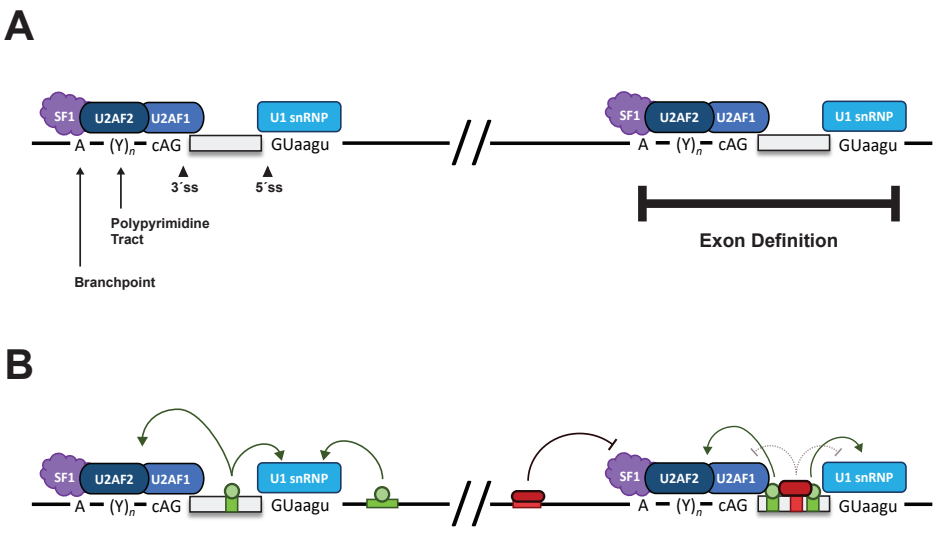


Figure 1.2 - *Cis* and *trans* regulators facilitate exon definition
(A) Schematic showing the key consensus splicing signals (on left), which collectively form the exon definition complex when bound by cognate spliceosomal components (on right).
(B) Schematic showing how *cis* and *trans* regulators can promote or inhibit the binding of spliceosomal components to respective consensus splicing signals. Splicing enhancers which promote splicing are depicted in green, and splicing silencers which inhibit splicing are depicted in red.

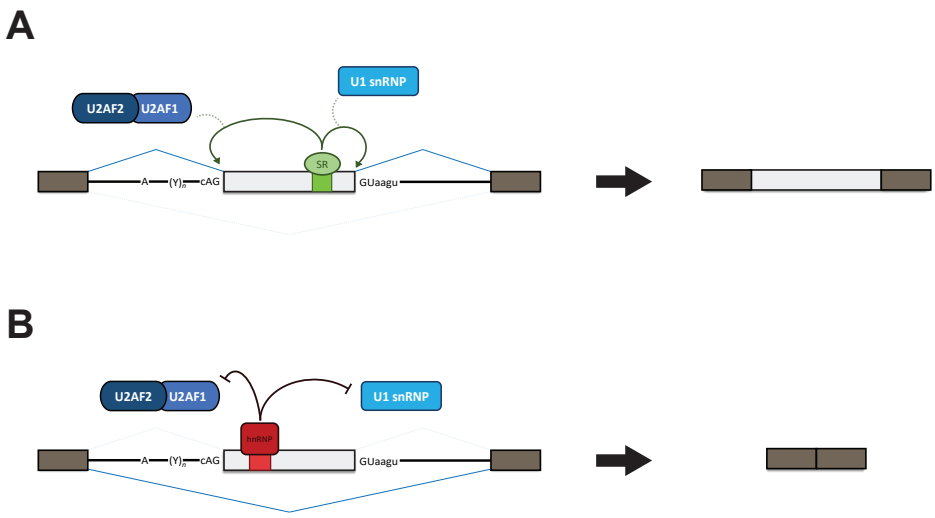


Figure 1.3 - Function of splicing enhancers and silencers
(A) A schematic showing how a splicing enhancer typically functions. An exonic splicing enhancer (ESE) bound by a cognate RNA-binding protein (RBP) partner on an exon is shown. Once bound to the ESE, specific domains on the RBP interact with spliceosomal components and bring them closer in proximity to their cognate binding sites. This strengthes exon definition, promoting the inclusion of this exon into a spliced mRNA.
(B) A schematic showing how a splicing silencer typically functions. A exonic splicing silencer (ESS) bound by a cognate RNA-binding protein (RBP) partner on an exon is shown. Once bound to the ESS, the RBP can sterically block and repel spliceosomal components from binding to their cognate binding sites. This weakens exon definition, inhibiting the inclusion of this exon into a spliced mRNA.

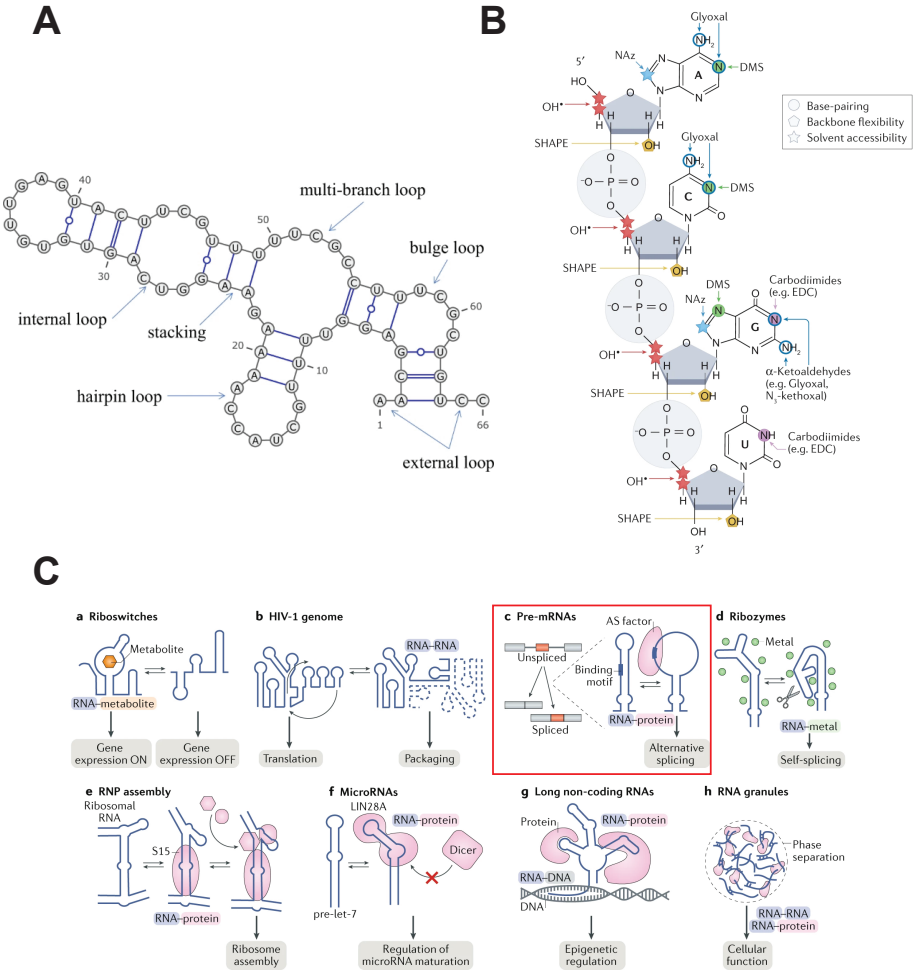


Figure 1.4 - RNA secondary structures and their potential functions in regulating pre-mRNA splicing
(A) A schematic depicting the types of RNA secondary structures that can form under optimal base pairing conditions. The names for each type of structure that can form are indicated. Adapted from Sato, K., Akiyama, M., and Sakakibara, Y., 2021, Nature Communications.
(B) A schematic depicting the various targets of different chemical probes. The mechanism of action for a given chemical probe is indicated by their shape (as shown in the box legend). The types of chemical probes based on their chemical properities, as well as their site of modification, are indicated by annotated names and colors. Adapted from Spitale, R. C., and Incarnato, D., 2022, Nature Reviews Genetics.
(C) A schematic showing the different types of functions RNA secondary structures can have in regulating gene expression, including pre-mRNA splicing (highlighted in red box). Adapted from Ganser, L. R., Kelly, M. L., Hershlag, D., and Al-Hashimi, H. M., 2019, Nature Reviews Molecular Cell Biology.

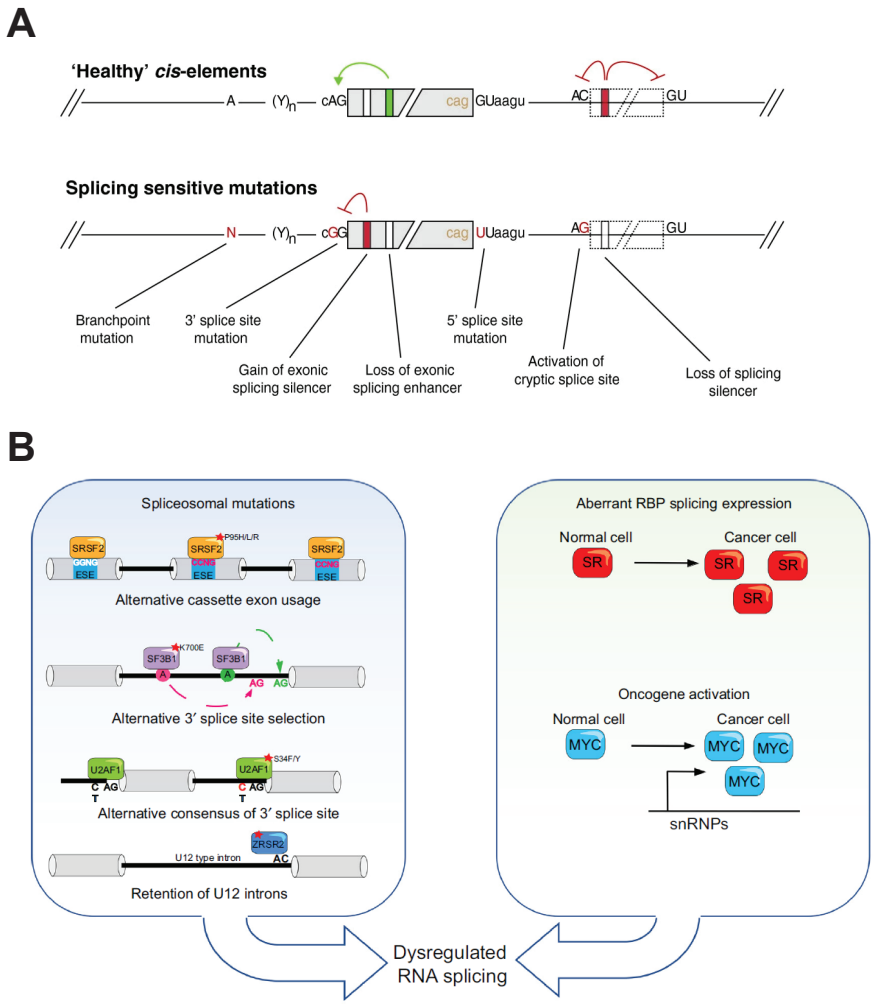


Figure 1.5 - Mutations can dysregulate pre-mRNA splicing mechanisms to cause disease phenotypes
(A) A schematic depicting the impact splicing-sensitive mutations can have on splicing regulatory sequences relative to the the "healthy" wild-type text. Adapted from Sterne-Weiler, T., and Sanford, J. R., 2014, Genome Biology. **(B)** (On left) A schematic showing how mutations in various splicing regulatory RNA-binding proteins (RBPs) can dysregulate splicing mechanisms. (On right) A schematic showing how mutations in splicing regulatory RBPs can also drive its aberrant expression, potentially promoting oncogenesis, and further dysregulating splicing mechanisms. Adapted from Wang, E., and Aifantis, I., 2020, Trends in Cancer.

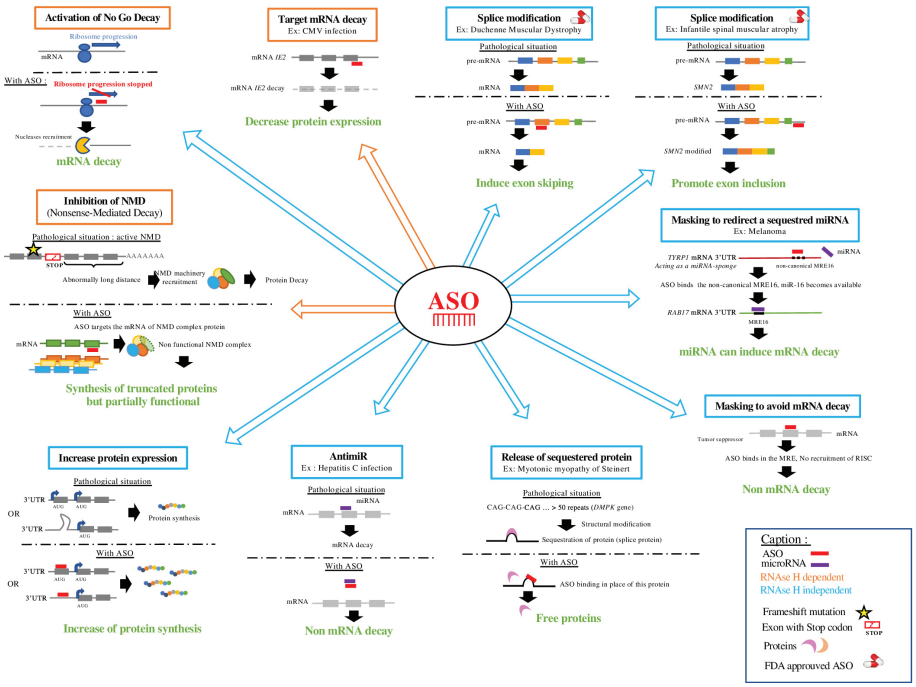


Figure 1.6 - Antisense oligonucleotides as multifaceted tools and precision
A schematic showing the plethora of regulatory processes that antisense oligonucleotides (ASOs) can modulate as a result of precise base pairing to its nucleic acid sequence. A legend is shown at the bottom right of the schematic to indicate the type of specific mechanisms the ASOs may promote or inhibit. The legend also indicates which ASOs have been FDA-approved as precision RNA drugs. Adapted from Quemener, A. M., *et al.*, 2020, Wiley Interdisciplinary Reviews RNA.

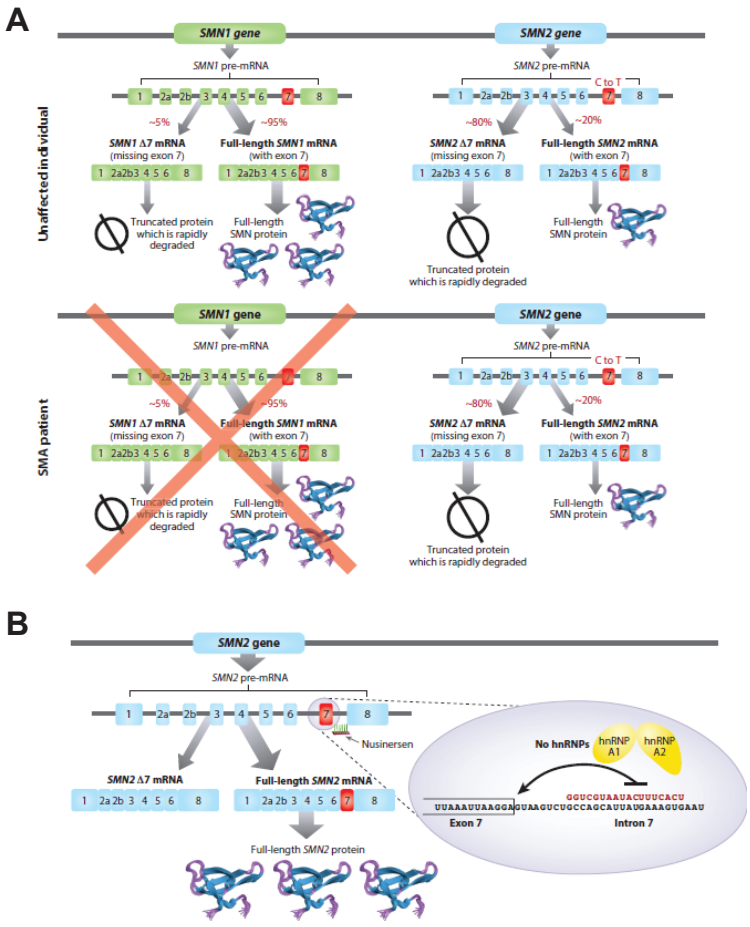


Figure 1.7 - Spinal muscular atrophy (SMA) is a quintessential model underscoring aberrant pre-mRNA splicing as an etiology of disease and ASOs as precision medicine
(A) A schematic showing how exons from paralogs *SMN1* and *SMN2* under pre-mRNA splicing, and how they differ in mRNA isoform production and subsequent SMN protein function. Note the difference in exon 7 pre-mRNA splicing between the two paralogs, as indicated by the red color. Note how the total absence of functional *SMN1* alleles drives phenotypes linked to the etiology of SMA (i.e., insufficient production of functional SMN proteins by *SMN2*).
(B) A schematic showing how FDA-approved ASO, nusinersen (marketed Spinraza) enhances pre-mRNA splicing of *SMN2* exon 7, rescuing the production of functional SMN proteins.
Both panels Adapted from Bennett, C. F., Krainer, A. R., and Cleveland, D. W., 2019, Annual Review of Neuroscience.

Chapter 2: An intronic RNA element sensitizes Factor VIII exon-16 to aberrant splicing

Publication Status: Published in *Nucleic Acids Research*. Published Online on November 14, 2023; Published in Print on January 11, 2024

Authors: **Victor Tse**[1, 2, †], Guillermo Chacaltana[3, 2, †], Martin Gutierrez[1, 2], Nicholas M Forino[1, 2], Arcelia G Jimenez[3], Hanzhang Tao[1], Phong H Do[1], Catherine Oh[1], Priyanka Chary[1], Isabel Quesada[1], Antonia Hamrick[1], Sophie Lee[1], Michael D Stone[3, 2, *], Jeremy R Sanford[1, 2, *]

1. Department of Molecular, Cell and Developmental Biology, University of California Santa Cruz, Santa Cruz, CA, 95064, USA.
2. Center for Molecular Biology of RNA, University of California Santa Cruz, Santa Cruz, CA, 95064, USA.
3. Department of Chemistry and Biochemistry, University of California Santa Cruz, Santa Cruz, CA, 95064, USA.

† The authors wish it to be known that, in their opinion, these authors should be regarded as Joint First Authors.

* To whom correspondence should be addressed. Tel: +1 831 459 1822; Email:

jsanfor2@ucsc.edu

Correspondence may also be addressed to Michael D. Stone. Tel: +1 831 459 2845; Email: mds@ucsc.edu

Abstract

Pathogenic variants in the human Factor VIII (*F8*) gene cause Hemophilia A (HA). Here, we investigated the impact of 97 HA-causing single-nucleotide variants on the splicing of 11 exons from *F8*. For the majority of *F8* exons, splicing was insensitive to the presence of HA-causing variants. However, splicing of several exons, including exon-16, was impacted by variants predicted to alter exonic splicing regulatory sequences. Using exon-16 as a model, we investigated the structure–function relationship of HA-causing variants on splicing. Intriguingly, RNA chemical probing analyses revealed a three-way junction structure at the 3′-end of intron-15 (TWJ-3–15) capable of sequestering the polypyrimidine tract. We discovered antisense oligonucleotides (ASOs) targeting TWJ-3–15 partially rescue splicing-deficient exon-16 variants by increasing accessibility of the polypyrimidine tract. The apical stem loop region of TWJ-3–15 also contains two hnRNPA1-dependent intronic splicing silencers (ISSs). ASOs blocking these ISSs also partially rescued splicing. When used in combination, ASOs targeting both the ISSs and the region sequestering the polypyrimidine tract, fully rescue pre-mRNA splicing of multiple HA-linked variants of exon-16. Together, our data reveal a putative RNA structure that sensitizes *F8* exon-16 to aberrant splicing.

My Contributions:

I led a team to generate the heterologous *F8* splicing reporters used in this study. I led the *in vitro* cell-based splicing assays used in this study. I developed, established, and led the application of the two-step RT-qPCR assay used in this study. I developed and established the automation platform used in this study. I designed all antisense oligonucleotides targeting F8 exon-16, and led efforts to assay them using the automation platform I developed. I helped design the *in vitro* RNA templates for SHAPE-MaP-seq. I led the hnRNPA1-ASO competition assay to experimentally validate hnRNPA1 binding sites identified in this study. I led the design, generation, and testing of the *F8* exon-16 minigene reporters used in this study. I performed and produced a majority of the data analysis and visualization presented in this study. I wrote the manuscript and produced all main and supplemental figures and legends presented, with input and suggestions by co-authors.

Additional Information

All supplemental figures and tables for this publication can be found online.

Nucleic Acids Research, 2024, **52**, 300–315
https://doi.org/10.1093/nar/gkad1034
Advance access publication date: 14 November 2023
Molecular Biology

An intronic RNA element modulates Factor VIII exon-16 splicing

Victor Tse[1,2,†], Guillermo Chacaltana[3,2,†], Martin Gutierrez[1,2], Nicholas M. Forino[1,2], Arcelia G. Jimenez[3], Hanzhang Tao[1], Phong H. Do[1], Catherine Oh[1], Priyanka Chary[1], Isabel Quesada[1], Antonia Hamrick[1], Sophie Lee[1], Michael D. Stone[3,2,*] and Jeremy R. Sanford[1,2,*]

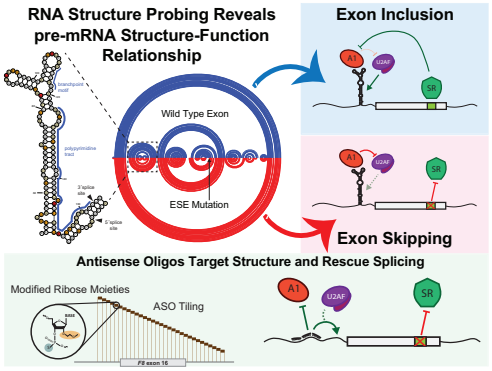
[1]Department of Molecular, Cell and Developmental Biology, University of California Santa Cruz, Santa Cruz, CA, 95064, USA
[2]Center for Molecular Biology of RNA, University of California Santa Cruz, Santa Cruz, CA, 95064, USA
[3]Department of Chemistry and Biochemistry, University of California Santa Cruz, Santa Cruz, CA, 95064, USA

[*]To whom correspondence should be addressed. Tel: +1 831 459 1822; Email: jsanfor2@ucsc.edu
Correspondence may also be addressed to Michael D. Stone. Tel: +1 831 459 2845; Email: mds@ucsc.edu
[†]The authors wish it to be known that, in their opinion, these authors should be regarded as Joint First Authors.

Abstract

Pathogenic variants in the human Factor VIII (*F8*) gene cause Hemophilia A (HA). Here, we investigated the impact of 97 HA-causing single-nucleotide variants on the splicing of 11 exons from *F8*. For the majority of *F8* exons, splicing was insensitive to the presence of HA-causing variants. However, splicing of several exons, including exon-16, was impacted by variants predicted to alter exonic splicing regulatory sequences. Using exon-16 as a model, we investigated the structure–function relationship of HA-causing variants on splicing. Intriguingly, RNA chemical probing analyses revealed a three-way junction structure at the 3′-end of intron-15 (TWJ-3–15) capable of sequestering the polypyrimidine tract. We discovered antisense oligonucleotides (ASOs) targeting TWJ-3–15 partially rescue splicing-deficient exon-16 variants by increasing accessibility of the polypyrimidine tract. The apical stem loop region of TWJ-3–15 also contains two hnRNPA1-dependent intronic splicing silencers (ISSs). ASOs blocking these ISSs also partially rescued splicing. When used in combination, ASOs targeting both the ISSs and the region sequestering the polypyrimidine tract, fully rescue pre-mRNA splicing of multiple HA-linked variants of exon-16. Together, our data reveal a putative RNA structure that sensitizes *F8* exon-16 to aberrant splicing.

Graphical abstract



Introduction

Noncoding sequences (introns) interrupt protein-coding information (exons) in most human genes. Conserved sequences known as splice sites (ss) demarcate exon-intron boundaries (1). Messenger RNA (mRNA) biogenesis requires intron removal from precursor mRNA (pre-mRNA) and exon ligation; this step of gene expression is known as pre-mRNA splicing (2,3). The spliceosome, a multi-megadalton ribonucleoprotein (RNP) complex, assembles *de novo* on every intron to catalyze splicing reactions. This process involves the stepwise assembly

Received: March 31, 2023. **Revised:** October 16, 2023. **Editorial Decision:** October 18, 2023. **Accepted:** October 23, 2023

of five uracil-rich small nuclear ribonucleoprotein particles (U snRNPs) and hundreds of accessory RNA-binding proteins on a pre-mRNA (4,5). To facilitate spliceosome assembly, exon definition is a critical initial step where exon-intron boundaries are defined (6). In this early spliceosome complex, U1 snRNP recognizes the 5′ss while the U2 snRNP auxiliary factor (U2AF) binds the 3′ss and polypyrimidine (poly-Y) tract (7–10). This initial step is highly regulated in cells by the presence or absence of sequences that can function as exonic splicing enhancers or silencers (11–16). Additionally, structured features within pre-mRNA can regulate splicing (17). For example, structured RNA elements influence splice site accessibility (18,19) by modulating protein–RNA interactions (20–22).

Aberrant splicing contributes to the etiology of many inherited diseases (23). Pathogenic variants impact pre-mRNA splicing through a variety of mechanisms. Most notably, variants remodel the *cis*-regulatory landscape of pre-mRNAs by ablation or creation of splice sites, and auxiliary splicing regulatory sequences such as exonic or intronic splicing enhancers (ESE and ISE, respectively) and splicing silencers (ESS and ISS, respectively). Splicing-sensitive variants cripple the integrity of the gene, resulting in the production of a faulty message that is either unstable or encodes an internally deleted protein (24–26). Antisense oligonucleotides (ASOs) are a promising therapeutic modality for rescuing pathogenic aberrant splicing patterns as their direct base pairing abilities make them highly customizable and specific to targets. Although challenges such as toxicity, delivery and stability represent barriers to the clinical translation of ASOs (27), solutions to these challenges exist, as exemplified by the recent FDA approval of multiple ASO drugs (28–31).

Our previous work implicated thousands of disease-causing variants in aberrant splicing of both constitutive and alternative exons (23,24). This work exploited the sequence bias of ESEs and ESSs to discover pathogenic genetic variants that cause the loss or gain of these putative functional elements relative to common, high allele frequency polymorphisms. Among these, the *F8* gene had the highest frequency of variants predicted to affect splicing regulatory sequences (23). The *F8* gene encodes for a procoagulant called Factor VIII (FVIII), which is required for initiating the coagulation cascade to form blood clots in response to wounds (32). Through this cascade, activated FVIII must bind to another procoagulant called Factor IX in its activated form. This initial interaction with FVIII sets off a series of biochemical reactions with additional procoagulants that forms a fibrin scaffold, serving as the basis for blood clotting. Thus, genetic variants can inactivate *F8* and cause Hemophilia A (HA), a deadly, X-linked recessive bleeding disorder.

The *F8* gene encodes 2351 amino acids split across 26 exons. Although the majority of these exons are constitutively spliced into a canonical 9032 nucleotide transcript, GENCODE (Version 43) contains several isoforms arising from alternative promoter usage and alternative 3′ss selection in exon-16 and intron-22. Despite modest support for alternative splicing, several constitutively spliced *F8* exons are aberrantly spliced in HA patients, (33–36). These data suggest the hypothesis that a broad array of pathogenic missense variants in the *F8* gene may manifest as splicing defects due to loss of exon identity or activation of cryptic splice sites. To test this hypothesis, we investigated 97 HA-causing variants on *F8* pre-mRNA splicing using a heterologous reporter assay. We discovered that among the eleven exons studied, splicing of exon-16 was the most sensitive to HA-causing variants. RNA structure probing revealed a predicted intronic three-way junction that sequesters the exon-16 3′ss. ASOs targeting this structured element rescue diverse splicing-sensitive HA-causing variants of *F8* exon-16, in both a heterologous and endogenous splicing context. Together, these data suggest an unexpected role for RNA structure in modulating exon identity and provide a rationale for the development of novel RNA therapeutics to potentially treat a subset of HA patients.

Materials and methods

F8 splicing reporters

The sequences of wild-type (WT) *F8* exons and 100–250 nucleotides flanking each exon were amplified from human genomic DNA (Promega) using WT PCR primers shown in Supplementary Table S1. Following gel purification, PCR products were ligated into pACT7_SC14 (*HBB* minigene reporter as previously described (37)) using homology-based cloning technology (In-Fusion HD Cloning kit, Takara Bio). Following sequence verification, each plasmid was then used as a template for site-directed mutagenesis via overlap-extension PCR using mutagenesis primers shown in Supplementary Table S1. Splicing reporter plasmids were then sequence-validated using Sanger sequencing to confirm successful cloning and identity of splicing reporters. *F8* minigene splicing reporters were derived using a similar approach, where the test exon in addition to flanking introns and exons, are cloned in between a strong promoter and polyadenylation site. The naming designation for each *F8* variant investigated in this study is based on the Human Genome Variation Society (HGVS) nomenclature. Therefore, each pathogenic variant of an *F8* exon presented in this study is based on the nucleotide being mutated (e.g. A > C), and its position within the coding DNA sequence context tested relative to the first ATG start codon (38). Each pathogenic variant's designation based on the HGVS nomenclature is shown in Supplementary Table S1.

Cell-based *in vitro* splicing assays

HEK293T cells (ATCC) were cultured in 6-well tissue culture plates (CytoOne, USA Scientific) using Dulbecco's Modified Eagle Medium (Gibco, supplemented with 10% FBS) at 37°C, 5% CO_2. The cells were transiently transfected at ∼60–80% confluency with 2.5 μg of each *F8* splicing reporter using Lipofectamine 2000 (Invitrogen). Total RNA was harvested from cells 24 h post-transfection using the Direct-zol RNA Miniprep kits (Zymo Research). Each splicing assay was performed with three independent/biological replicates. To ensure rigor and reproducibility for splicing-sensitive variants identified, splicing assays were performed with a total of nine independent/biological replicates, with the exception of exon-16$^{c.5562G>T}$ which contains 18 independent/biological replicates to assess variability.

ASO walk and combinatorial ASO experiments

2′-methoxyethyl (2′MOE) phosphorothioate substituted ASOs complementary to *F8* exon-16 and flanking introns were designed from the reverse complement of the *F8* sense sequence, creating non-overlapping 18-mers as shown in Supplementary Table S2. *F8* exon-16 ASOs were designed to contiguously tile across the exon and its flanking introns.

ASOs were synthesized by Integrated DNA Technologies (IDT). HEK293T cells (ATCC) were cultured in 96-well tissue culture plates (Perkin Elmer) as described above. Cells were transiently transfected with 250 ηg of WT or pathogenic variant splicing reporter and 10 ℘mol of each ASO (final concentration: ∼0.43 ηg/μl) using Lipofectamine 2000 (Invitrogen). 24 h post-transfection, cells were harvested and prepared for total RNA purification using the Quick-DNA/RNA Viral MagBead kit from Zymo Research and an Agilent Bravo NGS A liquid handler (39). Each experiment type (e.g. ASO walk or combinatorial ASO assays) was performed with three independent/biological replicates.

hnRNPA1 overexpression and western blot analysis

HEK293T cells (ATCC) were cultured in 6-well tissue culture plates as described above. Cells were co-transfected with 1.25 μg of the WT splicing reporter, 1.25 μg of either an empty expression vector or a T7-tagged hnRNPA1 expression vector, and 10 ℘mol of ASO(s) as described above. Total RNA and protein were isolated 24 h post-transfection using a RSB lysis buffer (10 mM Tris pH 7.0, 100 mM NaCl, 5 mM MgCl2, 0.5% NP40, 0.5% Triton X-100, and EDTA-free Protease Inhibitor Cocktail [Roche]). Following a 10 min incubation on ice, the cell lysate was then centrifuged at 10 000 × *g* for 10 min at 4°C. The supernatant was then collected and aliquoted for two separate applications. The first aliquot, comprising ∼90% of the cell lysate, was prepared for total RNA purification using the Direct-zol RNA Miniprep kits from Zymo Research. The remaining ∼10% of the cell lysate was then homogenized into a denaturing buffer solution containing 4X NuPAGE™ LDS Sample Buffer in preparation for polyacrylamide gel electrophoresis (Invitrogen™ NuPAGE™, 4–12%, Bis-Tris, 1.0–1.5 mm, Mini Protein Gels) and subsequent western blots. For western blots, protein samples were transferred to a Immobilon NC membrane (Millipore) using a Genie Blotter (Idea Scientific). Membranes were probed with anti-HSP90 (Santa Cruz Biotech) and anti-T7 (Novagen) monoclonal antibodies and visualized by HRP conjugated secondary antibodies and chemiluminescence (Pierce). These experiments were performed with three independent/biological replicates.

Two-step RT-qPCR and analysis of splicing reporter assays

A minimum of 500 ηg of purified total RNA was used as input for all first-strand cDNA synthesis using random primers and Multiscribe Reverse Transcriptase (Applied Biosystems). The resulting cDNA was then used as a template for endpoint PCR amplification using specific primers that detect our mRNA splicing reporter isoforms, Globin F: 5′-CGCAACCTCAAACAGACACC-3′; Globin R: 5′-TGAGCTGCACTGTGACAAGC-3′. The forward primer of the pair contains a 5′-FAM modification. The resulting amplicons were then analyzed using agarose gel electrophoresis to empirically evaluate mRNA isoforms detected. Intron-spanning primers against *SRSF3* mRNA were used as an endogenous internal control (SRSF3 F: 5′-GTAAGAGTGGAACTGTCGAATGG-3′; SRSF3 R: 5′-CGATCTCTCTCTTCTCCTATCTCTAG-3′). The abundance of each 5′-FAM labeled mRNA isoform is quantified using capillary electrophoresis and fragment analysis (UC Berkeley, DNA Sequencing Center). For fragment analysis, each sample is suspended in a formamide solution that contains a proper size standard for sizing detected fragments (GeneScan 1200 Liz, Applied Biosystems). Analysis was performed in PeakScanner (Thermofisher). Quantification of splicing efficiency is achieved by comparing relative fluorescence units (RFU) between 5′-FAM labeled reporter isoforms that include or exclude an exon of interest. The RFU detected for each reporter isoform is then plugged into the following formula to calculate the PSI index, which reflects the splicing efficiency of an exon in either the WT or pathogenic variant context:

$$PSI = \frac{Included\ Isoform\ RFU}{Included\ Isoform\ RFU + Excluded\ Isoform\ RFU}$$

The mean PSI for a given reporter context is then calculated using all its respective replicates for a corresponding experiment. Statistical significance in the differences between the mean PSI of the control group(s) versus the experimental group(s) is determined using analysis of variance (ANOVA), and Dunett's post-hoc test. All statistical tests for PSI analysis were done in GraphPad Prism 9. Values are determined to be statistically significant if the calculated *P*-value is below an alpha value of ≤0.05.

In vitro transcribed RNA for *F8* exon-16 WT and pathogenic variants

Templates for WT or pathogenic variants of *F8* exon-16 pre-mRNA sequences corresponding to the reporter plasmid inserts were synthesized by a primer assembly reaction designed using Primerize (40). RNA was purified by denaturing PAGE and eluted from gel slices overnight in 10 mM Tris pH 7.5, 480 mM sodium acetate, 1 mM EDTA, 0.1% SDS. Following ethanol precipitation, transcripts were resuspended in ddH2O and quantified by UV spectrophotometry.

In vitro SHAPE-MaP-seq of *F8* targets

F8 exon-16 *in vitro* transcribed pre-mRNA sequences were first denatured by incubating at 95°C for 3 min in 65 mM Na-HEPES (pH 8.0). The denatured RNA was then allowed to slowly cool to room temperature (RT) for 15 min, after which $MgCl_2$ was supplemented to 1 mM for a total volume of 15 μl and incubated at RT for an additional 5 min. To chemically modify RNA, 2-aminopyridine-3-carboxylic acid imidazolide (2A3) was added to a final concentration of 100 mM and incubated for 2 min at 37°C (41,42). The reaction was then quenched using dithiothreitol (DTT) to a final concentration of 500 mM at RT for 10 min. Reactions which substituted anhydrous dimethyl sulfoxide (DMSO) for 2A3 were used as negative controls. All modified RNAs, including negative controls, were then purified using RNA Clean & Concentrator-5 (Zymo Research). Modified RNAs were fragmented to a median size of 200 nucleotides by incubation at 94°C for 1 min using NEBNext® Magnesium RNA Fragmentation Kit and then purified using NEB's recommended ethanol precipitation protocol. Purified RNA was then prepared for reverse transcription, incubating the RNA with 1 μl of 10 mM dNTPs and 2 μl of 20 μM random hexamers at 70°C for 5 min, followed by immediate transfer to ice. Reverse transcription reactions were then supplemented with 4 μl of 5X RT buffer (250 mM Tris-HCl pH 8.3, 375 mM KCl), 2 μl of 0.1 M DTT, 1 μl of 120 mM $MnCl_2$, 10 U of SUPERase RNase Inhibitor, and 200 U of Superscript II Reverse Transcriptase (SSII, ThermoFisher Scientific, cat. 18064014) to a final volume of 20 μl. These reactions were then incubated at 25°C for 10 min to allow for

partial primer extension, followed by incubation at 42°C for 3 h to enable efficient extension. SSII was then heat-inactivated by incubation at 75°C for 20 min. Reverse transcription reactions were then supplemented with EDTA to a final concentration of 6 mM to chelate Mn^{2+} ions and incubated at RT for 5 min. $MgCl_2$ was then added to a final concentration of 6 mM for each reaction (41). Reverse transcription reactions were then used as input material for NEBNext® Ultra™ II DNA library Prep Kit for Illumina® (New England Biolabs, cat. E7645L), using NEBNext Multiplex Oligos for Illumina® (Unique Dual index UMI Adaptors DNA Set 1, cat. E7395). Subsequent reactions were performed following manufacturer instructions. All sequencing was performed on an Illumina iSeq100 instrument using a paired-end 2 × 150 sequencing reagent cartridge and flow cell. All SHAPE-MaP-seq experiments were performed in duplicate.

SHAPE-MaP data analysis and RNA structure prediction

All the relevant data analysis steps were conducted using RNA Framework v2.6.9 (43). Reads produced from Illumina libraries were pre-processed and mapped using the rf-map module (parameters: -b2 -mp '–no-mixed –no-discordant' -bs), ensuring only paired-end mates with expected mate orientation were considered with Bowtie2. The mutational signal was obtained using the rf-count module (parameters: -m -pp -nd -ni), enabling mutation counts of reads produced from properly paired mates. Mutational signal was normalized relative to an unmodified control using parameters (-sm 3 -nm 1 -mu 0.05) and further normalized using the 2–8% normalization approach provided by RNA Framework (44). Normalized reactivities were then supplied to RNAstructure to generate data-driven predicted structure models (45).

In silico analysis of splice site strength and protein-RNA interactions

MaxEntScan was used to analyze splice 5′ and 3′ site sequences of exons used in this study (46). To determine 3′ss strength, 23-mer sequences encompassing the last 20 intronic positions and first three exonic positions were selected for analysis. For 5′ss strength, 9-mers consisting of the last three positions of each exon and the first six positions of the downstream intron were selected for analysis. The tool RBPmap was used to identify putative protein-RNA interactions sites within *F8* pre-mRNAs (47). In this analysis, we used sequences corresponding to *F8* exon-16 and flanking introns from the splicing reporter. We included a high stringency constraint to match known RBP motifs within our input sequence, as well as a conservation filter to selectively identify motifs that best match sequences from the human and mouse genomes.

Results

Discovery of splicing-sensitive HA-causing variants in *F8*

We previously described inherited disease-causing variants with the potential to alter the landscape of ESEs or ESSs (24). Among all candidate genes, *F8* had the highest number of variants per exon and total number of putative splicing-sensitive point variants (23). To determine whether these HA-causing variants can induce aberrant splicing, we analyzed 97 distinct variants across eleven *F8* exons by generating heterologous splicing reporters where the wild-type (WT) or pathogenic variant for each *F8* exon, plus 100–250 bp of flanking intron sequence, were cloned into the human *beta globin (HBB)* minigene splicing reporter (Figure 1A and Supplementary Table S1). Following transient transfections of each reporter into HEK293T cells, *HBB* reporter mRNA isoform levels are quantified by a two-step, semi-quantitative end-labeled RT-qPCR assay and fragment analysis via capillary electrophoresis.

Out of the eleven *F8* exons tested, HA-causing variants across seven *F8* exons did not cause exon skipping (Supplementary Figure S1). In contrast, we found that HA-causing variants in four exons (exon-7, exon-11, exon-16 and exon-18) caused exon skipping in the heterologous reporter context, indicating that the splicing fidelity of these exons may be particularly sensitive to variants (Figure 1B and Supplementary Figure S2). For example, Figure 1B shows splicing assays for 16 HA-causing variants of exon-16. In comparison to the WT exon-16 splicing reporter which is efficiently spliced (lane 3), we found that six pathogenic variants, namely exon-16$^{c.5561G>A}$, exon-16$^{c.5558C>T}$, exon-16$^{c.5531C>A}$, exon-16$^{c.5531C>T}$, exon-16$^{c.5389C>T}$ and exon-16$^{c.5543A>G}$ significantly reduced exon-16 inclusion (Figure 1B,C). These results show that the *in vitro* splicing of *F8* exon-7, -11, -16 and -18 are sensitive to HA-causing variants.

A predicted three-way junction RNA structure occludes the 3′ss of *F8* exon-16

Pathogenic variants have previously been reported to disrupt native RNA structures and as a consequence alter their biological function (48). To determine how splicing-sensitive HA-causing variants influence the structure of fragile exons like *F8* exon-16, we performed selective 2′-hydroxyl acylation analyzed by primer extension and mutational profiling coupled to high-throughput sequencing (SHAPE-MaP-seq) on *in vitro* transcribed (IVT) RNA. Both the WT and pathogenic variant containing IVT exon-16 RNA encompasses the same sequence context as the splicing reporters. As such, for all SHAPE probing data presented, all nucleotide position numbering shown is based relative to the IVT RNA template (Supplementary Figure S3A). Accessible nucleotides strongly react with 2-aminopyridine-3-carboxylic acid imidazolide (2A3) in SHAPE-MaP-seq assays (41). Nucleotides acylated by 2A3 can vary in their degree of modification, or SHAPE reactivity, reflecting how accessible the nucleotides may be in the context of the global folding architecture of the RNA (41,44,49).

Using 2A3 in our SHAPE-MaP-seq assays, we first generated SHAPE reactivity profiles for WT exon-16 and the highly splicing-sensitive exon-16$^{c.5543A>G}$ variant (Figure 2A,B). We used arc diagrams to compare SHAPE-driven folding predictions of WT and exon-16$^{c.5543A>G}$ transcripts (Figure 2C). The exon-16$^{c.5543A>G}$ variant induces subtle RNA structural changes, creating several long-range base pairing interactions within the exon and between the flanking introns. Like exon-16$^{c.5543A>G}$, additional SHAPE probing demonstrates that other splicing-sensitive HA-causing variants in exon-16 also exhibited modest structural changes compared to the WT transcript (Supplementary Figures 3B–G and S4). Taken together, the SHAPE-MaP-seq experiments suggest that splicing-sensitive HA-causing variants modestly impact the secondary structure of exon-16 compared to the WT context.

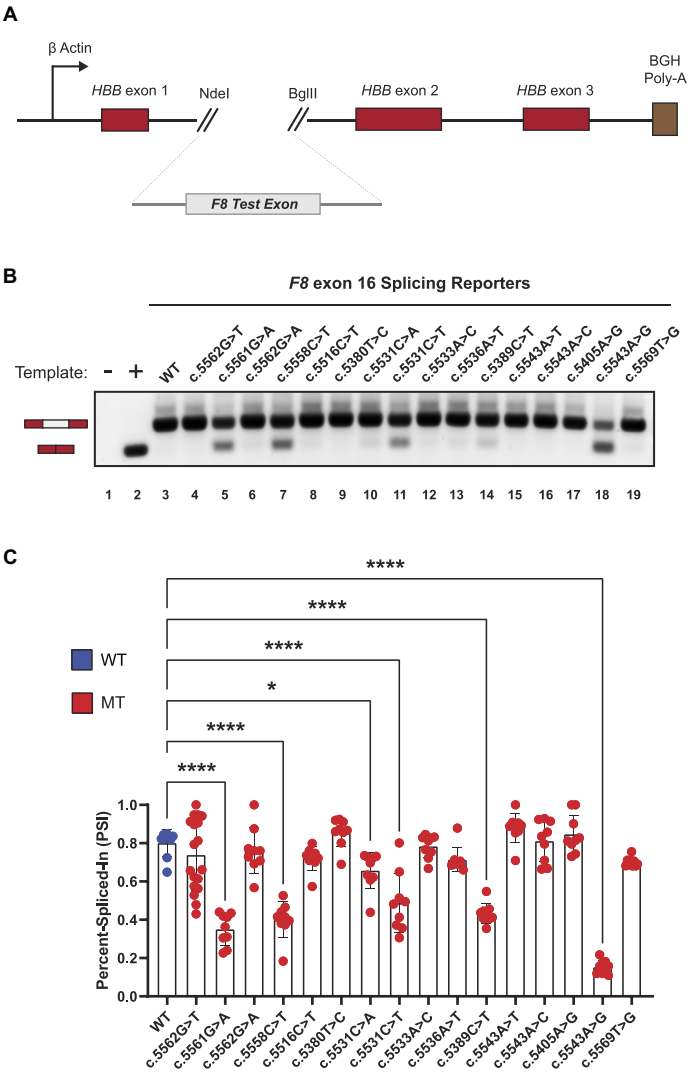


Figure 1. *In vitro* cell-based splicing reporter assays reveal *F8* exon-16 as a highly fragile exon susceptible to pathogenic variant-induced aberrant splicing. (**A**) Schematic of the heterologous splicing reporter used to assess the impacts of variants on splicing. Each test exon and flanking intronic sequence of *F8* gene was cloned between exon-1 and exon-2 of the *HBB* minigene. (**B**) A representative agarose RNA gel showing the effects on splicing by various HA-causing variants in a panel of *F8* exon-16 splicing reporters. Controls include a no template reaction (lane 1) and a positive control for exon skipping (lane 2). (**C**) Quantification of various HA-causing variants on the splicing extent of *F8* exon-16. Percent-spliced-in (PSI) refers to the ratio of test exon skipped to test exon included in mRNA. Statistical significance between comparisons is denoted by asterisks that represent *P*-values with the following range of significance: *****$P \leq 0.05$, and ********$P \leq 0.0001$. Statistical significance was determined using analysis of variance (ANOVA), and Dunett's post-hoc test. Each exon-16 splicing reporter context tested contains nine independent/biological replicates; only exon-16$^{c.5562G>T}$ contains 18 independent/biological replicates to assess variability.

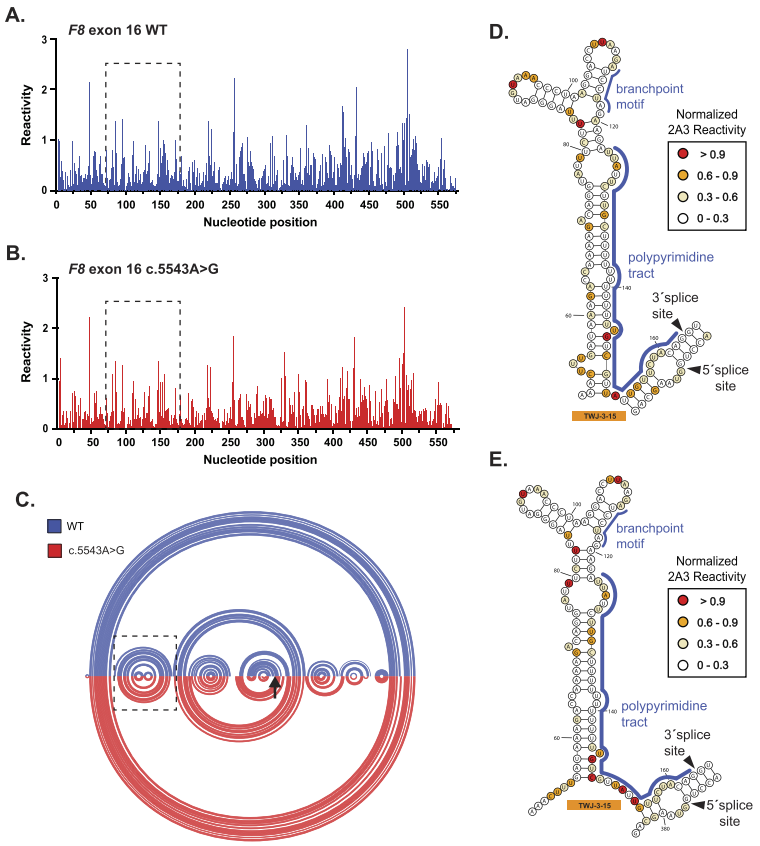


Figure 2. SHAPE probing identifies a native RNA structure (TWJ-3–15) that is uniquely positioned at the 3′ss of *F8* exon-16. (**A,B**) A normalized SHAPE reactivity plot for WT exon-16 (blue) and the exon-16$^{c.5543A>G}$ pathogenic variant (red), respectively. (**C**) Intramolecular base pairing interactions, constrained by normalized SHAPE reactivity are represented by arcs joining different regions of the transcript. Arc diagrams for WT exon-16 and exon-16$^{c.5543A>G}$ transcripts are depicted in blue and red, respectively. The broken box indicates the position of TWJ-3–15. The black arrow signifies the position of the c.5543A > G variant (**D,E**) SHAPE-driven secondary structure prediction of TWJ-3–15 depicted in its 2D state for WT exon-16 and exon-16$^{c.5543A>G}$ transcripts, respectively. Core splicing signals are annotated within the structure. All SHAPE probing data ($N = 2$) presented were generated *in vitro* using the SHAPE reagent 2A3, and all subsequent data analysis was performed in RNA Framework. All nucleotide positions numbering shown are based on the IVT RNA template used for SHAPE probing, from the 5′ to 3′ orientation.

When analyzing both the WT and all pathogenic variant SHAPE-driven structure predictions together, we discovered a three-way junction RNA structure involving the 3′ss of *F8* exon-16 that is supported by the experimental reactivity patterns but not from *in silico* RNA folding alone (Figure 2D,E; Supplementary Figures S5 and 6). Our SHAPE-driven structural predictions suggest that an upstream region of *F8* intron-15 base pairs with the branchpoint and poly-Y tract, potentially occluding their accessibility to splicing factor 1 (SF1) and U2AF. This 3′ss three-way junction structure, which we will refer to as TWJ-3–15 (Three-Way Junction at the 3′-end of intron-15), may be a potential target for splice-modulating compounds to enhance splice site strength and therefore increase exon-16 inclusion during splicing.

TWJ-3-15 attenuates F8 exon-16 3′ss strength

To determine if splicing-sensitive HA-causing variants of exon-16 can be rescued by ASOs, we designed non-overlapping, phosphorothioate-substituted, 2′-methoxyethyl modified 18-mer ASOs that span exon-16 and its flanking introns in a splicing reporter context (Figure 3A; Supplementary Table S2 and Supplementary Figure S7). We include an ASO that has no complementarity to *F8* sequences, serving as our non-targeting (NT) control (Figure 3B, lane 1), and an ASO 'blocker' that specifically targets the 5′ss of exon-16 to directly inhibit its splicing (Figure 3B, lane 2). Co-transfection of ASO_{32}, ASO_{30}, ASO_{29}, ASO_{27}, ASO_6, ASO_4 or ASO_3 with the exon-16$^{c.5543A>G}$ reporter resulted in a significant increase in exon-16 inclusion relative to the control (Figure 3C). Of note, these ASOs target the upstream and downstream

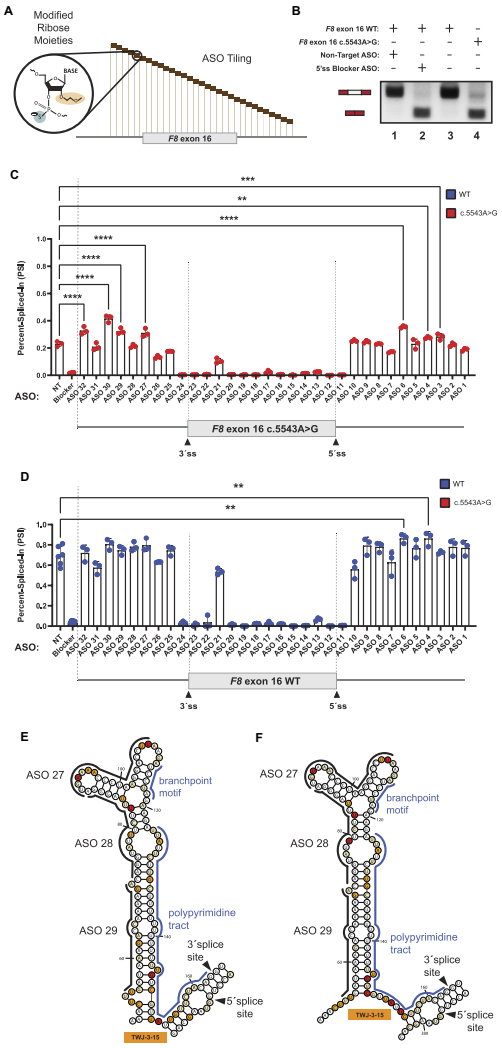


Figure 3. ASO walk reveals splice-modulating ASOs for the highly splicing-sensitive exon-16$^{c.5543A>G}$ variant. (**A**) A mock schematic of an ASO walk. Each ASO used in our walks are 18 nucleotides in length and are designed using ribose sugars that are modified with a 2′-methoxyethyl group (2′-MOE, highlighted in light orange), and the phosphate backbone is modified to a phosphorothioate backbone (highlighted in light blue). Each 18-mer ASO is contiguous by design, tiling across exon-16 and its flanking introns with no overlaps between each ASO. (**B**) Proof-of-concept demonstrating how our ASOs are expected to work in the ASO walk experiments. As shown in the annotative matrix above a representative agarose gel, the first two controls consist of our 5′ss blocker ASO (positive control) and our non-targeting ASO (negative control) being co-transfected with our WT exon-16 splicing reporter to demonstrate that our designed ASOs can modulate splicing. The last two controls consist of our WT exon-16 and exon-16$^{c.5543A>G}$ splicing reporters without ASOs co-transfected to illustrate the typical splicing ratios we may expect to see from their splicing. Expected mRNA isoforms including or excluding the test exon are also annotated to the left of the agarose gel. (**C**) and (**D**) show our ASO walk data for the exon-16$^{c.5543A>G}$ variant and WT exon-16, respectively. Data corresponding to the exon-16$^{c.5543A>G}$ variant is annotated by a red color whereas the WT is annotated by a blue color. ASO walk results for both (**C**) and (**D**) are quantified using the PSI ratio. Statistical significance between comparisons are denoted by asterisks that represent *P*-values with the following range of significance: ns, $P > 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, and ****$P \leq 0.0001$. Statistical significance was determined using analysis of variance (ANOVA), and Dunett's post-hoc test. Each exon-16 splicing reporter context and condition tested contains three independent/biological replicates. A schematic model of exon-16 and its flanking introns are shown at the bottom of each plot to illustrate relative positions of ASOs. (**E**) and (**F**) respectively depict SHAPE-driven secondary structure predictions of TWJ-3–15 for WT exon-16 and exon-16$^{c.5543A>G}$, where each (E) and (F) show where promising ASOs are hybridizing to within the structure. The sequence is numbered according to the nucleotide positions of the heterologous splicing reporter, from the 5′ to 3′ orientation.

intronic regions of exon-16$^{c.5543A>G}$. ASO_{32}, ASO_{30}, ASO_{29} and ASO_{27} target regions upstream of the 3′ss, including sequences comprising TWJ-3–15. These ASOs significantly increase inclusion of exon-16$^{c.5543A>G}$ (Figure 3C, *P*-value < 0.0001). ASO_6, ASO_4 and ASO_3 target a region downstream of the 5′ss and also significantly increase exon-16$^{c.5543A>G}$ inclusion (Figure 3C, *P*-value < 0.0008). We also observed that ASOs targeting exon-16, including the site of the variant at c.5543A > G, inhibited splicing. As a control we repeated the ASO walk on the WT *F8* exon-16 reporter. As observed from the exon-16$^{c.5543A>G}$ walk, ASOs directly targeting the exon, except for ASO_{21}, strongly inhibited the inclusion of WT exon-16 during splicing (Figure 3D). By contrast, individual ASOs targeting the flanking introns had little impact on WT exon-16 splicing relative to the NT control. Interestingly, ASO_6 and ASO_4 which target the intronic region downstream of the 5′ss enhanced exon-16 splicing relative to the NT control (Figure 3D, *P*-value < 0.0011). Taken together, these ASO walks indicate that targeting regions adjacent to *F8* exon-16 with ASOs, but not the exon itself, may rescue inclusion of splicing-sensitive HA-causing exon-16 variants by perturbing the influence of inhibitory elements found in the flanking introns.

Comparison of our SHAPE and ASO walk data revealed that ASOs which individually and most significantly improved splicing of exon-16$^{c.5543A>G}$ achieved this effect by directly targeting TWJ-3–15 (Figure 3E,F). These ASOs include ASO_{29} and ASO_{27}, which had a statistically significant improvement on exon-16$^{c.5543A>G}$ splicing, and ASO_{28} which also targets the structure but was unable to significantly rescue splicing of exon-16$^{c.5543A>G}$. Thus, our data supports the hypothesis that co-transfecting more than one ASO targeting TWJ-3–15 may further destabilize this structure to increase 3′ss accessibility and thereby enhance proper splicing of exon-16.

Splice-modulating ASOs alter the predicted structure and function of TWJ-3-15

To test the hypothesis that a combination of ASOs can additively enhance splicing of *F8* exon-16$^{c.5543A>G}$ compared to a single ASO when targeting TWJ-3–15, we performed additional cell-based splicing assays where exon-16$^{c.5543A>G}$ was co-transfected with ASO_{29}, ASO_{28} and ASO_{27}. Due

to the observation that WT exon-16 is inefficiently spliced (Figure 1C), we also co-transfected the WT reporter with the trio ASO combination in parallel. Remarkably, for both the WT and exon-16$^{c.5543A>G}$ reporters, we discover that the trio ASO combination had a significant effect on exon-16 inclusion (Figure 4A,B). Relative to WT exon-16 co-transfected with the NT control (Figure 4A,B), the trio ASO combination can enhance the splicing efficiency of WT exon-16 (Figure 4A,B), significantly increasing exon-16 inclusion (*P*-value < 0.0001). Similarly, we observed that the trio combination significantly rescued exon-16$^{c.5543A>G}$ splicing (Figure 4A,B, *P*-value < 0.0001) and completely restores splicing of exon-16$^{c.5543A>G}$ to WT levels (Figure 4A,B, compare lanes 1 and 6). Additionally, in comparing this result to a preliminary screen consisting of a duo ASO combination that includes ASO_{29} and ASO_{28} which worked the best (Supplementary Figure S8), the trio combination comprising ASO_{29}, ASO_{28} and ASO_{27} had a stronger additive effect on rescuing exon-16 splicing, also demonstrating that ASO_{27} is necessary to improve splicing.

To test the hypothesis that ASOs targeting TWJ-3–15 increases the accessibility of the poly-Y tract by antagonizing the TWJ-3–15 structure, we performed additional chemical probing experiments. Briefly, we adapted the same SHAPE-MaP-seq approach as previously described with the exception that prior to SHAPE probing, the *in vitro* RNA template was unfolded and then re-folded in the presence of ASOs. In order to draw comparisons between the probing condition with ASOs present and the control condition without ASOs, we calculated the average SHAPE reactivity for each nucleotide from the respective datasets and plotted them together (Figure 4C). Differences between each dataset's SHAPE reactivities are represented by their distinct color annotation, whereas predominant admixing of colors represent minimal differences. Accordingly, the results from this experiment show that there is an increased shift in SHAPE reactivities for nucleotides that surround or comprise the poly-Y tract in the condition with ASO_{29} and ASO_{28} present, compared to the condition with no ASOs present (Figure 4C). The reduced SHAPE reactivities in the regions with ASO_{29} and ASO_{28} complementarity provide a direct measure of ASO binding to the intended target sites (Figure 4C). Additionally, a single ASO such as ASO_{27} is also capable of increasing the SHAPE activities for nucleotides comprising the poly-Y tract (Supplementary Figure S9). These results indicate that appropriately designed ASOs enhance exon-16 splicing in part by destabilizing TWJ-3–15 to increase the accessibility of the 3′ss, likely increasing the accessibility of the poly-Y tract to U2AF.

TWJ-3-15 harbors hnRNPA1-dependent splicing silencers

Results of our RNA structure probing experiments in the presence and absence of specific ASO combinations suggest that TWJ-3–15 reduces the strength of *F8* exon-16 definition by occluding the poly-Y tract. To determine if TWJ-3–15 may also contain any functional binding sites for RNA-binding proteins (RBPs), we used RBPmap to identify RBP consensus motifs within the structure (47). We found two binding sites for the splicing repressor hnRNPA1 within TWJ-3–15 (Figure 5A, indicated in red). Based on nucleotide position numbering relative to the IVT RNA template used for SHAPE probing (Supplementary Figures S3A, 6 and 7), and subsequent analysis of TWJ-3–15, the first potential binding site is found at nucleotide positions 84–90 (UUAGGGA) and the second motif is found at nucleotide positions 99–105 (CUAAGGA). We term these predicted hnRNPA1 binding sites as ISS-15–1 and ISS-15–2, respectively. Based on published research, these predicted binding sites harbor motifs either identical or highly similar to the hnRNPA1 consensus motif, 'UAGG' (50,51). These predicted hnRNPA1 binding sites are positioned within the apical stem loop region of TWJ-3–15. Intriguingly, ASO_{28} and ASO_{27} directly bind ISS-15–1 and ISS-15–2, respectively. These ASOs, when used individually or in combination, antagonized the aberrant splicing of the exon-16$^{c.5543A>G}$ variant (Figure 3C and Supplementary Figure S8A). We hypothesize that this structured element weakens the *F8* exon-16 3′ss by sequestering the poly-Y tract and recruiting the splicing repressor protein hnRNPA1. To determine if TWJ-3–15 is sufficient to induce hnRNPA1-dependent exon skipping we replaced the robust 3′ss of *F8* exon-15 with the 3′ss of exon-16, which includes TWJ-3–15 (see Supplementary Figure S1). By contrast to the WT exon-15 reporter, we found that overexpression of hnRNPA1 induced significant levels of exon skipping in the TWJ-3–15 chimeric reporter (Supplementary Figure S10). These data suggest that TWJ-3–15 is sufficient to induce hnRNPA1-dependent exon skipping in a heterologous context. If hnRNPA1 recognizes ISS-15–1 and ISS-15–2, we predict that ASO_{28} and ASO_{27} will block hnRNPA1-dependent exon skipping. To test this hypothesis, we co-transfected the WT *F8* exon-16 splicing reporter into HEK293T cells with or without hnRNPA1 overexpression. We performed these experiments with several ASO conditions: (1) transfecting ASO_{28} and ASO_{29} to mask one putative hnRNPA1 binding site, ISS-15–1, while exposing the poly-Y track, (2) or transfecting ASO_{27}, ASO_{28} and ASO_{29} to mask both putative hnRNPA1 binding sites, ISS-15–1 and ISS-15–2, while simultaneously exposing the poly-Y track in TWJ-3–15. As expected, overexpression of hnRNPA1 strongly inhibits splicing of WT exon-16 (compare lane 4 to lane 1, Figure 5B,C). By contrast, we observed that the combination of ASO_{28} and ASO_{29} partially attenuated the effect of hnRNPA1 overexpression, and the combination of all three ASO resulted in the strongest attenuation hnRNPA1-dependent exon-16 skipping (compare lane 4 to lanes 5 and 6). Taken together, these data suggest that ISS-15–1 and ISS-15–2 are recognized by hnRNPA1 and that masking both sites is required to rescue splicing efficiently.

ASO_{27} contains a potential high affinity hnRNPA1 binding site (5′-AGGTCCTTAGGGTTTACA-3′), inviting the hypothesis that ASO_{27} may attenuate exon-16 splicing in our hnRNPA1-ASO competition assay by directly binding to and sponging hnRNPA1. To test this hypothesis, we repeated the hnRNPA1-ASO competition assay using an orthogonal hnRNPA1-responsive splicing reporter that does not bind ASO_{27}. This splicing reporter contains *SRSF6* exon-6, a reporter previously validated to exhibit splicing suppression upon hnRNPA1 overexpression (52). We observe that, relative to the empty expression vector, overexpression of hnRNPA1 caused skipping of *SRSF6* exon-6, as expected (Figure 5D, compare lane 1 to lanes 2 and 3; Figure 5E). The effect of hnRNPA1 overexpression is maintained in the presence of ASO_{27} (Figure 5D, lane 2), suggesting that ASO_{27} does not attenuate hnRNPA1-directed inhibition of splicing by binding to the splicing factor itself. Taken together, these experiments support our hypothesis that the trio ASO combination rescues exon-16 splicing in part by blocking *bona fide*

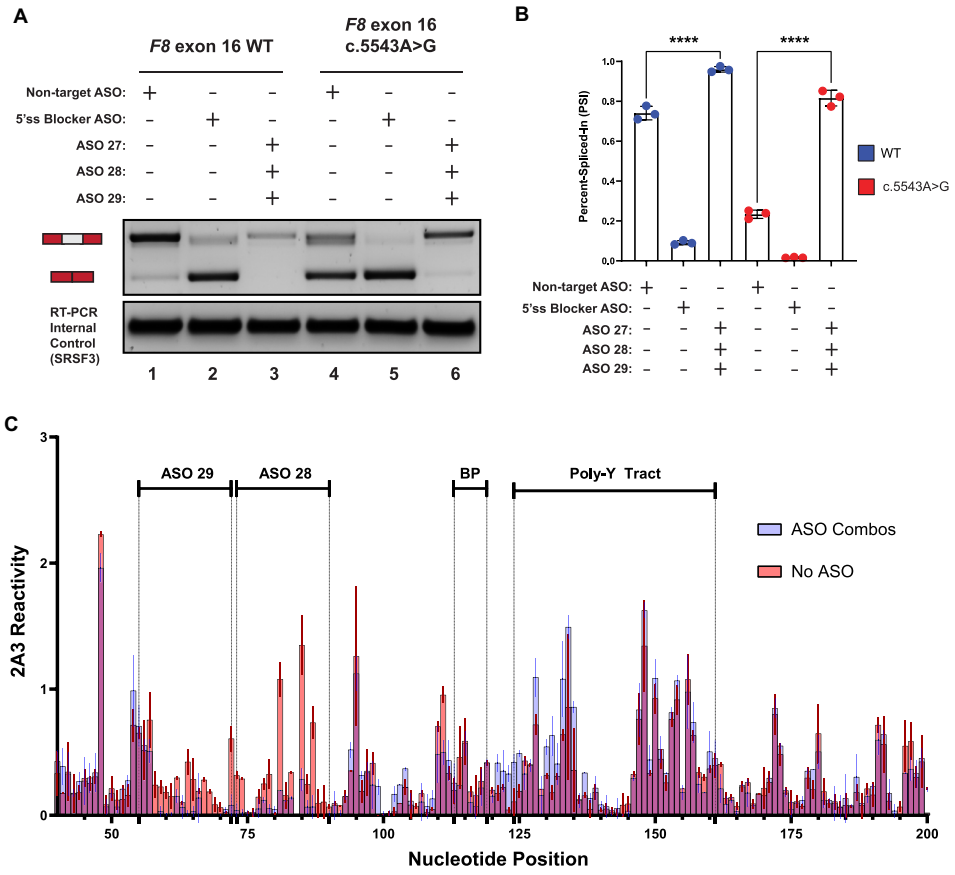


Figure 4. A combination of ASOs targeting TWJ-3–15 can rescue splicing of the highly splicing-sensitive exon-16$^{c.5543A>G}$ variant by increasing 3′ss accessibility. (**A**) A representative agarose gel depicting the results from our *in vitro* cell-based splicing assays testing duo and trio ASO combinations' ability to modulate reporter splicing (upper panel). The lower panel depicts an internal control corresponding to the SRSF3 mRNA (lower panel). Each splicing assay condition is annotated as shown in the matrix above the gel. Expected mRNA isoforms including or excluding the test exon are also annotated to the left of the agarose gel. (**B**) A plot quantifying the results from (A) using the PSI ratio. The WT context is annotated by a blue color whereas the exon-16$^{c.5543A>G}$ pathogenic variant is annotated by a red color. The same annotative matrix seen in (A) is used under the plot to label each ASO condition tested for each context. Statistical significance between comparisons are denoted by asterisks that represent *P*-values with the following range of significance: ns, $P > 0.05$, and **** $P \leq 0.0001$. Statistical significance was determined using analysis of variance (ANOVA) and Dunett's *post-hoc* test. Each exon-16 splicing reporter context and condition tested contains three independent/biological replicates. (**C**) An overlay plot comparing normalized 2A3 reactivities between two distinct SHAPE probing conditions used to probe the exon-16$^{c.5543A>G}$ variant. One SHAPE condition probes exon-16$^{c.5543A>G}$ with ASOs present (annotated light blue), and the other condition probes exon-16$^{c.5543A>G}$ without ASOs present (annotated light red). Admixing of colors (indicated by purple hue) where this is indistinguishable overlap represents similar SHAPE reactivity values between the two probing conditions at that nucleotide position. The nucleotide positions where the ASOs bind, in addition to important splicing signals, are annotated in the plot. All SHAPE probing data presented were generated *in vitro* using the SHAPE reagent 2A3, and all subsequent data analysis was performed in RNA Framework. The sequence is numbered according to the nucleotide positions of the heterologous splicing reporter, from the 5′ to 3′ orientation.

hnRNPA1-dependent silencers found within TWJ-3–15 and increasing accessibility of the 3′ss poly-Y tract.

Targeting TWJ-3-15 rescues multiple splicing-sensitive HA-causing variants of *F8* exon-16

Our experiments indicate that the splicing fidelity of *F8* exon-16 appears to be regulated in part by an RNA structure that weakens the 3′ss and recruits hnRNPA1 to further suppress exon definition (Figure 6A). Because this feature is shared across all exon-16 HA-causing variants (Supplementary Figures S11–16) and is targeted by ASOs capable of rescuing exon-16$^{c.5543A>G}$ splicing, we reasoned that targeting TWJ-3–15 might rescue splicing of other splicing-sensitive exon-16 variants. To test this hypothesis, we co-transfected WT exon 16, exon-16$^{c.5561G>A}$, exon-16$^{c.5558C>T}$, exon-16$^{c.5531C>A}$, exon-16$^{c.5531C>T}$ and

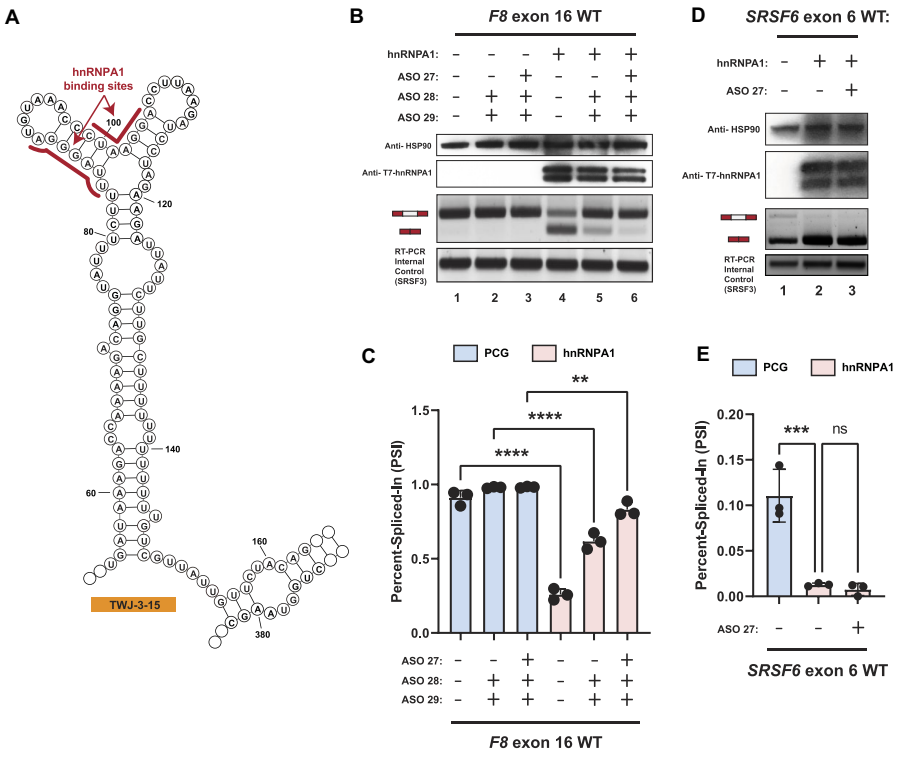


Figure 5. hnRNPA1 cooperates with TWJ-3–15 to amplify inhibitory effects at the 3′ss of *F8* exon-16. (**A**) Secondary structure model of TWJ-3–15 showing predicted hnRNPA1 binding motifs underscored in red. The sequence is numbered according to the nucleotide positions of the heterologous splicing reporter, from the 5′ to 3′ orientation. (**B**) Representative Western blot and agarose gel depicting results from our hnRNPA1-ASO competition assay. Upper two panels depict western blots for HSP90 and T7-epitope tagged hnRNPA1, respectively. Lower two panels depict the HBB splicing assay and the SRSF3 internal control for the RT-PCR reaction, respectively. Each condition tested in the assay is annotated as shown in the matrix above the gel. (**C**) A plot quantifying the results from (B) using the PSI ratio. Co-transfection of the WT exon-16 splicing reporter with either the empty expression vector (PCG) or the hnRNPA1 expression vector is indicated by a light blue or light red color, respectively. The same annotative matrix seen in (B) is used under the plot to label each ASO condition tested for each context. (**D**) Representative Western blot and agarose gel electrophoresis depicting results from our *SRSF6* splicing assay. Each condition tested in the assay is annotated as shown in the matrix above the gel. (**E**) A plot quantifying the results from (D) using the PSI ratio. Co-transfection of the *SRSF6* exon 6 splicing reporter with either the empty expression vector (PCG) or the hnRNPA1 expression vector is indicated by a light blue or light red color, respectively. The same annotative matrix seen in (D) is used under the plot to label each ASO condition tested for each context. Epitopes targeted by specific antibodies in the western blots are indicated to the left of their respective blots. Expected mRNA isoforms including or excluding the test exon are also annotated to the left of the agarose gel. Statistical significance between comparisons are denoted by asterisks that represent *P*-values with the following range of significance: ns, $P > 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, and ****$P \leq 0.0001$. Statistical significance was determined using analysis of variance (ANOVA), and Dunett's *post-hoc* test. Each exon-16 splicing reporter context and condition tested contains three independent/biological replicates.

exon-16$^{c.5389C>T}$ splicing reporters with the NT ASO control or the trio ASO combination. Similar to exon-16$^{c.5543A>G}$, we observed that co-transfecting ASO_{29}, ASO_{28} and ASO_{27} together strongly promotes inclusion of other splicing-sensitive HA-causing pathogenic variants of exon-16 (Figure 6B,C). Taken together, these data demonstrate that ASO_{29}, ASO_{28} and ASO_{27} interfere with the function of TWJ-3–15 and can generally rescue a broad array of splicing-sensitive pathogenic alleles implicated in HA.

Splicing reporters, like those employed in this study, enable the rapid screening and identification of splicing-sensitive variants across a wide array of exons. However, the heterologous sequence contexts do not always accurately reflect the splicing patterns of the endogenous gene. We therefore established an orthogonal splicing reporter to determine whether exon-16$^{c.5543A>G}$ induces exon skipping in a more native context. We substituted the *HBB* locus in our reporter plasmid with the region of the *F8* gene spanning from exon-15 to exon-17, including the full length introns flanking exon-16 (Figure 6D). To make appropriate comparisons, we generated minigenes for both WT exon-16 and the exon-16$^{c.5543A>G}$ pathogenic variant that induces the most severe exon skipping. Once the reporters were generated and sequence-validated, we then repeated splicing assays to determine if the exon-16$^{c.5543A>G}$ variant induces exon skipping, and if the trio ASO combination also rescued this splicing defect. This minigene

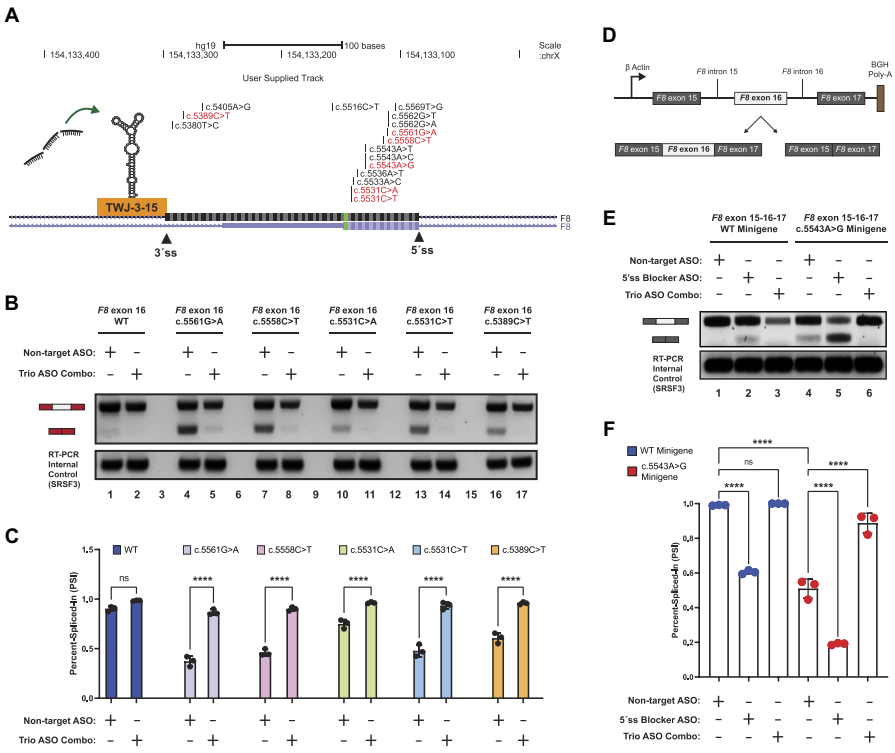


Figure 6. A combination of ASOs targeting TWJ-3–15 in a heterologous and endogenous context can rescue splicing for a broad array of HA-associated variants of exon-16 by increasing 3′ss accessibility and blocking hnRNPA1 binding. (**A**) A UCSC Genome Browser screenshot depicting the *F8* exon-16 locus and the positions of HA-causing variants tested in this study. Pathogenic variants demonstrated to be splicing-sensitive from our assays are shown in red, whereas non-sensitive variants are shown in black. The 3′ and 5′ splice sites are annotated in addition to TWJ-3–15. Successful ASOs targeting TWJ-3–15. (**B**) A representative agarose gel depicting the results from our cell-based splicing assays testing the trio ASO combinations' ability to rescue splicing of other HA-linked splicing-sensitive exon-16 variants (upper panel). The SRSF3 internal control for the RT-PCR reaction is also shown (lower panel). Each splicing assay condition included in this specific assay is annotated as shown in the matrix above the gel. Expected mRNA isoforms including or excluding the test exon are also annotated to the left of the agarose gel. (**C**) A plot quantifying the results from (B) using the PSI ratio. Each sequence context tested (WT or pathogenic variant) is annotated by a distinct color. The same annotative matrix seen in (B) is used under the plot to label each ASO condition tested for each context. (**D**) A schematic depicting the *F8* exon-16 minigene splicing reporter to validate aberrant splicing defects in an endogenous context. *F8* exon-16, along with its neighboring introns and exons as annotated, are cloned in between a strong promoter and polyadenylation signal. Expected isoforms generated from the splicing reporter are also shown and annotated. (**E**) A representative agarose gel depicting the results from our exon-16 minigene splicing assays validating aberrant splicing and trio ASO rescue effects in an endogenous context. Each assay condition included in this experiment is annotated as shown in the matrix above the gel. Expected mRNA isoforms including or excluding exon-16 are also annotated to the left of the agarose gel. (**F**) A plot quantifying the results from (E) using the PSI ratio. Each sequence context tested (WT minigene or exon-16$^{c.5543A>G}$ minigene) is annotated by a distinct color. The same annotative matrix seen in (E) is used under the plot to label each ASO condition tested for each context. Statistical significance between comparisons are denoted by asterisks that represent *P*-values with the following range of significance: ns, $P > 0.05$, and ********$P \leq 0.0001$. Statistical significance was determined using analysis of variance (ANOVA), and Dunett's post-hoc test. Each exon-16 splicing reporter context and condition tested contains three independent/biological replicates.

assay demonstrates that c.5543A > G induces exon-16 skipping in a native context (Figure 6E, compare lanes 1 and 4). As expected, the exon-16 5′ss blocking ASO reduces inclusion of the WT and pathogenic variant minigene reporters (Figure 6E, compare lanes 1 to 2 and 4 to 5), whereas the trio ASO combination significantly increased exon-16$^{c.5543A>G}$ inclusion (Figure 6E, compare lanes 4 and 6, *P*-value < 0.0001). Together, these data demonstrate that our heterologous reporter assay fully recapitulates the results of an orthogonal model system presenting exon-16 and the c.5543A > G pathogenic variant in an native context.

Discussion

RNA structure–function relationships in exon definition

Our results suggest that a putative RNA structure, TWJ-3–15, sensitizes exon-16 to aberrant splicing by attenuating the strength of the 3′ss. This interpretation is supported by several lines of evidence. First, SHAPE data demonstrates that upstream intronic sequences base pair with the poly-Y tract. This base pairing interaction could reduce accessibility of the poly-Y tract to U2AF. A similar mechanism was demonstrated

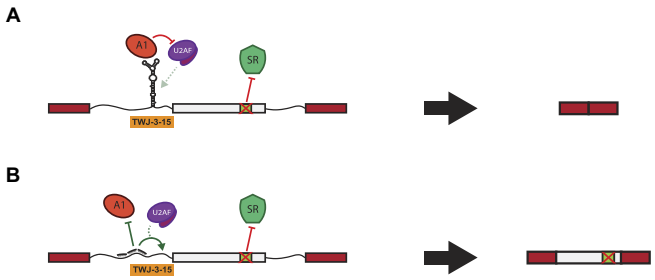


Figure 7. The loss of a critical ESE in *F8* exon-16 is hypothesized to amplify the inhibitory nature of TWJ-3–15 to alter exon definition and splicing fidelity. The following models depict experimentally-driven mechanisms on how *F8* exon-16 (in light gray) may be aberrantly spliced and rescued as determined in this study. (**A**) A schematic depicting the loss of a critical ESE in exon-16 due to the A > G pathogenic variant in the exon-16$^{c.5543A>G}$ variant. Losing the ESE diminishes the ability to recruit a positive splicing factor that likely regulates TWJ-3–15 and hnRNPA1 antagonism at the 3′ss of *F8* exon-16, leading to decreased 3′ss strength. (**B**) A schematic depicting the trio ASO combinations' ability to rescue splicing of exon-16$^{c.5543A>G}$ by destabilizing TWJ-3–15, and preventing the recruitment of hnRNPA1 to the 3′ss. Collectively, our data-supported model indicates that the trio ASOs block the recruitment of a negative splicing factor and increases the accessibility of the 3′ss to the splicing machinery. TWJ-3–15 is annotated by a simplified depiction of the 'Y-shaped' RNA secondary structure at the 3′ss of exon-16. RBPs binding to TWJ-3–15 and this region such as hnRNPA1 and U2AF are respectively annotated. The predicted ESE is annotated in light green within exon-16, and its binding partner, presumably an RBP like SR proteins that are known to enhance splicing, is depicted as well.

for the 5′ss of Tau exon 10 (53). In this case, the RNA helicase p68 controls accessibility of the 5′ss to U1 snRNP and ASOs stabilizing this structure inhibit splicing (54). Similarly, at least two putative hnRNPA1 binding sites are present in the apical stem of TWJ-3–15. These sequences appear to be functional binding sites as overexpression of hnRNPA1 inhibits splicing of both wild-type and splicing defective exon-16 variants. TWJ-3–15 also imparts hnRNPA1-dependent exon skipping in a heterologous constitutive exon, demonstrating the presence of splicing silencer activity (Supplementary Figure S10). Finally, we also discovered splice-modulating ASOs that rescue exon-16 inclusion by disrupting the function of this RNA structure. ASO_{27}, ASO_{28} and ASO_{29} antagonize TWJ-3–15 by occluding hnRNPA1-responsive splicing silencers and by exposing the poly-Y tract, underscoring the functional significance of this structure and its sequence. Despite strong splice site signals (Supplementary Table S3), our data suggest that TWJ-3–15 sensitizes exon-16 to HA-causing variants that cripple ESEs or create ESSs (Figure 7A). Modulating this RNA structure–function relationship with ASOs rescues splicing of a diverse array of splicing-sensitive HA-linked variants of exon-16 (Figure 7B). We note that in the absence of *in cellulo* chemical probing data, TWJ-3–15 remains a structural model that requires additional experimental validation. However, the impact of our work does not rest solely on this RNA structure. Collectively, our data strongly suggest that this sequence functions as an intronic splicing silencer and that ASOs occluding this region rescues an array of splicing-sensitive variants in exon-16.

By contrast to TWJ-3–15, we found that ASOs hybridizing to the exon-16 itself, spare one, inhibit pre-mRNA splicing (Figure 3D). While ASO_{21} hybridizes to a sequence predicted to function as a splicing silencer and therefore might not be expected to interfere with splicing, it is less clear why other ASOs targeting the exon are deleterious to splicing. One possibility is that the secondary structure of exon-16 is important for exon definition. Indeed, SHAPE-MaP-seq suggests extensive intramolecular pairing juxtaposes the 5′ss and 3′ss (Supplementary Figures S11–16). Perhaps ASOs targeting exon-16 interfere with base pairing interactions within the exon and disrupt splicing. Alternatively, *F8* exon-16 may be strongly dependent on exonic splicing enhancers. In this model, ASOs targeting the exon could mask binding sites for splicing factors.

The interplay of genetic variation and exon definition is complex

Interpreting the impact of genetic variation on pre-mRNA splicing is fundamental to understanding genotype-phenotype relationships (25,55,56). Recognition of exon-intron boundaries requires the combinatorial interaction of splicing regulatory proteins with ESEs and ESSs, and is required for both constitutive and alternative splicing (57). Recent high-throughput saturating mutagenesis screens of model human exons demonstrate that constitutive exons are more resilient to substitutions than alternative exons (58). In their study, Baeza-Centurion *et al.* use human genetic variation to argue that alternative exons are more prone to aberrant splicing than constitutive exons. Surprisingly, our work demonstrates that splicing of *F8* exon-16 is vulnerable to pathogenic genetic variants, despite limited evidence for regulated alternative splicing (Figure 1 and Supplementary Figure S2). Interestingly, pathogenic genetic variants in exon-16 and the 5′ss of intron 15 cause aberrant exon-16 splicing in HA patients (59,60).

An important consideration for interpreting this result is that common polymorphisms (minor allele frequency $\geq$ 5%) are fixed in the population and unlikely to negatively affect fitness. By contrast, rare pathogenic variants can induce aberrant splicing of constitutive exons (61–71). We previously demonstrated that pathogenic variants are more likely to cause the loss or gain of ESEs or ESSs, respectively (24). Despite this observation, we were surprised that for a majority of exons tested in our study, pathogenic variants tested here did not significantly alter splicing efficiency relative to the WT reporters (Supplementary Figure S1). Clearly, additional layers of reg-

ulatory information such as splicing factor occupancy maps, chromatin structure, and RNA polymerase kinetics need to be considered in future predictive models (25,55,56,72).

Aberrant splicing of *F8* in Hemophilia A

An HA diagnosis causes a severe life expectancy disadvantage. The global prevalence of HA is ∼17.5 cases per 100 000 males, while the prevalence at birth is 26.4 per 100 000 males (73). These prevalence estimates suggest as many as 1 million patients globally are afflicted with HA. Based on the Coagulation Factor Variant Database, exon-16 contains 4.5% of the 3052 unique pathogenic variants in the *F8* gene. These data suggest that >40 000 HA patients could harbor exon-16 pathogenic variants. Our work suggests that ∼37.5% of variants tested in exon-16 induce aberrant splicing, suggesting that globally, >15 480 patients could benefit from ASOs targeting TWJ-3–15. In the future, it will be important to create patient-derived models to confirm the role for TWJ-3–15 in exon-16 aberrant splicing *in vivo*. Better models will facilitate development of singular ASOs targeting TWJ-3–15 capable of fully rescuing exon-16 splicing. *F8* exon-16 encodes a portion of the A3 domain which is required for efficient blood clotting and is frequently mutated in HA (74–76). Because exon-16 is divisible by 3, the aberrant skipped isoform is predicted to express an internally deleted protein, rather than undergoing nonsense mediated decay.

Limitations of the study and future directions

There are several limitations of our study. Because we did not have access to patient-derived cell lines and cell lines expressing the canonical full length *F8* mRNA isoform are not readily available, we employed the well-established *HBB* reporter system transfected into HEK293T cells to assay the impact of *F8* pathogenic variants on exon identity (77). This minimal system exploits the context dependence of exon definition as most auxiliary *cis*-regulatory signals are near splice sites (78). However, because this reporter system lacks most of the endogenous *F8* sequence, it remains possible that results obtained with the heterologous reporter may differ from results obtained from the endogenous locus. Additionally, other processes such as chromatin modification and transcriptional elongation can regulate splicing (79). While the artificial system used here may differ from the endogenous locus, exon-16 and intron-15 variants that cause aberrant exon-16 splicing have been reported in HA patients (59,60). Despite this uncertainty, our study represents an important step towards correcting splicing-sensitive variants in *F8* exon-16. Additional work is required to validate or invalidate aspects of this potential structure–function relationship in exon-16 identity. It will also be important to learn how ESEs antagonize the function of this putative regulatory element.

In the absence of patient-derived cell models, other groups developed CRISPR edited induced pluripotent stem cell models (77,80). In the future, translating our findings into a preclinical setting will require the testing of variants and ASOs in the context of the endogenous *F8* locus in a disease-in-a-dish preclinical model. Although we have not tested our hypotheses in a patient-derived cell line or in engineered cell models, evidence supporting the fragile nature of exon-16 can be found in HA patient registries. The European Coagulation Factor Variant Database contains at least four unrelated patients with three different synonymous pathogenic variants at position c.5586 causing severe to modest HA phenotypes (75,81,82). Silent pathogenic variants most frequently induce aberrant splicing but can also alter RNA structure and mRNA stability, leading to changes in gene expression. Additionally, variants in both intron-15 and exon-16 cause aberrant splicing of exon-16 in HA patient samples (59,60). These published clinical findings, combined with our *in vitro* analysis of HA-pathogenic variants in exon-16, support the hypothesis that aberrant splicing of *F8* exon-16 occurs in HA patients and represents a potential therapeutic target. Finally, it will be important to determine how the *F8* protein tolerates missense variants in exon-16. Recent studies demonstrated that non-conservative substitutions in exon-19, including those associated with aberrant splicing, have modest effects on protein activity when expressed from a cDNA (83). Thus, the potential translation of ASOs targeting TWJ-3–15 requires a similar study of *F8* protein function and secretion using a pre-clinical model system focusing on the endogenous *F8* locus.

Data availability

The sequencing data underlying this article are available in the Gene Expression Omnibus at https://www.ncbi.nlm.nih.gov/geo/ and can be accessed under accession code GSE230495.

Supplementary data

Supplementary Data are available at NAR Online.

Acknowledgements

We thank the Toxic RNA Lab (TRL), a curricular undergraduate research experience led by J.R.S., for their help in generating the *F8* splicing reporters. We thank Alexander J. Ritter for writing the Python script that expedited our capacity to quickly design and test antisense oligonucleotides presented in this study. We thank Danny Incarnato for the gift of 2A3 and for their help in getting the RNA Framework analysis pipeline up and running. We thank the Sanford lab, and Gina Mawla for their feedback on manuscript drafts.

Author contributions: V.T., G.C., M.D.S. and J.R.S. conceptualized and led the experiments presented in this study. V.T., H.T., P.H.D., C.O., P.C., I.Q., A.H. and S.L generated the heterologous *F8* splicing reporters. V.T. and M.G. performed all *in vitro* cell-based splicing assays. V.T. developed and established the automation platform used in this study. V.T. and M.G. designed and assayed all antisense oligonucleotides targeting *F8* exon-16 using the automation platform. G.C., A.G.J. and V.T. generated the *in vitro* RNA templates for SHAPE-MaP-seq. G.C. and N.M.F. designed, performed, prepared and analyzed all *in vitro* SHAPE-MaP-seq experiments and sequencing libraries for subsequent RNA structure predictions. V.T. and M.G. performed the hnRNPA1-ASO competition assay to experimentally validate hnRNPA1 binding sites identified in this study. V.T. and M.G. designed, generated, and assayed the *F8* exon-16 minigene reporters used in this study. V.T. and G.C. performed all of the data analysis and visualization presented in this study. V.T. wrote the manuscript, with input and review of successive manuscript drafts by G.C., M.D.S., and J.R.S.

Funding

National Institutes of Health [R35GM130361 to J.R.S.; R01GM095850 to M.D.S.]; Santa Cruz Cancer Benefit Group to M.D.S.; UCSC Office of Research Seed Funding for the Center for Open Access Splicing Therapeutics. Funding for open access charge: National Institutes of Health [R35GM130361].

Conflict of interest statement

The authors affirm that no conflict of interest exists for this work.

References

1. Breathnach,R., Benoist,C., O'Hare,K., Gannon,F. and Chambon,P. (1978) Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *Proc. Natl. Acad. Sci. USA*, **75**, 4853–4857.
2. Chow,L.T., Gelinas,R.E., Broker,T.R. and Roberts,R.J. (1977) An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell*, **12**, 1–8.
3. Berget,S.M., Moore,C. and Sharp,P.A. (1977) Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. USA*, **74**, 3171–3175.
4. Konarska,M.M. and Sharp,P.A. (1987) Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. *Cell*, **49**, 763–774.
5. Kastner,B., Will,C.L., Stark,H. and Lührmann,R. (2019) Structural insights into nuclear pre-mRNA splicing in higher eukaryotes. *Cold Spring Harb. Perspect. Biol.*, **11**, https://doi.org/10.1101/cshperspect.a032417.
6. Berget,S.M. (1995) Exon Recognition in Vertebrate Splicing (∗). *J. Biol. Chem.*, **270**, 2411–2414.
7. Kohtz,J.D., Jamison,S.F., Will,C.L., Zuo,P., Lührmann,R., Garcia-Blanco,M.A. and Manley,J.L. (1994) Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature*, **368**, 119–124.
8. Graveley,B.R., Hertel,K.J. and Maniatis,T. (2001) The role of U2AF35 and U2AF65 in enhancer-dependent splicing. *RNA*, **7**, 806–818.
9. Ruskin,B., Zamore,P.D. and Green,M.R. (1988) A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell*, **52**, 207–219.
10. Wu,S., Romfo,C.M., Nilsen,T.W. and Green,M.R. (1999) Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature*, **402**, 832–835.
11. Reed,R. and Maniatis,T. (1987) A Role for Exon Sequences and Splice-Site Proximity in Splice Site Selection. *Cell*, **46**, 681–690.
12. Michaud,S. and Reed,R. (1993) A functional association between the 5' and 3' splice site is established in the earliest prespliceosome complex (E) in mammals. *Genes Dev.*, **7**, 1008–1020.
13. Watakabe,A., Tanaka,K. and Shimura,Y. (1993) The role of exon sequences in splice site selection. *Genes Dev.*, **7**, 407–418.
14. Wang,Z., Rolish,M.E., Yeo,G., Tung,V., Mawson,M. and Burge,C.B. (2004) Systematic identification and analysis of exonic splicing silencers. *Cell*, **119**, 831–845.
15. Fairbrother,W.G., Yeh,R.-F., Sharp,P.A. and Burge,C.B. (2002) Predictive identification of exonic splicing enhancers in human genes. *Science*, **297**, 1007–1013.
16. Ke,S., Shang,S., Kalachikov,S.M., Morozova,I., Yu,L., Russo,J.J., Ju,J. and Chasin,L.A. (2011) Quantitative evaluation of all hexamers as exonic splicing elements. *Genome Res.*, **21**, 1360–1374.
17. Warf,M.B. and Berglund,J.A. (2010) Role of RNA structure in regulating pre-mRNA splicing. *Trends Biochem. Sci.*, **35**, 169–178.
18. Graveley,B.R. (2005) Mutually exclusive splicing of the insect Dscam pre-mRNA directed by competing intronic RNA secondary structures. *Cell*, **123**, 65–73.
19. Shepard,P.J. and Hertel,K.J. (2008) Conserved RNA secondary structures promote alternative splicing. *RNA*, **14**, 1463–1469.
20. Muro,A.F., Caputi,M., Pariyarath,R., Pagani,F., Buratti,E. and Baralle,F.E. (1999) Regulation of fibronectin EDA exon alternative splicing: possible role of RNA secondary structure for enhancer display. *Mol. Cell. Biol.*, **19**, 2657–2671.
21. Buratti,E., Muro,A.F., Giombi,M., Gherbassi,D., Iaconcig,A. and Baralle,F.E. (2004) RNA folding affects the recruitment of SR proteins by mouse and human polypurinic enhancer elements in the fibronectin EDA exon. *Mol. Cell. Biol.*, **24**, 1387–1400.
22. Jones,A.N., Graß,C., Meininger,I., Geerlof,A., Klostermann,M., Zarnack,K., Krappmann,D. and Sattler,M. (2022) Modulation of pre-mRNA structure by hnRNP proteins regulates alternative splicing of MALT1. *Sci. Adv.*, **8**, eabp9153.
23. Sterne-Weiler,T. and Sanford,J.R. (2014) Exon identity crisis: disease-causing mutations that disrupt the splicing code. *Genome Biol.*, **15**, 201.
24. Sterne-Weiler,T., Howard,J., Mort,M., Cooper,D.N. and Sanford,J.R. (2011) Loss of exon identity is a common mechanism of human inherited disease. *Genome Res.*, **21**, 1563–1571.
25. Mort,M., Sterne-Weiler,T., Li,B., Ball,E.V., Cooper,D.N., Radivojac,P., Sanford,J.R. and Mooney,S.D. (2014) MutPred Splice: machine learning-based prediction of exonic variants that disrupt splicing. *Genome Biol.*, **15**, R19.
26. Barash,Y., Calarco,J.A., Gao,W., Pan,Q., Wang,X., Shai,O., Blencowe,B.J. and Frey,B.J. (2010) Deciphering the splicing code. *Nature*, **465**, 53–59.
27. Gagliardi,M. and Ashizawa,A.T. (2021) The challenges and strategies of antisense oligonucleotide drug delivery. *Biomedicines*, **9**, 433.
28. Hill,S.F. and Meisler,M.H. (2021) Antisense oligonucleotide therapy for neurodevelopmental disorders. *Dev. Neurosci.*, **43**, 247–252.
29. Rigo,F., Chun,S.J., Norris,D.A., Hung,G., Lee,S., Matson,J., Fey,R.A., Gaus,H., Hua,Y., Grundy,J.S., *et al.* (2014) Pharmacology of a central nervous system delivered 2'-O-methoxyethyl-modified survival of motor neuron splicing oligonucleotide in mice and nonhuman primates. *J. Pharmacol. Exp. Ther.*, **350**, 46–55.
30. Hua,Y., Sahashi,K., Hung,G., Rigo,F., Passini,M.A., Bennett,C.F. and Krainer,A.R. (2010) Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes Dev.*, **24**, 1634–1644.
31. Kim,J., Hu,C., Moufawad El Achkar,C., Black,L.E., Douville,J., Larson,A., Pendergast,M.K., Goldkind,S.F., Lee,E.A., Kuniholm,A., *et al.* (2019) Patient-customized oligonucleotide therapy for a rare genetic disease. *N. Engl. J. Med.*, **381**, 1644–1652.
32. Ronayne,E.K., Peters,S.C., Gish,J.S., Wilson,C., Spencer,H.T., Doering,C.B., Lollar,P., Spiegel,P.C. Jr and Childers,K.C. (2021) Structure of Blood Coagulation Factor VIII in Complex With an Anti-C2 Domain Non-Classical, Pathogenic Antibody Inhibitor. *Front. Immunol.*, **12**, 697602.
33. Liang,Q., Xiang,M., Lu,Y., Ruan,Y., Ding,Q., Wang,X., Xi,X. and Wang,H. (2015) Characterisation and quantification of F8 transcripts of ten putative splice site mutations. *Thromb. Haemost.*, **113**, 585–592.
34. Jourdy,Y., Fretigny,M., Nougier,C., Négrier,C., Bozon,D. and Vinciguerra,C. (2019) Splicing analysis of 26 F8 nucleotide variations using a minigene assay. *Haemophilia*, **25**, 306–315.
35. Zimmermann,M.A., Gehrig,A., Oldenburg,J., Müller,C.R. and Rost,S. (2013) Analysis of F8 mRNA in haemophilia A patients with silent mutations or presumptive splice site mutations. *Haemophilia*, **19**, 310–317.
36. Famà,R., Borroni,E., Zanolini,D., Merlin,S., Bruscaggin,V., Walker,G.E., Olgasi,C., Babu,D., Agnelli Giacchello,J., Valeri,F., *et al.* (2020) Identification and functional characterization of a

novel splicing variant in the F8 coagulation gene causing severe hemophilia A. *J. Thromb. Haemost.*, **18**, 1050–1064.
37. Rothrock,C., Cannon,B., Hahm,B. and Lynch,K.W. (2003) A conserved signal-responsive sequence mediates activation-induced alternative splicing of CD45. *Mol. Cell*, **12**, 1317–1324.
38. den Dunnen,J.T. and Antonarakis,S.E. (2000) Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Hum. Mutat.*, **15**, 7–12.
39. Chang,T., Draper,J.M., Van den Bout,A., Kephart,E., Maul-Newby,H., Vasquez,Y., Woodbury,J., Randi,S., Pedersen,M., Nave,M., *et al.* (2021) A method for campus-wide SARS-CoV-2 surveillance at a large public university. *PLoS One*, **16**, e0261230.
40. Tian,S., Yesselman,J.D., Cordero,P. and Das,R. (2015) Primerize: automated primer assembly for transcribing non-coding RNA domains. *Nucleic Acids Res.*, **43**, W522–W526.
41. Marinus,T., Fessler,A.B., Ogle,C.A. and Incarnato,D. (2021) A novel SHAPE reagent enables the analysis of RNA structure in living cells with unprecedented accuracy. *Nucleic Acids Res.*, **49**, e34.
42. Morandi,E., van Hemert,M.J. and Incarnato,D. (2022) SHAPE-guided RNA structure homology search and motif discovery. *Nat. Commun.*, **13**, 1722.
43. Incarnato,D., Morandi,E., Simon,L.M. and Oliviero,S. (2018) RNA Framework: an all-in-one toolkit for the analysis of RNA structures and post-transcriptional modifications. *Nucleic Acids Res.*, **46**, e97.
44. Siegfried,N.A., Busan,S., Rice,G.M., Nelson,J.A.E. and Weeks,K.M. (2014) RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods*, **11**, 959–965.
45. Reuter,J.S. and Mathews,D.H. (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinf.*, **11**, 129.
46. Yeo,G. and Burge,C.B. (2004) Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol.*, **11**, 377–394.
47. Paz,I., Kosti,I., Ares,M. Jr, Cline,M. and Mandel-Gutfreund,Y. (2014) RBPmap: a web server for mapping binding sites of RNA-binding proteins. *Nucleic Acids Res.*, **42**, W361–W367.
48. Waldern,J.M., Kumar,J. and Laederach,A. (2022) Disease-associated human genetic variation through the lens of precursor and mature RNA structure. *Hum. Genet.*, **141**, 1659–1672.
49. Cordero,P. and Das,R. (2015) Rich RNA Structure Landscapes Revealed by Mutate-and-Map Analysis. *PLoS Comput. Biol.*, **11**, e1004473.
50. An,P. and Grabowski,P.J. (2007) Exon silencing by UAGG motifs in response to neuronal excitation. *PLoS Biol.*, **5**, e36.
51. Han,K., Yeo,G., An,P., Burge,C.B. and Grabowski,P.J. (2005) A combinatorial code for splicing silencing: UAGG and GGGG motifs. *PLoS Biol.*, **3**, e158.
52. Howard,J.M., Lin,H., Wallace,A.J., Kim,G., Draper,J.M., Haeussler,M., Katzman,S., Toloue,M., Liu,Y. and Sanford,J.R. (2018) HNRNPA1 promotes recognition of splice site decoys by U2AF2 in vivo. *Genome Res.*, **28**, 689–698.
53. Kar,A., Fushimi,K., Zhou,X., Ray,P., Shi,C., Chen,X., Liu,Z., Chen,S. and Wu,J.Y. (2011) RNA helicase p68 (DDX5) regulates tau exon 10 splicing by modulating a stem-loop structure at the 5' splice site. *Mol. Cell. Biol.*, **31**, 1812–1821.
54. Peacey,E., Rodriguez,L., Liu,Y. and Wolfe,M.S. (2012) Targeting a pre-mRNA structure with bipartite antisense molecules modulates tau alternative splicing. *Nucleic Acids Res.*, **40**, 9836–9849.
55. Leung,M.K.K., Xiong,H.Y., Lee,L.J. and Frey,B.J. (2014) Deep learning of the tissue-regulated splicing code. *Bioinformatics*, **30**, i121–i129.
56. Cygan,K.J., Sanford,C.H. and Fairbrother,W.G. (2017) Spliceman2: a computational web server that predicts defects in pre-mRNA splicing. *Bioinformatics*, **33**, 2943–2945.
57. Hertel,K.J. (2008) Combinatorial control of exon recognition. *J. Biol. Chem.*, **283**, 1211–1215.
58. Baeza-Centurion,P., Miñana,B., Valcárcel,J. and Lehner,B. (2020) Mutations primarily alter the inclusion of alternatively spliced exons. *eLife*, **9**, e59959.
59. Tavassoli,K., Eigel,A., Pollmann,H. and Horst,J. (1997) Mutational analysis of ectopic factor VIII transcripts from hemophilia A patients: identification of cryptic splice site, exon skipping and novel point mutations. *Hum. Genet.*, **100**, 508–511.
60. Yenchitsomanus,P., Thanootarakul,P., Akkarapatumwong,V., Oranwiroon,S., Pung-Amritt,P., Veerakul,G. and Mahasandana,C. (2001) Mutation causing exon 15 skipping and partial exon 16 deletion in factor VIII transcript, and a method for direct mutation detection. *Haemophilia*, **7**, 335–338.
61. Holm,L.L., Doktor,T.K., Hansen,M.B., Petersen,U.S.S. and Andresen,B.S. (2022) Vulnerable exons, like ACADM exon 5, are highly dependent on maintaining a correct balance between splicing enhancers and silencers. *Hum. Mutat.*, **43**, 253–265.
62. Cartegni,L., Hastings,M.L., Calarco,J.A., de Stanchina,E. and Krainer,A.R. (2006) Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. *Am. J. Hum. Genet.*, **78**, 63–77.
63. Ars,E., Serra,E., García,J., Kruyer,H., Gaona,A., Lázaro,C. and Estivill,X. (2000) Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1. *Hum. Mol. Genet.*, **9**, 237–247.
64. Kelly,D.P., Whelan,A.J., Ogden,M.L., Alpers,R., Zhang,Z.F., Bellus,G., Gregersen,N., Dorland,L. and Strauss,A.W. (1990) Molecular characterization of inherited medium-chain acyl-CoA dehydrogenase deficiency. *Proc. Natl. Acad. Sci. USA*, **87**, 9236–9240.
65. Gregersen,N., Andresen,B.S., Bross,P., Winter,V., Rüdiger,N., Engst,S., Christensen,E., Kelly,D., Strauss,A.W. and Kølvraa,S. (1991) Molecular characterization of medium-chain acyl-CoA dehydrogenase (MCAD) deficiency: identification of a lys329 to glu mutation in the MCAD gene, and expression of inactive mutant enzyme protein in *E. coli*. *Hum. Genet.*, **86**, 545–551.
66. Tubeuf,H., Caputo,S.M., Sullivan,T., Rondeaux,J., Krieger,S., Caux-Moncoutier,V., Hauchard,J., Castelain,G., Fiévet,A., Meulemans,L., *et al.* (2020) Calibration of pathogenicity due to variant-induced leaky splicing defects by using BRCA2 Exon 3 as a Model System. *Cancer Res.*, **80**, 3593–3605.
67. Rave-Harel,N., Kerem,E., Nissim-Rafinia,M., Madjar,I., Goshen,R., Augarten,A., Rahat,A., Hurwitz,A., Darvasi,A. and Kerem,B. (1997) The molecular basis of partial penetrance of splicing mutations in cystic fibrosis. *Am. J. Hum. Genet.*, **60**, 87–94.
68. Aznarez,I., Chan,E.M., Zielenski,J., Blencowe,B.J. and Tsui,L.-C. (2003) Characterization of disease-associated mutations affecting an exonic splicing enhancer and two cryptic splice sites in exon 13 of the cystic fibrosis transmembrane conductance regulator gene. *Hum. Mol. Genet.*, **12**, 2031–2040.
69. Dietz,H.C., Valle,D., Francomano,C.A., Kendzior,R.J. Jr, Pyeritz,R.E. and Cutting,G.R. (1993) The skipping of constitutive exons in vivo induced by nonsense mutations. *Science*, **259**, 680–683.
70. D'Souza,I., Poorkaj,P., Hong,M., Nochlin,D., Lee,V.M., Bird,T.D. and Schellenberg,G.D. (1999) Missense and silent tau gene mutations cause frontotemporal dementia with parkinsonism-chromosome 17 type, by affecting multiple alternative RNA splicing regulatory elements. *Proc. Natl. Acad. Sci. USA*, **96**, 5598–5603.
71. Liu,H.X., Cartegni,L., Zhang,M.Q. and Krainer,A.R. (2001) A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. *Nat. Genet.*, **27**, 55–58.
72. Holm,L.L., Doktor,T.K., Flugt,K.K., Petersen,U.S.S., Pedersen,R. and Andresen,B.S. (2023) All exons are not created equal - Exon vulnerability determines the effect of exonic mutations on splicing. bioRxiv doi: https://doi.org/10.1101/2023.06.14.544306, 14 June 2023, preprint: not peer reviewed.

73. Iorio,A., Stonebraker,J.S., Chambost,H., Makris,M., Coffin,D., Herr,C., Germini,F. and Data,and Demographics Committee of the World Federation of Hemophilia (2019) Establishing the prevalence and prevalence at birth of hemophilia in males: a meta-analytic approach using national registries. *Ann. Intern. Med.*, **171**, 540–546.
74. Giansily-Blaizot,M., Rallapalli,P.M., Perkins,S.J., Kemball-Cook,G., Hampshire,D.J., Gomez,K., Ludlam,C.A. and McVey,J.H. (2020) The EAHAD blood coagulation factor VII variant database. *Hum. Mutat.*, **41**, 1209–1219.
75. McVey,J.H., Rallapalli,P.M., Kemball-Cook,G., Hampshire,D.J., Giansily-Blaizot,M., Gomez,K., Perkins,S.J. and Ludlam,C.A. (2020) The European Association for Haemophilia and Allied Disorders (EAHAD) Coagulation Factor Variant Databases: important resources for haemostasis clinicians and researchers. *Haemophilia*, **26**, 306–313.
76. Gouw,S.C., Van Der Bom,J.G., Van Den Berg,H.M., Zewald,R.A., Ploos Van Amstel,J.K. and Mauser-Bunschoten,E.P. (2011) Influence of the type of F8 gene mutation on inhibitor development in a single centre cohort of severe haemophilia A patients. *Haemophilia*, **17**, 275–281.
77. Zhou,H., Arechavala-Gomeza,V. and Garanto,A. (2023) Experimental model systems used in the preclinical development of nucleic acid therapeutics. *Nucleic Acid Ther.*, **33**, 238–247.
78. Fu,X.-D. and Ares,M. Jr (2014) Context-dependent control of alternative splicing by RNA-binding proteins. *Nat. Rev. Genet.*, **15**, 689–701.
79. Gehring,N.H. and Roignant,J.-Y. (2021) Anything but ordinary – emerging splicing mechanisms in eukaryotic gene regulation. *Trends Genet.*, **37**, 355–372.
80. Tomkiewicz,T.Z., Nieuwenhuis,S.E., Cremers,F.P.M., Garanto,A. and Collin,R.W.J. (2022) Correction of the splicing defect caused by a recurrent variant in ABCA4 (c.769-784C>T) that underlies stargardt disease. *Cells*, **11**, 3947.
81. Higuchi,M., Kazazian,H.H. Jr, Kasch,L., Warren,T.C., McGinniss,M.J., Phillips,J.A. 3rd, Kasper,C., Janco,R. and Antonarakis,S.E. (1991) Molecular characterization of severe hemophilia A suggests that about half the mutations are not within the coding regions and splice junctions of the factor VIII gene. *Proc. Natl. Acad. Sci. USA*, **88**, 7405–7409.
82. Johnsen,J.M., Fletcher,S.N., Huston,H., Roberge,S., Martin,B.K., Kircher,M., Josephson,N.C., Shendure,J., Ruuska,S., Koerper,M.A., *et al.* (2017) Novel approach to genetic analysis and results in 3000 hemophilia patients enrolled in the My Life, Our Future initiative. *Blood Adv.*, **1**, 824–834.
83. Lombardi,S., Leo,G., Merlin,S., Follenzi,A., McVey,J.H., Maestri,I., Bernardi,F., Pinotti,M. and Balestra,D. (2021) Dissection of pleiotropic effects of variants in and adjacent to F8 exon 19 and rescue of mRNA splicing and protein function. *Am. J. Hum. Genet.*, **108**, 1512–1525.

Received: March 31, 2023. **Revised:** October 16, 2023. **Editorial Decision:** October 18, 2023. **Accepted:** October 23, 2023

Chapter 3: OpenASO: RNA Rescue — designing splice-modulating antisense oligonucleotides through community science

Publication Status: Manuscript in Preparation. Ready for Submission.

Authors: **Victor Tse**[1, 3], Martin Guiterrez[1, 3], Jill Townley[6], Guillermo Chacaltana[2, 3], Eterna Players[6], Rhiju Das[4, 5, 6, 7], Jeremy R. Sanford[1, 3, *], and Michael D. Stone[2, 3, *]

1. Department of Molecular, Cell and Developmental Biology, University of California Santa Cruz, Santa Cruz, CA, 95064, USA.
2. Department of Chemistry and Biochemistry, University of California Santa Cruz, Santa Cruz, CA, 95064, USA.
3. Center for Molecular Biology of RNA, University of California Santa Cruz, Santa Cruz, CA, 95064, USA.
4. Department of Biochemistry, Stanford University, Stanford, CA 94305, USA
5. Department of Physics, Stanford University, Stanford, CA 94305, USA
6. Eterna Massive Open Laboratory
7. Howard Hughes Medical Institute, Stanford, CA 94305, USA

* To whom correspondence should be addressed: Michael Stone; Tel: 831-459-2845; Email: mds@ucsc.edu. Correspondence may also be addressed to: Jeremy Sanford; Tel: 831-459-1822; Email: jsanfor2@ucsc.edu

Abstract

Splice-modulating antisense oligonucleotides (ASOs) are precision RNA-based drugs that are becoming an established modality to treat human disease. Previously, we reported the discovery of a novel, putative RNA structure-function mechanism for ASOs to target to rescue splicing of multiple pathogenic variants of *F8* exon 16 that cause hemophilia A. However, the conventional approach to discovering splice-modulating ASOs is laborious and expensive. Here, we describe an alternative paradigm that integrates RNA structure data and community science to gamify the discovery of splice-modulating ASOs. Using a splicing-deficient pathogenic variant of *F8* exon 16 as a model in a new puzzle challenge we call OpenASO, we show that 25% of player-designed ASOs have a statistically significant impact on enhancing exon 16 splicing. Additionally, we show that a distinct combination of ASOs designed by Eterna players can additively enhance the inclusion of the splicing-deificent exon 16 variant. Together, our data suggests that crowdsourcing designs from a community of citizen scientists may accelerate the discovery of lead splice-modulating ASOs.

My Contributions:

I led the coordination of data sharing and outreach engagement efforts between UC Santa Cruz researchers and the Eterna community. I led a team to generate the heterologous *F8* exon-16 splicing reporters used in this study. I led the *in vitro* cell-based splicing assays used in this study. I developed, established, and led the application of the two-step RT-qPCR assay used in this study. I led efforts to assay all antisense oligonucleotides targeting *F8* exon-16 designed by the Eterna community presented in this study, and led efforts to assay them using the automation platform I developed. I performed and produced all of the data analysis and visualization presented in this study. I wrote the manuscript and produced all main and supplemental figures and legends presented, with input and suggestions by co-authors.

Introduction

Just as how puzzle pieces need to be placed in the correct position to create an intended picture, protein-coding sequences must be treated the same way to properly express a gene. Interspersed, non-protein-coding sequences (introns) must be removed, and protein-coding sequences (exons) must be joined together in a specific order to create a functional messenger RNA (mRNA); this process is known as precursor messenger RNA (pre-mRNA) splicing (Berget, Moore, and Sharp 1977; Chow et al. 1977; Weber, Jelinek, and Darnell 1977; Berk and Sharp 1977). Splicing occurs through two phosphoryl transesterification reactions that are catalyzed by the ribonucleoprotein known as the spliceosome (Padgett et al. 1984). The spliceosome is

a multi-megadalton complex composed of five essential uracil-rich small nuclear ribonucleoprotein particles (snRNPs), and hundreds of accessory RNA-binding proteins (RBPs), that assembles *de novo* on each intron in a stepwise manner (Wilkinson, Charenton, and Nagai 2020; Wan et al. 2020; Konarska and Sharp 1987). The formation of a catalytic spliceosome on a pre-mRNA substrate is only possible if the earliest conformation, known as the E complex, assembles on correct exon-intron boundaries (Michaud and Reed 1993; Plaschka et al. 2018).

Exon definition is a critical early step that facilitates E complex formation by demarcating exon-intron boundaries during pre-mRNA splicing (De Conti, Baralle, and Buratti 2013; Robberson, Cote, and Berget 1990). Conserved consensus sequence motifs within a pre-mRNA comprise the bare essence of exon definition. Motifs at the 5′ and 3′ end of an intron establish splice sites (5′ss and 3′ss) that define exon-intron boundaries (Mount et al. 1983; Zorio and Blumenthal 1999; Michaud and Reed 1993; M. S. Wong, Kinney, and Krainer 2018), in addition to a stretch of pyrimidines and a branchpoint sequence upstream of the 3′ss that is essential for E complex formation and subsequent remodeling to form the catalytic spliceosome (Berglund et al. 1997; Sheth et al. 2006). U1 snRNP binds to the 5′ss while U2 snRNP auxiliary factor (U2AF) binds to the 3′ss and polypyrimidine (poly-Y) tract. Auxiliary cis-acting splicing regulatory sequences residing within or adjacent to exonic sequences also contribute to the strength of these consensus motifs (Fu and Ares 2014; Izquierdo and Valcárcel 2006; Spingola et al. 1999; Qin et al. 2016; Ke et al. 2011; Reed and

Maniatis 1986; Fairbrother et al. 2002; Z. Wang et al. 2004; Berget 1995). Exonic and intronic splicing enhancers (ESEs and ISEs, respectively) and splicing silencers (ESSs and ISSs, respectively) are activated when they are bound by their cognate *trans*-acting RBP, inhibiting or promoting the assembly of spliceosomal components. Together, a dynamic and balanced interplay between splicing regulatory sequences and RBPs determines an exon's identity and governs its splicing fidelity .

Aberrant splicing is a hallmark of numerous human diseases, from rare genetic disorders to cancers (Faustino and Cooper 2003a; Cáceres and Kornblihtt 2002; Garcia-Blanco, Baraniak, and Lasda 2004b; R. K. Singh and Cooper 2012b; Sterne-Weiler et al. 2011b; Sterne-Weiler and Sanford 2014b; Lord and Baralle 2021b; G.-S. Wang and Cooper 2007b; E. Wang and Aifantis 2020b). Disease-causing mutations that perturb splicing regulatory sequences can induce aberrant splicing. It is known that at least 10% of these pathogenic mutations will typically ablate the 5′ or 3′ss (Krawczak et al. 2007b). Approximately, one-third of pathogenic mutations can perturb the *cis*-regulatory landscape of an exon, disrupting splicing enhancers or creating splicing silencers (Sterne-Weiler et al. 2011b; Kian Huat Lim et al. 2011; Fredericks et al. 2015b). The consequence of mutations can alter the reading frame of an mRNA transcript, leading to the synthesis of truncated, non-functional, or antagonistic proteins that can be deleterious for cells (Puisac et al. 2013; S. H. Kim et al. 2009; Fackenthal and Godley 2008). Thus, when an exon's identity is dysregulated, aberrant splicing of an exon can lead to molecular

phenotypes that cause human disease. However, an ongoing challenge for researchers is identifying where functional RNA elements reside within exons and flanking introns and what precise mechanisms are perturbed by mutations to induce aberrant splicing.

To better understand what triggers aberrant splicing, we previously used human genetic variation to identify *cis*-regulatory RNA elements and predict the impact mutations may have on them (Sterne-Weiler et al. 2011a; Sterne-Weiler and Sanford 2014a). Relative to single nucleotide polymorphisms, our previous work indicates that a large proportion of disease-causing mutations are more likely to perturb splicing regulatory sequences and are therefore more common in inducing aberrant splicing. Recently, we also demonstrated that a variety of hemophilia A (HA) causing mutations across multiple exons in the *F8* gene can readily induce their aberrant splicing (Tse et al. 2023a). Of the exons screened, we discover that *F8* exon 16 is particularly susceptible to aberrant splicing induced by a wide array of pathogenic mutations. Intriguingly, our experimental model indicates that an inhibitory RNA structure, which we term TWJ-3-15 (**t**hree-**w**ay **j**unction at the **3**' end of intron **15**), sequesters the poly-Y tract and harbors hnRNPA-1 dependent splicing silencers that weaken its 3′ss, sensitizing exon 16 to mutation-induced aberrant splicing. We showed that we could rescue splicing of multiple splicing-deficient pathogenic variants of exon 16 by targeting this novel, putative RNA structure-function mechanism with antisense oligonucleotides (ASOs). This work

added to the growing observation that some exons are indeed more fragile and sensitive to mutation-induced aberrant splicing (Holm et al. 2024; Glidden et al. 2021). Collectively, our work has suggested that RNA structures may have larger roles in exon identity, and that RNA structures may serve as therapeutic targets for splice-modulating drugs.

Splice-modulating ASOs are RNA-based drugs that have seen wide application and success over the last decades beyond the bench, as exemplified by Spinraza, Eteplirsen, and Milasen (M. A. Havens, Duelli, and Hastings 2013b; Corey 2017b; Mendell et al. 2013b; J. Kim et al. 2019). ASOs are therefore both a powerful tool and an emerging choice of drug modality to manipulate gene expression by exploiting the simple chemical logic of base pairing. Conventionally, splice-modulating ASOs are first identified following a brute-force approach known as an "ASO walk." An ASO walk is a preliminary screening process that involves the contiguous tiling of ASOs across a target sequence to uncover splicing regulatory sequences that *trans*-acting splicing factors bind to control exon identity (Yimin Hua et al. 2007, 2008). Typically, ASOs are designed as 18-mers with nucleotides that overlap with the preceding ASO to enable higher resolution screening of splicing regulatory sequences (Mallory A. Havens and Hastings 2016). As RNA-based compounds, the ribose sugar allows ASOs to support chemical modifications at the 2′-OH position which can confer greater stability and nuclease resistance to the oligomer *in vivo (Egli and Manoharan 2023)*. However, the costs required for an

ASO walk can rapidly increase as a result of the sequence coverage, the types of chemical modifications desired on the ASOs, and the molar yield desired.

Hundreds or even thousands of ASOs may need to be synthesized and screened before leads are identified. An underlying challenge from our previous study, and an overall challenge that is pervasive across studies from other groups developing splice-modulating ASOs, was the ability to identify promising ASOs in a timely and cost-effective manner. ASO walks often end up identifying only a handful of splice-modulating ASOs, and thus, these walks can be both a laborious and expensive project that is not economically tenable for basic research and ASO drug development. Together, it is clear that an alternative approach is needed to expedite the discovery of splice-modulating ASOs at a systematic scale.

Here, we evaluated whether community science might accelerate ASO drug discovery. Eterna is an open science platform that engages the collective intelligence and creativity of its players to solve puzzle game challenges related to RNA folding and design that current computer algorithms have trouble with. The goal of Eterna is to return experimental data from players' puzzle solutions to determine or correct models of RNA folding to better predict and manipulate the structure of RNAs *in vitro* and *in vivo*, potentially solving complex biology challenges to expedite the invention of medicine (J. Lee et al. 2014). Altogether since Eterna's conception a decade ago, the Eterna citizen science community has significantly contributed

towards refining RNA structure folding principles and experimental methods (Anderson-Lee et al. 2016). As an example of their commitment and the broader impacts of their work, Eterna players have recently made striking discoveries showing they can design novel mRNA vaccines that are more stable and more effective in protein production than designs based on conventional mRNA codon optimization (Wayment-Steele et al. 2021). Eterna, and other video games of the like (Cooper et al. 2010; Aneni et al. 2023; Wais et al. 2021; Anguera et al. 2013; Johannes, Vuorre, and Przybylski 2021; Kollins et al. 2020), have therefore demonstrated that gamification has the potential to advance the sciences, including drug discovery.

Using *F8* exon 16 and its pathogenic splicing-deficient variant as a model, c.5543A>G, we conceptualized the OpenASO:RNA Rescue challenge for Eterna players to design ASOs that can rescue exon $16^{c.5543A>G}$ splicing. Players generated and voted on the top 12 OpenASO designs to experimentally test. We show that 25% of the top ASOs voted by the Eterna players have a statistically significant impact on enhancing splicing of the exon $16^{c.5543A>G}$ variant. Subsequent experiments show that a distinct combination of the top scoring designs can significantly rescue splicing of exon $16^{c.5543A>G}$ to near wild-type levels. Together, we demonstrate an alternative paradigm for discovering splice-switching ASOs that may circumvent conventional ASO screening, and additionally validate our previous data indicating RNA structure may be a key target for precision medicines.

Results

Uniting RNA structure and citizen science can identify splice-modulating ASO designs

The central goal for the OpenASO:RNA Rescue challenge was to leverage RNA structure data and citizen science to accelerate the design and discovery of splice-modulating ASOs (Fig 1A). The OpenASO challenge involved creating a puzzle using our exon $16^{c.5543A>G}$ reporter's sequence context as a model system to design splice-modulating ASOs (Fig. 1B). Players were asked to design ASOs with a length of 18 nucleotides. An ASO design constraint was asking players not to directly interfere with key splicing signals, such as the 5′ss and 3′ss. The challenge also provided Eterna players with the essential information needed to understand the biology behind pre-mRNA splicing mechanisms, in addition to sharing SHAPE-guided RNA structure models (GEO Accession: GSE230495) and predicted splicing regulatory elements we previously acquired for exon 16 (Supplemental Fig. 1). Results from an ASO walk on exon $16^{c.5543A>G}$ involving 32 ASOs, conducted in parallel and later published in (Tse et al. 2023a), were kept blinded to players. To provide a simple game objective, players were tasked to design an ASO that could either refold the pathogenic exon $16^{c.5543A>G}$ variant to look more like the wild type (WT) context or alter the accessibility of any predicted splicing regulatory sequences to potentially enhance splicing.

300 hundred independent OpenASO designs were submitted by the Eterna community, with players engaging amongst each other to vote for the top 12 designs to test (Fig. 2A; Table 1). To determine the impact each OpenASO might have on enhancing exon $16^{c.5543A>G}$ splicing, we employed our established cell-based splicing reporter assay as previously described (Fig. 2B) (Tse et al. 2023a). We co-transfected each OpenASO with our splicing-deficient exon $16^{c.5543A>G}$ splicing reporter into HEK293T cells, extracted total RNA, and performed an end-labeled two-step RT-qPCR that quantifies the splicing reporter isoforms detected using fragment analysis via capillary electrophoresis. We observe that, relative to the non-target ASO condition (lane 1), three of the twelve designs, OpenASO 3 (lane 5), OpenASO 10 (lane 12), and OpenASO 11 (lane 13), have a statistically significant impact in enhancing splicing of exon $16^{c.5543A>G}$ (Fig. 3A-B). Together, our data demonstrates that 3/12 (25%) of ASOs designed by Eterna players were capable of significantly modulating splicing.

A combination of lead OpenASOs can additively enhance splicing of the pathogenic exon $16^{c.5543A>G}$ variant

Intriguingly, when analyzing where promising OpenASO designs hybridized to within the *F8* exon-16 sequence context, we discovered that they targeted TWJ-3-15 (Fig. 3C), an inhibitory RNA structure-function mechanism that we previously identified to sensitize exon-16 to mutation-induced aberrant splicing (Tse et al. 2023a). All 3 of the 12 OpenASOs that gave significant effects in enhancing

exon 16 inclusion, target this TWJ-3-15 RNA structure. OpenASO 3 solely targets ISS-15-1 with full complementarity. On the contrary, OpenASO 10 targets ISS-15-1 and ISS-15-2 simultaneously, though it only has partial complementarity to ISS-15-1, but full complementarity to ISS-15-2. Since there is general overlap in sequence composition between OpenASO 3 and 10, and because OpenASO 10 has higher coverage in occluding both ISS-15-1 and ISS-15-2 and performs better in enhancing exon 16 inclusion, we reasoned OpenASO 10 would be more effective for further experimentation. OpenASO 11, which was also significant in enhancing exon 16 inclusion, appears to strictly hybridize to the partner strand that occludes the poly-Y tract. Since these top-voted OpenASOs target TWJ-3-15, we hypothesized that a distinct combination comprising OpenASOs 10 and 11 would further enhance splicing of the pathogenic exon $16^{c.5543A>G}$ variant.

To compare the additive effectiveness of OpenASO designs, we repeated our cell-based splicing assays where our WT and exon $16^{c.5543A>G}$ splicing reporters were co-transfected with a combination of OpenASOs 10 and 11, or a non-targeting (NT) ASO control. To measure total rescue impact, we also included conditions where these reporters were co-transfected with our prior study's trio ASO combination that fully rescues exon 16 splicing. Relative to the pathogenic $16^{c.5543A>G}$ variant co-transfected with the NT ASO (lane 1), we observe that combining OpenASOs 10 and 11 had a statistically significant effect in additively enhancing exon 16 splicing (lane 4) (Fig. 4A-B), bringing the inclusion level of the exon $16^{c.5543A>G}$ variant to

levels near WT. However, in ascertaining total rescue impact, the OpenASO duo combination does not fully rescue splicing of the exon $16^{c.5543A>G}$ variant to WT levels (compare lanes 4 to 6), as observed from comparing the NT ASO to our trio ASO combination as previously discovered in *Tse et al.* (Fig. 4A-B; compare lanes 5 to 6). Nonetheless, our data collectively reinforces the hypothesis that RNA structures may be key therapeutic targets for splice-modulating ASO drugs, as demonstrated by the additive rescue effect elicited by the duo combination of OpenASOs 10 and 11 targeting TWJ-3-15. As a whole, we also show that gamification and crowdsourcing of citizen science can yield impactful splice-modulating ASO designs.

Discussion

An intriguing aspect of leveraging citizen science in ASO discovery is that we can potentially expedite the identification and subsequent optimization of promising ASO designs and sidestep conventional brute-force strategies. For instance, OpenASO 10, the most significant ASO from this challenge, is a promising ASO that individually performs almost as well as the lead ASO identified from our ASO walk in *Tse et al.*, based on comparisons of mean PSI values (compare highlighted rows between Table 2 and Table 3). The design premise of OpenASO 10 has similarities to the mechanism-of-action of our prior study's trio ASO combination, which elicits to achieve full rescue of exon 16 splicing across a wide array of pathogenic exon 16 variants (Tse et al. 2023a). OpenASO 10 was designed by a player with the intention to refold the pre-mRNA of exon 16 to enhance the accessibility of the poly-Y tract

while also simultaneously blocking two putative splicing silencers, a mechanism validated through our recent work done in parallel to the OpenASO challenge. Relative to conventional ASO walks, this discovery of a promising ASO through predictive modeling and community engagement indicates that crowdsourcing citizen science coupled to a data-driven strategy can yield promising leads efficiently.

Crowdsourcing citizen science also results in creative hypotheses and novel ASO designs. From our screen of the 12 top-voted OpenASO designs, we observe that 9/12 (75%) of the OpenASO designs by Eterna players are capable of modulating exon 16 splicing (Fig. 3A; Table 2). Although most of these designs did not significantly enhance splicing of *F8* exon 16, there were some designs whose proposed mechanisms were unusual and could prove to be effective, given more optimization and testing. For example, OpenASO 5 was designed to act as a "staple" to bring splicing regulatory sequences in closer proximity (Fig. 3A-B, lane 7). The premise of this OpenASO design was to open the poly-Y tract while subduing the accessibility of the silencers, while also bringing an enhancer closer to the poly-Y tract. Though OpenASO 5 did not have a statistically significant effect in enhancing exon 16 splicing, this idea of designing an ASO to physically remodel and stabilize an RNA's structure to optimally position splicing regulatory sequences is an intriguing hypothesis that would not be discovered in a standard ASO walk study and merits further investigation.

Our study also emphasizes that there are indeed limitations to ASO design strategies when contextual and positionally-dependent splicing regulatory mechanisms that control exon definition are not fully considered or understood. Although Eterna players discovered new ASO designs targeting *F8* exon 16 that can individually and additively enhance its splicing, we note that it is critical to consider how well an ASO sterically blocks the binding site for its cognate RBP to modulate canonical regulation of exon identity. For example, OpenASOs 10 and 11 as a combination do not fully rescue splicing of the pathogenic exon 16 variant to WT levels. We reiterate that OpenASO 10 targets ISS-15-1 and ISS-15-2 simultaneously, having only partial complementarity to ISS-15-1 yet full complementarity to ISS-15-2. Our prior study demonstrates that both ISS-15-1 and ISS-15-1 are hnRNPA1-dependent silencers that function to inhibit exon identity, and that masking both elements completely is required to augment rescue of exon 16 splicing (Tse et al. 2023a). Thus, because OpenASO 10 does not fully mask ISS-15-1, perhaps this may explain why we do not see a full rescue of exon 16 splicing. Additionally, in respect to this duo combination of OpenASOs relative to our trio ASO combination from our prior study, we observe that there is still some base pairing that may occlude the accessibility of the poly-Y tract (Fig. 4D), further explaining the inability of the duo combination of OpenASOs to fully rescue splicing.

5/12 of the top-voted OpenASOs to be tested by players also appear to strongly inhibit splicing of exon 16 (Fig. 3A; Table 2); these comprise OpenASO 2

(lane 4), OpenASO 6 (lane 8), OpenASO 7 (lane 9), OpenASO 8 (lane 10), and OpenASO 9 (lane 11). Although we do not know the precise *cis*- and *trans*- acting regulators that are perturbed by these ASOs, it is clear that these sequences play some degree of enhancing the definition of exon 16. Amalgamating various ASO designs that account for these factors may be favorable to conceive a singular ASO that can optimally modulate the accessibility of these splicing regulatory sequences to fully rescue inclusion of a splicing-deficient exon 16 variant. Moreover, incorporating better predictions of where splicing regulatory elements reside, as well as the experimental validation of these elements, may help inform design decisions for future OpenASO challenges, and in principle, the design of splice-modulating ASOs that elicits a full splicing rescue effect for any exon with dysregulated identity.

We demonstrate that Eterna can crowdsource the design of ASOs to modulate pre-mRNA splicing. Our results suggest that citizen science can independently design effective ASOs based on a simple game objective and RNA folding rules, with minimal or novice understanding of pre-mRNA splicing mechanisms. This approach based on RNA structure may sidestep the laborious and uneconomical ASO walks conventionally done to identify splice-modulating ASOs. We present here a proof-of-concept where gamification of ASO discovery may replace conventional brute-force approaches.

Materials and Methods

Wild-type (WT) and mutant F8 splicing reporters

Heterologous splicing reporters containing the sequence contexts corresponding to WT exon 16 or the pathogenic splicing-deficient variant, exon $16^{c.5543A>G}$, were previously generated and validated as described in *Tse et al. 2023*. The naming designation for the pathogenic variant presented in this study is based on the Human Genome Variation Society (HGVS) nomenclature. Splicing regulatory sequences predicted for *F8* exon 16 are also annotated in Supplemental Fig. 1, information that was shared with Eterna players.

Designing antisense oligonucleotides (ASOs) using citizen science

18-mer nucleic acid sequences complementary to *F8* exon-16 were designed by Eterna video game players. These ASOs will be henceforth referred to as "OpenASOs." The 12 OpenASOs with the highest player votes were synthesized by Integrated DNA Technologies (IDT) as 2′-methoxyethyl (2′MOE) phosphorothioate substituted oligonucleotides. The sequence of each OpenASO and relevant corresponding information can be found in Table 1.

Cell-based splicing assays to measure ASOs' ability to correct aberrant splicing

HEK293T cells (ATCC) were cultured in 6-well tissue culture plates (CytoOne, USA Scientific) using Dulbecco's Modified Eagle Medium (Gibco, supplemented with 10% FBS) at 37°C, 5% CO_2. The cells were transiently transfected with 1.25 μg of

WT or pathogenic variant of the *F8* exon-16 splicing reporter, and 10 ℘mol of each ASO using Lipofectamine 2000 (Invitrogen). Total RNA was then harvested from cells 24 h post-transfection using the Direct-zol RNA Miniprep kits (Zymo Research). To ensure rigor and reproducibility, each experiment type (e.g. single ASO or combinatorial ASO testing) was performed with a minimum of three independent/biological replicates.

Two-step RT-qPCR and analysis of splicing reporter assays

1.00 μg of purified total RNA was used as input for all first-strand cDNA synthesis using random primers and Multiscribe Reverse Transcriptase (Applied Biosystems). The resulting cDNA was then used as a template for endpoint PCR amplification using specific primers that detect our mRNA splicing reporter isoforms, Globin F: 5′-CGCAACCTCAAACAGACACC-3′; Globin R: 5′-AGCTTGTCACAGTGCAGCTC-3′. The forward primer of the pair contains a 5′-FAM modification. The resulting amplicons were then analyzed using agarose gel electrophoresis to empirically evaluate mRNA isoforms detected. Intron-spanning primers against SRSF3 mRNA were used as an endogenous internal control (SRSF3 F: 5′-GTAAGAGTGGAACTGTCGAATGG-3′; SRSF3 R: 5′-CGATCTCTCTCTTCTCCTATCTCTAG-3′). The abundance of each 5′-FAM labeled mRNA isoform was quantified using capillary electrophoresis and fragment analysis (UC Berkeley, DNA Sequencing Center). For fragment analysis, each sample was suspended in a formamide solution that contains a proper size standard for sizing

detected fragments (GeneScan 1200 Liz, Applied Biosystems). Analysis was performed in PeakScanner (Thermofisher). Quantification of splicing efficiency was achieved by comparing relative fluorescence units (RFU) between 5′-FAM labeled reporter isoforms that include or exclude an exon of interest. The RFU detected for each reporter isoform was then plugged into the following formula to calculate the PSI index, which reflects the splicing efficiency of an exon in either the WT or pathogenic variant context:

$$PSI = \frac{Included\ Isoform\ RFU}{Included\ Isoform\ RFU + Excluded\ Isoform\ RFU}$$

The mean PSI for a given reporter context is then calculated using all its respective replicates for a corresponding experiment. Statistical significance in the differences between the mean PSI of the control group(s) versus the experimental group(s) is determined using analysis of variance (ANOVA), and Dunett's post-hoc test. All statistical tests for PSI analysis were done in GraphPad Prism 9. Values are determined to be statistically significant if the calculated *P*-value is below an alpha value of ≤0.05.

Generation of RNA Structure Probing Data using SHAPE-MaP-seq Data

RNA structure data produced from SHAPE-MaP probing of *in vitro* transcribed *F8* exon 16 RNA, of both the WT context and the highly splicing-deficient exon $16^{c.5543A>G}$ variant, was previously generated and described in *Tse et al. 2023*.

Acknowledgements

We thank the Toxic RNA Lab (TRL), a curricular undergraduate research experience led by J.R.S., for their help in generating the *F8* exon 16 splicing reporter presented in this study. We thank Olena Vaske for participating in the video production that was made to introduce the objective and broader implications of the OpenASO: RNA Rescue challenge.

Author Contributions

V.T., J.T., R.D., J.R.S., and M.D.S. conceptualized and led the experiments presented in this study. V.T. , J.T., and G.C. coordinated the sharing of RNA structures and preliminary splicing data for *F8* exon 16 with the Eterna community. V.T. and J.T. led outreach engagement efforts between UC Santa Cruz researchers and the Eterna community. E.P. designed all antisense oligonucleotides targeting *F8* exon 16 presented in this study. V.T. and M.G. performed all cell-based splicing assays presented in this study. V.T. performed all of the data analysis and visualization presented in this study. V.T. wrote the manuscript, with input and review of successive manuscript drafts by co-authors.

Figure and Table Legends

Figure 1. Splice-modulating ASO discovery through community science. **(A)** A schematic depicting the OpenASO: RNA Rescue challenge for *F8* exon $16^{c.5543A>G}$.

The process entails inviting Eterna video game players to come up with an ASO design, submit them to the community database, and then vote to select the best designs to experimentally test. **(B)** A snapshot depicting the typical player environment in the Eterna video game. Specifically shown is the RNA structure profile of the splicing-deficient exon $16^{c.5543A>G}$ variant. Features of this exon and its flanking intronic sequences are annotated, and can be pinpointed and highlighted within the game environment. Nucleotides are depicted as the following colors: Yellow - A, Red - G, Green - C, and Blue - U. Players are tasked to design ASOs that target this structure, with the simple objective of remodeling the RNA to make certain features more accessible or inaccessible.

Figure 2. Experimentally testing the impact of top-voted ASO designs generated by Eterna players **(A)** The OpenASO: RNA Rescue challenge received 293 ASO designs targeting *F8* exon $16^{c.5543A>G}$. Players engaged and voted amongst themselves to select the top 12 OpenASOs to experimentally test. **(B)** A schematic depicting our established workflow to assay an ASOs ability to modulate pre-mRNA splicing using a cell-based splicing reporter system and two-step end-labeled RT-qPCR assay.

Figure 3. Top-voted OpenASOs can modulate the inclusion of the pathogenic splicing-deficient *F8* exon $16^{c.5543A>G}$ variant. **(A)** A representative agarose gel showing the effects OpenASOs have on the splicing of exon $16^{c.5543A>G}$. Controls include a non-targeting ASO control (lane 1) and a positive ASO control targeting the

5′ss to inhibit splicing (lane 2). Each of the top 12 voted OpenASOs tested (lane 3-14) is annotated as shown in the matrix above the gel. Expected mRNA reporter isoforms including or excluding exon 16 are also annotated to the left of the agarose gel. **(B)** A plot quantifying the OpenASOs' impact on exon $16^{c.5543A>G}$ splicing as shown in (A). Quantification of splicing efficiency is determined using the percent-spliced-in (PSI) ratio. PSI refers to the fraction of mRNA reporter isoforms that include the exon of interest, relative to the total population of mRNA reporter isoforms. Statistical significance between comparisons is denoted by asterisks that represent *P*-values with the following range of significance: * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $P \leq 0.0001$. Statistical significance was determined using analysis of variance (ANOVA), and Dunett's post-hoc test. Each OpenASO condition tested and presented contains three independent/biological replicates. (**C**) SHAPE-driven secondary structure prediction of TWJ-3-15 depicted in its two dimensional state for exon $16^{c.5543A>G}$. Core splicing signals are annotated within the structure. OpenASOs shown to be statistically significant in enhancing exon $16^{c.5543A>G}$ inclusion are annotated with a specific color (i.e., dark blue, green, and purple), and are depicted hybridizing to their complementary sequence within TWJ-3-15. The sequence is numbered according to the nucleotide positions of the heterologous splicing reporter, from the 5′ to 3′ orientation. The SHAPE-driven structure prediction is derived from *Tse et al. 2023*.

Figure 4. A duo combination of OpenASOs can additively enhance the inclusion of the pathogenic splicing-deficient exon $16^{c.5543A>G}$ variant. **(A)** A representative agarose gel showing the combinatorial effects OpenASO 10 and 11 has on the exon 16 splicing, relative to our study's prior trio ASO combination, in both the WT and exon $16^{c.5543A>G}$ sequence contexts. Each condition tested is annotated as shown in the matrix above the gel. Expected mRNA reporter isoforms including or excluding exon 16 are also annotated to the left of the agarose gel. **(B)** A plot quantifying ASOs impact on exon 16 splicing as shown in (A). Quantification of splicing efficiency is determined using the percent-spliced-in (PSI) ratio. PSI refers to the fraction of mRNA reporter isoforms that include the exon of interest, relative to the total population of mRNA reporter isoforms. Statistical significance between comparisons is denoted by asterisks that represent *P*-values with the following range of significance: * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $P \leq 0.0001$. Statistical significance was determined using analysis of variance (ANOVA), and Dunett's post-hoc test. Each ASO condition tested and presented for the WT context (annotated in blue) and the exon $16^{c.5543A>G}$ variant (annotated in red) contains a minimum of three independent/biological replicates. **(C)** and **(D)** depict SHAPE-driven secondary structure predictions of TWJ-3-15 for exon-$16^{c.5543A>G}$, where (C) shows where OpenASOs hybridize to and (D) show where our prior study's ASOs are hybridizing to within the structure. Core splicing signals are annotated within the structure. The sequence is numbered according to the nucleotide

positions of the heterologous splicing reporter, from the 5′ to 3′ orientation. The SHAPE-driven structure prediction is derived from *Tse et al. 2023*.

Table 1. Design descriptions and sequence compositions for the Top 12 OpenASOs voted by Eterna Players.

Table 2. Statistical results from OpenASO challenge.

Table 3. Statistical results from *Tse et al. 2023*.

Supplemental Figure 1. A schematic depicting the sequence context for *F8* exon 16 that has been studied in a splicing reporter system from *Tse et al. 2023*. Key splicing signals, such as: the branchpoint motif, the poly-Y tract and 3′ss, and 5′ss are indicated. Predicted splicing regulatory sequences are also annotated above the sequences by a red or green bar (represented to inhibit or enhance splicing, respectively).

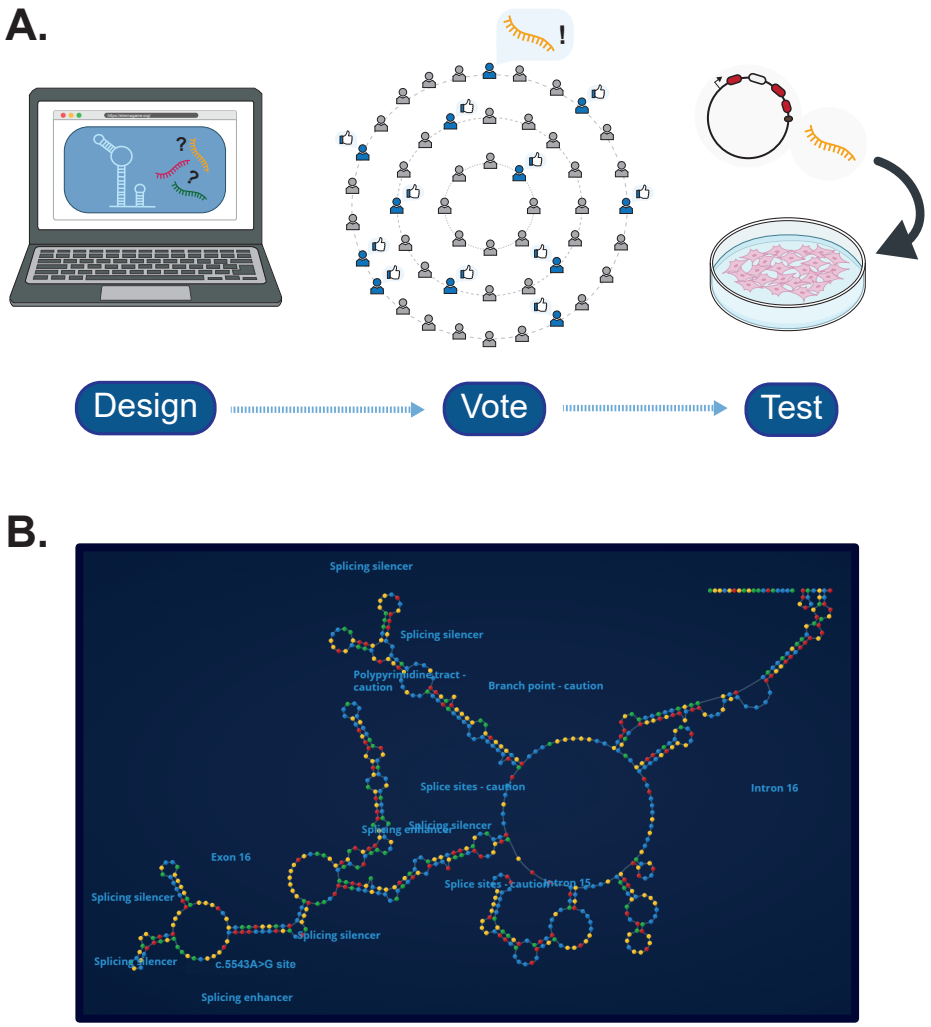


Figure 1. Gamifying splice-modulating ASO discovery. (A) A schematic depicting the OpenASO: RNA Rescue challenge for F8 exon 16c.5543A>G. The process entails inviting Eterna video game players to come up with an ASO design, submit them to the community database, and have players vote amongst themselves to select the best designs to experimentally test. (B) A snapshot depicting the typical player environment in the Eterna video game. Specifically shown is the RNA structure profile of the splicing-deficient exon 16c.5543A>G variant. Features of this exon and its flanking intronic sequences are annotated, and can be pinpointed and highlighted within the game environment. Nucleotides are depicted as the following colors: Yellow - A, Red - G, Green - C, and Blue - U. Players are tasked to design ASOs that target this structure, with the simple objective of remodeling the RNA to make certain features more accessible or inaccessible.

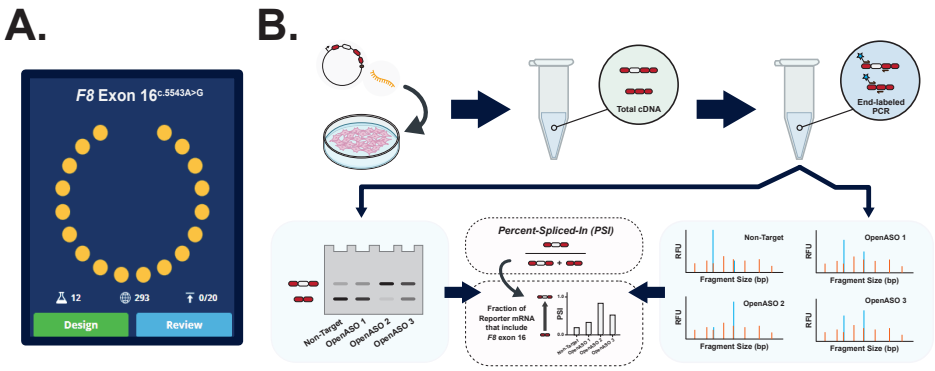


Figure 2. Experimentally testing the impact of top-voted ASO designs generated by Eterna players (A) The OpenASO: RNA Rescue challenge received 293 ASO designs targeting F8 exon 16c.5543A>G. Players engaged and voted amongst themselves to select the top 12 OpenASOs to experimentally test. (B) A schematic depicting our established workflow to assay an ASOs ability to modulate pre-mRNA splicing using a cell-based splicing reporter system and two-step end-labeled RT-qPCR assay.

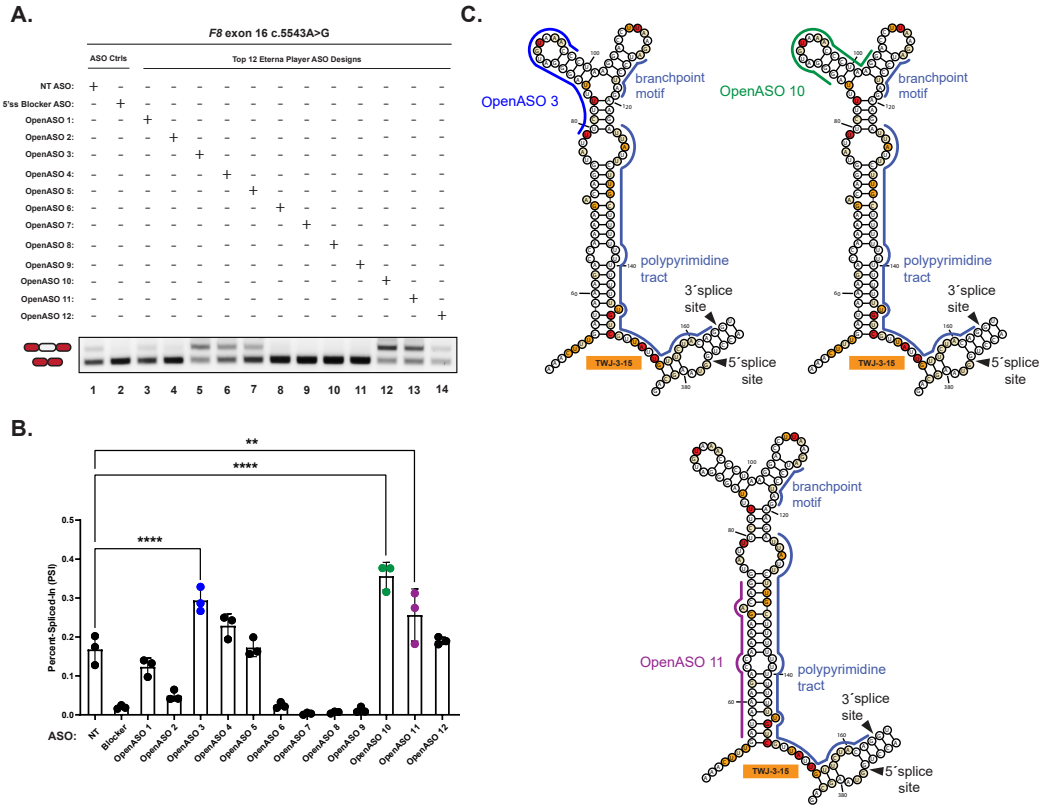


Figure 3. Top-voted OpenASOs can modulate the inclusion of the pathogenic splicing-deficient F8 exon 16c.5543A>G variant. (A) A representative agarose gel showing the effects OpenASOs have on the splicing of exon 16c.5543A>G. Controls include a non-targeting ASO control (lane 1) and a positive ASO control targeting the 5'ss to inhibit splicing (lane 2). Each of the top 12 voted OpenASOs tested (lane 3-14) is anno tated as shown in the matrix above the gel. Expected mRNA reporter isoforms including or excluding exon 16 are also annotated to the left of the agarose gel. (B) A plot quantifying the OpenASOs' impact on exon 16c.5543A>G splicing as shown in (A). Quantification of splicing efficiency is determined using the per cent-spliced-in (PSI) ratio. PSI refers to the fraction of mRNA reporter isoforms that include the exon of inter est, relative to the total population of mRNA reporter isoforms. Statistical significance between comparisons is denoted by asterisks that represent P-values with the following range of significance: * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = P ≤ 0.0001. Statistical significance was determined using analysis of variance (ANOVA), and Dunett's post-hoc test. Each OpenASO condition tested and presented contains three indepen dent/biological replicates. (C) SHAPE-driven secondary structure prediction of TWJ-3-15 depicted in its two dimensional state for exon 16c.5543A>G. Core splicing signals are annotated within the structure. OpenASOs shown to be statistically significant in enhancing exon 16c.5543A>G inclusion are annotated with a specific color (i.e., dark blue, green, and purple), and are depicted hybridizing to their complementary sequence within TWJ-3-15. The sequence is numbered according to the nucleotide positions of the heterologous splicing reporter, from the 5´ to 3´ orientation. The SHAPE-driven structure prediction is derived from Tse et al. 2023.

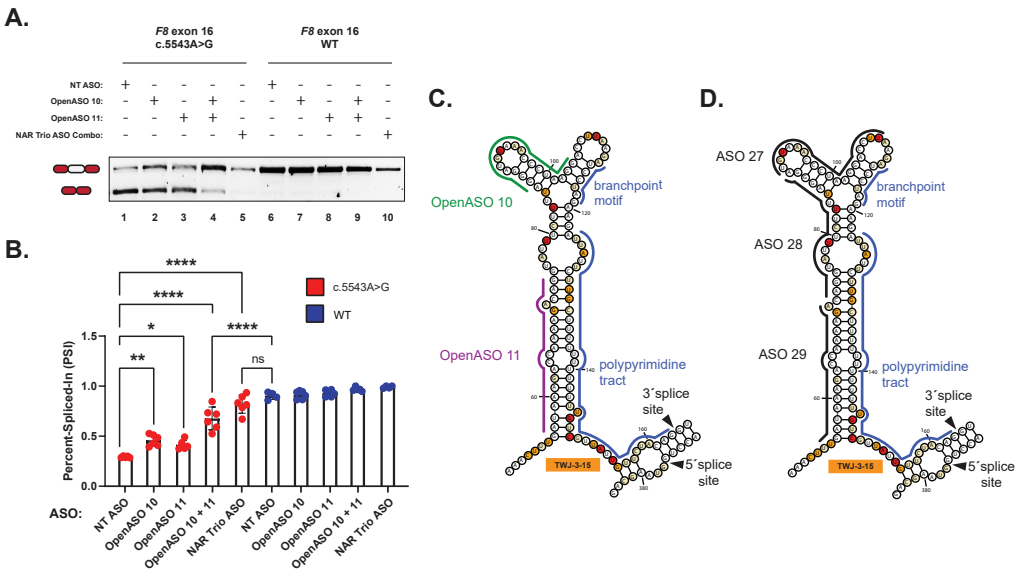


Figure 4. A duo combination of OpenASOs can additively enhance the inclusion of the pathogenic splicing-deficient exon 16c.5543A>G variant. (A) A representative agarose gel showing the combinatorial effects OpenASO 10 and 11 has on the exon 16 splicing, relative to our study's prior trio ASO combination, in both the WT and exon 16c.5543A>G sequence contexts. Each condition tested is annotated as shown in the matrix above the gel. Expected mRNA reporter isoforms including or excluding exon 16 are also annotated to the left of the agarose gel. (B) A plot quantifying ASOs impact on exon 16 splicing as shown in (A). Quantification of splicing efficiency is determined using the percent-spliced-in (PSI) ratio. PSI refers to the fraction of mRNA reporter isoforms that include the exon of interest, relative to the total population of mRNA reporter isoforms. Statistical significance between comparisons is denoted by asterisks that represent P-values with the following range of significance: * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = P ≤ 0.0001. Statistical significance was determined using analysis of variance (ANOVA), and Dunett's post-hoc test. Each ASO condition tested and presented for the WT context (annotated in blue) and the exon 16c.5543A>G variant (annotated in red) contains a minimum of three independent/biological replicates. (C) and (D) depict SHAPE-driven secondary structure predictions of TWJ-3-15 for exon-16c.5543A>G, where (C) shows where OpenASOs hybridize to and (D) show where our prior study's ASOs are hybridizing to within the structure. Core splicing signals are annotated within the structure. The sequence is numbered according to the nucleotide positions of the heterologous splicing reporter, from the 5´ to 3´ orientation. The SHAPE-driven structure prediction is derived from Tse et al. 2023.

Table 1. Design descriptions and sequence compositions for the Top 12 OpenASOs voted by Eterna Players.

Eterna ASO ID	Eterna ASO Name	Eterna ASO Description (positions based on Eterna seq-context)	Eterna Player	votes	UCSC ASO Name	ASO Sequence (reverese comp/antisense of *F8* CDS in 5' to 3' directionality)	"T" nucleobase substitution (for IDT)	IDT-formatted ASO Sequence
11557880	Astros-branchy-pioint-Pluss-P-tract-mod1	This design has the branch point in a multi-loop, the P-tract well opened, and Splice sites in a hairpin within a multiloop unpaired. The difference between this design and the original here https://eternagame.org/game/browse/11461132/?filter1=std&filter1_arg1=11549901&filter1_arg2=11549901 is this design's ASO sequence does not bind with itself in any current engine.	Astromon	9	**OpenASO1**	GGGUCUUAAGGUCUUUAG	GGGTCTTAAGGTCTTTAG	/52MOErG/*/i2MOErG/*/i2MOErG/*/i2MOEbT/*/i2MOErC/*/i2MOErT/*/i2MOErT/*/i2MOEbA/*/i2MOErA/*/i2MOErG/*/i2MOErG/*/i2MOErT/*/i2MOErC/*/i2MOErT/*/i2MOErT/*/i2MOErT/*/i2MOErA/*/32MOErG/
11567889	JL STS Eli Mod #NuPACK #Vienna2	Mod of Eli's Silencing the Silencer #NuPACK#Vienna2. Still works in both models. Added 7 GUs to weaken the oglio association to make unbonding easier. The oglio does not tend to fold with itself nor bond with itself in either Vienna2 nor EternaFold.	jandersonlee	7	**OpenASO2**	CAAUGUUAGGGAGGUGAG	CAATGTTAGGGAGGTGAG	/52MOErC/*/i2MOErA/*/i2MOErA/*/i2MOErT/*/i2MOErG/*/i2MOErT/*/i2MOErT/*/i2MOEbA/*/i2MOErG/*/i2MOErG/*/i2MOErG/*/i2MOErA/*/i2MOErG/*/i2MOErG/*/i2MOErT/*/i2MOErG/*/i2MOErA/*/32MOErG/
11604851	On loan 2	Leaves a fairly long section of PP tract unpaired (9 bases in V2 and 10 in NuPACK) and branch point is unpaired in both V2 and NuPACK. Sequence from WC OpenASO Matches sheet. Credit to Eli, jandersonlee, and Omei.	Merida	6	**OpenASO3**	GUUUACACUCCCUAAAAGA	GTTTACATCCCTAAAAGA	/52MOErG/*/i2MOErT/*/i2MOErT/*/i2MOEbT/*/i2MOErA/*/i2MOErC/*/i2MOErA/*/i2MOErT/*/i2MOErC/*/i2MOErC/*/i2MOErC/*/i2MOErT/*/i2MOErA/*/i2MOErA/*/i2MOErA/*/i2MOErA/*/i2MOErG/*/32MOErA/
11575019	SCREAMING silencer 1	Digital embrace asked me to glob the oligo onto both silencers around INTRON 15 This begins at 92 the tail end of the (86-92) silencer and ends at 109 fully covering it and going over two bases of the other silencer (101-107)	Jieux	6	**OpenASO4**	CCCUAAAAGAAAUACCUG	CCCTAAAAGAAATACCTG	/52MOErC/*/i2MOErC/*/i2MOErC/*/i2MOEbT/*/i2MOErA/*/i2MOErA/*/i2MOErA/*/i2MOErA/*/i2MOErG/*/i2MOErA/*/i2MOErA/*/i2MOErA/*/i2MOEbT/*/i2MOErA/*/i2MOErC/*/i2MOErC/*/i2MOErT/*/i2MOErG/*/32MOErG/
11552267	Staple 56-77	Finally was able to make a staple. Even works in v2.	DigitalEmbrace	6	**OpenASO5**	UCAAGUAAUACCCUGGA	TCAAGTAATACCCTGGA	/52MOEbT/*/i2MOErC/*/i2MOErC/*/i2MOErA/*/i2MOErA/*/i2MOErA/*/i2MOErG/*/i2MOErT/*/i2MOErA/*/i2MOErA/*/i2MOErT/*/i2MOErA/*/i2MOErA/*/i2MOErC/*/i2MOEbC/*/i2MOErT/*/i2MOErG/*/i2MOErG/*/32MOErA/
11548171	AK 6.1	NuPACK masks splicing silencer 341-346 and exposes the splicing enhancer 348-354 and mutation in a bulge	Andrewkae	6	**OpenASO6**	GCAGUCAAGUUAGUGGGU	GCAGTCAAGTTAGTGGGT	/52MOErG/*/i2MOErC/*/i2MOErA/*/i2MOErG/*/i2MOETf/*/i2MOErC/*/i2MOErA/*/i2MOErA/*/i2MOErG/*/i2MOETf/*/i2MOErT/*/i2MOErA/*/i2MOErG/*/i2MOEbT/*/i2MOErG/*/i2MOErG/*/i2MOErG/*/32MOETf/
11548080	submission 9	exposes enhancer around 190, also exposes both splice sites	hi1000	6	**OpenASO7**	AGGUCAACACAGAGAAA	AGGTCAACATCAGAGAAA	/52MOErA/*/i2MOErG/*/i2MOErG/*/i2MOErT/*/i2MOErC/*/i2MOErA/*/i2MOErA/*/i2MOErC/*/i2MOErA/*/i2MOETf/*/i2MOErC/*/i2MOErA/*/i2MOErG/*/i2MOErA/*/i2MOErG/*/i2MOErA/*/i2MOErA/*/i2MOErA/
11548066	whbob_i16_401 to 418	ASO bound 5' of i16 but left splice site exposed.	whbob	6	**OpenASO8**	AGCACAAUAGACACCUGC	AGCACAATAGACACCTGC	/52MOErA/*/i2MOErG/*/i2MOErC/*/i2MOErA/*/i2MOErC/*/i2MOErA/*/i2MOErA/*/i2MOETf/*/i2MOErA/*/i2MOErG/*/i2MOErA/*/i2MOErC/*/i2MOErA/*/i2MOErC/*/i2MOErC/*/i2MOETf/*/i2MOErG/*/32MOErC/
11547018	#Intermolecular pseudoknot	I believe a potential pseudoknot went missing due to the mutation. I'm trying to recreate it with my oligo. I bound my oligo in a split way where it binds	Eli Fisker	6	**OpenASO9**	AGCCCAGGCCUUUAGUGG	AGCCCAGGCCTTTAGTGG	/52MOErA/*/i2MOErG/*/i2MOErC/*/i2MOErC/*/i2MOErC/*/i2MOErA/*/i2MOErG/*/i2MOErG/*/i2MOErC/*/i2MOErC/*/i2MOErT/*/i2MOEbT/*/i2MOErT/*/i2MOErA/*/i2MOErG/*/i2MOErT/*/i2MOErG/*/32MOErG/
11544785	AK 1	partially masks 2 splicing silencers in intron 15 and opens part of the polypyrimidine tract	Andrewkae	6	**OpenASO10**	CCUUAGGGUUUACAUCCC	CCTTAGGGTTTACATCCC	/52MOEbC/*/i2MOErC/*/i2MOErT/*/i2MOEbT/*/i2MOErA/*/i2MOErG/*/i2MOEbG/*/i2MOErG/*/i2MOETf/*/i2MOErT/*/i2MOETf/*/i2MOErA/*/i2MOErC/*/i2MOErA/*/i2MOErT/*/i2MOEbC/*/i2MOErC/*/32MOErC/
11600061	Mar26 ASO design	Attached to Intron 15 at 77-94 in Vienna2 and Nupack, silencers in both engines mostly bound, Branch point free in Nupack but not Vienna2	mjt	5	**OpenASO11**	CUGUCUUUUGGUCUUUAU	CTGTCTTTTGGTCTTTAT	/52MOErC/*/i2MOETf/*/i2MOErG/*/i2MOErT/*/i2MOEbC/*/i2MOErT/*/i2MOETf/*/i2MOErT/*/i2MOEbT/*/i2MOErG/*/i2MOErG/*/i2MOETf/*/i2MOErC/*/i2MOETf/*/i2MOErT/*/i2MOEbT/*/i2MOErA/*/32MOETf/
11554562	volt-4 ASO stuck a Splicing Silencer	ASO stuck a Splicing Silencer (N:119-125) this seems to open the loop.	voltor	5	**OpenASO12**	AGGUCCUUAGGGUUUACA	AGGTCCTTAGGGTTTACA	/52MOErA/*/i2MOErG/*/i2MOErG/*/i2MOErT/*/i2MOErC/*/i2MOEbC/*/i2MOErT/*/i2MOETf/*/i2MOErA/*/i2MOErG/*/i2MOErG/*/i2MOEbG/*/i2MOErT/*/i2MOETf/*/i2MOErT/*/i2MOErA/*/i2MOEbC/*/32MOErA/

Table 2. Statistical Results from OpenASO Challenge.

Number of families	1
Number of comparisons per family	13
Alpha	0.05

Dunnett's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Below threshold?	Summary	Adjusted P Value	A-?	
NT vs. Blocker	0.1484	0.08140 to 0.2153	Yes	****	<0.0001	B	Blocker
NT vs OpenASO 1	0.04478	-0.02218 to 0.1117	No	ns	0.3438	C	OpenASO 1
NT vs OpenASO 2	0.1189	0.05196 to 0.1859	Yes	***	0.0002	D	OpenASO 2
NT vs OpenASO 3	-0.126	-0.1930 to -0.05906	Yes	****	<0.0001	E	OpenASO 3
NT vs OpenASO 4	-0.06088	-0.1278 to 0.006072	No	ns	0.0899	F	OpenASO 4
NT vs OpenASO 5	-0.004753	-0.07171 to 0.06220	No	ns	0.9997	G	OpenASO 5
NT vs OpenASO 6	0.1439	0.07699 to 0.2109	Yes	****	<0.0001	H	OpenASO 6
NT vs OpenASO 7	0.165	0.09809 to 0.2320	Yes	****	<0.0001	I	OpenASO 7
NT vs OpenASO 8	0.1617	0.09472 to 0.2286	Yes	****	<0.0001	J	OpenASO 8
NT vs OpenASO 9	0.1556	0.08864 to 0.2226	Yes	****	<0.0001	K	OpenASO 9
NT vs OpenASO 10	-0.1883	-0.2553 to -0.1214	Yes	****	<0.0001	L	OpenASO 10
NT vs OpenASO 11	-0.08827	-0.1552 to -0.02131	Yes	**	0.005	M	OpenASO 11
NT vs OpenASO 12	-0.02215	-0.08911 to 0.04480	No	ns	0.9553	N	OpenASO 12

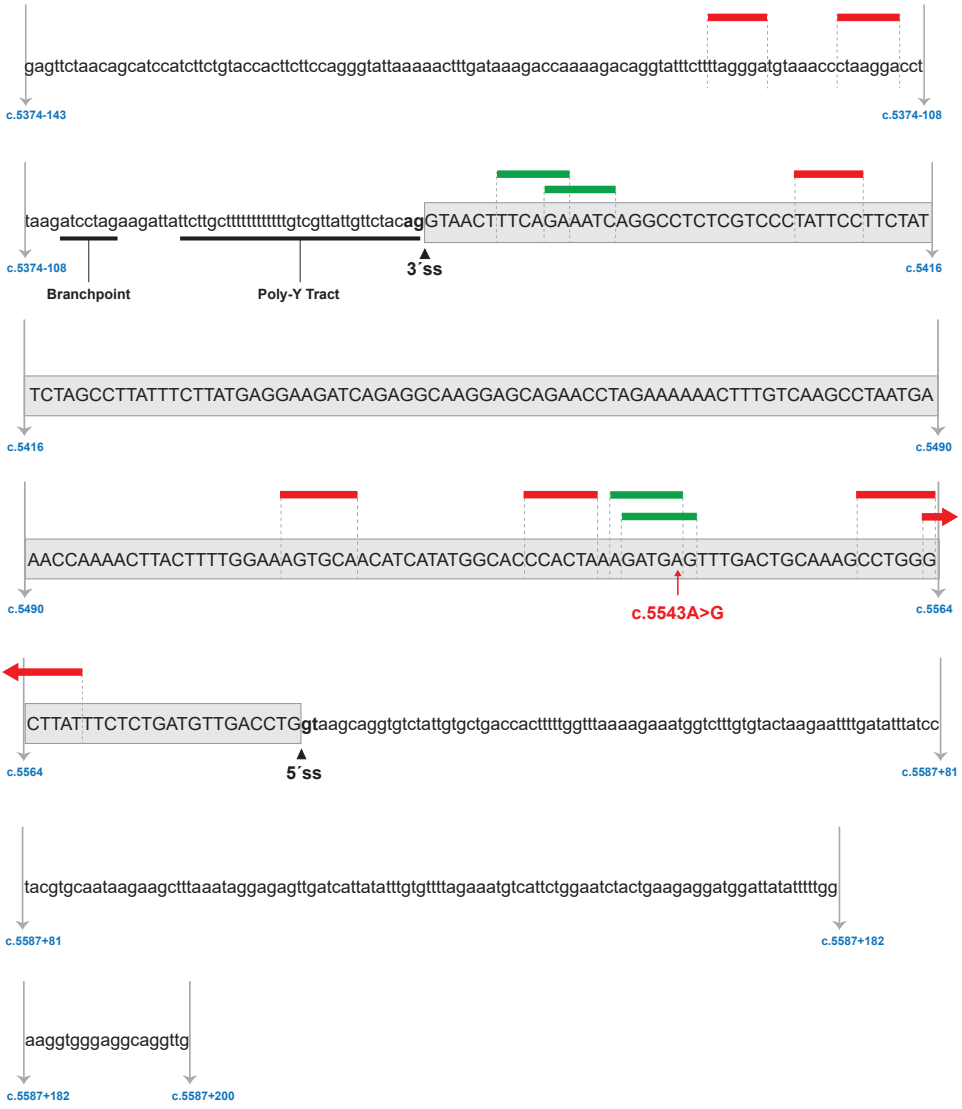
Test details	Mean 1	Mean 2	Mean Diff.	SE of diff.	n1	n2	q	DF
NT vs. Blocker	0.168	0.01961	0.1484	0.02241	3	3	6.62	28
NT vs OpenASO 1	0.168	0.1232	0.04478	0.02241	3	3	1.998	28
NT vs OpenASO 2	0.168	0.04905	0.1189	0.02241	3	3	5.306	28
NT vs OpenASO 3	0.168	0.294	-0.126	0.02241	3	3	5.623	28
NT vs OpenASO 4	0.168	0.2289	-0.06088	0.02241	3	3	2.717	28
NT vs OpenASO 5	0.168	0.1727	-0.004753	0.02241	3	3	0.2121	28
NT vs OpenASO 6	0.168	0.02402	0.1439	0.02241	3	3	6.423	28
NT vs OpenASO 7	0.168	0.002917	0.165	0.02241	3	3	7.365	28
NT vs OpenASO 8	0.168	0.006289	0.1617	0.02241	3	3	7.214	28
NT vs OpenASO 9	0.168	0.01237	0.1556	0.02241	3	3	6.943	28
NT vs OpenASO 10	0.168	0.3563	-0.1883	0.02241	3	3	8.403	28
NT vs OpenASO 11	0.168	0.2562	-0.08827	0.02241	3	3	3.939	28
NT vs OpenASO 12	0.168	0.1901	-0.02215	0.02241	3	3	0.9886	28

Table 3. Statistical Results from *Tse et al. 2023.*

Number of families	1
Number of comparisons per family	33
Alpha	0.05

Dunnett's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Below threshold?	Summary	Adjusted P Value	A-?	
NT vs. Blocker	0.2105	0.1735 to 0.2474	Yes	****	<0.0001	B	Blocker
NT vs. ASO 32	-0.09734	-0.1343 to -0.06042	Yes	****	<0.0001	C	ASO 32
NT vs. ASO 31	0.02045	-0.01647 to 0.05737	No	ns	0.7143	D	ASO 31
NT vs. ASO 30	-0.1859	-0.2228 to -0.1489	Yes	****	<0.0001	E	ASO 30
NT vs. ASO 29	-0.09389	-0.1308 to -0.05697	Yes	****	<0.0001	F	ASO 29
NT vs. ASO 28	0.01598	-0.02093 to 0.05290	No	ns	0.9425	G	ASO 28
NT vs. ASO 27	-0.08087	-0.1178 to -0.04396	Yes	****	<0.0001	H	ASO 27
NT vs. ASO 26	0.09605	0.05913 to 0.1330	Yes	****	<0.0001	I	ASO 26
NT vs. ASO 25	0.05407	0.01715 to 0.09098	Yes	***	0.0006	J	ASO 25
NT vs. ASO 24	0.2265	0.1896 to 0.2634	Yes	****	<0.0001	K	ASO 24
NT vs. ASO 23	0.226	0.1891 to 0.2629	Yes	****	<0.0001	L	ASO 23
NT vs. ASO 22	0.2245	0.1876 to 0.2614	Yes	****	<0.0001	M	ASO 22
NT vs. ASO 21	0.1238	0.08691 to 0.1607	Yes	****	<0.0001	N	ASO 21
NT vs. ASO 20	0.2243	0.1873 to 0.2612	Yes	****	<0.0001	O	ASO 20
NT vs. ASO 19	0.2272	0.1903 to 0.2641	Yes	****	<0.0001	P	ASO 19
NT vs. ASO 18	0.2266	0.1896 to 0.2635	Yes	****	<0.0001	Q	ASO 18
NT vs. ASO 17	0.2054	0.1684 to 0.2423	Yes	****	<0.0001	R	ASO 17
NT vs. ASO 16	0.2248	0.1879 to 0.2618	Yes	****	<0.0001	S	ASO 16
NT vs. ASO 15	0.2285	0.1915 to 0.2654	Yes	****	<0.0001	T	ASO 15
NT vs. ASO 14	0.2165	0.1796 to 0.2534	Yes	****	<0.0001	U	ASO 14
NT vs. ASO 13	0.205	0.1681 to 0.2419	Yes	****	<0.0001	V	ASO 13
NT vs. ASO 12	0.2298	0.1929 to 0.2667	Yes	****	<0.0001	W	ASO 12
NT vs. ASO 11	0.2273	0.1904 to 0.2642	Yes	****	<0.0001	X	ASO 11
NT vs. ASO 10	-0.02321	-0.06012 to 0.01371	No	ns	0.5334	Y	ASO 10
NT vs. ASO 9	-0.01609	-0.05301 to 0.02083	No	ns	0.9391	Z	ASO 9
NT vs. ASO 8	8.129e-005	-0.03683 to 0.03700	No	ns	>0.9999	AA	ASO 8
NT vs. ASO 7	0.05945	0.02253 to 0.09636	Yes	***	0.0001	AB	ASO 7
NT vs. ASO 6	-0.1254	-0.1623 to -0.08849	Yes	****	<0.0001	AC	ASO 6
NT vs. ASO 5	-0.0003113	-0.03723 to 0.03660	No	ns	>0.9999	AD	ASO 5
NT vs. ASO 4	-0.04613	-0.08305 to -0.009216	Yes	**	0.0053	AE	ASO 4
NT vs. ASO 3	-0.05301	-0.08993 to -0.01610	Yes	***	0.0008	AF	ASO 3
NT vs. ASO 2	0.005936	-0.03098 to 0.04285	No	ns	0.9993	AG	ASO 2
NT vs. ASO 1	0.04116	0.004247 to 0.07808	Yes	*	0.0188	AH	ASO 1

Test details	Mean 1	Mean 2	Mean Diff.	SE of diff.	n1	n2	q	DF
NT vs. Blocker	0.2298	0.01937	0.2105	0.01118	3	3	17.84	68
NT vs. ASO 32	0.2298	0.3272	-0.09734	0.01118	3	3	8.251	68
NT vs. ASO 31	0.2298	0.2094	0.02045	0.01118	3	3	1.733	68
NT vs. ASO 30	0.2298	0.4157	-0.1859	0.01118	3	3	15.75	68
NT vs. ASO 29	0.2298	0.3237	-0.09389	0.01118	3	3	7.958	68
NT vs. ASO 28	0.2298	0.2138	0.01598	0.01118	3	3	1.355	68
NT vs. ASO 27	0.2298	0.3107	-0.08087	0.01118	3	3	6.855	68
NT vs. ASO 26	0.2298	0.1338	0.09605	0.01118	3	3	8.141	68
NT vs. ASO 25	0.2298	0.1758	0.05407	0.01118	3	3	4.583	68
NT vs. ASO 24	0.2298	0.003298	0.2265	0.01118	3	3	19.2	68
NT vs. ASO 23	0.2298	0.003848	0.226	0.01118	3	3	19.15	68
NT vs. ASO 22	0.2298	0.005301	0.2245	0.01118	3	3	19.03	68
NT vs. ASO 21	0.2298	0.106	0.1238	0.01118	3	3	10.5	68
NT vs. ASO 20	0.2298	0.005566	0.2243	0.01118	3	3	19.01	68
NT vs. ASO 19	0.2298	0.002596	0.2272	0.01118	3	3	19.26	68
NT vs. ASO 18	0.2298	0.00326	0.2266	0.01118	3	3	19.2	68
NT vs. ASO 17	0.2298	0.02447	0.2054	0.01118	3	3	17.41	68
NT vs. ASO 16	0.2298	0.004986	0.2248	0.01118	3	3	19.06	68
NT vs. ASO 15	0.2298	0.00136	0.2285	0.01118	3	3	19.37	68
NT vs. ASO 14	0.2298	0.01334	0.2165	0.01118	3	3	18.35	68
NT vs. ASO 13	0.2298	0.02485	0.205	0.01118	3	3	17.37	68
NT vs. ASO 12	0.2298	0	0.2298	0.01118	3	3	19.48	68
NT vs. ASO 11	0.2298	0.0025	0.2273	0.01118	3	3	19.27	68
NT vs. ASO 10	0.2298	0.253	-0.02321	0.01118	3	3	1.967	68
NT vs. ASO 9	0.2298	0.2459	-0.01609	0.01118	3	3	1.364	68
NT vs. ASO 8	0.2298	0.2297	8.129e-005	0.01118	3	3	0.00689	68
NT vs. ASO 7	0.2298	0.1704	0.05945	0.01118	3	3	5.039	68
NT vs. ASO 6	0.2298	0.3552	-0.1254	0.01118	3	3	10.63	68
NT vs. ASO 5	0.2298	0.2301	-0.0003113	0.01118	3	3	0.02639	68
NT vs. ASO 4	0.2298	0.276	-0.04613	0.01118	3	3	3.91	68
NT vs. ASO 3	0.2298	0.2828	-0.05301	0.01118	3	3	4.494	68
NT vs. ASO 2	0.2298	0.2239	0.005936	0.01118	3	3	0.5031	68
NT vs. ASO 1	0.2298	0.1887	0.04116	0.01118	3	3	3.489	68



Supplemental Figure 1. A schematic depicting the sequence context for F8 exon 16 that has been studied in a splicing reporter system from Tse et al. 2023. Key splicing signals, such as: the branchpoint motif, the poly-Y tract and 3´ss, and 5´ss are indicated. Predicted splicing regulatory sequences are also annotated above the sequences by a red or green bar (represented to inhibit or enhance splicing, respectively).

Chapter 4: Conclusions and future directions

My research characterizing the splicing patterns of multiple exons across the *F8* gene contributes to the emerging understanding that some exons are fragile. This fragile nature of certain exons makes them more predisposed and sensitive to aberrant splicing induced by pathogenic mutations. However, the mechanism by which fragility is imparted onto an exon remains unclear. Recent research suggests that certain exons are fragile because they are enhancer-dependent (Holm et al. 2024). *Holm et al.* suggests that all exons are not created equal to have strong determinants of exon identity based on their extensive computational analysis on the impact of exonic mutations on splicing regulatory sequences. Data from *Holm et al.* support the notion that each exon's identity is intrinsically dependent on a central determinant of exon identity, such as enhancer dependency, and so when this feature is ablated by mutations, this makes the exon more likely to be aberrantly spliced. Additionally, this hypothesis has previously been alluded to in prior computational work describing "hotspot" exons that are more likely to accrue mutations in RBP binding sites, perturbing an exon's *cis*-regulatory landscape and increasing its likelihood to be aberrantly spliced (Glidden et al. 2021). However, one feature that these studies did not fully consider was the role of RNA structures in imparting fragility to an exon's identity.

My research suggests that RNA structures can have important roles in determining exon definition than previously appreciated. Firstly, we show that a bifunctional intronic RNA structure serves to weaken the 3′ss of *F8* exon 16 by occluding the poly-Y tract and in harboring binding sites for the repressive splicing factor hnRNP A1 (Tse et al. 2023b). My work coalesce with work from other labs that have shown similar mechanisms where intronic RNA structures can have a predominant inhibitory effect on an exon's identity and its splicing potential, as seen in: *Tau* exon 10 (Kar et al. 2011; Grover et al. 1999; Varani et al. 1999, 2000), *SMN2* exon 7 (Natalia N. Singh, Androphy, and Singh 2004; Garcia-Lopez et al. 2018), and *MALT1* exon 7 (Jones et al. 2022). Together, my thesis research on *F8* exon 16 indicates that intronic RNA structures that function to inhibit splicing fidelity can sensitize an exon to aberrant splicing when splicing enhancers are ablated by mutations, imparting a state of fragility to such exons that make them more likely to have dysfunctional *cis*-regulatory landscapes.

Secondly, I believe my work on *F8* exon 16 also sheds some light on how the folding of exonic sequences may impact exon definition (Tse et al. 2023b). Our ASO walks on both the WT and MT context of exon 16 shows that ASOs hybridizing to any exonic sequence significantly reduces its splicing efficiency to approximately ≤10% inclusion into mRNA. This data adds support to the hypothesis that exon 16 is naturally highly dependent on splicing enhancers, and is therefore fragile and sensitive to the inhibitory effects of silencing elements like that of our intronic RNA

structure when key enhancers are lost. Additionally, our SHAPE data shows that the WT and pathogenic sequence contexts of exon 16 are highly structured. Intriguingly, when taking into account both ASO and SHAPE data, there is a picture being painted where perhaps some exons require secondary structures for exon definition. Our data from exon 16 suggest that perhaps a highly structured exon is favorable for exon definition, at least in part by forming RNA structures that brings splice sites close to each other, and perhaps that a specific structural conformation for a fragile exon is optimal for enhancer function in exon definition. However, more work is needed to refine a model that directly connects exonic RNA structure to exon definition fidelity.

I believe my work has expanded our understanding of contextual and positional dependencies that may universally govern exon identity by implicating RNA structure as a determinant of exon definition. I believe *F8* exon 16 presents itself as a good model for future studies to better connect the relationship between RNA structure and canonical splicing mechanisms. Future experiments studying *F8* exon 16 using more sophisticated methods, such as fluorescent resonance energy transfer (FRET), can reveal intricate RNA folding dynamics and key structural conformations and sequences that may be imperative to establish structure-function relationships (Fairman, Lever, and Kenyon 2021). Based on the use of appropriate FRET dye pairs at various positions, such as the splice sites and exonic sequences that enhance splicing based on our splicing reporter assays, I believe this approach of study can enable us to better understand why destabilizing the exonic structure of *F8*

exon 16 apparently depletes its splicing, and perhaps improve our understanding of how enhancers may function. Thus, this work may reveal what properties of RNA structures may broadly control exon definition fidelity, particularly how RNA folding affects the function of splicing regulatory sequences and their accessibility to RBPs.

An outstanding question and therefore immediate future direction for my research is to understand what and where RBPs bind to on *F8* exon 16 and its flanking introns. An UV crosslinking approach using labeled *in vitro* RNA baits corresponding to the WT and MT contexts of *F8* exon 16, that are then incubated in nuclear extracts, RNase-treated to digest unbound regions of RNA, and immunoprecipitated out using validated antibodies targeting known splicing factors, can inform us what RBPs are key for the definition of *F8* exon 16 by comparing the readout between the WT and MT, as well as where these RBP binding sites are (Huang and Yu 2013; Ramanathan, Porter, and Khavari 2019). Defining these canonical *cis*- and *trans*- regulatory interactions, and linking them to RNA structural studies, will be paramount to better understand the mechanisms that may impart fragility to exons such as *F8* exon 16.

In conclusion, I believe my thesis expands our understanding of pre-mRNA splicing mechanisms. I feel optimistic that if we can pinpoint the underlying universal controls behind exon definition for any given exon, this can advance our

understanding of pre-mRNA splicing mechanisms in rare disease and enable the development of precision medicines to serve patients with unmet needs.

Appendix: A novel loss-of-function splice site variant of *DEGS1* exon 2 drives hypomyelinating leukodystrophy

Publication Status: Manuscript in Preparation.

Authors: Authorship list is incomplete. The following lists known contributing authors based on their institutional affiliation (authors listed do not represent any specific ordering):

UC Santa Cruz - Holly C. Beale, Allison Cheney, Katrina Learned, Olena M. Vaske, **Victor Tse**, Martin Gutierrez, Guillermo Chacaltana, Michael D. Stone, and Jeremy R. Sanford

UC San Francisco - Chansonette Badduke, Jessica V. Ziffle, Anne M. Slavotinek, Jessica Tenney, Alexander Fay, Ugur Hodoglugil, Yusuph Mavura, Joanna Y. Lee, Pierre M. Martin, Neil J. Risch, Julie D. Saba, Tiffany Yip, and Walter P. Devine

Abstract

Myelination of neurons is paramount for establishing cognitive and motor functions all complex life requires. *DEGS1* encodes for a sphingolipid desaturase that converts dihydroceramides into ceramides, a process that is critical for maintaining the myelin sheath that protects and impart sensory functions to neurons. Genetic

variants of *DEGS1* can cause hypomyelinating leukodystrophy (HL), a lethal neurodegenerative disease. Here, we investigate an unreported novel 5′ splice site variant of *DEGS1* exon 2 from two unrelated patients diagnosed with HL. Using patients' samples, whole exome sequencing identified this as an indel variant that converts the +4/+5 AG to TT, and RNA-seq shows less inclusion of the HL-linked exon 2 variant in spliced mRNA. Using splicing reporters, we demonstrate that this variant is sufficient to induce complete skipping of exon 2, an unusually large internal coding exon. RNA structure probing of the HL-linked variant revealed significant structural changes to the 5′ and 3′ splice site regions, suggesting that the variant drives unfavorable intramolecular base pairing interactions. Lipidomic analysis also indicates this variant is sufficient to inactivate *DEGS1* desaturase function. Additionally, we discover multiple regions that inhibit splicing of exon 2 using antisense oligonucleotides, exposing its dependence on splicing enhancers, which might explain our inability to successfully rescue splicing of the HL-linked variant. Our data suggests an aberrant splicing mechanism where the HL-linked variant has a weakened 5′ splice site, and adopts an unfavorable structural conformation due to its unusual size, making it inaccessible to spliceosomal components and splicing factors, and potential oligo-based drugs. Taken together, we show that aberrant splicing is an etiology of HL and that alternative therapeutic strategies require further investigation to rescue splicing of this novel variant.

My Contributions

I conceptualized and led this project, particularly on aspects directly related to splicing and indirectly on RNA structure. I designed and generated the heterologous *DEGS1* exon 2 splicing reporters used in this study. I led the *in vitro* cell-based splicing assays used in this study. I developed, established, and led the application of the two-step RT-qPCR assay used in this study. I developed and established the automation platform used in this study. I designed all antisense oligonucleotides targeting *DEGS1* exon-2, and led efforts to assay them using the automation platform I developed. I performed the *in silico* analysis to identify any binding sites for RNA-binding proteins. I performed and produced a majority of the data analysis specifically relating to the molecular biology and biochemistry perspective. I wrote the manuscript (as presented here in this appendix), and produced the main and supplemental figures using data generated by all authors. Co-authors provided text for describing methodologies and figure legends presented in this manuscript.

Introduction

Accurate gene expression in humans requires precursor messenger RNA (pre-mRNA) splicing. Splicing removes superfluous sequences (introns) from pre-mRNA and ligates protein-coding regions (exons) to create mRNA (Chow et al. 1977; Berget, Moore, and Sharp 1977). To catalyze this reaction, the spliceosome, a large ribonucleoprotein complex, must assemble *de novo* on each intron. The spliceosome is composed of five core uracil-rich small nuclear RNAs (U snRNAs) and hundreds of associated proteins (Konarska and Sharp 1987; Kastner et al. 2019).

A critical early step in the spliceosome assembly pathway for mammalian genes is the recognition of exon-intron boundaries, a process known as exon-definition (Berget 1995). These landmarks are defined by conserved GU and AG splice site sequences at the 5′ and 3′ ends of an intron (Breathnach et al. 1978). The assembly of the spliceosome is then initiated by the recruitment of U1 small nuclear ribonucleoprotein (snRNP) to the 5′ss, followed by the recruitment of heterodimeric splicing factor U2AF to the polypyrimidine (poly-Y) tract and 3′ss (Kohtz et al. 1994; Graveley, Hertel, and Maniatis 2001b; Ruskin, Zamore, and Green 1988; S. Wu et al. 1999). Collectively, these interactions form the early spliceosome (E complex), the stage of the spliceosome assembly pathway that commits an exon of a pre-mRNA to splicing.

Beyond the conserved splicing signals, auxiliary *cis*-regulatory elements and their cognate RNA-binding proteins (RBPs) can strongly influence spliceosome assembly (Reed and Maniatis 1986; Michaud and Reed 1993; Watakabe, Tanaka, and Shimura 1993; Z. Wang et al. 2004; Fairbrother et al. 2002; Ke et al. 2011). Splicing enhancers and silencers can be found within exons or introns, typically as short degenerate motifs that interact with specific residues of RBPs. Serine-arginine rich (SR) proteins and heterogeneous ribonucleoprotein particles (hnRNPs) are two classes of RBPs that are known to promote or inhibit the recruitment of spliceosomal components by binding exonic or intronic splicing enhancers (ESEs, ISEs) and exonic or intronic splicing silencers (ESSs, ISSs), respectively (J. Y. Wu and Maniatis 1993; Stark et al. 1998; Mayeda, Helfman, and Krainer 1993b; Caputi et al. 1999b; Busch

and Hertel 2012b). When these RBPs are bound to their respective *cis*-regulatory sequence, they can aid or antagonize spliceosome assembly (Zhu, Mayeda, and Krainer 2001; Mayeda and Krainer 1992b; Expert-Bezançon et al. 2004b). The consequence of these binding events can influence the efficiency of spliceosome assembly and resultantly splicing fidelity. As such, the assembly of the spliceosome is a highly dynamic and regulated process that must act precisely to identify exons from the sea of flanking introns.

Pathogenic variants can induce aberrant splicing, a phenomenon that contributes to the etiology of numerous diseases (Sterne-Weiler and Sanford 2014b; R. K. Singh and Cooper 2012b; Faustino and Cooper 2003b; Lord and Baralle 2021b; E. Wang and Aifantis 2020b; Anna and Monika 2018b). A primary mode by which variants disrupt splicing is by perturbing consensus splicing signals, such as the 5′ or 3′ splice sites. It is well established that about 10% of pathogenic variants that alter consensus donor or acceptor splice sites can lead to aberrant splicing (Krawczak et al. 2007b). Moreover, one-third of pathogenic variants are predicted to cause aberrant splicing by perturbing auxiliary *cis*-regulatory elements that are critical for splicing fidelity (Kian Huat Lim et al. 2011; Sterne-Weiler et al. 2011b). The loss of ESEs or gain of ESSs through variants can destabilize exon-definition, resulting in the spliceosomes inability to correctly identify exons and thereby lead to aberrant splicing (Cartegni et al. 2006b; Cartegni and Krainer 2002b; Kashima and Manley 2003b; Yimin Hua et al. 2008; Natalia N. Singh and Singh 2011). Recent research

also showed that pathogenic variants can directly stall spliceosome assembly, preventing the formation of the catalytic spliceosome complex (Soemedi et al. 2017). Together, it is clear that the dysregulation of splicing mechanisms by pathogenic variants can play underlying roles in promoting disease phenotypes. However, there remains a significant challenge in being able to predict and identify pathogenic variants that cause aberrant splicing, particularly in a clinical context.

Hypomyelinating leukodystrophy (HL) is a neurodegenerative white matter disease that can lead to severe deficits of myelin that can affect the function of neurons, impairing cognitive and motor skills (Köhler, Curiel, and Vanderver 2018; Parikh et al. 2015). Clinical research has previously demonstrated that loss of DEGS1 desaturase function is an etiology of HL (Pant et al. 2019). *DEGS1* is a three exon gene that encodes for a sphingolipid desaturase, an enzyme that catalyzes the conversion of sphingolipids such as dihydroceramides (DhCer) into ceramides (Cer) (Ternes et al. 2002). *DEGS1* exon 2 comprises a large majority of this gene's coding capacity. This conversion of saturated to unsaturated sphingolipids is an essential process that maintains the myelin that surrounds and protects the axons of neurons (Schmitt, Castelvetri, and Simons 2015). At the protein-level, genetic variants of *DEGS1* have been shown to disrupt the genetic code and thereby desaturase activity, causing HL (Pant et al. 2019). However, there is currently no peer-reviewed research that investigates the impact of pathogenic *DEGS1* variants to gene function at the RNA-level.

Here, we identify a homozygous non-canonical splice site variant of *DEGS1* exon 2 using whole exome sequencing on two unrelated probands who present symptoms consistent with HL. Based on current available information on ClinVar and other databases, this variant has not been previously identified to be pathogenic, or associated with HL, and is therefore novel. RNA-seq of proband samples indicates this HL-linked exon 2 variant primarily induces its exclusion from spliced *DEGS1* mRNA. Indeed, using a heterologous splicing reporter assay, we demonstrate this novel HL-linked variant is sufficient to induce complete exon 2 skipping. Additionally, RNA structure probing suggests this HL-linked variant reduces U1 snRNP and U2AF accessibility to splice sites. Lipidomic work shows this splicing-deficient HL-linked variant is sufficient to dysregulate desaturase function, classifying it as a loss-of-function (LOF) variant. Attempts to rescue splicing of this HL-linked variant using antisense oligonucleotides (ASOs) indicates that the splicing fidelity of *DEGS1* exon 2 is likely reliant on an enhancer-rich *cis*-regulatory landscape, which is consistent with previous research investigating unusually large internal exons. Together, our data shows that aberrant splicing induced by a novel LOF splice site variant of *DEGS1* exon 2 is an etiology of HL, and that the HL-linked exon 2 variant likely becomes inaccessible to essential splicing factors involved in exon-definition, and possibly ASOs, suggesting alternative strategies may need to be consider to develop a precision therapy.

Methods

Whole exome sequencing workflow and analysis

Exon regions were targeted in extracted genomic DNA from submitted probands and family members using the xGen Whole Exome Panel kit (Integrated DNA Technologies). Targeted regions were sequenced using the Illumina sequencing system with 100bp paired-end reads. The resulting DNA sequences were mapped to and analyzed in comparison with the published human genome (UCSC hg19 reference sequence). Sequence variants were filtered, ranked, and annotated using Opal Clinical (Fabric Genomics) and Moon Diploid, and compared to other provided family members to search for causes of Mendelian genetic disease (de novo, homozygous, compound heterozygous and inherited heterozygous disease-causing variants). Key word(s) and/or HPO terms used to filter and rank variants in proband 1 included: Severe global developmental delay (HP:0011344), Contractures (HP:0001371), Seizures (HP:0001250), Generalized hypotonia (HP:0001290), Cortical visual impairment (HP:0100704), Optic atrophy (HP:0000648). Key word(s) and/or HPO terms used to filter and rank variants in proband 2 included: Seizures (HP:0001250), Global developmental delay (HP:0001263), Hypertonia (HP:0001276), Microcephaly (HP:0000252), Spastic quadriparesis (HP:0001285), Cerebral palsy (HP:0100021), Facial dysmorphism (HP:0001999), Hirsutism (HP:0001007), Cerebellar atrophy (HP:0001272), White matter abnormalities (HP:0002500), Hypomyelination (HP:0003429). The targeted coding exons and splice junctions of known protein-coding RefSeq genes were assessed for the mean

depth of coverage and data quality threshold values. Mean depth of coverage refers to the sequencing mean read depth across the targeted region, defined as coding exons and splice junctions of the IDT xGen Whole Exome Panel capture reagent kit targeting protein-coding RefSeq genes. The quality threshold refers to the percentage of the defined target region where read depth was at least 10X coverage, which permits high quality exome variant base calling, annotation and evaluation. Average quality thresholds may range from >90-99% of the targeted region, indicating that only a small portion of the target region may not be covered with sufficient depth or quality to call variant positions confidently. Mean depth of coverage and quality thresholds for probands 1 and 2 were 94 reads and 102 reads, and 99.5% and 99.6%, respectively.

Proband 1 Mean Depth of coverage: 94 reads, Quality threshold: 99.5%
Proband 2 Mean Depth of coverage: 102 reads, Quality threshold: 99.6%

Identified variants were confirmed by Sanger sequencing with a BigDye Terminator 3.1 Cycle Sequencing Kit and Genetic Analyzer 3500 (ThermoFisher Scientific).

RNA sequencing workflow and analysis

Illumina RNA-Seq data was generated from RNA extracted from whole blood (Maxwell® RSC simplyRNA Blood Kit) from the proband and an unrelated individual with no pathological or VUS in DEGS1. hisat2:2.2.1 was used to align

reads to GRCh38 (with the index grch38_snp_tran.tar.gz downloaded from ftp://ftp.ccb.jhu.edu/pub/infphilo/hisat2/data/). Novel and reference transcripts were identified using Stringtie 2.1.6 with and without Gencode v38 as a guide. Additionally, Stringtie was run on aligned reads after duplicates were removed with samblaster 0.1.26. All outputs were merged to generate a GTF file with consistent identifiers containing all reference and de novo transcripts, and transcript quantification was repeated with Stringtie using the merged GTF as a guide. The fraction of DEGS1 expression accounted for by each isoform was calculated based on TPM of each isoform. We report isoform abundance using duplicate-free data; the abundances based on data inclusive of duplicate reads are similar.

Cell-based in vivo splicing assay

Inserts corresponding to wild-type (WT) *DEGS1* exon 2 and its HL-linked variant, in addition to 150 nucleotides upstream and 350 nucleotides downstream of the exon, were derived from genomic DNA (Promega) using PCR techniques and locus-specific primers (Supplementary Table 1). The HL-linked insert was fully assembled using site-directed mutagenesis and overlap extension PCR using primers as shown in Supplementary Table 1. Prepared inserts were then cloned into pACT7_SC14 (*HBB* minigene reporter as previously described (Rothrock et al. 2003)) using homology-directed cloning technology (In-Fusion HD Cloning kit from Takara Bio). Splicing reporters were then sequence-validated using Sanger sequencing to confirm successful cloning and identity of reporters.

HEK293T cells (ATCC) were cultured in 6-well tissue culture plates (CytoOne, USA Scientific) using Dulbecco's Modified Eagle Medium (Gibco, supplemented with 10% FBS) at 37°C, 5% CO2. The cells were transiently transfected at ~60–80% confluency with 2.5 μg of each *DEGS1* splicing reporter using Lipofectamine 2000 (Invitrogen). Total RNA was harvested from cells 24 h post-transfection using the Direct-zol RNA Miniprep kits (Zymo Research). Each splicing assay was performed with nine independent/biological replicates.

Two-step end-labeled RT-qPCR analysis of splicing assays

1.00 μg of purified total RNA was used as input for all first-strand cDNA synthesis using random primers and Multiscribe Reverse Transcriptase (Applied Biosystems). The resulting cDNA was then used as a template for endpoint PCR amplification using specific primers that detect our mRNA splicing reporter isoforms, Globin F: 5′-CGCAACCTCAAACAGACACC-3′; Globin R: 5′-AGCTTGTCACAGTGCAGCTC-3′. The forward primer of the pair contains a 5′-FAM modification. The resulting amplicons were then analyzed using agarose gel electrophoresis to empirically evaluate mRNA isoforms detected. Intron-spanning primers against SRSF3 mRNA were used as an endogenous internal control (SRSF3 F: 5′-GTAAGAGTGGAACTGTCGAATGG-3′; SRSF3 R: 5′-CGATCTCTCTCTTCTCCTATCTCTAG-3′). The abundance of each 5′-FAM labeled mRNA isoform is quantified using capillary electrophoresis and fragment

analysis (UC Berkeley, DNA Sequencing Center). For fragment analysis, each sample is suspended in a formamide solution that contains a proper size standard for sizing detected fragments (GeneScan 1200 Liz, Applied Biosystems). Analysis was performed in PeakScanner (Thermofisher). Quantification of splicing efficiency is achieved by comparing relative fluorescence units (RFU) between 5′-FAM labeled reporter isoforms that include or exclude an exon of interest. The RFU detected for each reporter isoform is then plugged into the following formula to calculate the PSI index, which reflects the splicing efficiency of an exon in either the WT or pathogenic variant context:

$$PSI = \frac{Included\ Isoform\ RFU}{Included\ Isoform\ RFU + Excluded\ Isoform\ RFU}$$

The mean PSI for a given reporter context is then calculated using all its respective replicates for a corresponding experiment. Statistical significance in the differences between the mean PSI of the control group(s) versus the experimental group(s) is determined using analysis of variance (ANOVA), and Dunett's post-hoc test. All statistical tests for PSI analysis were done in GraphPad Prism 9. Values are determined to be statistically significant if the calculated P-value is below an alpha value of ≤0.05.

SHAPE-MaP-seq RNA structure probing and analysis workflow

In vitro transcribed (IVT) RNA corresponding to the WT or the HL-linked variant sequence contexts of *DEGS1* exon 2, as cloned into the splicing reporter plasmid, were synthesized using a T7 RNA polymerase approach. IVT RNA was purified using standard agarose gel extraction followed by overnight ethanol precipitation.

1.2 pmol of *DEGS1* RNA in 65mM Na-HEPES pH 8.0 was heated to 95°C for 3 min and then allowed to slowly cool at 24c for 15 min, afterwhich 10mM MgCl2 was added to a final volume of 15µL and left at 24c for an additional 5 min. 1M 2A3 was added to a final concentration of 50mM and incubated at 37°C for a maximum of 2 min. The reaction was then quenched with DTT to a final concentration of 500mM at 24°C for 10 min. Modified and unmodified RNA samples were then purified using Zymogen RNA Clean & Concentrator-5 Kit. RNA(s) were then incubated at 94°C for 1 min and fragmented to a median size of 200nt using NEBNext Magnesium RNA Fragmentation Kit. Fragmented RNA(s) were purified via standard ethanol precipitation. Purified fragmented RNA(s) were supplemented with 2µL of 20µM random hexamers and 1µL of 10mM dNTPs then incubated at 70°C for 5 min and then immediately placed back on ice. Reverse transcription reactions were set up using SSII (200U) in 4µL of 5X RT buffer (250mM Tris-HCl pH 8.3, 375mM KCl), 2µL of 100mM DTT, 1µL of 120mM MnCl2, and 1µL SUPERase RNase inhibitor (10U), first incubated at 25°C for 10 min, then 42°C for 3 hours followed by heat inactivated at 75°C for 20 min. EDTA was then added to RT reactions to a final

concentration of 6mM and left to incubate at 24°C for 5min, followed by addition of MgCl2 to a final concentration of 6mM. Reverse transcription reactions were then used as input for the NEBNext® Ultra II Non-Directional RNA Second Strand Synthesis Module (New England Biolabs, cat. E6111L). Second strand synthesis was performed by incubating reactions at 16°C for 1 hour, as per manufacturer instructions, followed by cleanup using AMPure XP beads. DsDNA was used as input for the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (New England Biolabs, cat. E7645L), following manufacturer instructions. All sequencing was performed on an Illumina iSeq100 instrument using a paired-end 2 × 150 sequencing reagent cartridge and flow cell. All SHAPE-MaP-seq experiments were performed in duplicate.

All of the relevant data analysis was conducted using RNA Framework v2.6.9 (Incarnato et al. 2018). Reads from Illumina libraries were pre-processed and mapped using the rf-map module (parameters: -b2 -mp '–no-mixed –no-discordant' -bs), ensuring only paired-end mates with expected mate orientation were considered with Bowtie2. The mutational signal was obtained using the rf-count module (parameters: -m -pp -nd -ni), enabling mutation counts of reads produced from properly paired mates. Mutational signal was normalized relative to an unmodified control using parameters (-sm 3 -nm 1 -mu 0.05) and further normalized using the 2–8% normalization approach provided by RNA Framework (Siegfried et al. 2014).

Normalized reactivities were then supplied to RNAstructure to generate data-driven predicted structure models (Reuter and Mathews 2010).

Sphingolipid lipidomics analysis

Plasma sphingolipids were extracted as described in reference (Bielawski et al. 2006). Briefly, 100uL plasma samples were mixed with 1.9 mL phosphate buffered-saline (PBS, pH7.4) containing 100mM DTPA. Ten microliter (10μL) of internal standard mixture was added to each sample and the sample mixtures were briefly vortexed. Plasma extraction mixture containing isopropanol:ethyl acetate (15:85, v:v, 2mL) were added to each plasma sample and briefly vortexed, and centrifuged at 2500xg at 4°C for 10 minutes. The upper organic phase was transferred into a new glass tube. Concentrated formic acid (3μL) was added to the bottom layer and briefly vortexed. The bottom layer samples were re-extracted with the additional 2mL of extraction buffer followed by briefly vortex and centrifugation at 2500xg at 4°C for 10 minutes. The upper organic phase from the second extraction was combined with the first extracted samples and vortexed. Approximately 2mL aliquot of the combined samples was transferred into a new glass tube and dried down completely under a constant stream of nitrogen. Dried samples were stored at -20°C until the date for HPLC injection. On day of HPLC injection, dried samples were reconstituted with 200 μL of Mobile Phase B (1 mM ammounium formate in methanol with 0.2% formic acid), and 70μL was transferred into a HPLC vial with glass insert for injection.

Liquid Chromatography Mass Spectrometry (LC-MS/MS) was used to detect sphingolipids, and data processing methods were used as described in reference (Suh, J.H. et al., Inflamm. Bowel Dis. 24:1321-1334, 2018). Briefly, a 1290 ultra-high performance LC(UHPLC) system coupled to an Agilent 6495 Triple Quadrupole MS equipped with Agilent Jet Stream (AJS)-electrospray ionization interface was used. Sphingolipid metabolites were resolved on a Zorbax RRHD Eclipse Plus C18 column (2.1x50mm; 1.8 micron). Samples were eluted from the column using a binary gradient composed of Mobile Phase A (2 mM ammonium formate in 18mΩ water with 0.2% formic acid) and Mobile Phase B (1 mM ammonium formate in methanol with 0.2% formic acid) at a flow rate of 1mL/min. The instrument was operated by Mass Hunter Workstation software. Mass detection was used in the positive electrospray ionization (ESI+) mode with multiple reaction monitoring (MRM). Specific MRM transitions and optimized conditions used for sphingolipid analysis are listed in the Supplementary Data section. The general source settings in the positive ionization mode were as follows: gas temperature 200°C; gas flow, 14L/min; nebulizer 20psi; sheath gas temperature 250°C; sheath gas flow 11L/min; capillary voltage 3000V; and nozzle voltage 0V. The fragmentor voltage of 380V and a dwell time of 15ms were used for all mass transitions, and both Q1 and Q3 resolutions were set to nominal mass unit resolution. All samples were analyzed using triplicates. All data are expressed as the mean ± SD. Differences were examined for significance using the two-tailed Student's test (t test), with $p < 0.05$ as the cutoff for significance.

Antisense Oligonucleotide (ASO) Walks

DEGS1 exon 2 ASOs were designed by taking the reverse complement of the coding sequence, specifying sequences of k-mer length which are then annotated with desired modifications to the ribose sugar. To infer nuclease resistance and in vivo stability to ASOs, the 2′-OH contains a methoxyethyl modification (2′-MOE) and the phosphate backbone was modified to a phosphorothioate backbone. All *DEGS1* ASOs presented in this study were designed to be 18 nucleotides in length, and were synthesized by Integrated DNA Technologies (IDT). The first ten ASO (i.e., ASO 1-10) hybridize to the intronic sequences downstream the 5′ splice site of *DEGS1* exon 2, with 10 nucleotide overlaps between each one to enable high resolution screening. ASO 1 was designed to start tiling inwards towards the 5′ splice site of *DEGS1* exon 2, positioned just before the first detectable short interspersed nuclear element (SINE) as annotated in the UCSC Genome Browser (hg19). The remaining ASOs (i.e., ASO 11-58) were designed to be contiguous in order, with no overlaps between the proceeding and preceding ASO. All ASO sequences can be found in Supplementary Table 2.

Each ASO is co-transfected with either the WT or HL-linked variant splicing reporter, using the same methodology as described in our cell-based splicing assays. HEK293T cells (ATCC) were cultured in 96-well tissue culture plates (Perkin Elmer) using Dulbecco's Modified Eagle Medium (Gibco, supplemented with 10% FBS) at 37°C

with a CO2 level of 5%. Prior to the time of performing the assays, cells were grown to a cell confluency of ~60-80%. 250 ng of the WT or HL-linked splicing reporter were transiently co-transfected with 10 pmol of each *DEGS1* ASO into HEK293T cells using lipofection technology (Lipofectamine 2000). After 24-hours post transfection, cells were then harvested and prepared for total RNA purification using the Quick-DNA/RNA Viral MagBead kit from Zymo Research, in which this workflow has been automated on the Agilent Bravo. Each ASO walk for either the WT or HL-linked variant was performed with three independent/biological replicates.

Results

Next-generation sequencing reveals a novel LOF mutation in the 5′ splice site of *DEGS1* exon 2 from two unrelated HL patients

Two unrelated probands were diagnosed with HL and present phenotypes consistent with the lethal neurodegenerative disease. To identify molecular signatures that may underlie the cause of these HL patients' phenotypes, we employed next-generation sequencing approaches to identify any genetic variations unique to the patients. Using exome sequencing, we detected an intronic indel in proband-derived samples at the 5′ end of *DEGS1* intron 2 (chromosomal position chr1:224378025-224378026 on hg19; chr1:224190323-224190324 on hg38) (Fig. 1A). *DEGS1* is a three exon gene that predominantly produces a single canonical mRNA transcript (NM_003676.3). The indel creates a non-canonical 5′ splice site variant of *DEGS1* exon 2, c.825+4_825+5delAGinsTT (Fig. 1B). This variant was

detected in the heterozygous state for available parents, and was detected to be in the homozygous state for both of the two unrelated probands using Sanger sequencing (Fig. 1C). According to the American College of Medical Genetics and Genomics (ACMG) guidelines, this 5′ splice site variant would be classified as a variant of uncertain significance (VUS) as it is: (1) absent from population databases (e.g., gnomAD and dbSNP), (2) has not previously been associated with disease databases such as ClinVar, and (3) this variant is not unanimously predicted by *in silico* algorithms to have a deleterious effect on splicing (e.g., Alamut) (Supplemental Fig. 1). As such, we report this as a novel pathogenic variant that has not been documented or known to cause HL. In contrast to the majority of splice site inactivating mutations that ablate the consensus splice site donor or acceptor (Krawczak et al. 2007b), this mutation converts the +4/+5 AG to TT, but leaves the 5′ splice site GT intact. Nonetheless, exome sequencing indicates this variant possibly weakens the 5′ splice site of exon 2 in comparison to the general consensus motif for a strong 5′ splice site (M. S. Wong, Kinney, and Krainer 2018). Thus, this change to a critical splicing regulatory sequence might suggest this HL-linked variant dysregulates the identity of exon 2 to inhibit its inclusion in spliced mRNA, potentially resulting in non-functional *DEGS1* mRNA isoforms that may explain HL phenotypes.

To better understand the functional impact this pathogenic splice site variant may have in respect to the type of *DEGS1* mRNA transcripts produced, we performed

transcriptome-wide RNA-seq on a control patient and on proband-derived samples. Alignment and visualization of splice site junction reads from the control patient show high coverage for all three exons from *DEGS1* mRNA (Fig. 2A). In comparison to the control, the proband-derived sample shows high coverage of splice site junction reads for only exons 1 and 3, with considerably less coverage for exon 2 (Fig 2B). We observe that 89.62% of the *DEGS1* mRNA isoforms detected in the control patient correspond to the canonical mRNA transcript (NM_003676.3) (Fig. 2C). Additionally, 6.29% of isoforms detected in the control correspond to a transcript that is not present in reference gene models, and thus is considered novel. The remaining isoforms detected in the control patient appear to correspond to another reference transcript. On the contrary, we see a stark contrast in the types of isoforms detected in the proband relative to the control (Fig. 2D). We observe that 68.74% of *DEGS1* mRNA isoforms detected in the proband corresponds to a transcript that skips exon 2 (MSTRG.2800.2). We consider MSTRG.2800.2 a novel transcript since the splice junction characterizing MSTRG.2800.2, chr1:224183418-224192332, was not present in reference gene models, the control patient, or GTEX (Supplemental Fig. 2). The second and third most common isoforms that make up approximately 11.6-11.8% of transcripts detected in our proband, MSTRG.2800.3 and MSTRG.2800.4, are only present in the proband as well and are novel due to their absence from available data. Only 5.61% of the *DEGS1* mRNA isoforms detected in the proband correspond to the canonical NM_003676.3 transcript. Together, our RNA-seq data substantiates the hypothesis that this novel 5′ splice site likely induces aberrant splicing of *DEGS1*

exon 2 to produce a novel non-functional mRNA, MSTRG.2800.2, which consequently depletes DEGS1 desaturase activity to promote phenotypes associated with HL.

HL-linked 5′ splice site variant induces complete skipping of *DEGS1* exon 2 and drives refolding of RNA secondary structures

The base pairing strength of a 5′ splice site to the spliceosomal RNA component of U1 snRNP dictates the formation of the E complex, and therefore, an exon's splicing potential. *In silico* tools such as MaxEntScan score how well any given 5′ splice site of an exon matches to the general consensus motif, where an an ideal 5′ splice site has a score of 11.81 or higher (Yeo and Burge 2004; Eng et al. 2004). MaxEntScan analysis of *DEGS1* exon 2 in its WT context suggests that its 5′ splice site is relatively strong (Table 1; MaxEntScan::score5ss: 10.65). In contrast, MaxEntScan indicates that the HL-linked variant substantially weakens the 5′ splice site of exon 2 (Table 1; MaxEntScan::score5ss: 3.5). This computational prediction compliments our RNA-seq data from proband samples, supporting the hypothesis that the HL-variant likely induces *DEGS1* exon 2 skipping through the weakening of 5′ splice site strength.

To determine if the HL-linked 5′ splice site variant is sufficient to induce skipping of *DEGS1* exon 2, we generated heterologous splicing reporters examining the wild-type (WT) and HL-linked sequence-contexts of *DEGS1* exon 2 using a

cell-based splicing assay (Fig. 3A). After transfecting these reporters into HEK293T cells, the levels of mRNA reporter isoforms detected were then quantified by a two-step end-labeled RT-qPCR assay coupled to fragment analysis via capillary electrophoresis to measure *DEGS1* exon 2 inclusion. Relative to the WT (lanes 1-3), the HL-linked variant is sufficient to significantly cause total skipping of *DEGS1* exon 2 from the splicing reporter (lanes 6-8) (Fig. 3B-C). Our experimental results showing *DEGS1* exon 2 skipping recapitulates and validates MSTRG.2800.2 as the predominant transcript variant detected from RNA-seq of proband-derived samples (Fig. 2B, D). Together, we demonstrate that this novel HL-linked 5′ splice site variant is sufficient by itself to cause significant exon skipping of *DEGS1* exon 2, indicating that aberrant splicing is an etiology of HL.

DEGS1 exon 2 is an unusually large exon with potential to form local and long-range RNA secondary structures. To test the hypothesis that this HL-linked 5′ splice site variant can affect the RNA folding landscape of *DEGS1* exon 2, we used SHAPE-MaP-seq (**s**elective 2'-**h**ydroxyl **a**cylation analyzed by **p**rimer **e**xtension and **m**uta**t**ional **p**rofiling coupled to high-throughput **seq**uencing) to chemically probe the accessibility of the WT and HL-linked pre-mRNA context corresponding to *DEGS1* exon 2 (Siegfried et al. 2014). SHAPE probing reveals striking differences between the RNA structure profiles of the WT and HL-linked variants (Fig. 3D, compare blue to red, respectively; Supplemental Fig. 3; Supplemental Fig. 4). In addition to further weakening of the 5′ splice site by the +4/+5 AG>TT through its sequestration in a

long-range structure, we now also see that the 3′ splice site in the HL-linked context is now structured, instead of being in an apical loop as seen in the WT (Supplemental Fig. 5; take note of the boxed features). Our RNA folding models suggests this splice site variant rearranges the RNA structure profile of the HL-linked context, weakening the base pairing potential between U1 snRNA and the 5′ splice site, and U2AF heterodimer from potentially accessing the 3′ splice site.

HL-linked 5′ splice site variant is sufficient to result in the LOF of desaturase function.

Having identified the genetic variation that underlies the phenotypes presented in these HL patients, we next wanted to validate that this novel splice site variant indeed inactivates *DEGS1* desaturase function. We performed liquid chromatography-mass spectrometry (LC-MS) to detect the type of sphingolipids present in the HL patients relative to a control using patient plasma. As *DEGS1* encodes for a sphingolipid desaturase that converts DhCer into Cer, we aim to measure the ratio between DhCer:Cer to infer desaturase activity. Our lipidomic analysis shows that, relative to the control patient, the HL probands have higher ratios of DhCer:Cer, reflecting increased accumulation of DhCer which indicates the lack of desaturase activity (Fig. 4A). This phenotype of high DhCer accumulation and DhCer:Cer ratios due to the loss of *DEGS1* is consistent with previous studies showing similar results in *DEGS1* deficient mice (Holland et al. 2007). Together, our genomics, molecular RNA biology, and lipidomics analysis indicates that this

HL-linked 5′ splice site variant of *DEGS1* exon 2, which results in its aberrant splicing by the mode of exon skipping, underlies basis behind the loss of *DEGS1* desaturase function and how HL manifests in patients with this pathogenic variant (Fig. 4B).

Antisense Oligonucleotide (ASO) walks reveal the enhancer-dependence of *DEGS1* exon 2 which suggests further drug strategization is required

To see if we can identify splicing regulatory sequences that we can target to rescue splicing of the HL-linked exon 2 variant, we employed ASO walks on the WT and HL-linked contexts. ASOs are a proven drug modality that can ameliorate disease phenotypes by modulating splicing outcomes to elicit a therapeutic effect (M. A. Havens, Duelli, and Hastings 2013c; Rigo, Seth, and Bennett 2014; Y. Hua et al. 2010b; Corey 2017b; J. Kim et al. 2019; Mendell et al. 2013b). We designed a large set of 18-mer ASOs that tiled across *DEGS1* exon 2 and its upstream and flanking introns (Fig. 5A). We also designed a non-target ASO with no specificity for *DEGS1* exon 2 and demonstrated that, relative to the conditions with no ASOs present (lane 1 and 3), there are no significant off-target effects or difference on splicing with an appropriately designed ASO co-transfected with either the WT or HL-linked exon 2 variant (lane 2 and 4) (Fig. 5B-C). Our ASO walk on the WT context revealed a diverse array of ASOs that effectively modulated splicing of exon 2 (Fig. 5D). Relative to the control, we discover a striking abundance of ASOs that have a statistically significant effect in inhibiting splicing of *DEGS1* exon 2, indicating it is

highly dependent on splicing enhancers for exon identity and splicing fidelity. This observation is consistent with previous findings that large internal exons contain an abundance of ESEs that are required for their splicing (Bolisetty and Beemon 2012). Additionally, we identify ASO 1 and ASO 10 to have a statistically significant impact on enhancing exon 2 splicing in the WT context.

In respect to the ASO walk on the HL-linked variant, we unfortunately did not identify any ASOs that can enhance or rescue its splicing (Fig. 5E). That is, we observe that none of the same ASOs employed in the WT ASO walk had any splice-modulating effects on the HL-linked exon 2 variant. This peculiar result may be explained by the SHAPE-MaP-seq experiments that revealed a long-range intramolecular interaction which essentially folds HL-linked exon 2 on itself, possibly occluding the accessibility of splicing regulatory sequences. Moreover, our WT ASO walk on *DEGS1* exon 2 indicates that a majority of its exonic sequence context contains splicing enhancers that this unusually large exon appears highly dependent on (Fig. 5D). As such, we hypothesize that masking any of these regions will likely not elicit any rescue of splicing for the HL-linked exon 2 variant. Additionally, due to the nature of the 5′ splice site variant exhibiting weakened base pairing potential to U1 snRNP, this suggests that a stable spliceosomal E complex is unlikely to form on *DEGS1* exon 2 to catalyze its splicing into mRNA. Our data support the hypothesis that oligo-based drugs require further design strategies and optimization, or that

alternative drug modalities need to be tested, to determine the best therapeutic intervention to rescue splicing of the HL-linked variant.

Discussion

We report that a novel LOF 5′ splice site variant of *DEGS1* exon 2 is sufficient to cause its aberrant splicing and diminish DEGS1 desaturase function in two unrelated HL patients. Being the bulk of the genetic information in this three exon gene, it is striking to see that *DEGS1* exon 2 is a 743bp long internal exon. Thus, it is unsurprising to surmise that skipping of *DEGS1* exon 2 can result in disease phenotypes associated with HL since a majority of the protein sequence is missing. A pitfall and limitation of our presented model connecting aberrant splicing to phenotypes observed in our probands is our use of a splicing reporter system. Despite how this minimal system exploits the context dependence of exon definition, as most auxiliary *cis*-regulatory elements required for splicing fidelity are near splice sites (Fu and Ares 2014), our reporter's sequence context may not capture the full endogenous environment *DEGS1* exon 2 may require for its splicing regulation. However, what is clear is that loss of *DEGS1* exon 2 is deleterious and can result in phenotypes consistent with the loss of desaturase activity and symptoms of HL. A splice-blocking morpholino oligonucleotide targeting the 5′ splice site of *DEGS1* exon 2 was used to induce its aberrant splicing in zebrafish, and clearly demonstrates the loss of *DEGS1* exon 2 is sufficient to lose desaturase function and impair locomotion (Pant et al. 2019). As such, a key future direction is to create mice models or patient-derived

disease-in-a-dish models that contain this novel LOF 5′ splice site variant of *DEGS1* exon 2 to recapitulate our findings presented in this study, and use it as a model system to better connect aberrant splicing to HL phenotypes.

How large exons are defined for splicing remains unclear, especially for *DEGS1* exon 2, as there is minimal peer-reviewed research that deciphers any splicing regulation that takes place in *DEGS1*. This large size for an exon is highly unusual in respect to gene architecture, where internal exons are typically 200bp long in mammalian genes (Movassat et al. 2019). Prior research illuminates a mechanistic model where large internal exons are heavily dependent on splicing enhancers (Bolisetty and Beemon 2012). Using RBPmap (Paz et al. 2014), an *in silico* tool that predicts of RNA-protein interactions, we observe that there are predicted binding sites in *DEGS1* exon 2 for SR proteins such as SRSF5 and SRSF7 (Supplemental Fig. 6), splicing factors that are known to bind splicing enhancers. Our data shows that, when an ASO masks these putative binding sites in the WT context, we strongly inhibit *DEGS1* exon 2 splicing. Additionally, connecting the RBPmap analysis to our SHAPE-driven model shows that these putative RBP binding sites are residing within or adjacent to structural bulges. This observation follows previously published models that indicate RNA structures with bulges serve as a recognition element that appear important for RBP binding specificity (H. N. Wu and Uhlenbeck 2002; Hermann and Patel 2000; Fukunaga et al. 2014), as shown for Tat protein binding to HIV-I TAR RNA for example (Roy et al. 1990). More intriguingly, however, is when

we consider these RNA-protein interactions in the context of how our ASOs might be perturbing the *cis*-regulatory landscape of *DEGS1* exon 2.

Our results from the ASO walk on the HL-linked variant (Fig. 5E), coupled to our RBPmap analysis (Supplemental Fig. 6), coalesce with the notion that we cannot rescue its splicing because our ASOs are masking splicing enhancers required for splicing of *DEGS1* exon 2. Computational analyses and experimental validation on a candidate target from *Bolisetty and Beemon* show that large internal exons have a high density of "C"-rich motifs that appear to function as splicing enhancers (Bolisetty and Beemon 2012). *DEGS1* exon 2 appears "C"-rich, and our experimental data from our ASO walk on the WT context clearly demonstrates that there is a high density of splicing enhancers required for exon 2 splicing (Fig. 5C). The primary mechanism by which ASOs work to ameliorate disease phenotypes connected to aberrant splicing is by targeting *cis*-regulatory elements to modulate the production of specific mRNA transcripts, and thereby the abundance of respective proteins. Thus, a typical strategy would be to design an ASO that masks a splicing silencer and block the silencing splicing factor to enhance an exon's definition, increasing it's likelihood to be recognized for splicing; such a strategy is exemplified by the FDA-approved ASO drug Spinraza (Yimin Hua et al. 2008). Masking splicing enhancers with ASOs is therefore a counterintuitive strategy that will induce more inhibition of splicing, rather than enhancing it.

Clearly, we demonstrate that aberrant splicing induced by this novel LOF 5′ splice site variant of *DEGS1* exon 2 drives HL phenotypes in these two unrelated patients. Our current model and the existing literature indicates that additional work needs to be done to totally decipher the splicing regulation *DEGS1* exon 2, an unusually large exon, in order to fine-tune the design of a viable splice-modulating ASO for HL. Key criterias to further examine are the: site(s) of ASO hybridization, optimal oligo length and composition, specific chemical modifications, or perhaps a conjugated-drug strategy that delivers a splicing factor known to enhance splicing. It may also be possible that a combination of ASOs may be required to necessitate a complete rescue effect, as we recently demonstrated for *F8* exon 16 (Tse et al. 2023a). Together, our preliminary data towards rescuing HL phenotypes insinuate that assessing a wider range of drug modalities, including alternative ASO designs not tested in this study, may also need to be explored to discover a viable therapy for these HL patients with this novel variant.

Figure and Table Legends

Figure 1. Homozygous non-canonical *DEGS1* splice site variant detected in two unrelated probands. **(A)** IGV plots from whole exome sequencing showing homozygous variants in two unrelated probands, and heterozygous variants in available parents, that map to the 5′ splice site of exon 2 from the *DEGS1* gene (NM_003676.3, chromosomal position chr1:224378025-224378026 on hg19). The position of the HL-linked variant is boxed. **(B)** A schematic depicting the gene

architecture of *DEGS1*. The position of the non-canonical *DEGS1* splice site variant, c.825+4_825+5delAGinsTT, is also shown at a single-nucleotide resolution. **(C)** Sanger sequencing readouts, in forward and reverse directions, confirming the identity and presence of homozygous 5′ splice site variants in both unrelated probands presenting symptoms consistent with HL, relative to synthetic wild-type controls. The position of the HL-linked variant is boxed.

Figure 2. 5′ splice site variant of *DEGS1* exon 2 produces novel transcripts in HL probands. **(A, B)** IGV plots of RNA-seq data showing splice junction reads from one control patient and one proband aligning to the *DEGS1* gene, respectively in (A) and (B). Density of read coverage across each coding-sequence of *DEGS1* is displayed and indicated. Track indicating reads supporting the canonical *DEGS1* mRNA transcript is indicated (blue arrow). Track indicating reads supporting a novel *DEGS1* mRNA transcript that is predominant and unique to the HL-linked variant is indicated (red arrow). **(C, D)** Isoform plot depicting the relative abundance of each *DEGS1* mRNA transcript detected in the control patient and proband, respectively. Relative expression of each transcript is represented on the Y-axis. *DEGS1* exon 2 boundaries are indicated with vertical dashed black lines. The canonical reference transcript (blue) is indicated in blue. All novel transcripts that do not align to reference gene models and available gene annotations are indicated in red. Isoforms are only shown on the plot if they comprise more than 5% of *DEGS1* transcripts detected.

Figure 3. HL-linked splice site variant is sufficient to induce aberrant splicing and drive intramolecular refolding of RNA secondary structures of *DEGS1* exon 2. **(A)** A schematic depicting the heterologous splicing reporter system used to assay the functional impact the novel splice site variant has on *DEGS1* exon 2 splicing, relative to the wild-type (WT). *DEGS1* exon 2 and flanking intronic sequences corresponding to either the WT or HL-linked variant sequence context were cloned in between exon 1 and exon 2 of the *HBB* minigene **(B)** A representative agarose gel showing the effects the HL-linked variant has on *DEGS1* exon 2 splicing. As shown in the annotation above the representative agarose gel, controls include a no template reaction (lane 1) and a positive control for exon skipping (lane 2). Lanes corresponding to the splicing reporters assaying the WT or HL-linked variant sequence context of *DEGS1* exon 2 are indicated respectively. Expected mRNA isoforms including or excluding the *DEGS1* exon 2 are also annotated to the left of the agarose gel. **(C)** Percent-spliced-in (PSI) plot for splicing reporters shown in (B) measuring the splice site variant's impact on *DEGS1* exon 2 inclusion. PSI refers to the fraction of mRNA reporter isoforms that include the exon of interest, relative to the total population of mRNA reporter isoforms. Statistical significance between comparisons is denoted by asterisks that represent *P*-values with the following range of significance: * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $P \leq 0.0001$. Statistical significance was determined using analysis of variance (ANOVA), and Dunett's post-hoc test. Each condition tested and presented contains nine independent/biological replicates. **(D)** A normalized SHAPE reactivity vs structure

prediction plot comparing the RNA folding profiles for the WT and HL-linked variant sequence context of *DEGS1* exon 2 (depicted in blue and red, respectively). The top part of the plot shows the normalized SHAPE reactivity for each nucleotide. The bottom part of the plot shows SHAPE-constrained structure predictions represented by intramolecular base pairing interactions; secondary structure formation is denoted by arcs joining different regions of the RNA sequence context. A schematic model of *DEGS1* exon 2 and its flanking introns is also shown at the bottom of the plot to illustrate relative positions of RNA structure data. All SHAPE data analysis was performed in RNA Framework.

Figure 4. HL-linked splice site variant of *DEGS1* exon 2 is sufficient to inactivate DEGS1 desaturase function in probands. **(A)** A plot quantifying saturated:unsaturated (DhCer:Cer) ratios of various sphingolipids detected from control patient samples or proband-derived samples (depicted in red and blue, respectively). Statistical significance was determined using the two-tailed Student's test (t-test), with a $P \leq 0.05$ as the cutoff for significance. **(B)** A schematic depicting a data-supported model of how this novel HL-linked splice site variant of *DEGS1* exon 2 drives HL pathogenicity.

Figure 5. *DEGS1* exon 2 is highly dependent on exonic splicing enhancers for its splicing. **(A)** A mock schematic of the ASO walk conducted in this study. All ASOs tile across *DEGS1* exon 2 and its flanking introns. Each ASO used in our walks are

18 nucleotides in length and are designed using ribose sugars that are modified with a 2′-methoxyethyl group (2′-MOE, highlighted in light orange), and the phosphate backbone is modified to a phosphorothioate backbone (highlighted in light blue). Black boxes represent 18-mer ASOs that are contiguous by design, whereas purple boxes represent 18-mer ASOs that have 10nt overlaps between the preceding and proceeding ASO. **(B)** A representative agarose gel demonstrating no significant difference between conditions with and without a non-targeting ASO being co-transfected with our WT or HL-linked splicing reporter. Expected mRNA isoforms including or excluding *DEGS1* exon 2 are also annotated to the left of the agarose gel. **(C)** PSI plot quantifying the non-targeting ASO's impact on *DEGS1* exon 2 splicing as shown in (B). **(D, E)** PSI plots quantifying our ASO walk data on the sequence context corresponding to the WT or HL-linked variant of *DEGS1* exon 2, respectively. A schematic model of *DEGS1* and its flanking introns are shown at the bottom of each PSI plot to illustrate relative positions of ASOs, and for which sequence context. ASOs that significantly inhibit splicing are indicated in red, whereas those that significantly enhance splicing are indicated in green. The control ASO is depicted in yellow, and non-significant ASOs are depicted in black. Statistical significance between comparisons is denoted by asterisks that represent P-values with the following range of significance: * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $P \leq 0.0001$. Statistical significance was determined using analysis of variance (ANOVA), and Dunett's post-hoc test. Each condition tested and presented in this figure contains a minimum of three independent/biological replicates.

Table 1. MaxEntScan analysis of 5′ and 3′ splice site strength of *DEGS1* exon 2.

Supplemental Figure 1. Region search for novel HL-linked splice site variant using ClinVar and gnomAD database. The pathogenic variant detected from HL probands, c.825+4_825+5delAGinsTT (NM_003676.3, chromosomal position chr1:224378025-224378026 on hg19), was not observed in either ClinVar or gnomAD results.

Supplemental Figure 2. Isoform and Bulk Tissue Expression of *DEGS1* (ENSG00000143753.12; GTEX). **(A)** Isoform expression of *DEGS1*, delta-4-desaturase, sphingolipid 1 [HGNC:13709]. **(B)** Bulk tissue gene expression for *DEGS1*. **(C)** Transcript table for DEGS1 (DECIPHER).

Supplemental Figure 3. SHAPE-derived secondary structural models for WT *DEGS1* exon 2 and flanking intron sequences, as assayed in splicing reporter assays. Bases are colored according to their normalized 2A3 SHAPE reactivity. All nucleotide position numbering shown is based on the IVT RNA template used for SHAPE probing, from the 5′ to 3′ orientation.

Supplemental Figure 4. SHAPE-derived secondary structural models for the HL-linked variant of *DEGS1* exon 2 and flanking intron sequences, as assayed in

splicing reporter assays. Bases are colored according to their normalized 2A3 SHAPE reactivity. All nucleotide position numbering shown is based on the IVT RNA template used for SHAPE probing, from the 5′ to 3′ orientation.

Supplemental Figure 5. Examining the structural accessibility of splice sites between the WT and HL-linked sequence context corresponding to *DEGS1* exon 2. The figure zooms into the splice sites for the WT, as shown in Panel **(A)**, and for the HL-linked variant, as shown in Panel **(B)**. Respective splice sites are indicated with a black arrow and are also boxed in red.

Supplemental Figure 6. Cross-referencing RBPmap analysis results to SHAPE and ASO data. RBPmap prediction results for two putative binding sites for splicing factors known to enhance splicing are depicted, showing their motif confidence score and their position within our sequence context assayed (as indicated following the light blue arrow), and the effect an ASO has in interfering with these putative splicing enhancers (as indicated following the red arrow).

Supplementary Table 1. Primer sequences used to generate DEGS1 exon 2 WT and HL-linked variant inserts for cloning into *HBB* splicing reporters.

Supplementary Table 2. The sequence context and chemical modifications of antisense oligonucleotides targeting *DEGS1* exon 2.

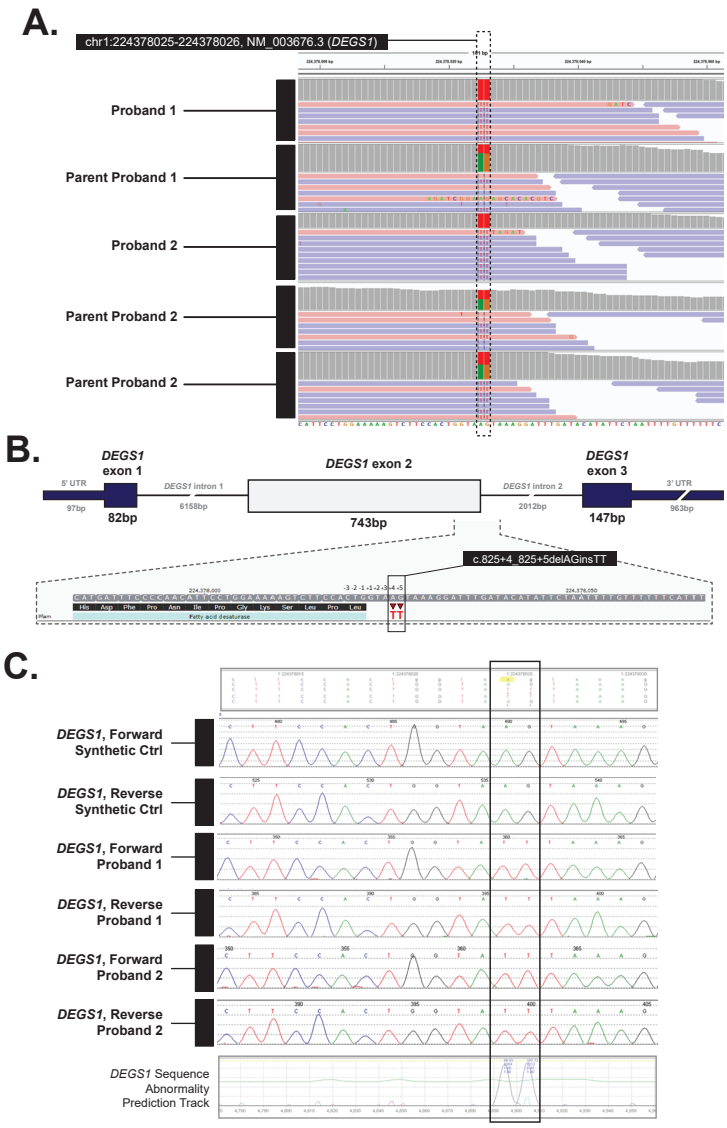


Figure 1. Homozygous non-canonical DEGS1 splice site variant detected in two unrelated probands. (A) IGV plots from whole exome sequencing showing homozygous variants in two unre lated probands, and heterozygous variants in available parents, that map to the 5′ splice site of exon 2 from the DEGS1 gene (NM_003676.3, chromosomal position chr1:224378025-224378026 on hg19). The position of the HL-linked variant is boxed. (B) A schematic depicting the gene archi tecture of DEGS1. The position of the non-canonical DEGS1 splice site variant, c.825+4_825+5de lAGinsTT, is also shown at a single-nucleotide resolution. (C) Sanger sequencing readouts, in forward and reverse directions, confirming the identity and presence of homozygous 5′ splice site variants in both unrelated probands presenting symptoms consistent with HL, relative to synthetic wild-type controls. The position of the HL-linked variant is boxed.

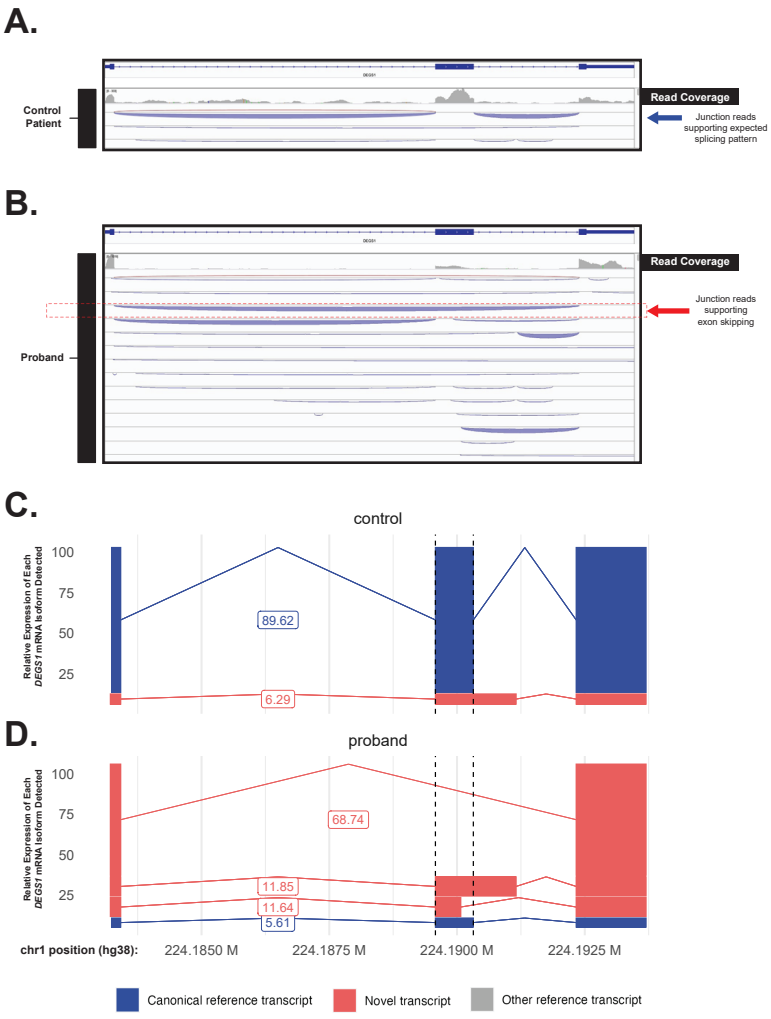


Figure 2. 5′ splice site variant of DEGS1 exon 2 produces novel transcripts in HL probands. (A, B) IGV plots of RNA-seq data showing splice junction reads from one control patient and one proband aligning to the DEGS1 genes, respectively in (A) and (B). Density of read coverage across each coding-sequence of DEGS1 is displayed and indicated. Track indicating reads supporting the canonical DEGS1 mRNA transcript is indicated (blue arrow). Track indicating reads supporting a novel DEGS1 mRNA transcript that is predominant and unique to the HL-linked variant is indicated (red arrow). (C, D) Isoform plot depicting the relative abundance of each DEGS1 mRNA transcript detected in the control patient and proband, respectively. Relative expression of each transcript is represented on the Y--axis. DEGS1 exon 2 boundaries are indicated with vertical dashed black lines. The canonical reference transcript (blue) is indicated in blue. All novel transcripts that do not align to reference gene models and available gene annotations are indicated in red. Isoforms are only shown on the plot if they comprise more than 5% of DEGS1 transcripts detected.

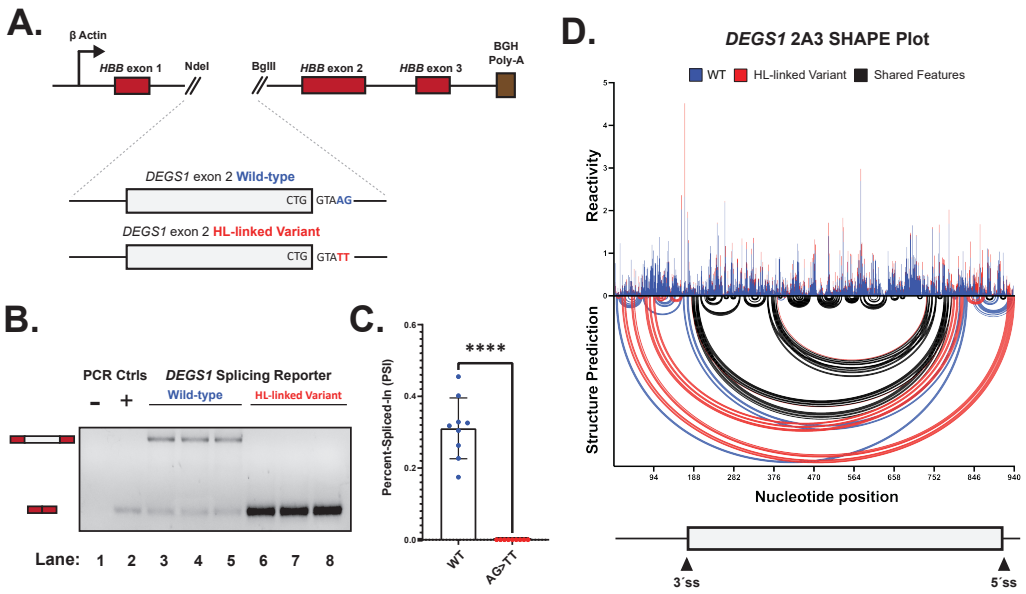


Figure 3. HL-linked splice site variant is sufficient to induce aberrant splicing and drive intramolecular refolding of RNA secondary structures of DEGS1 exon 2. (A) A schematic depicting the heterologous splicing reporter system used to assay the functional impact the novel splice site variant has on DEGS1 exon 2 splicing, relative to the wild-type (WT). DEGS1 exon 2 and flanking intronic sequences corresponding to either the WT or HL-linked variant sequence context were cloned in between exon 1 and exon 2 of the HBB minigene (B) A representative agarose gel showing the effects the HL-linked variant has on DEGS1 exon 2 splicing. As shown in the annotation above the representative agarose gel, controls include a no template reaction (lane 1) and a positive control for exon skipping (lane 2). Lanes corresponding to the splicing reporters assaying the WT or HL-linked variant sequence context of DEGS1 exon 2 are indicated respectively. Expected mRNA isoforms including or excluding the DEGS1 exon 2 are also annotated to the left of the agarose gel. (C) Percent-spliced-in (PSI) plot for splicing reporters shown in (B) measuring the splice site variant's impact on DEGS1 exon 2 inclusion. PSI refers to the fraction of mRNA reporter isoforms that include the exon of interest, relative to the total population of mRNA reporter isoforms. Statistical significance between comparisons is denoted by asterisks that represent P-values with the following range of significance: * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = P ≤ 0.0001. Statistical significance was determined using analysis of variance (ANOVA), and Dunett's post-hoc test. Each condition tested and presented contains nine independent/biological replicates. (D) A normalized SHAPE reactivity vs structure prediction plot comparing the RNA folding profiles for the WT and HL-linked variant sequence context of DEGS1 exon 2 (depicted in blue and red, respectively). The top part of the plot shows the normalized SHAPE reactivity for each nucleotide. The bottom part of the plot shows SHAPE-constrained structure predictions represented by intramolecular base pairing interactions; secondary structure formation is denoted by arcs joining different regions of the RNA sequence context. A schematic model of DEGS1 exon 2 and its flanking introns is also shown at the bottom of the plot to illustrate relative positions of RNA structure data. All SHAPE data analysis was performed in RNA Framework.

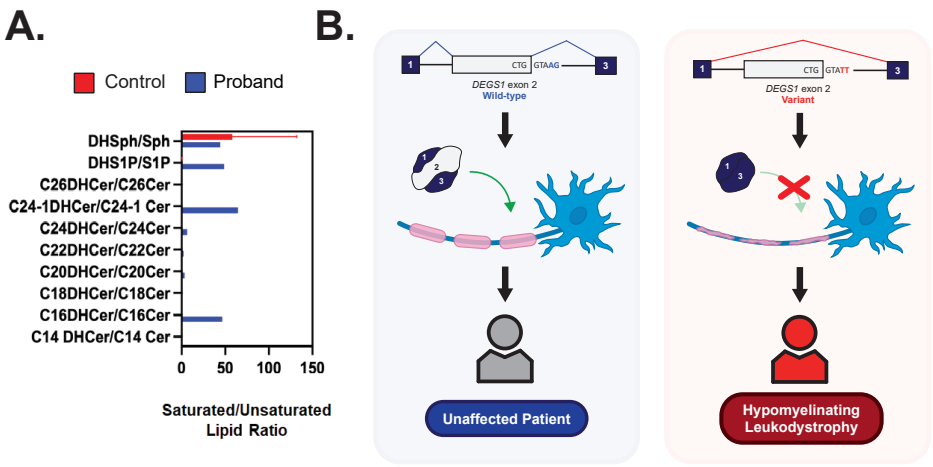


Figure 4. HL-linked splice site variant of DEGS1 exon 2 is sufficient to inactivate DEGS1 desatu rase function in probands. (A) A plot quantifying saturated:unsaturated (DhCer:Cer) ratios of various sphingolipids detected from control patient samples or proband-derived samples (depicted in red and blue, respectively). Statistical significance was determined using the two-tailed Student's test (t-test), with a $P \leq 0.05$ as the cutoff for significance. (B) A schematic depicting a data-supported model of how this novel HL-linked splice site variant of DEGS1 exon 2 drives HL pathogenicity.

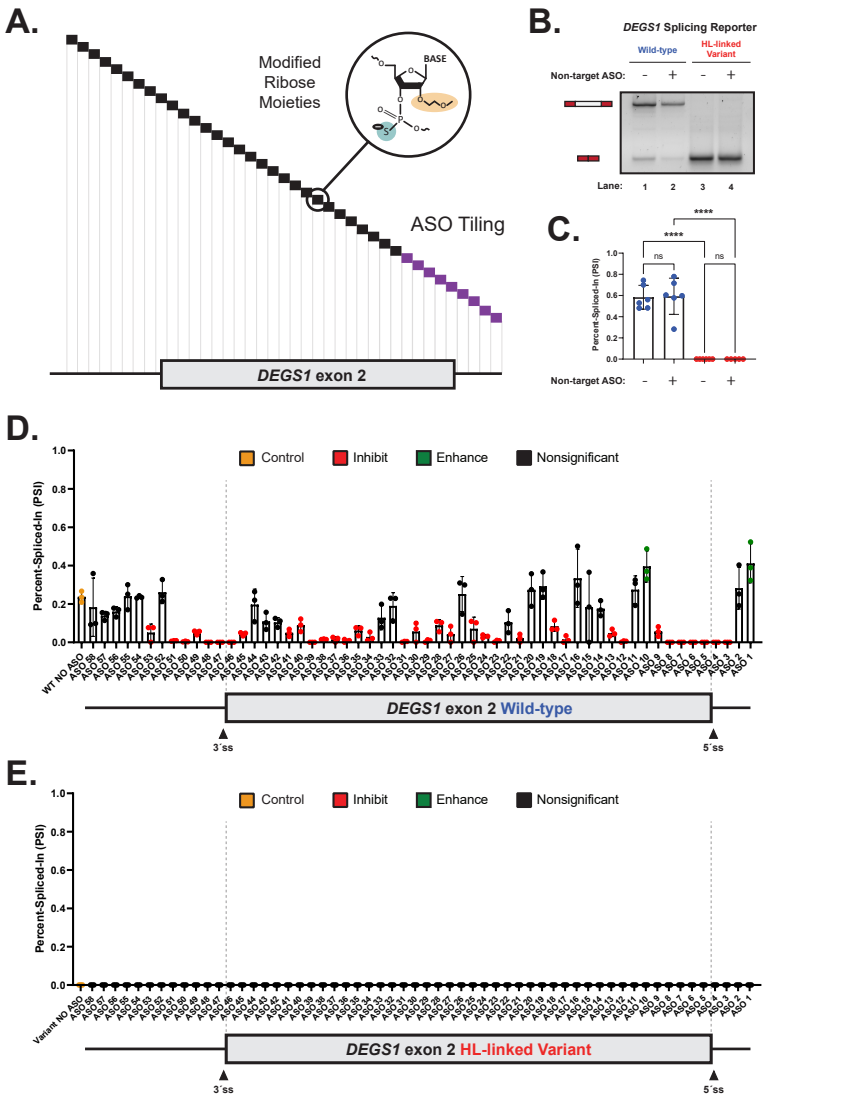
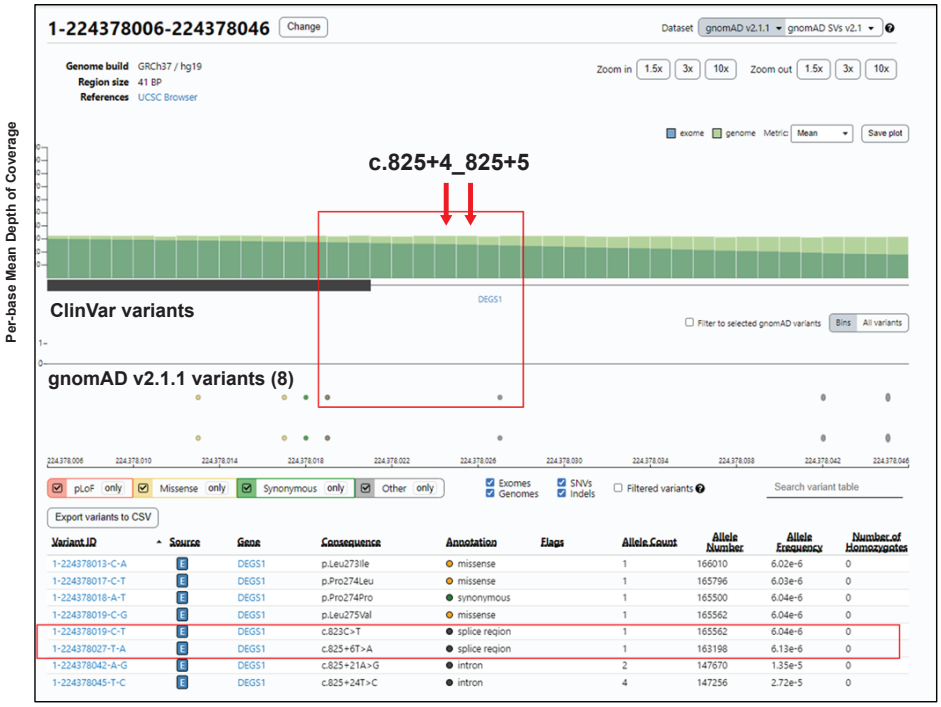


Figure 5. DEGS1 exon 2 is highly dependent on exonic splicing enhancers for its splicing. (A) A mock schematic of the ASO walk conducted in this study. All ASOs tile across DEGS1 exon 2 and its flanking introns. Each ASO used in our walks are 18 nucleotides in length and are designed using ribose sugars that are modified with a 2′-methoxyethyl group (2′-MOE, highlighted in light orange), and the phosphate backbone is modified to a phosphorothioate backbone (highlighted in light blue). Black boxes represent 18-mer ASOs that are contiguous by design, whereas purple boxes represent 18-mer ASOs that have 10nt overlaps between the preceding and proceeding ASO. (B) A representative agarose gel demonstrating no significant difference between conditions with and without a non-targeting ASO being co-transfected with our WT or HL-linked splicing reporter. Expected mRNA isoforms including or excluding DEGS1 exon 2 are also annotated to the left of the agarose gel. (C) PSI plot quantifying the non-targeting ASO's impact on DEGS1 exon 2 splicing as shown in (B). (D, E) PSI plots quantifying our ASO walk data on the sequence context corresponding to the WT or HL-linked variant of DEGS1 exon 2, respectively. A schematic model of DEGS1 and its flanking introns are shown at the bottom of each PSI plot to illustrate relative positions of ASOs, and for which sequence context. ASOs that significantly inhibit splicing are indicated in red, whereas those that significantly enhance splicing are indicated in green. The control ASO is depicted in yellow, and non-significant ASOs are depicted in black. Statistical significance between comparisons is denoted by asterisks that represent P-values with the following range of significance: * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = P ≤ 0.0001. Statistical significance was determined using analysis of variance (ANOVA), and Dunett's post-hoc test. Each condition tested and presented in this figure contains a minimum of three independent/biological replicates.

Table 1. MaxEntScan analysis of 5′ and 3′ splice site strength of *DEGS1* exon 2.

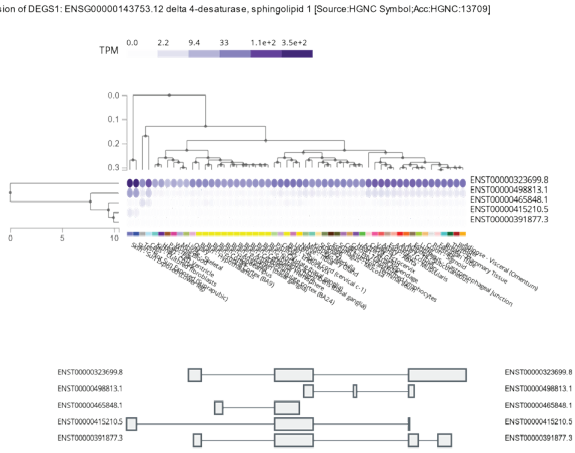
gene	Exon	mutation	cDNA Nomenclature Identity	5'ss Sequence Input (6 intronic nts, 3 exonic nts)	5'ss Strength Score (Max Entropy Model)	5'ss Strength Score (PWM)	3'ss Sequence Input (20 intronic nts, 3 exonic nts)	3'ss Strength Score (Max Entropy Model)	3'ss Strength Score (PWM)
DEGS1	**Exon 2**	------	WT	CTGgtaagt	**10.65**	**11.04**	tttttgttttttctttacagCAA	**11.24**	**14.91**
DEGS1	**Exon 2**	AG>TT	c.825+4_825+5delAGinsTT	CTGgtattt	**3.5**	**4.19**	tttttgttttttctttacagCAA	**11.24**	**14.91**

A.

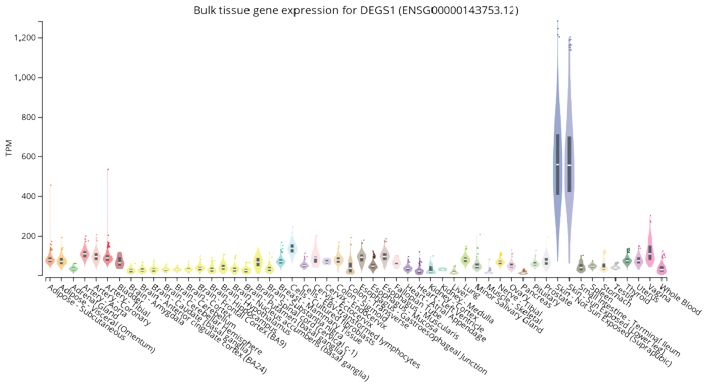


Supplemental Figure 1. Region search for novel HL-linked splice site variant using ClinVar and gnomAD database. The pathogenic variant detected from HL probands, c.825+4_825+5delAGinsTT (NM_003676.3, chromosomal position chr1:224378025-224378026 on hg19), was not observed in either ClinVar or gnomAD results.

A.



B.



C.

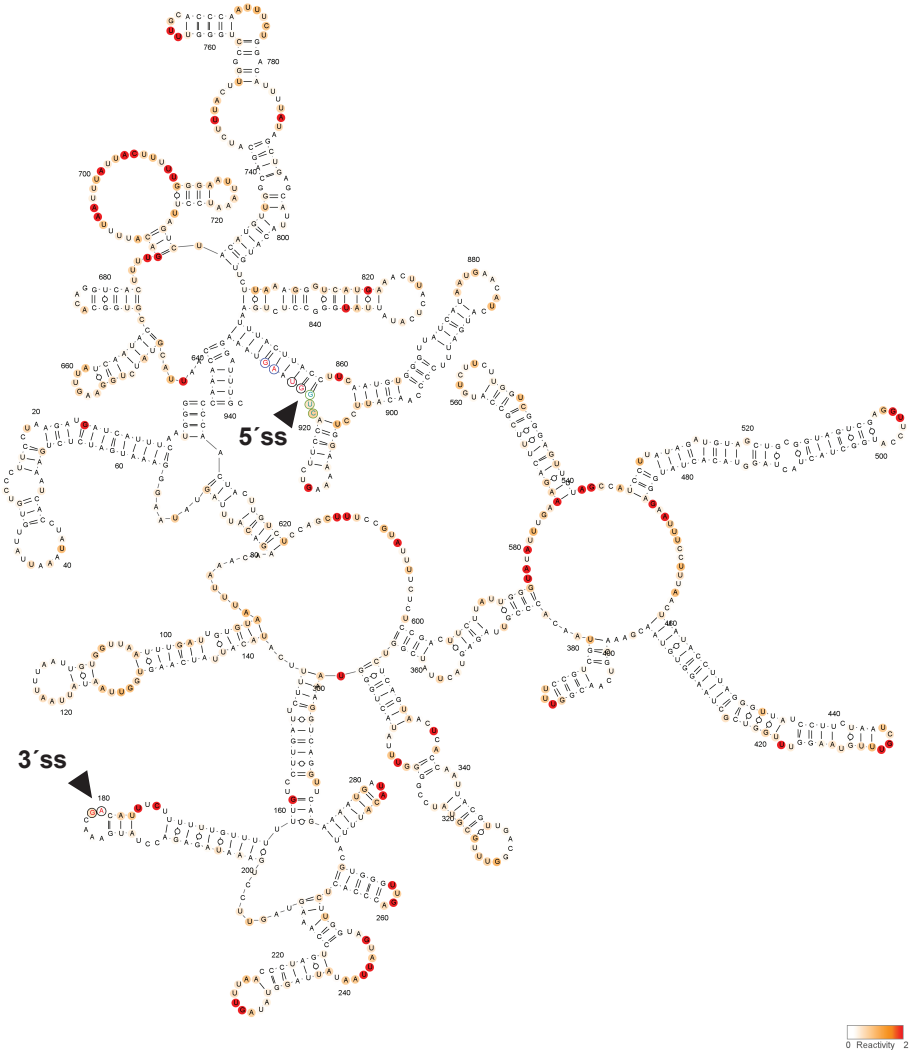
Transcripts for DEGS1: 1 to 5 of 5

Filter...

Transcript	Location	Size	Type	Protein	Exons
ENST00000323699.9 - DEGS1-201 - NM_003676.4 MANE Select NM_001321541.2 NM_001321542.2	1 224183240 224193441	10.20 kb	Protein coding	323aa	3
ENST00000391877.3 - DEGS1-202	1 224183266 224193157	9.89 kb	Protein coding	323aa	4
ENST00000415210.5 - DEGS1-203	1 224175756 224192353	16.60 kb	Protein coding	261aa	3
ENST00000465848.1 - DEGS1-204	1 224183675 224190052	6.38 kb	Processed transcript		2
ENST00000498813.1 - DEGS1-205	1 224190136 224192440	2.31 kb	Processed transcript		3

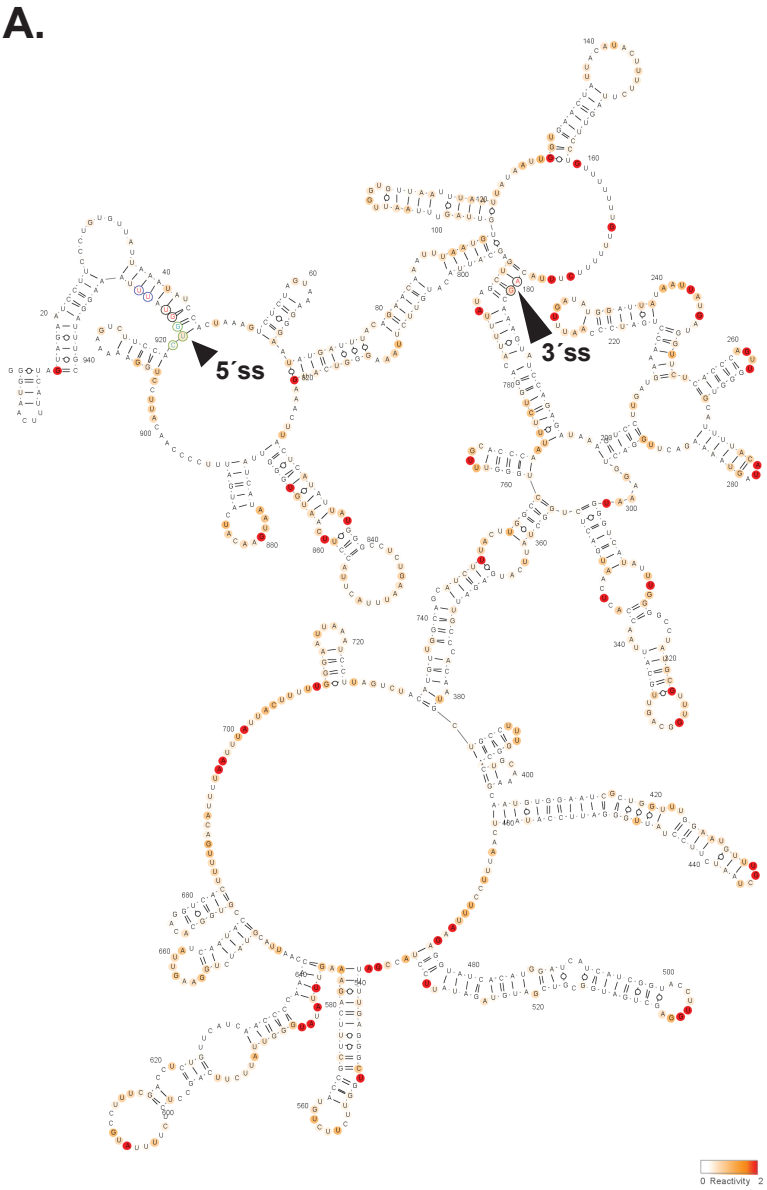
Supplemental Figure 2. Isoform and Bulk Tissue Expression of DEGS1 (ENSG00000143753.12; GTEX). (A) Isoform expression of DEGS1, delta-4-desaturase, sphingolipid 1 [HGNC:13709]. (B) Bulk tissue gene expression for DEGS1. (C) Transcript table for DEGS1 (DECIPHER).

A.



DEGS1 average 2A3 50mM reactivity >2

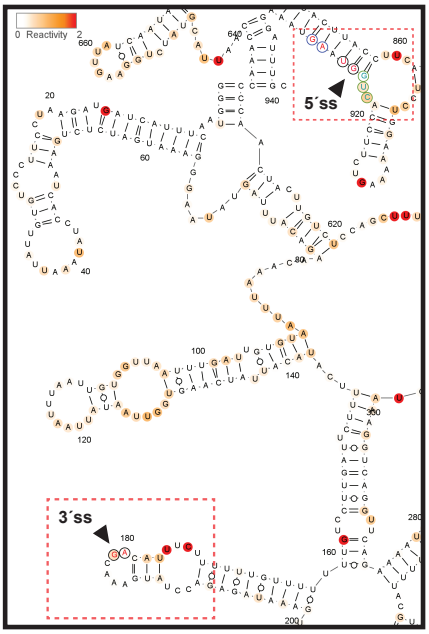
Supplemental Figure 3. SHAPE-derived secondary structural models for WT DEGS1 exon 2 and flanking intron sequences, as assayed in splicing reporter assays. Bases are colored according to their normalized 2A3 SHAPE reactivity. All nucleotide position numbering shown is based on the IVT RNA template used for SHAPE probing, from the 5′ to 3′ orientation.



DEGS1MT average 2A3 50mM reactivity>2

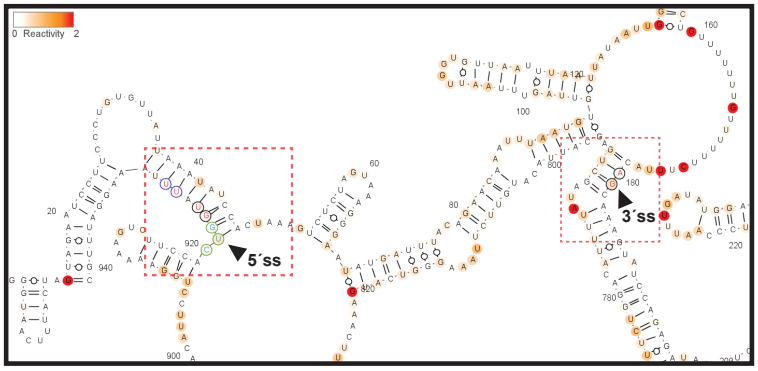
Supplemental Figure 4. SHAPE-derived secondary structural models for the HL-linked variant of DEGS1 exon 2 and flanking intron sequences, as assayed in splicing reporter assays. Bases are colored according to their normalized 2A3 SHAPE reactivity. All nucleotide position numbering shown is based on the IVT RNA template used for SHAPE probing, from the 5′ to 3′ orientation.

A.



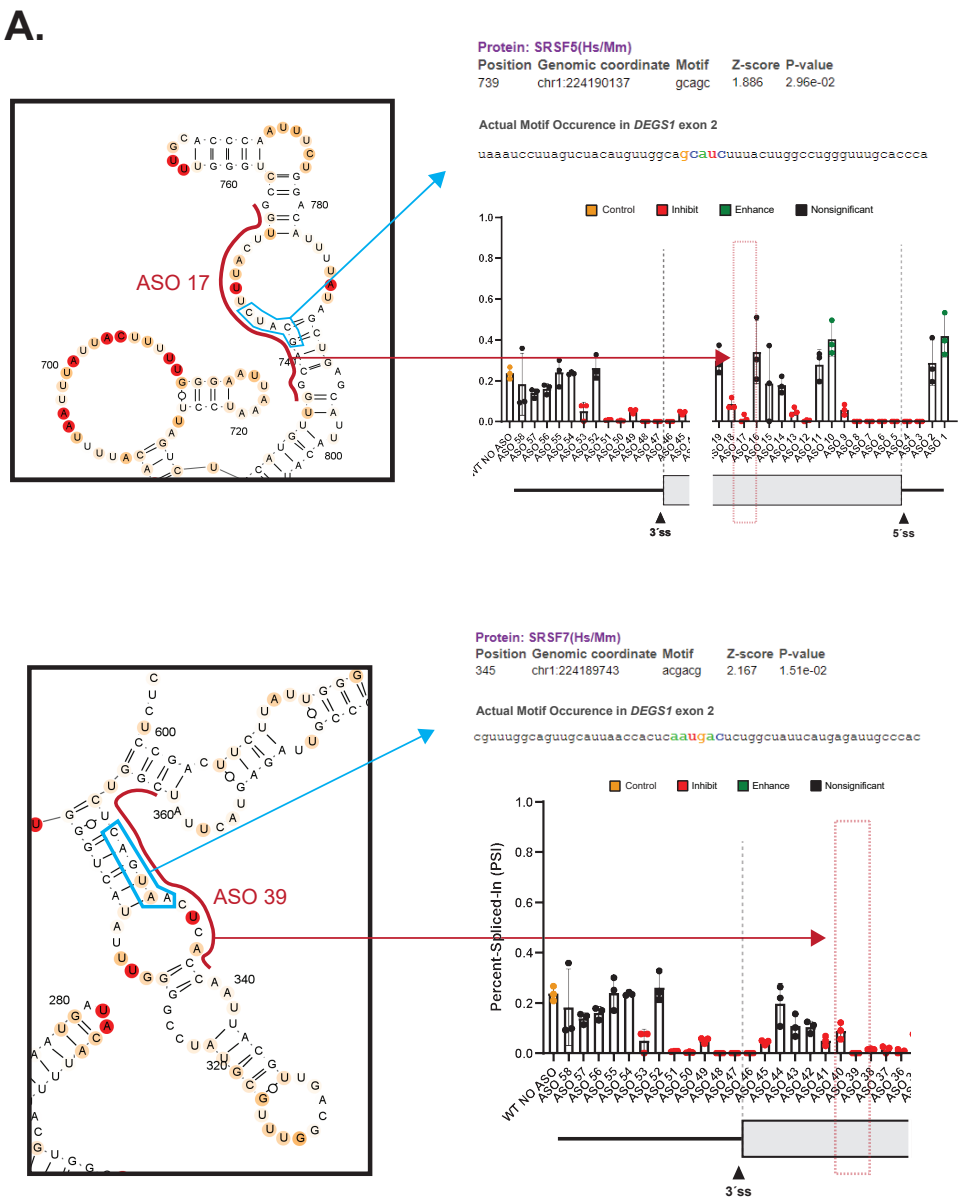
***DEGS1* exon 2 WT**

B.



***DEGS1* exon 2 HL-linked Variant**

Supplemental Figure 5. Examining the structural accessibility of splice sites between the WT and HL-linked sequence context corresponding to DEGS1 exon 2. The figure zooms into the splice sites for the WT, as shown in Panel (A), and for the HL-linked variant, as shown in Panel (B). Respective splice sites are indicated with a black arrow and are also boxed in red.



Supplemental Figure 6. Cross-referencing RBPmap analysis results to SHAPE and ASO data. RBPmap prediction results for two putative binding sites for splicing factors known to enhance splicing are depicted, showing their motif confidence score and their position within our sequence context assayed (as indicated following the light blue arrow), and the effect an ASO has in interfering with these putative splicing enhancers (as indicated following the red arrow).

Supplementary Table 1. Primer sequences used to generate DEGS1 exon 2 WT and HL-linked variant inserts for cloning into *HBB* splicing reporters.

Sequence context	forward_primer (5' to 3')	reverse_primer (5' to 3')	mutagenesis_forward_primer (5' to 3')	mutagenesis_reverse_primer (5' to 3')
DEGS1 exon 2	agaaactgggCATATGTAACTTTACTAGTAGAATCC	ccacatgcccAGATCTGAACCCTTTGTAAAAGTAGG	TTCCACTGGTATTTAAAGGATTTG	CAAATCCTTTAAATACCAGTGGAA

Supplementary Table 2. The sequence context and chemical modifications of antisense oligonucleotides targeting *DEGS1* exon 2.

Compound Name	Backbone Modification Type	2' Sugar Modification Type	Oligo Length	ASO Sequence (5' -> 3')	IDT Formatted ASO Sequence
DEGS1_10ntOvlp_ASO 1	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	ttagaatatgtatccaaat	/5 2MOE/T/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/
DEGS1_10ntOvlp_ASO 2	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	tgtatcaaatcctttaaa	/5 2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/
DEGS1_10ntOvlp_ASO 3	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	atccttaaatatacCAG TG	/5 2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/
DEGS1_10ntOvlp_ASO 4	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	aatacCAGTGGAAGACTT	/5 2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/T/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/
DEGS1_10ntOvlp_ASO 5	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TGGAAGACTTTTTCCAGG	/5 2MOE/T/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/C/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/
DEGS1_10ntOvlp_ASO 6	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TTTTTCCAGGAATGTTGG	/5 2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/G/
DEGS1_10ntOvlp_ASO 7	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	GGAATGTTGGGGAAATCA	/5 2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/32MOE/A/
DEGS1_10ntOvlp_ASO 8	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	GGGGAAATCATGATGTTC	/5 2MOE/G/*/i2MOE/G/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/*/32MOE/C/
DEGS1_10ntOvlp_ASO 9	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	CATGATGTTCATTATGAT	/5 2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/32MOE/T/
DEGS1_10ntOvlp_ASO 10	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TCATTATGATAACCCACA	/5 2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/C/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/C/*/32MOE/A/
DEGS1_NoOvlp_ASO 11	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TTGAAGGTAAGTAAATTC	/5 2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/32MOE/C/
DEGS1_NoOvlp_ASO 12	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	AGAGGCCCATAATATGAG	/5 2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/C/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/G/
DEGS1_NoOvlp_ASO 13	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TAAGTTTCATGACCCTTT	/5 2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/i2MOE/C/*/i2MOE/C/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/32MOE/T/
DEGS1_NoOvlp_ASO 14	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	AAGAACATGTAATGCTCA	/5 2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/C/*/i2MOE/T/*/i2MOE/C/*/32MOE/A/
DEGS1_NoOvlp_ASO 15	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	GCTATAAAATGTCCAGAA	/5 2MOE/G/*/i2MOE/C/*/i2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/
DEGS1_NoOvlp_ASO 16	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	ATTGGGTGCAAACCCAGG	/5 2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/G/*/i2MOE/G/*/i2MOE/T/*/i2MOE/G/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/32MOE/G/
DEGS1_NoOvlp_ASO 17	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	CCAAGTAAAGATGCTGCC	/5 2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/C/*/i2MOE/T/*/i2MOE/G/*/i2MOE/C/*/i2MOE/C/
DEGS1_NoOvlp_ASO 18	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	AACATGTAGACTAAGGAT	/5 2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/C/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/T/
DEGS1_NoOvlp_ASO 19	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TTAATTCCCAAAAGTAA	/5 2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/
DEGS1_NoOvlp_ASO 20	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TAAATTAAATGTCAAAA	/5 2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/
DEGS1_NoOvlp_ASO 21	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	GTGACCTGTGCCACGGTA	/5 2MOE/G/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/i2MOE/C/*/i2MOE/C/*/i2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/G/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/T/*/i2MOE/A/
DEGS1_NoOvlp_ASO 22	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TTGATA ACTTCCAGATAC	/5 2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/C/
DEGS1_NoOvlp_ASO 23	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	GTAATTGGTTTGGGGTTG	/5 2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/G/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/G/*/i2MOE/G/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/32MOE/G/
DEGS1_NoOvlp_ASO 24	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	ATGAACAGAGGTCGAAAG	/5 2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/T/*/i2MOE/C/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/
DEGS1_NoOvlp_ASO 25	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	GCATAAAGAGAGGCTGA	/5 2MOE/G/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/G/*/i2MOE/C/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/
DEGS1_NoOvlp_ASO 26	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	AGAATAACCCATATAAAC	/5 2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/
DEGS1_NoOvlp_ASO 27	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TTTCTGAAAGCGGTACAG	/5 2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/C/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/
DEGS1_NoOvlp_ASO 28	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	AAGAACCAGCCCTCAAAA	/5 2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/C/*/i2MOE/C/*/i2MOE/C/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/
DEGS1_NoOvlp_ASO 29	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TCGGTAGGAATATCTACA	/5 2MOE/T/*/i2MOE/C/*/i2MOE/G/*/i2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/A/
DEGS1_NoOvlp_ASO 30	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TCGAGGCCATCAGCTCCA	/5 2MOE/T/*/i2MOE/C/*/i2MOE/G/*/i2MOE/A/*/i2MOE/C/*/i2MOE/G/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/C/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/*/32MOE/A/
DEGS1_NoOvlp_ASO 31	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	AGGTACCGATGATGATCC	/5 2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/C/*/i2MOE/C/*/i2MOE/G/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/
DEGS1_NoOvlp_ASO 32	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	ATGTGATACCTCTTAAAG	/5 2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/C/*/i2MOE/C/*/i2MOE/T/*/i2MOE/C/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/
DEGS1_NoOvlp_ASO 33	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	GAAATTGAATATGGAATC	/5 2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/
DEGS1_NoOvlp_ASO 34	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	CCAATAGGAAGATTAGCA	/5 2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/C/*/i2MOE/A/
DEGS1_NoOvlp_ASO 35	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	AACATTCCAACCAGCGGA	/5 2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/
DEGS1_NoOvlp_ASO 36	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TTCCACATTGCTTTGCAG	/5 2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/C/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/C/*/i2MOE/A/*/32MOE/G/
DEGS1_NoOvlp_ASO 37	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TTGCCAAGGCAGCATTG	/5 2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/
DEGS1_NoOvlp_ASO 38	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TGGGCAATCTCATGAATA	/5 2MOE/T/*/i2MOE/G/*/i2MOE/G/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/
DEGS1_NoOvlp_ASO 39	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	GCCAGAGTCATTGAGTGG	/5 2MOE/G/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/G/
DEGS1_NoOvlp_ASO 40	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TTAATGCAACTGCCAAAC	/5 2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/C/*/i2MOE/T/*/i2MOE/G/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/
DEGS1_NoOvlp_ASO 41	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	GCATAGGCCCCAAATATG	/5 2MOE/G/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/C/*/i2MOE/C/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/
DEGS1_NoOvlp_ASO 42	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	ACCCATTTCCAGTCCAAG	/5 2MOE/A/*/i2MOE/C/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/
DEGS1_NoOvlp_ASO 43	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	TCTTTTACTATGTAAAAT	/5 2MOE/T/*/i2MOE/C/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/C/*/i2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/
DEGS1_NoOvlp_ASO 44	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	GCACCCAACTGGGTGAGA	/5 2MOE/G/*/i2MOE/C/*/i2MOE/A/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/G/*/i2MOE/G/*/i2MOE/T/*/i2MOE/G/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/
DEGS1_NoOvlp_ASO 45	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	ACCATCATAATTATAATC	/5 2MOE/A/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/
DEGS1_NoOvlp_ASO 46	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	CATATCAAATTGGGATCA	/5 2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/
DEGS1_NoOvlp_ASO 47	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	GGTTTCATCAAGGACTTT	/5 2MOE/G/*/i2MOE/G/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/
DEGS1_NoOvlp_ASO 48	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	ATCTCTGGATACTTTGct	/5 2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/T/*/i2MOE/C/*/i2MOE/T/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/C/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/C/*/i2MOE/C/*/i2MOE/T/
DEGS1_NoOvlp_ASO 49	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	gtaagagaaaaacaaaaa	/5 2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/
DEGS1_NoOvlp_ASO 50	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	aacagcgaactaagaaaag	/5 2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/
DEGS1_NoOvlp_ASO 51	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	tatgtaagtagttcacctaa	/5 2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/C/*/i2MOE/C/*/i2MOE/A/
DEGS1_NoOvlp_ASO 52	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	ttataaattaaatcaacac	/5 2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/C/*/i2MOE/A/*/i2MOE/C/
DEGS1_NoOvlp_ASO 53	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	caattaaactaacacatt	/5 2MOE/C/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/A/*/i2MOE/C/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/
DEGS1_NoOvlp_ASO 54	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	aaatttgttctgtaaatc	/5 2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/T/*/i2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/T/*/i2MOE/C/
DEGS1_NoOvlp_ASO 55	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	atattccctttactagag	/5 2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/C/*/i2MOE/T/*/i2MOE/A/*/i2MOE/G/*/i2MOE/A/*/i2MOE/G/
DEGS1_NoOvlp_ASO 56	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	acttagatggatatttaa	/5 2MOE/A/*/i2MOE/C/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/
DEGS1_NoOvlp_ASO 57	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	taacacaggg aaggattc	/5 2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/C/*/i2MOE/A/*/i2MOE/C/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/
DEGS1_NoOvlp_ASO 58	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	attctactag taaagtta	/5 2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/C/*/i2MOE/T/*/i2MOE/A/*/i2MOE/C/*/i2MOE/T/*/i2MOE/A/*/i2MOE/G/*/i2MOE/T/*/i2MOE/A/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/T/*/i2MOE/T/*/32MOE/A/
Non-target ASO (F8Ex16_NoOvlp_ASO 27)	Phosphorothioate	2'-O-Methoxyethyl (**2'-MOE**)	**18**	aggtcctta gggtttaca	/5 2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/T/*/i2MOE/C/*/i2MOE/C/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/G/*/i2MOE/G/*/i2MOE/G/*/i2MOE/A/*/i2MOE/T/*/i2MOE/T/*/i2MOE/T/*/i2MOE/A/*/i2MOE/C/*/32MOE/A/

Bibliography

Aartsma-Rus, Annemieke, and David R. Corey. 2020. “The 10th Oligonucleotide Therapy Approved: Golodirsen for Duchenne Muscular Dystrophy.” *Nucleic Acid Therapeutics* 30 (2): 67.

Alt, F. W., A. L. Bothwell, M. Knapp, E. Siden, E. Mather, M. Koshland, and D. Baltimore. 1980. “Synthesis of Secreted and Membrane-Bound Immunoglobulin Mu Heavy Chains Is Directed by mRNAs That Differ at Their 3’ Ends.” *Cell* 20 (2). https://doi.org/10.1016/0092-8674(80)90615-7.

Anastassiou, D., H. Liu, and V. Varadan. 2006. “Variable Window Binding for Mutually Exclusive Alternative Splicing.” *Genome Biology* 7 (1). https://doi.org/10.1186/gb-2006-7-1-r2.

Anderson-Lee, J., E. Fisker, V. Kosaraju, M. Wu, J. Kong, J. Lee, M. Lee, M. Zada, A. Treuille, and R. Das. 2016. “Principles for Predicting RNA Secondary Structure Design Difficulty.” *Journal of Molecular Biology* 428 (5 Pt A). https://doi.org/10.1016/j.jmb.2015.11.013.

Aneni, Kammarauche, Claudia-Santi F. Fernandes, Lily A. Hoerner, Claire Szapary, Tyra M. Pendergrass Boomer, and Lynn E. Fiellin. 2023. “A Video Game Intervention to Prevent Opioid Misuse Among Older Adolescents: Development and Preimplementation Study.” *JMIR Serious Games* 11. https://doi.org/10.2196/46912.

Anguera, J. A., J. Boccanfuso, J. L. Rintoul, O. Al-Hashimi, F. Faraji, J. Janowich, E. Kong, et al. 2013. “Video Game Training Enhances Cognitive Control in Older Adults.” *Nature* 501 (7465): 97–101.

Anna, A., and G. Monika. 2018a. “Splicing Mutations in Human Genetic Disorders: Examples, Detection, and Confirmation.” *Journal of Applied Genetics* 59 (3). https://doi.org/10.1007/s13353-018-0444-7.

———. 2018b. “Splicing Mutations in Human Genetic Disorders: Examples, Detection, and Confirmation.” *Journal of Applied Genetics* 59 (3). https://doi.org/10.1007/s13353-018-0444-7.

Avila, J., F. Lim, F. Moreno, C. Belmonte, and A. C. Cuello. 2002. “Tau Function and Dysfunction in Neurons: Its Role in Neurodegenerative Disorders.” *Molecular Neurobiology* 25 (3). https://doi.org/10.1385/MN:25:3:213.

Barbosa-Morais, N. L., M. Irimia, Q. Pan, H. Y. Xiong, S. Gueroussov, L. J. Lee, V. Slobodeniuc, et al. 2012. “The Evolutionary Landscape of Alternative Splicing in Vertebrate Species.” *Science* 338 (6114). https://doi.org/10.1126/science.1230612.

Berezovski, M. V., M. Lechmann, M. U. Musheev, T. W. Mak, and S. N. Krylov. 2008. “Aptamer-Facilitated Biomarker Discovery (AptaBiD).” *Journal of the American Chemical Society* 130 (28). https://doi.org/10.1021/ja801951p.

Berget, S. M. 1995. “Exon Recognition in Vertebrate Splicing.” *The Journal of Biological Chemistry* 270 (6). https://doi.org/10.1074/jbc.270.6.2411.

Berget, S. M., C. Moore, and P. A. Sharp. 1977. “Spliced Segments at the 5’ Terminus of Adenovirus 2 Late mRNA.” *Proceedings of the National Academy of Sciences of the United States of America* 74 (8). https://doi.org/10.1073/pnas.74.8.3171.

Berget, S. M., and P. A. Sharp. 1979. “Structure of Late Adenovirus 2 Heterogeneous Nuclear RNA.” *Journal of Molecular Biology* 129 (4): 547–65.

Berglund, J. A., K. Chua, N. Abovich, R. Reed, and M. Rosbash. 1997. “The Splicing Factor BBP Interacts Specifically with the Pre-mRNA Branchpoint Sequence UACUAAC.” *Cell* 89 (5). https://doi.org/10.1016/s0092-8674(00)80261-5.

Berg, Matthew D., and Christopher J. Brandl. 2021. “Transfer RNAs: Diversity in Form and Function.” *RNA Biology* 18 (3): 316.

Berk, Arnold J., and Phillip A. Sharp. 1977. “Sizing and Mapping of Early Adenovirus mRNAs by Gel Electrophoresis of S1 Endonuclease-Digested Hybrids.” *Cell* 12 (3): 721–32.

Bielawski, J., Z. M. Szulc, Y. A. Hannun, and A. Bielawska. 2006. “Simultaneous Quantitative Analysis of Bioactive Sphingolipids by High-Performance Liquid Chromatography-Tandem Mass Spectrometry.” *Methods* 39 (2). https://doi.org/10.1016/j.ymeth.2006.05.004.

Bilodeau, Patricia S., Jeffrey K. Domsic, Akila Mayeda, Adrian R. Krainer, and C. Martin Stoltzfus. 2001. “RNA Splicing at Human Immunodeficiency Virus Type 1 3′ Splice Site A2 Is Regulated by Binding of hnRNP A/B Proteins to an Exonic Splicing Silencer Element.” *Journal of Virology* 75 (18): 8487.

Blencowe, B. J. 2000. “Exonic Splicing Enhancers: Mechanism of Action, Diversity and Role in Human Genetic Diseases.” *Trends in Biochemical Sciences* 25 (3). https://doi.org/10.1016/s0968-0004(00)01549-8.

Bolisetty, Mohan T., and Karen L. Beemon. 2012. “Splicing of Internal Large Exons Is Defined by Novel Cis -Acting Sequence Elements.” *Nucleic Acids Research* 40 (18): 9244–54.

Both, G. W., Y. Furuichi, S. Muthukrishnan, and A. J. Shatkin. 1975. “Ribosome Binding to Reovirus mRNA in Protein Synthesis Requires 5’ Terminal 7-Methylguanosine.” *Cell* 6 (2). https://doi.org/10.1016/0092-8674(75)90009-4.

Breathnach, R., C. Benoist, K. O’Hare, F. Gannon, and P. Chambon. 1978. “Ovalbumin Gene: Evidence for a Leader Sequence in mRNA and DNA Sequences at the Exon-Intron Boundaries.” *Proceedings of the National Academy of Sciences of the United States of America* 75 (10). https://doi.org/10.1073/pnas.75.10.4853.

Bringmann, P., and R. Lührmann. 1986. “Purification of the Individual snRNPs U1, U2, U5 and U4/U6 from HeLa Cells and Characterization of Their Protein Constituents.” *The EMBO Journal* 5 (13). https://doi.org/10.1002/j.1460-2075.1986.tb04676.x.

Brody, E., and J. Abelson. 1985. “The ‘Spliceosome’: Yeast Pre-Messenger RNA Associates with a 40S Complex in a Splicing-Dependent Reaction.” *Science* 228 (4702). https://doi.org/10.1126/science.3890181.

Busch, A., and K. J. Hertel. 2012a. “Evolution of SR Protein and hnRNP Splicing Regulatory Factors.” *Wiley Interdisciplinary Reviews. RNA* 3 (1). https://doi.org/10.1002/wrna.100.

———. 2012b. “Evolution of SR Protein and hnRNP Splicing Regulatory Factors.” *Wiley Interdisciplinary Reviews. RNA* 3 (1). https://doi.org/10.1002/wrna.100.

Byrnes, A. E., F. Roudnicky, A. Gogineni, A. L. Soung, M. Xiong, M. Hayne, T. Heaster-Ford, et al. 2024. “A Fluorescent Splice-Switching Mouse Model Enables High-Throughput, Sensitive Quantification of Antisense Oligonucleotide Delivery and Activity.” *Cell Reports Methods* 4 (1): 100673.

Cáceres, J. F., and A. R. Kornblihtt. 2002. “Alternative Splicing: Multiple Control Mechanisms and Involvement in Human Disease.” *Trends in Genetics: TIG* 18 (4). https://doi.org/10.1016/s0168-9525(01)02626-9.

Caputi, M., A. Mayeda, A. R. Krainer, and A. M. Zahler. 1999a. “hnRNP A/B Proteins Are Required for Inhibition of HIV-1 Pre-mRNA Splicing.” *The EMBO Journal* 18 (14): 4060.

———. 1999b. “hnRNP A/B Proteins Are Required for Inhibition of HIV-1 Pre-mRNA Splicing.” *The EMBO Journal* 18 (14): 4060.

Cartegni, Luca, Michelle L. Hastings, John A. Calarco, Elisa de Stanchina, and Adrian R. Krainer. 2006a. “Determinants of Exon 7 Splicing in the Spinal Muscular Atrophy Genes, SMN1 and SMN2.” *American Journal of Human Genetics* 78 (1): 63–77.

———. 2006b. “Determinants of Exon 7 Splicing in the Spinal Muscular Atrophy Genes, SMN1 and SMN2.” *American Journal of Human Genetics* 78 (1): 63–77.

Cartegni, Luca, and Adrian R. Krainer. 2002a. “Disruption of an SF2/ASF-Dependent Exonic Splicing Enhancer in SMN2 Causes Spinal Muscular Atrophy in the Absence of SMN1.” *Nature Genetics* 30 (4): 377–84.

———. 2002b. “Disruption of an SF2/ASF-Dependent Exonic Splicing Enhancer in SMN2 Causes Spinal Muscular Atrophy in the Absence of SMN1.” *Nature Genetics* 30 (4): 377–84.

Chen, C. D., R. Kobayashi, and D. M. Helfman. 1999. “Binding of hnRNP H to an Exonic Splicing Silencer Is Involved in the Regulation of Alternative Splicing of the Rat β-Tropomyosin Gene.” *Genes & Development* 13 (5): 593.

Cheng, Wenyu, Guohua Chen, Huaijie Jia, Xiaobing He, and Zhizhong Jing. 2018. “DDX5 RNA Helicases: Emerging Roles in Viral Infection.” *International Journal of Molecular Sciences* 19 (4). https://doi.org/10.3390/ijms19041122.

Chen, Jiunn-Liang, Maria A. Blasco, and Carol W. Greider. 2000. “Secondary Structure of Vertebrate Telomerase RNA.” *Cell* 100 (5): 503–14.

Chen, Jiunn-Liang, and Carol W. Greider. 2004. “An Emerging Consensus for Telomerase RNA Structure.” *Proceedings of the National Academy of Sciences* 101 (41): 14683–84.

Chen, J. L., K. K. Opperman, and C. W. Greider. 2002. “A Critical Stem-Loop Structure in the CR4-CR5 Domain of Mammalian Telomerase RNA.” *Nucleic Acids Research* 30 (2). https://doi.org/10.1093/nar/30.2.592.

Chiriboga, C. A., K. J. Swoboda, B. T. Darras, S. T. Iannaccone, J. Montes, D. C. De Vivo, D. A. Norris, C. F. Bennett, and K. M. Bishop. 2016. “Results from a Phase 1 Study of Nusinersen (ISIS-SMN(Rx)) in Children with Spinal Muscular Atrophy.” *Neurology* 86 (10). https://doi.org/10.1212/WNL.0000000000002445.

Chow, L. T., R. E. Gelinas, T. R. Broker, and R. J. Roberts. 1977. “An Amazing

Sequence Arrangement at the 5’ Ends of Adenovirus 2 Messenger RNA.” *Cell* 12 (1). https://doi.org/10.1016/0092-8674(77)90180-5.

Cobb, Matthew. 2017. “60 Years Ago, Francis Crick Changed the Logic of Biology.” *PLoS Biology* 15 (9). https://doi.org/10.1371/journal.pbio.2003243.

Cooper, Seth, Firas Khatib, Adrien Treuille, Janos Barbero, Jeehyung Lee, Michael Beenen, Andrew Leaver-Fay, David Baker, Zoran Popović, and 000 Foldit Players >57. 2010. “Predicting Protein Structures with a Multiplayer Online Game.” *Nature* 466 (7307): 756.

Cordero, P., and R. Das. 2015. “Rich RNA Structure Landscapes Revealed by Mutate-and-Map Analysis.” *PLoS Computational Biology* 11 (11). https://doi.org/10.1371/journal.pcbi.1004473.

Corey, D. R. 2007. “RNA Learns from Antisense.” *Nature Chemical Biology* 3 (1). https://doi.org/10.1038/nchembio0107-8.

———. 2017a. “Nusinersen, an Antisense Oligonucleotide Drug for Spinal Muscular Atrophy.” *Nature Neuroscience* 20 (4). https://doi.org/10.1038/nn.4508.

———. 2017b. “Nusinersen, an Antisense Oligonucleotide Drug for Spinal Muscular Atrophy.” *Nature Neuroscience* 20 (4). https://doi.org/10.1038/nn.4508.

Cramer, P., J. F. Cáceres, D. Cazalla, S. Kadener, A. F. Muro, F. E. Baralle, and A. R. Kornblihtt. 1999. “Coupling of Transcription with Alternative Splicing: RNA Pol II Promoters Modulate SF2/ASF and 9G8 Effects on an Exonic Splicing Enhancer.” *Molecular Cell* 4 (2). https://doi.org/10.1016/s1097-2765(00)80372-x.

Cramer, P., C. G. Pesce, F. E. Baralle, and A. R. Kornblihtt. 1997. “Functional Association between Promoter Structure and Transcript Alternative Splicing.” *Proceedings of the National Academy of Sciences of the United States of America* 94 (21). https://doi.org/10.1073/pnas.94.21.11456.

Crick, Francis. 1970. “Central Dogma of Molecular Biology.” *Nature* 227 (5258): 561–63.

Crooke, S. T., J. L. Witztum, C. F. Bennett, and B. F. Baker. 2018. “RNA-Targeted Therapeutics.” *Cell Metabolism* 27 (4): 714–39.

Darnell, James E., Jr. 1978. “Implications of RNA·RNA Splicing in Evolution of Eukaryotic Cells.” *Science*, December. https://doi.org/10.1126/science.364651.

Darnell, J. E., R. Wall, and R. J. Tushinski. 1971. “An Adenylic Acid-Rich Sequence in Messenger RNA of HeLa Cells and Its Possible Relationship to Reiterated Sites in DNA.” *Proceedings of the National Academy of Sciences of the United States of America* 68 (6): 1321.

De Conti, Laura, Marco Baralle, and Emanuele Buratti. 2013. “Exon and Intron Definition in Pre-mRNA Splicing.” *Wiley Interdisciplinary Reviews. RNA* 4 (1): 49–60.

Del Gatto-Konczak, F., M. Olive, M. C. Gesnel, and R. Breathnach. 1999. “hnRNP A1 Recruited to an Exon in Vivo Can Function as an Exon Splicing Silencer.” *Molecular and Cellular Biology* 19 (1). https://doi.org/10.1128/MCB.19.1.251.

Deutekom, J. C. van, M. Bremmer-Bout, A. A. Janson, I. B. Ginjaar, F. Baas, J. T. den Dunnen, and G. J. van Ommen. 2001. “Antisense-Induced Exon Skipping Restores Dystrophin Expression in DMD Patient Derived Muscle Cells.” *Human Molecular Genetics* 10 (15). https://doi.org/10.1093/hmg/10.15.1547.

Dirksen, W. P., Q. Sun, and F. M. Rottman. 1995. “Multiple Splicing Signals Control Alternative Intron Retention of Bovine Growth Hormone Pre-mRNA.” *The Journal of Biological Chemistry* 270 (10). https://doi.org/10.1074/jbc.270.10.5346.

Disset, A., C. F. Bourgeois, N. Benmalek, M. Claustres, J. Stevenin, and S. Tuffery-Giraud. 2006. “An Exon Skipping-Associated Nonsense Mutation in the Dystrophin Gene Uncovers a Complex Interplay between Multiple Antagonistic Splicing Elements.” *Human Molecular Genetics* 15 (6). https://doi.org/10.1093/hmg/ddl015.

Disterer, P., A. Kryczka, Y. Liu, Y. E. Badi, J. J. Wong, J. S. Owen, and B. Khoo. 2014. “Development of Therapeutic Splice-Switching Oligonucleotides.” *Human Gene Therapy* 25 (7). https://doi.org/10.1089/hum.2013.234.

Dominski, Z., and R. Kole. 1993. “Restoration of Correct Splicing in Thalassemic Pre-mRNA by Antisense Oligonucleotides.” *Proceedings of the National Academy of Sciences of the United States of America* 90 (18). https://doi.org/10.1073/pnas.90.18.8673.

Dong, Haiyang, Lei Li, Xiaohua Zhu, Jilong Shi, Ying Fu, Bingbing Xu, Jian Zhang, and Yongfeng Jin. 2021. “Complex RNA Secondary Structures Mediate Mutually Exclusive Splicing of Coleoptera Dscam1.” *Frontiers in Genetics* 12 (March): 644238.

Draper, David E. 2004. “A Guide to Ions and RNA Structure.” *RNA* 10 (3): 335–43.

Dunckley, M. G., M. Manoharan, P. Villiet, I. C. Eperon, and G. Dickson. 1998. "Modification of Splicing in the Dystrophin Gene in Cultured Mdx Muscle Cells by Antisense Oligoribonucleotides." *Human Molecular Genetics* 7 (7). https://doi.org/10.1093/hmg/7.7.1083.

Early, P., J. Rogers, M. Davis, K. Calame, M. Bond, R. Wall, and L. Hood. 1980. "Two mRNAs Can Be Produced from a Single Immunoglobulin Mu Gene by Alternative RNA Processing Pathways." *Cell* 20 (2). https://doi.org/10.1016/0092-8674(80)90617-0.

Echigoya, Y., and T. Yokota. 2014. "Skipping Multiple Exons of Dystrophin Transcripts Using Cocktail Antisense Oligonucleotides." *Nucleic Acid Therapeutics* 24 (1). https://doi.org/10.1089/nat.2013.0451.

Edmonds, M., M. H. Vaughan, and H. Nakazato. 1971. "Polyadenylic Acid Sequences in the Heterogeneous Nuclear RNA and Rapidly-Labeled Polyribosomal RNA of HeLa Cells: Possible Evidence for a Precursor Relationship." *Proceedings of the National Academy of Sciences of the United States of America* 68 (6). https://doi.org/10.1073/pnas.68.6.1336.

Egli, Martin, and Muthiah Manoharan. 2023. "Chemistry, Structure and Function of Approved Oligonucleotide Therapeutics." *Nucleic Acids Research* 51 (6): 2529–73.

Eldridge, Adam G., Yu Li, Phillip A. Sharp, and Benjamin J. Blencowe. 1999. "The SRm160/300 Splicing Coactivator Is Required for Exon-Enhancer Function." *Proceedings of the National Academy of Sciences of the United States of America* 96 (11): 6125.

Ellington, A. D., and J. W. Szostak. 1990. "In Vitro Selection of RNA Molecules That Bind Specific Ligands." *Nature* 346 (6287). https://doi.org/10.1038/346818a0.

Eng, L., G. Coutinho, S. Nahas, G. Yeo, R. Tanouye, M. Babaei, T. Dörk, C. Burge, and R. A. Gatti. 2004. "Nonclassical Splicing Mutations in the Coding and Noncoding Regions of the ATM Gene: Maximum Entropy Estimates of Splice Junction Strengths." *Human Mutation* 23 (1). https://doi.org/10.1002/humu.10295.

Eperon, Lucy P., Ian R. Graham, Andrew D. Griffiths, and Ian C. Eperon. 1988. "Effects of RNA Secondary Structure on Alternative Splicing of Pre-mRNA: Is Folding Limited to a Region behind the Transcribing RNA Polymerase?" *Cell* 54 (3): 393–401.

Expert-Bezançon, A., A. Sureau, P. Durosay, R. Salesse, H. Groeneveld, J. P. Lecaer, and J. Marie. 2004a. “hnRNP A1 and the SR Proteins ASF/SF2 and SC35 Have Antagonistic Functions in Splicing of Beta-Tropomyosin Exon 6B.” *The Journal of Biological Chemistry* 279 (37). https://doi.org/10.1074/jbc.M405377200.

———. 2004b. “hnRNP A1 and the SR Proteins ASF/SF2 and SC35 Have Antagonistic Functions in Splicing of Beta-Tropomyosin Exon 6B.” *The Journal of Biological Chemistry* 279 (37). https://doi.org/10.1074/jbc.M405377200.

Fackenthal, James D., and Lucy A. Godley. 2008. “Aberrant RNA Splicing and Its Functional Consequences in Cancer Cells.” *Disease Models & Mechanisms* 1 (1): 37.

Fairbrother, W. G., R. F. Yeh, P. A. Sharp, and C. B. Burge. 2002. “Predictive Identification of Exonic Splicing Enhancers in Human Genes.” *Science* 297 (5583). https://doi.org/10.1126/science.1073774.

Fairman, Connor W., Andrew M. L. Lever, and Julia C. Kenyon. 2021. “Evaluating RNA Structural Flexibility: Viruses Lead the Way.” *Viruses* 13 (11). https://doi.org/10.3390/v13112130.

Faustino, N. A., and T. A. Cooper. 2003a. “Pre-mRNA Splicing and Human Disease.” *Genes & Development* 17 (4). https://doi.org/10.1101/gad.1048803.

———. 2003b. “Pre-mRNA Splicing and Human Disease.” *Genes & Development* 17 (4). https://doi.org/10.1101/gad.1048803.

Finkel, R. S., C. A. Chiriboga, J. Vajsar, J. W. Day, J. Montes, D. C. De Vivo, M. Yamashita, et al. 2016. “Treatment of Infantile-Onset Spinal Muscular Atrophy with Nusinersen: A Phase 2, Open-Label, Dose-Escalation Study.” *Lancet* 388 (10063). https://doi.org/10.1016/S0140-6736(16)31408-8.

Fire, Andrew, Siqun Xu, Mary K. Montgomery, Steven A. Kostas, Samuel E. Driver, and Craig C. Mello. 1998. “Potent and Specific Genetic Interference by Double-Stranded RNA in Caenorhabditis Elegans.” *Nature* 391 (6669): 806–11.

Fong, N., H. Kim, Y. Zhou, X. Ji, J. Qiu, T. Saldi, K. Diener, K. Jones, X. D. Fu, and D. L. Bentley. 2014. “Pre-mRNA Splicing Is Facilitated by an Optimal RNA Polymerase II Elongation Rate.” *Genes & Development* 28 (23). https://doi.org/10.1101/gad.252106.114.

Fredericks, Alger M., Kamil J. Cygan, Brian A. Brown, and William G. Fairbrother. 2015a. “RNA-Binding Proteins: Splicing Factors and Disease.” *Biomolecules* 5 (2): 893–909.

———. 2015b. “RNA-Binding Proteins: Splicing Factors and Disease.” *Biomolecules* 5 (2): 893–909.

Frendewey, D., and W. Keller. 1985. “Stepwise Assembly of a Pre-mRNA Splicing Complex Requires U-snRNPs and Specific Intron Sequences.” *Cell* 42 (1). https://doi.org/10.1016/s0092-8674(85)80131-8.

Fukunaga, Tsukasa, Haruka Ozaki, Goro Terai, Kiyoshi Asai, Wataru Iwasaki, and Hisanori Kiryu. 2014. “CapR: Revealing Structural Specificities of RNA-Binding Protein Target Recognition Using CLIP-Seq Data.” *Genome Biology* 15 (1): 1–15.

Fu, X. D. 1995. “The Superfamily of Arginine/serine-Rich Splicing Factors.” *RNA* 1 (7): 663.

Fu, X. D., and M. Ares Jr. 2014. “Context-Dependent Control of Alternative Splicing by RNA-Binding Proteins.” *Nature Reviews. Genetics* 15 (10): 689–701.

Gagliardi, M., and A. T. Ashizawa. 2021. “The Challenges and Strategies of Antisense Oligonucleotide Drug Delivery.” *Biomedicines* 9 (4). https://doi.org/10.3390/biomedicines9040433.

Galej, W. P., M. E. Wilkinson, S. M. Fica, C. Oubridge, A. J. Newman, and K. Nagai. 2016. “Cryo-EM Structure of the Spliceosome Immediately after Branching.” *Nature* 537 (7619). https://doi.org/10.1038/nature19316.

Garber, K. 2016. “Big Win Possible for Ionis/Biogen Antisense Drug in Muscular Atrophy.” *Nature Biotechnology* 34 (10). https://doi.org/10.1038/nbt1016-1002.

Garcia-Blanco, M. A., A. P. Baraniak, and E. L. Lasda. 2004a. “Alternative Splicing in Disease and Therapy.” *Nature Biotechnology* 22 (5). https://doi.org/10.1038/nbt964.

———. 2004b. “Alternative Splicing in Disease and Therapy.” *Nature Biotechnology* 22 (5). https://doi.org/10.1038/nbt964.

Garcia-Lopez, A., F. Tessaro, H. R. A. Jonker, A. Wacker, C. Richter, A. Comte, N. Berntenis, et al. 2018. “Targeting RNA Structure in SMN2 Reverses Spinal Muscular Atrophy Molecular Phenotypes.” *Nature Communications* 9 (1). https://doi.org/10.1038/s41467-018-04110-1.

Gardner, Paul P., and Robert Giegerich. 2004. “A Comprehensive Comparison of Comparative RNA Structure Prediction Approaches.” *BMC Bioinformatics* 5 (1):

1–18.

Geary, R. S. 2009. “Antisense Oligonucleotide Pharmacokinetics and Metabolism.” *Expert Opinion on Drug Metabolism & Toxicology* 5 (4). https://doi.org/10.1517/17425250902877680.

Gilbert, Walter. 1978a. “Why Genes in Pieces?” Nature Publishing Group UK. February 1, 1978. https://doi.org/10.1038/271501a0.

———. 1978b. “Why Genes in Pieces?” Nature Publishing Group UK. February 1, 1978. https://doi.org/10.1038/271501a0.

Glidden, David T., Jeramiah L. Buerer, Camillo F. Saueressig, and William G. Fairbrother. 2021. “Hotspot Exons Are Common Targets of Splicing Perturbations.” *Nature Communications* 12 (1): 1–10.

Goguel, V., and M. Rosbash. 1993. “Splice Site Choice and Splicing Efficiency Are Positively Influenced by Pre-mRNA Intramolecular Base Pairing in Yeast.” *Cell* 72 (6). https://doi.org/10.1016/0092-8674(93)90578-e.

Grabowski, P. J., S. R. Seiler, and P. A. Sharp. 1985. “A Multicomponent Complex Is Involved in the Splicing of Messenger RNA Precursors.” *Cell* 42 (1). https://doi.org/10.1016/s0092-8674(85)80130-6.

Graveley, B. R. 2000. “Sorting out the Complexity of SR Protein Functions.” *RNA* 6 (9). https://doi.org/10.1017/s1355838200000960.

———. 2005. “Mutually Exclusive Splicing of the Insect Dscam Pre-mRNA Directed by Competing Intronic RNA Secondary Structures.” *Cell* 123 (1). https://doi.org/10.1016/j.cell.2005.07.028.

Graveley, B. R., K. J. Hertel, and T. Maniatis. 2001a. “The Role of U2AF35 and U2AF65 in Enhancer-Dependent Splicing.” *RNA* 7 (6). https://doi.org/10.1017/s1355838201010317.

———. 2001b. “The Role of U2AF35 and U2AF65 in Enhancer-Dependent Splicing.” *RNA* 7 (6). https://doi.org/10.1017/s1355838201010317.

Greider, C. W., and E. H. Blackburn. 1985. “Identification of a Specific Telomere Terminal Transferase Activity in Tetrahymena Extracts.” *Cell* 43 (2 Pt 1). https://doi.org/10.1016/0092-8674(85)90170-9.

———. 1987. “The Telomere Terminal Transferase of Tetrahymena Is a Ribonucleoprotein Enzyme with Two Kinds of Primer Specificity.” *Cell* 51 (6).

https://doi.org/10.1016/0092-8674(87)90576-9.

———. 1989. “A Telomeric Sequence in the RNA of Tetrahymena Telomerase Required for Telomere Repeat Synthesis.” *Nature* 337 (6205). https://doi.org/10.1038/337331a0.

Grover, A., H. Houlden, M. Baker, J. Adamson, J. Lewis, G. Prihar, S. Pickering-Brown, K. Duff, and M. Hutton. 1999. “5’ Splice Site Mutations in Tau Associated with the Inherited Dementia FTDP-17 Affect a Stem-Loop Structure That Regulates Alternative Splicing of Exon 10.” *The Journal of Biological Chemistry* 274 (21): 15134–43.

Guil, Sònia, Renata Gattoni, Montserrat Carrascal, Joaquín Abián, James Stévenin, and Montse Bach-Elias. 2003. “Roles of hnRNP A1, SR Proteins, and p68 Helicase in c-H-Ras Alternative Splicing Regulation.” *Molecular and Cellular Biology* 23 (8): 2927.

Gutell, R. R. 2014. “Ten Lessons with Carl Woese about RNA and Comparative Analysis.” *RNA Biology* 11 (3). https://doi.org/10.4161/rna.28718.

Haché, M., K. J. Swoboda, N. Sethna, A. Farrow-Gillespie, A. Khandji, S. Xia, and K. M. Bishop. 2016. “Intrathecal Injections in Children With Spinal Muscular Atrophy: Nusinersen Clinical Trial Experience.” *Journal of Child Neurology* 31 (7). https://doi.org/10.1177/0883073815627882.

Halvorsen, M., J. S. Martin, S. Broadaway, and A. Laederach. 2010. “Disease-Associated Mutations That Alter the RNA Structural Ensemble.” *PLoS Genetics* 6 (8). https://doi.org/10.1371/journal.pgen.1001074.

Havens, M. A., D. M. Duelli, and M. L. Hastings. 2013a. “Targeting RNA Splicing for Disease Therapy.” *Wiley Interdisciplinary Reviews. RNA* 4 (3). https://doi.org/10.1002/wrna.1158.

———. 2013b. “Targeting RNA Splicing for Disease Therapy.” *Wiley Interdisciplinary Reviews. RNA* 4 (3). https://doi.org/10.1002/wrna.1158.

———. 2013c. “Targeting RNA Splicing for Disease Therapy.” *Wiley Interdisciplinary Reviews. RNA* 4 (3). https://doi.org/10.1002/wrna.1158.

Havens, Mallory A., and Michelle L. Hastings. 2016. “Splice-Switching Antisense Oligonucleotides as Therapeutic Drugs.” *Nucleic Acids Research* 44 (14): 6549.

Hermann, T., and D. J. Patel. 2000. “RNA Bulges as Architectural and Recognition Motifs.” *Structure* 8 (3). https://doi.org/10.1016/s0969-2126(00)00110-6.

Hicks, Martin J., Chin-Rang Yang, Matthew V. Kotlajich, and Klemens J. Hertel. 2006. “Linking Splicing to Pol II Transcription Stabilizes Pre-mRNAs and Influences Splicing Patterns.” *PLoS Biology* 4 (6). https://doi.org/10.1371/journal.pbio.0040147.

Hollander, D., S. Naftelberg, G. Lev-Maor, A. R. Kornblihtt, and G. Ast. 2016. “How Are Short Exons Flanked by Long Introns Defined and Committed to Splicing?” *Trends in Genetics: TIG* 32 (10). https://doi.org/10.1016/j.tig.2016.07.003.

Holland, William L., Joseph T. Brozinick, Li-Ping Wang, Eric D. Hawkins, Katherine M. Sargent, Yanqi Liu, Krishna Narra, et al. 2007. “Inhibition of Ceramide Synthesis Ameliorates Glucocorticoid-, Saturated-Fat-, and Obesity-Induced Insulin Resistance.” *Cell Metabolism* 5 (3): 167–79.

Holm, Lise L., Thomas K. Doktor, Katharina K. Flugt, Ulrika S. S. Petersen, Rikke Petersen, and Brage S. Andresen. 2024. “All Exons Are Not Created Equal—exon Vulnerability Determines the Effect of Exonic Mutations on Splicing.” *Nucleic Acids Research*, February, gkae077.

Hong, W., Y. Shi, B. Xu, and Y. Jin. 2020. “RNA Secondary Structures in Dscam1 Mutually Exclusive Splicing: Unique Evolutionary Signature from the Midge.” *RNA* 26 (9). https://doi.org/10.1261/rna.075259.120.

Hong, W., J. Zhang, H. Dong, Y. Shi, H. Ma, F. Zhou, B. Xu, et al. 2021. “Intron-Targeted Mutagenesis Reveals Roles for Dscam1 RNA Pairing Architecture-Driven Splicing Bias in Neuronal Wiring.” *Cell Reports* 36 (2). https://doi.org/10.1016/j.celrep.2021.109373.

Hong, X., D. G. Scofield, and M. Lynch. 2006. “Intron Size, Abundance, and Distribution within Untranslated Regions of Genes.” *Molecular Biology and Evolution* 23 (12): 2392–2404.

Huang, Chao, and Yi-Tao Yu. 2013. “Synthesis and Labeling of RNA In Vitro.” *Current Protocols in Molecular Biology / Edited by Frederick M. Ausubel ... [et Al.]* 0 4 (April): Unit4.15.

Hua, Yimin, Timothy A. Vickers, Brenda F. Baker, C. Frank Bennett, and Adrian R. Krainer. 2007. “Enhancement of SMN2 Exon 7 Inclusion by Antisense Oligonucleotides Targeting the Exon.” *PLoS Biology* 5 (4). https://doi.org/10.1371/journal.pbio.0050073.

Hua, Yimin, Timothy A. Vickers, Hazeem L. Okunola, C. Frank Bennett, and Adrian R. Krainer. 2008. “Antisense Masking of an hnRNP A1/A2 Intronic Splicing

Silencer Corrects SMN2 Splicing in Transgenic Mice." *American Journal of Human Genetics* 82 (4): 834–48.

Hua, Y., K. Sahashi, G. Hung, F. Rigo, M. A. Passini, C. F. Bennett, and A. R. Krainer. 2010a. "Antisense Correction of SMN2 Splicing in the CNS Rescues Necrosis in a Type III SMA Mouse Model." *Genes & Development* 24 (15). https://doi.org/10.1101/gad.1941310.

———. 2010b. "Antisense Correction of SMN2 Splicing in the CNS Rescues Necrosis in a Type III SMA Mouse Model." *Genes & Development* 24 (15). https://doi.org/10.1101/gad.1941310.

Hua, Y., K. Sahashi, F. Rigo, G. Hung, G. Horev, C. F. Bennett, and A. R. Krainer. 2011. "Peripheral SMN Restoration Is Essential for Long-Term Rescue of a Severe Spinal Muscular Atrophy Mouse Model." *Nature* 478 (7367). https://doi.org/10.1038/nature10485.

Hu, J., J. Liu, and D. R. Corey. 2010. "Allele-Selective Inhibition of Huntingtin Expression by Switching to an miRNA-like RNAi Mechanism." *Chemistry & Biology* 17 (11). https://doi.org/10.1016/j.chembiol.2010.10.013.

Hutton, Mike, Corinne L. Lendon, Patrizia Rizzu, Matt Baker, Susanne Froelich, Henry Houlden, Stuart Pickering-Brown, et al. 1998. "Association of Missense and 5′-Splice-Site Mutations in Tau with the Inherited Dementia FTDP-17." *Nature* 393 (6686): 702–5.

Incarnato, D., E. Morandi, L. M. Simon, and S. Oliviero. 2018. "RNA Framework: An All-in-One Toolkit for the Analysis of RNA Structures and Post-Transcriptional Modifications." *Nucleic Acids Research* 46 (16). https://doi.org/10.1093/nar/gky486.

Iqbal, K., F. Liu, C. X. Gong, Adel C. Alonso, and I. Grundke-Iqbal. 2009. "Mechanisms of Tau-Induced Neurodegeneration." *Acta Neuropathologica* 118 (1). https://doi.org/10.1007/s00401-009-0486-3.

Izquierdo, José-María, and Juan Valcárcel. 2006. "A Simple Principle to Explain the Evolution of Pre-mRNA Splicing." *Genes & Development* 20 (13): 1679–84.

Jadhav, Vasant, Akshay Vaishnaw, Kevin Fitzgerald, and Martin A. Maier. 2024. "RNA Interference in the Era of Nucleic Acid Therapeutics." *Nature Biotechnology*, February, 1–12.

Jiang, Z., J. Cote, J. M. Kwon, A. M. Goate, and J. Y. Wu. 2000. "Aberrant Splicing of Tau Pre-mRNA Caused by Intronic Mutations Associated with the Inherited

Dementia Frontotemporal Dementia with Parkinsonism Linked to Chromosome 17." *Molecular and Cellular Biology* 20 (11): 4036–48.

Jin, Y., Y. Yang, and P. Zhang. 2011. "New Insights into RNA Secondary Structure in the Alternative Splicing of Pre-mRNAs." *RNA Biology* 8 (3). https://doi.org/10.4161/rna.8.3.15388.

Johannes, Niklas, Matti Vuorre, and Andrew K. Przybylski. 2021. "Video Game Play Is Positively Correlated with Well-Being." *Royal Society Open Science*, February. https://doi.org/10.1098/rsos.202049.

Jones, A. N., C. Graß, I. Meininger, A. Geerlof, M. Klostermann, K. Zarnack, D. Krappmann, and M. Sattler. 2022. "Modulation of Pre-mRNA Structure by hnRNP Proteins Regulates Alternative Splicing of MALT1." *Science Advances* 8 (31). https://doi.org/10.1126/sciadv.abp9153.

Juliano, R. L. 2016. "The Delivery of Therapeutic Oligonucleotides." *Nucleic Acids Research* 44 (14). https://doi.org/10.1093/nar/gkw236.

Jurica, M. S., and M. J. Moore. 2003. "Pre-mRNA Splicing: Awash in a Sea of Proteins." *Molecular Cell* 12 (1). https://doi.org/10.1016/s1097-2765(03)00270-3.

Kalbfuss, B., S. A. Mabon, and T. Misteli. 2001. "Correction of Alternative Splicing of Tau in Frontotemporal Dementia and Parkinsonism Linked to Chromosome 17." *The Journal of Biological Chemistry* 276 (46): 42986–93.

Kalinina, M., D. Skvortsov, S. Kalmykova, T. Ivanov, O. Dontsova, and D. D. Pervouchine. 2021. "Multiple Competing RNA Structures Dynamically Control Alternative Splicing in the Human ATE1 Gene." *Nucleic Acids Research* 49 (1). https://doi.org/10.1093/nar/gkaa1208.

Kanopka, A., O. Mühlemann, and G. Akusjärvi. 1996. "Inhibition by SR Proteins of Splicing of a Regulated Adenovirus Pre-mRNA." *Nature* 381 (6582): 535–38.

Kar, A., K. Fushimi, X. Zhou, P. Ray, C. Shi, X. Chen, Z. Liu, S. Chen, and J. Y. Wu. 2011. "RNA Helicase p68 (DDX5) Regulates Tau Exon 10 Splicing by Modulating a Stem-Loop Structure at the 5' Splice Site." *Molecular and Cellular Biology* 31 (9). https://doi.org/10.1128/MCB.01149-10.

Kashima, Tsuyoshi, and James L. Manley. 2003a. "A Negative Element in SMN2 Exon 7 Inhibits Splicing in Spinal Muscular Atrophy." *Nature Genetics* 34 (4): 460–63.

———. 2003b. “A Negative Element in SMN2 Exon 7 Inhibits Splicing in Spinal Muscular Atrophy.” *Nature Genetics* 34 (4): 460–63.

Kastner, Berthold, Cindy L. Will, Holger Stark, and Reinhard Lührmann. 2019. “Structural Insights into Nuclear Pre-mRNA Splicing in Higher Eukaryotes.” *Cold Spring Harbor Perspectives in Biology* 11 (11): a032417.

Kedes, D. H., and J. A. Steitz. 1987. “Accurate 5’ Splice-Site Selection in Mouse Kappa Immunoglobulin Light Chain Premessenger RNAs Is Not Cell-Type-Specific.” *Proceedings of the National Academy of Sciences of the United States of America* 84 (22): 7928.

———. 1988. “Correct in Vivo Splicing of the Mouse Immunoglobulin Kappa Light-Chain Pre-mRNA Is Dependent on 5’ Splice-Site Position Even in the Absence of Transcription.” *Genes & Development* 2 (11). https://doi.org/10.1101/gad.2.11.1448.

Ke, S., S. Shang, S. M. Kalachikov, I. Morozova, L. Yu, J. J. Russo, J. Ju, and L. A. Chasin. 2011. “Quantitative Evaluation of All Hexamers as Exonic Splicing Elements.” *Genome Research* 21 (8). https://doi.org/10.1101/gr.119628.110.

Kim, H. K., M. H. C. Pham, K. S. Ko, B. D. Rhee, and J. Han. 2018. “Alternative Splicing Isoforms in Health and Disease.” *Pflugers Archiv: European Journal of Physiology* 470 (7). https://doi.org/10.1007/s00424-018-2136-x.

Kim, J., C. Hu, Moufawad El Achkar C, L. E. Black, J. Douville, A. Larson, M. K. Pendergast, et al. 2019. “Patient-Customized Oligonucleotide Therapy for a Rare Genetic Disease.” *The New England Journal of Medicine* 381 (17). https://doi.org/10.1056/NEJMoa1813279.

Kim, Sang Hyon, Olga A. Koroleva, Dominika Lewandowska, Ali F. Pendle, Gillian P. Clark, Craig G. Simpson, Peter J. Shaw, and John W. S. Brown. 2009. “Aberrant mRNA Transcripts and the Nonsense-Mediated Decay Proteins UPF2 and UPF3 Are Enriched in the Arabidopsis Nucleolus.” *The Plant Cell* 21 (7): 2045.

Kitamura, Koji, and Keisuke Nimura. 2021. “Regulation of RNA Splicing: Aberrant Splicing Regulation and Therapeutic Targets in Cancer.” *Cells* 10 (4). https://doi.org/10.3390/cells10040923.

Köhler, Wolfgang, Julian Curiel, and Adeline Vanderver. 2018. “Adulthood Leukodystrophies.” *Nature Reviews. Neurology* 14 (2): 94–105.

Kohtz, J. D., S. F. Jamison, C. L. Will, P. Zuo, R. Lührmann, M. A. Garcia-Blanco,

and J. L. Manley. 1994. “Protein-Protein Interactions and 5’-Splice-Site Recognition in Mammalian mRNA Precursors.” *Nature* 368 (6467). https://doi.org/10.1038/368119a0.

Kolb, S. J., D. J. Battle, and G. Dreyfuss. 2007. “Molecular Functions of the SMN Complex.” *Journal of Child Neurology* 22 (8). https://doi.org/10.1177/0883073807305666.

Kolb, S. J., and J. T. Kissel. 2015. “Spinal Muscular Atrophy.” *Neurologic Clinics* 33 (4). https://doi.org/10.1016/j.ncl.2015.07.004.

Kole, R., R. R. Shukla, and S. Akhtar. 1991. “Pre-mRNA Splicing as a Target for Antisense Oligonucleotides.” *Advanced Drug Delivery Reviews* 6 (3): 271–86.

Kollins, S. H., D. J. DeLoss, E. Cañadas, J. Lutz, R. L. Findling, R. S. E. Keefe, J. N. Epstein, A. J. Cutler, and S. V. Faraone. 2020. “A Novel Digital Intervention for Actively Reducing Severity of Paediatric ADHD (STARS-ADHD): A Randomised Controlled Trial.” *The Lancet. Digital Health* 2 (4). https://doi.org/10.1016/S2589-7500(20)30017-0.

Konarska, M. M., and P. A. Sharp. 1987. “Interactions between Small Nuclear Ribonucleoprotein Particles in Formation of Spliceosomes.” *Cell* 49 (6). https://doi.org/10.1016/0092-8674(87)90614-3.

Kornblihtt, A. R., M. de la Mata, J. P. Fededa, M. J. Munoz, and G. Nogues. 2004. “Multiple Links between Transcription and Splicing.” *RNA* 10 (10). https://doi.org/10.1261/rna.7100104.

Krainer, A. R., T. Maniatis, B. Ruskin, and M. R. Green. 1984. “Normal and Mutant Human Beta-Globin Pre-mRNAs Are Faithfully and Efficiently Spliced in Vitro.” *Cell* 36 (4). https://doi.org/10.1016/0092-8674(84)90049-7.

Krawczak, M., N. S. Thomas, B. Hundrieser, M. Mort, M. Wittig, J. Hampe, and D. N. Cooper. 2007a. “Single Base-Pair Substitutions in Exon-Intron Junctions of Human Genes: Nature, Distribution, and Consequences for mRNA Splicing.” *Human Mutation* 28 (2). https://doi.org/10.1002/humu.20400.

———. 2007b. “Single Base-Pair Substitutions in Exon-Intron Junctions of Human Genes: Nature, Distribution, and Consequences for mRNA Splicing.” *Human Mutation* 28 (2). https://doi.org/10.1002/humu.20400.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, et al. 2001. “Initial Sequencing and Analysis of the Human Genome.” *Nature* 409 (6822): 860–921.

Larson, J. L., A. J. Silver, D. Chan, C. Borroto, B. Spurrier, and L. M. Silver. 2015. "Validation of a High Resolution NGS Method for Detecting Spinal Muscular Atrophy Carriers among Phase 3 Participants in the 1000 Genomes Project." *BMC Medical Genetics* 16 (October). https://doi.org/10.1186/s12881-015-0246-2.

Lavigueur, A., H. La Branche, A. R. Kornblihtt, and B. Chabot. 1993. "A Splicing Enhancer in the Human Fibronectin Alternate ED1 Exon Interacts with SR Proteins and Stimulates U2 snRNP Binding." *Genes & Development* 7 (12A). https://doi.org/10.1101/gad.7.12a.2405.

Lee, J., W. Kladwang, M. Lee, D. Cantu, M. Azizyan, H. Kim, A. Limpaecher, S. Yoon, A. Treuille, and R. Das. 2014. "RNA Design Rules from a Massive Open Laboratory." *Proceedings of the National Academy of Sciences of the United States of America* 111 (6). https://doi.org/10.1073/pnas.1313039111.

Lee, S. Y., J. Mendecki, and G. Brawerman. 1971. "A Polynucleotide Segment Rich in Adenylic Acid in the Rapidly-Labeled Polyribosomal RNA Component of Mouse Sarcoma 180 Ascites Cells." *Proceedings of the National Academy of Sciences of the United States of America* 68 (6). https://doi.org/10.1073/pnas.68.6.1331.

Lefebvre, S., L. Bürglen, S. Reboullet, O. Clermont, P. Burlet, L. Viollet, B. Benichou, C. Cruaud, P. Millasseau, and M. Zeviani. 1995. "Identification and Characterization of a Spinal Muscular Atrophy-Determining Gene." *Cell* 80 (1). https://doi.org/10.1016/0092-8674(95)90460-3.

Lerner, M. R., and J. A. Steitz. 1979. "Antibodies to Small Nuclear RNAs Complexed with Proteins Are Produced by Patients with Systemic Lupus Erythematosus." *Proceedings of the National Academy of Sciences of the United States of America* 76 (11). https://doi.org/10.1073/pnas.76.11.5495.

Lim, K. H., L. Ferraris, M. E. Filloux, B. J. Raphael, and W. G. Fairbrother. 2011. "Using Positional Distribution to Identify Splicing Elements and Predict Pre-mRNA Processing Defects in Human Genes." *Proceedings of the National Academy of Sciences of the United States of America* 108 (27): 11093–98.

Lim, Kian Huat, Luciana Ferraris, Madeleine E. Filloux, Benjamin J. Raphael, and William G. Fairbrother. 2011. "Using Positional Distribution to Identify Splicing Elements and Predict Pre-mRNA Processing Defects in Human Genes." *Proceedings of the National Academy of Sciences of the United States of America* 108 (27): 11093–98.

Liu, H. X., M. Zhang, and A. R. Krainer. 1998. “Identification of Functional Exonic Splicing Enhancer Motifs Recognized by Individual SR Proteins.” *Genes & Development* 12 (13): 1998–2012.

Liu, J., J. Hu, and D. R. Corey. 2012. “Expanding the Action of Duplex RNAs into the Nucleus: Redirecting Alternative Splicing.” *Nucleic Acids Research* 40 (3). https://doi.org/10.1093/nar/gkr780.

Lord, Jenny, and Diana Baralle. 2021a. “Splicing in the Diagnosis of Rare Disease: Advances and Challenges.” *Frontiers in Genetics* 12. https://doi.org/10.3389/fgene.2021.689892.

———. 2021b. “Splicing in the Diagnosis of Rare Disease: Advances and Challenges.” *Frontiers in Genetics* 12. https://doi.org/10.3389/fgene.2021.689892.

Lorson, C. L., E. Hahnen, E. J. Androphy, and B. Wirth. 1999. “A Single Nucleotide in the SMN Gene Regulates Splicing and Is Responsible for Spinal Muscular Atrophy.” *Proceedings of the National Academy of Sciences of the United States of America* 96 (11). https://doi.org/10.1073/pnas.96.11.6307.

Manley, J. L., and R. Tacke. 1996. “SR Proteins and Splicing Control.” *Genes & Development* 10 (13). https://doi.org/10.1101/gad.10.13.1569.

Marinus, T., A. B. Fessler, C. A. Ogle, and D. Incarnato. 2021. “A Novel SHAPE Reagent Enables the Analysis of RNA Structure in Living Cells with Unprecedented Accuracy.” *Nucleic Acids Research* 49 (6). https://doi.org/10.1093/nar/gkaa1255.

Martinez-Contreras, R., P. Cloutier, L. Shkreta, J. F. Fisette, T. Revil, and B. Chabot. 2007. “hnRNP Proteins and Splicing Control.” *Advances in Experimental Medicine and Biology* 623. https://doi.org/10.1007/978-0-387-77374-2_8.

Martinez-Contreras, Rebeca, Jean-François Fisette, Faiz-Ul Hassan Nasim, Richard Madden, Mélanie Cordeau, and Benoit Chabot. 2006. “Intronic Binding Sites for hnRNP A/B and hnRNP F/H Proteins Stimulate Pre-mRNA Splicing.” *PLoS Biology* 4 (2): e21.

Mathews, D. H., and D. H. Turner. 2006. “Prediction of RNA Secondary Structure by Free Energy Minimization.” *Current Opinion in Structural Biology* 16 (3). https://doi.org/10.1016/j.sbi.2006.05.010.

Mayeda, A., D. M. Helfman, and A. R. Krainer. 1993a. “Modulation of Exon Skipping and Inclusion by Heterogeneous Nuclear Ribonucleoprotein A1 and

Pre-mRNA Splicing Factor SF2/ASF." *Molecular and Cellular Biology* 13 (5). https://doi.org/10.1128/mcb.13.5.2993-3001.1993.

———. 1993b. "Modulation of Exon Skipping and Inclusion by Heterogeneous Nuclear Ribonucleoprotein A1 and Pre-mRNA Splicing Factor SF2/ASF." *Molecular and Cellular Biology* 13 (5). https://doi.org/10.1128/mcb.13.5.2993-3001.1993.

Mayeda, A., and A. R. Krainer. 1992a. "Regulation of Alternative Pre-mRNA Splicing by hnRNP A1 and Splicing Factor SF2." *Cell* 68 (2). https://doi.org/10.1016/0092-8674(92)90477-t.

———. 1992b. "Regulation of Alternative Pre-mRNA Splicing by hnRNP A1 and Splicing Factor SF2." *Cell* 68 (2). https://doi.org/10.1016/0092-8674(92)90477-t.

May, Gemma E., Sara Olson, C. Joel McManus, and Brenton R. Graveley. 2011. "Competing RNA Secondary Structures Are Required for Mutually Exclusive Splicing of the Dscam Exon 6 Cluster." *RNA* 17 (2): 222.

Meininger, I., R. A. Griesbach, D. Hu, T. Gehring, T. Seeholzer, A. Bertossi, J. Kranich, et al. 2016. "Alternative Splicing of MALT1 Controls Signalling and Activation of CD4(+) T Cells." *Nature Communications* 7 (April). https://doi.org/10.1038/ncomms11292.

Mendell, J. R., L. R. Rodino-Klapac, Z. Sahenk, K. Roush, L. Bird, L. P. Lowes, L. Alfano, et al. 2013a. "Eteplirsen for the Treatment of Duchenne Muscular Dystrophy." *Annals of Neurology* 74 (5). https://doi.org/10.1002/ana.23982.

———. 2013b. "Eteplirsen for the Treatment of Duchenne Muscular Dystrophy." *Annals of Neurology* 74 (5). https://doi.org/10.1002/ana.23982.

Meyer, Irmtraud M., and István Miklós. 2004. "Co-Transcriptional Folding Is Encoded within RNA Genes." *BMC Molecular Biology* 5 (1): 1–10.

Michaud, S., and R. Reed. 1993. "A Functional Association between the 5' and 3' Splice Site Is Established in the Earliest Prespliceosome Complex (E) in Mammals." *Genes & Development* 7 (6). https://doi.org/10.1101/gad.7.6.1008.

Monani, U. R., C. L. Lorson, D. W. Parsons, T. W. Prior, E. J. Androphy, A. H. Burghes, and J. D. McPherson. 1999. "A Single Nucleotide Difference That Alters Splicing Patterns Distinguishes the SMA Gene SMN1 from the Copy Gene SMN2." *Human Molecular Genetics* 8 (7). https://doi.org/10.1093/hmg/8.7.1177.

Mount, S. M., I. Pettersson, M. Hinterberger, A. Karmas, and J. A. Steitz. 1983. “The U1 Small Nuclear RNA-Protein Complex Selectively Binds a 5’ Splice Site in Vitro.” *Cell* 33 (2). https://doi.org/10.1016/0092-8674(83)90432-4.

Movassat, M., E. Forouzmand, F. Reese, and K. J. Hertel. 2019. “Exon Size and Sequence Conservation Improves Identification of Splice-Altering Nucleotides.” *RNA* 25 (12). https://doi.org/10.1261/rna.070987.119.

Muro, A. F., M. Caputi, R. Pariyarath, F. Pagani, E. Buratti, and F. E. Baralle. 1999. “Regulation of Fibronectin EDA Exon Alternative Splicing: Possible Role of RNA Secondary Structure for Enhancer Display.” *Molecular and Cellular Biology* 19 (4). https://doi.org/10.1128/MCB.19.4.2657.

Nilsen, Timothy W., and Brenton R. Graveley. 2010. “Expansion of the Eukaryotic Proteome by Alternative Splicing.” *Nature* 463 (7280): 457–63.

Ni, Xiaohua, Mark Castanares, Amarnath Mukherjee, and Shawn E. Lupold. 2011. “Nucleic Acid Aptamers: Clinical Applications and Promising New Horizons.” *Current Medicinal Chemistry* 18 (27): 4206.

Noels, H., G. van Loo, S. Hagens, V. Broeckx, R. Beyaert, P. Marynen, and M. Baens. 2007. “A Novel TRAF6 Binding Site in MALT1 Defines Distinct Mechanisms of NF-κB Activation by API2·MALT1 Fusions.” *The Journal of Biological Chemistry* 282 (14): 10180–89.

Noller, Harry F., and Carl R. Woese. 1981. “Secondary Structure of 16S Ribosomal RNA.” *Science*, April. https://doi.org/10.1126/science.6163215.

O’Neill, T. J., T. Seeholzer, A. Gewies, T. Gehring, F. Giesert, I. Hamp, C. Graß, et al. 2021. “TRAF6 Prevents Fatal Inflammation by Homeostatic Suppression of MALT1 Protease.” *Science Immunology* 6 (65). https://doi.org/10.1126/sciimmunol.abh2095.

Ottesen, Eric W. 2017. “ISS-N1 Makes the First FDA-Approved Drug for Spinal Muscular Atrophy.” *Translational Neuroscience* 8: 1.

Pabit, S. A., J. L. Sutton, H. Chen, and L. Pollack. 2013. “Role of Ion Valence in the Submillisecond Collapse and Folding of a Small RNA Domain.” *Biochemistry* 52 (9). https://doi.org/10.1021/bi3016636.

Padgett, R. A., M. M. Konarska, P. J. Grabowski, S. F. Hardy, and P. A. Sharp. 1984. “Lariat RNA’s as Intermediates and Products in the Splicing of Messenger RNA Precursors.” *Science* 225 (4665). https://doi.org/10.1126/science.6206566.

Pan, Q., O. Shai, L. J. Lee, B. J. Frey, and B. J. Blencowe. 2008. “Deep Surveying of Alternative Splicing Complexity in the Human Transcriptome by High-Throughput Sequencing.” *Nature Genetics* 40 (12). https://doi.org/10.1038/ng.259.

Pan, Tao, and Tobin Sosnick. 2006. “RNA FOLDING DURING TRANSCRIPTION.” *Annual Review of Biophysics* 35 (Volume 35, 2006): 161–75.

Pan, T., I. Artsimovitch, X. W. Fang, R. Landick, and T. R. Sosnick. 1999. “Folding of a Large Ribozyme during Transcription and the Effect of the Elongation Factor NusA.” *Proceedings of the National Academy of Sciences of the United States of America* 96 (17). https://doi.org/10.1073/pnas.96.17.9545.

Pant, Devesh C., Imen Dorboz, Agatha Schluter, Stéphane Fourcade, Nathalie Launay, Javier Joya, Sergio Aguilera-Albesa, et al. 2019. “Loss of the Sphingolipid Desaturase DEGS1 Causes Hypomyelinating Leukodystrophy.” *The Journal of Clinical Investigation* 129 (3): 1240.

Parikh, S., G. Bernard, R. J. Leventer, M. S. van der Knaap, J. van Hove, A. Pizzino, N. H. McNeill, et al. 2015. “A Clinical Approach to the Diagnosis of Patients with Leukodystrophies and Genetic Leukoencephelopathies.” *Molecular Genetics and Metabolism* 114 (4). https://doi.org/10.1016/j.ymgme.2014.12.434.

Passini, M. A., J. Bu, A. M. Richards, C. Kinnecom, S. P. Sardi, L. M. Stanek, Y. Hua, et al. 2011. “Antisense Oligonucleotides Delivered to the Mouse CNS Ameliorate Symptoms of Severe Spinal Muscular Atrophy.” *Science Translational Medicine* 3 (72). https://doi.org/10.1126/scitranslmed.3001777.

Paz, I., I. Kosti, M. Ares, M. Cline, and Y. Mandel-Gutfreund. 2014. “RBPmap: A Web Server for Mapping Binding Sites of RNA-Binding Proteins.” *Nucleic Acids Research* 42 (Web Server issue). https://doi.org/10.1093/nar/gku406.

Pellizzoni, L., J. Yong, and G. Dreyfuss. 2002. “Essential Role for the SMN Complex in the Specificity of snRNP Assembly.” *Science* 298 (5599). https://doi.org/10.1126/science.1074962.

Plaschka, C., P. C. Lin, C. Charenton, and K. Nagai. 2018. “Prespliceosome Structure Provides Insights into Spliceosome Assembly and Regulation.” *Nature* 559 (7714). https://doi.org/10.1038/s41586-018-0323-8.

Pollard, A. J., A. R. Krainer, S. C. Robson, and G. N. Europe-Finner. 2002. “Alternative Splicing of the Adenylyl Cyclase Stimulatory G-Protein G Alpha(s) Is Regulated by SF2/ASF and Heterogeneous Nuclear Ribonucleoprotein A1

(hnRNPA1) and Involves the Use of an Unusual TG 3'-Splice Site." *The Journal of Biological Chemistry* 277 (18). https://doi.org/10.1074/jbc.M109046200.

Prior, T. W., A. R. Krainer, Y. Hua, K. J. Swoboda, P. C. Snyder, S. J. Bridgeman, A. H. Burghes, and J. T. Kissel. 2009. "A Positive Modifier of Spinal Muscular Atrophy in the SMN2 Gene." *American Journal of Human Genetics* 85 (3). https://doi.org/10.1016/j.ajhg.2009.08.002.

Prior, T. W., P. J. Snyder, B. D. Rink, D. K. Pearl, R. E. Pyatt, D. C. Mihal, T. Conlan, et al. 2010. "Newborn and Carrier Screening for Spinal Muscular Atrophy." *American Journal of Medical Genetics. Part A* 152A (7). https://doi.org/10.1002/ajmg.a.33474.

Puisac, B., M. E. Teresa-Rodrigo, M. Arnedo, M. C. Gil-Rodríguez, C. Pérez-Cerdá, A. Ribes, A. Pié, G. Bueno, P. Gómez-Puertas, and J. Pié. 2013. "Analysis of Aberrant Splicing and Nonsense-Mediated Decay of the Stop Codon Mutations c.109G>T and c.504_505delCT in 7 Patients with HMG-CoA Lyase Deficiency." *Molecular Genetics and Metabolism* 108 (4). https://doi.org/10.1016/j.ymgme.2013.01.019.

Qin, Daoming, Lei Huang, Alissa Wlodaver, Jorge Andrade, and Jonathan P. Staley. 2016. "Sequencing of Lariat Termini in S. Cerevisiae Reveals 5′ Splice Sites, Branch Points, and Novel Splicing Events." *RNA* 22 (2): 237.

Ramanathan, Muthukumar, Douglas F. Porter, and Paul A. Khavari. 2019. "Methods to Study RNA-Protein Interactions." *Nature Methods* 16 (3): 225.

Reddy, Ram, and Harris Busch. 1988. "Small Nuclear RNAs: RNA Sequences, Structure, and Modifications." *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*, 1–37.

Reed, R., and T. Maniatis. 1986. "A Role for Exon Sequences and Splice-Site Proximity in Splice-Site Selection." *Cell* 46 (5). https://doi.org/10.1016/0092-8674(86)90343-0.

Reuter, Jessica S., and David H. Mathews. 2010. "RNAstructure: Software for RNA Secondary Structure Prediction and Analysis." *BMC Bioinformatics* 11 (1): 1–9.

Rigo, F., P. P. Seth, and C. F. Bennett. 2014. "Antisense Oligonucleotide-Based Therapies for Diseases Caused by Pre-mRNA Processing Defects." *Advances in Experimental Medicine and Biology* 825. https://doi.org/10.1007/978-1-4939-1221-6_9.

Rivas, Elena. 2021. "Evolutionary Conservation of RNA Sequence and Structure."

Wiley Interdisciplinary Reviews. RNA 12 (5). https://doi.org/10.1002/wrna.1649.

Robberson, B. L., G. J. Cote, and S. M. Berget. 1990. "Exon Definition May Facilitate Splice Site Selection in RNAs with Multiple Exons." *Molecular and Cellular Biology* 10 (1). https://doi.org/10.1128/mcb.10.1.84-94.1990.

Rochette, C. F., N. Gilbert, and L. R. Simard. 2001. "SMN Gene Duplication and the Emergence of the SMN2 Gene Occurred in Distinct Hominids: SMN2 Is Unique to Homo Sapiens." *Human Genetics* 108 (3). https://doi.org/10.1007/s004390100473.

Rothrock, C., B. Cannon, B. Hahm, and K. W. Lynch. 2003. "A Conserved Signal-Responsive Sequence Mediates Activation-Induced Alternative Splicing of CD45." *Molecular Cell* 12 (5). https://doi.org/10.1016/s1097-2765(03)00434-9.

Roy, S., U. Delling, C. H. Chen, C. A. Rosen, and N. Sonenberg. 1990. "A Bulge Structure in HIV-1 TAR RNA Is Required for Tat Binding and Tat-Mediated Trans-Activation." *Genes & Development* 4 (8): 1365–73.

Ruskin, B., A. R. Krainer, T. Maniatis, and M. R. Green. 1984. "Excision of an Intact Intron as a Novel Lariat Structure during Pre-mRNA Splicing in Vitro." *Cell* 38 (1): 317–31.

Ruskin, B., P. D. Zamore, and M. R. Green. 1988. "A Factor, U2AF, Is Required for U2 snRNP Binding and Splicing Complex Assembly." *Cell* 52 (2). https://doi.org/10.1016/0092-8674(88)90509-0.

Rw, Holley, J. Apgar, Everett Ga, Madison Jt, M. Marquisee, Merrill Sh, Penswick Jr, and A. Zamir. 1965. "STRUCTURE OF A RIBONUCLEIC ACID." *Science* 147 (3664). https://doi.org/10.1126/science.147.3664.1462.

Salari, Raheleh, Chava Kimchi-Sarfaty, Michael M. Gottesman, and Teresa M. Przytycka. 2013. "Sensitive Measurement of Single-Nucleotide Polymorphism-Induced Changes of RNA Conformation: Application to Disease Studies." *Nucleic Acids Research* 41 (1): 44.

Saldi, Tassa, Nova Fong, and David L. Bentley. 2018. "Transcription Elongation Rate Affects Nascent Histone Pre-mRNA Folding and 3′ End Processing." *Genes & Development* 32 (3-4): 297–308.

Saldi, T., M. A. Cortazar, R. M. Sheridan, and D. L. Bentley. 2016. "Coupling of RNA Polymerase II Transcription Elongation with Pre-mRNA Splicing." *Journal of Molecular Biology* 428 (12): 2623–35.

Saldi, T., K. Riemondy, B. Erickson, and D. L. Bentley. 2021. “Alternative RNA Structures Formed during Transcription Depend on Elongation Rate and Modify RNA Processing.” *Molecular Cell* 81 (8). https://doi.org/10.1016/j.molcel.2021.01.040.

Schmitt, S., L. C. Castelvetri, and M. Simons. 2015. “Metabolism and Functions of Lipids in Myelin.” *Biochimica et Biophysica Acta* 1851 (8). https://doi.org/10.1016/j.bbalip.2014.12.016.

Senapathy, P. 1986. “Origin of Eukaryotic Introns: A Hypothesis, Based on Codon Distribution Statistics in Genes, and Its Implications.” *Proceedings of the National Academy of Sciences of the United States of America* 83 (7): 2133.

———. 1988. “Possible Evolution of Splice-Junction Signals in Eukaryotic Genes from Stop Codons.” *Proceedings of the National Academy of Sciences of the United States of America* 85 (4): 1129.

Senapathy, Periannan. 1995. “Introns and the Origin of Protein-Coding Genes.” *Science*, June. https://doi.org/10.1126/science.7761858.

Sharma, V. K., and J. K. Watts. 2015. “Oligonucleotide Therapeutics: Chemistry, Delivery and Clinical Progress.” *Future Medicinal Chemistry* 7 (16). https://doi.org/10.4155/fmc.15.144.

Sharp, S. J., J. Schaack, L. Cooley, D. J. Burke, and D. Söll. 1985. “Structure and Transcription of Eukaryotic tRNA Genes.” *CRC Critical Reviews in Biochemistry* 19 (2). https://doi.org/10.3109/10409238509082541.

Shepard, Samuel, Mark McCreary, and Alexei Fedorov. 2009. “The Peculiarities of Large Intron Splicing in Animals.” *PloS One* 4 (11). https://doi.org/10.1371/journal.pone.0007853.

Sheth, N., X. Roca, M. L. Hastings, T. Roeder, A. R. Krainer, and R. Sachidanandam. 2006. “Comprehensive Splice-Site Analysis Using Comparative Genomics.” *Nucleic Acids Research* 34 (14). https://doi.org/10.1093/nar/gkl556.

Siegfried, Nathan A., Steven Busan, Greggory M. Rice, Julie A. E. Nelson, and Kevin M. Weeks. 2014. “RNA Motif Discovery by SHAPE and Mutational Profiling (SHAPE-MaP).” *Nature Methods* 11 (9): 959.

Singh, Natalia N., Elliot J. Androphy, and Ravindra N. Singh. 2004. “An Extended Inhibitory Context Causes Skipping of Exon 7 of SMN2 in Spinal Muscular Atrophy.” *Biochemical and Biophysical Research Communications* 315 (2):

381–88.

Singh, Natalia N., Matthew D. Howell, Elliot J. Androphy, and Ravindra N. Singh. 2017. “How the Discovery of ISS-N1 Led to the First Medical Therapy for Spinal Muscular Atrophy.” *Gene Therapy* 24 (9): 520.

Singh, Natalia N., Diou Luo, and Ravindra N. Singh. 2018. “Pre-mRNA Splicing Modulation by Antisense Oligonucleotides.” *Methods in Molecular Biology* 1828: 415.

Singh, Natalia N., and Ravindra N. Singh. 2011. “Alternative Splicing in Spinal Muscular Atrophy Underscores the Role of an Intron Definition Model.” *RNA Biology* 8 (4): 600–606.

Singh, Natalia N., Ravindra N. Singh, and Elliot J. Androphy. 2007. “Modulating Role of RNA Structure in Alternative Splicing of a Critical Exon in the Spinal Muscular Atrophy Genes.” *Nucleic Acids Research* 35 (2): 371–89.

Singh, Nirmal K., Natalia N. Singh, Elliot J. Androphy, and Ravindra N. Singh. 2006. “Splicing of a Critical Exon of Human Survival Motor Neuron Is Regulated by a Unique Silencer Element Located in the Last Intron.” *Molecular and Cellular Biology* 26 (4): 1333–46.

Singh, N. N., E. J. Androphy, and R. N. Singh. 2004. “In Vivo Selection Reveals Combinatorial Controls That Define a Critical Exon in the Spinal Muscular Atrophy Genes.” *RNA* 10 (8). https://doi.org/10.1261/rna.7580704.

Singh, N. N., M. N. Lawler, E. W. Ottesen, D. Upreti, J. R. Kaczynski, and R. N. Singh. 2013. “An Intronic Structure Enabled by a Long-Distance Interaction Serves as a Novel Target for Splicing Correction in Spinal Muscular Atrophy.” *Nucleic Acids Research* 41 (17). https://doi.org/10.1093/nar/gkt609.

Singh, N. N., J. Seo, E. W. Ottesen, M. Shishimorova, D. Bhattacharya, and R. N. Singh. 2011. “TIA1 Prevents Skipping of a Critical Exon Associated with Spinal Muscular Atrophy.” *Molecular and Cellular Biology* 31 (5). https://doi.org/10.1128/MCB.00945-10.

Singh, Ravindra N., Matthew D. Howell, Eric W. Ottesen, and Natalia N. Singh. 2017. “Diverse Role of Survival Motor Neuron Protein.” *Biochimica et Biophysica Acta* 1860 (3): 299.

Singh, R. K., and T. A. Cooper. 2012a. “Pre-mRNA Splicing in Disease and Therapeutics.” *Trends in Molecular Medicine* 18 (8). https://doi.org/10.1016/j.molmed.2012.06.006.

———. 2012b. “Pre-mRNA Splicing in Disease and Therapeutics.” *Trends in Molecular Medicine* 18 (8). https://doi.org/10.1016/j.molmed.2012.06.006.

Soemedi, Rachel, Kamil J. Cygan, Christy L. Rhine, Jing Wang, Charlston Bulacan, John Yang, Pinar Bayrak-Toydemir, Jamie McDonald, and William G. Fairbrother. 2017. “Pathogenic Variants That Alter Protein Code Often Disrupt Splicing.” *Nature Genetics* 49 (6): 848–55.

Spingola, M., L. Grate, D. Haussler, and M. Ares. 1999. “Genome-Wide Bioinformatic and Molecular Analysis of Introns in Saccharomyces Cerevisiae.” *RNA* 5 (2). https://doi.org/10.1017/s1355838299981682.

Spitale, Robert C., and Danny Incarnato. 2022. “Probing the Dynamic RNA Structurome and Its Functions.” *Nature Reviews. Genetics* 24 (3): 178–96.

Stark, Jeremy M., David P. Bazett-Jones, Manfred Herfort, and Mark B. Roth. 1998. “SR Proteins Are Sufficient for Exon Bridging across an Intron.” *Proceedings of the National Academy of Sciences of the United States of America* 95 (5): 2163.

Steitz, T. A., and J. A. Steitz. 1993. “A General Two-Metal-Ion Mechanism for Catalytic RNA.” *Proceedings of the National Academy of Sciences of the United States of America* 90 (14). https://doi.org/10.1073/pnas.90.14.6498.

Sterne-Weiler, T., J. Howard, M. Mort, D. N. Cooper, and J. R. Sanford. 2011a. “Loss of Exon Identity Is a Common Mechanism of Human Inherited Disease.” *Genome Research* 21 (10). https://doi.org/10.1101/gr.118638.110.

———. 2011b. “Loss of Exon Identity Is a Common Mechanism of Human Inherited Disease.” *Genome Research* 21 (10). https://doi.org/10.1101/gr.118638.110.

Sterne-Weiler, T., and J. R. Sanford. 2014a. “Exon Identity Crisis: Disease-Causing Mutations That Disrupt the Splicing Code.” *Genome Biology* 15 (1). https://doi.org/10.1186/gb4150.

———. 2014b. “Exon Identity Crisis: Disease-Causing Mutations That Disrupt the Splicing Code.” *Genome Biology* 15 (1). https://doi.org/10.1186/gb4150.

Sugarman, E. A., N. Nagan, H. Zhu, V. R. Akmaev, Z. Zhou, E. M. Rohlfs, K. Flynn, et al. 2012. “Pan-Ethnic Carrier Screening and Prenatal Diagnosis for Spinal Muscular Atrophy: Clinical Laboratory Analysis of >72,400 Specimens.” *European Journal of Human Genetics: EJHG* 20 (1). https://doi.org/10.1038/ejhg.2011.134.

Sullenger, B. A., H. F. Gallardo, G. E. Ungers, and E. Gilboa. 1990. “Overexpression of TAR Sequences Renders Cells Resistant to Human Immunodeficiency Virus Replication.” *Cell* 63 (3). https://doi.org/10.1016/0092-8674(90)90455-n.

Summerton, J. 1999. “Morpholino Antisense Oligomers: The Case for an RNase H-Independent Structural Type.” *Biochimica et Biophysica Acta* 1489 (1). https://doi.org/10.1016/s0167-4781(99)00150-5.

Sumner, C. J. 2007. “Molecular Mechanisms of Spinal Muscular Atrophy.” *Journal of Child Neurology* 22 (8). https://doi.org/10.1177/0883073807305787.

Sun, L., L. Deng, C. K. Ea, Z. P. Xia, and Z. J. Chen. 2004. “The TRAF6 Ubiquitin Ligase and TAK1 Kinase Mediate IKK Activation by BCL10 and MALT1 in T Lymphocytes.” *Molecular Cell* 14 (3). https://doi.org/10.1016/s1097-2765(04)00236-9.

Takeshima, Y., H. Nishio, H. Sakamoto, H. Nakamura, and M. Matsuo. 1995. “Modulation of in Vitro Splicing of the Upstream Intron by Modifying an Intra-Exon Sequence Which Is Deleted from the Dystrophin Gene in Dystrophin Kobe.” *The Journal of Clinical Investigation* 95 (2). https://doi.org/10.1172/JCI117693.

Ternes, P., S. Franke, U. Zähringer, P. Sperling, and E. Heinz. 2002. “Identification and Characterization of a Sphingolipid Delta 4-Desaturase Family.” *The Journal of Biological Chemistry* 277 (28). https://doi.org/10.1074/jbc.M202947200.

Tse, Victor, Guillermo Chacaltana, Martin Gutierrez, Nicholas M. Forino, Arcelia G. Jimenez, Hanzhang Tao, Phong H. Do, et al. 2023a. “An Intronic RNA Element Modulates Factor VIII Exon-16 Splicing.” *Nucleic Acids Research* 52 (1): 300–315.

———. 2023b. “An Intronic RNA Element Modulates Factor VIII Exon-16 Splicing.” *Nucleic Acids Research* 52 (1): 300–315.

Tuerk, C., and L. Gold. 1990. “Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase.” *Science* 249 (4968). https://doi.org/10.1126/science.2200121.

Varani, L., M. Hasegawa, M. G. Spillantini, M. J. Smith, J. R. Murrell, B. Ghetti, A. Klug, M. Goedert, and G. Varani. 1999. “Structure of Tau Exon 10 Splicing Regulatory Element RNA and Destabilization by Mutations of Frontotemporal Dementia and Parkinsonism Linked to Chromosome 17.” *Proceedings of the National Academy of Sciences of the United States of America* 96 (14): 8229–34.

Varani, L., M. G. Spillantini, M. Goedert, and G. Varani. 2000. “Structural Basis for Recognition of the RNA Major Groove in the Tau Exon 10 Splicing Regulatory Element by Aminoglycoside Antibiotics.” *Nucleic Acids Research* 28 (3): 710–19.

Wais, Peter E., Melissa Arioli, Roger Anguera-Singla, and Adam Gazzaley. 2021. “Virtual Reality Video Game Improves High-Fidelity Memory in Older Adults.” *Scientific Reports* 11 (1): 1–14.

Wang, E., and I. Aifantis. 2020a. “RNA Splicing and Cancer.” *Trends in Cancer Research* 6 (8). https://doi.org/10.1016/j.trecan.2020.04.011.
———. 2020b. “RNA Splicing and Cancer.” *Trends in Cancer Research* 6 (8). https://doi.org/10.1016/j.trecan.2020.04.011.

Wang, E. T., R. Sandberg, S. Luo, I. Khrebtukova, L. Zhang, C. Mayr, S. F. Kingsmore, G. P. Schroth, and C. B. Burge. 2008. “Alternative Isoform Regulation in Human Tissue Transcriptomes.” *Nature* 456 (7221). https://doi.org/10.1038/nature07509.

Wang, Guey-Shin, and Thomas A. Cooper. 2007a. “Splicing in Disease: Disruption of the Splicing Code and the Decoding Machinery.” *Nature Reviews. Genetics* 8 (10): 749–61.

———. 2007b. “Splicing in Disease: Disruption of the Splicing Code and the Decoding Machinery.” *Nature Reviews. Genetics* 8 (10): 749–61.

Wang, Xuebin, Guoli Li, Yun Yang, Wenfeng Wang, Wenjing Zhang, Huawei Pan, Peng Zhang, et al. 2012. “An RNA Architectural Locus Control Region Involved in Dscam Mutually Exclusive Splicing.” *Nature Communications* 3 (1): 1–11.

Wang, Z., H. M. Hoffmann, and P. J. Grabowski. 1995. “Intrinsic U2AF Binding Is Modulated by Exon Enhancer Signals in Parallel with Changes in Splicing Activity.” *RNA* 1 (1). https://pubmed.ncbi.nlm.nih.gov/7489484/.

Wang, Z., M. E. Rolish, G. Yeo, V. Tung, M. Mawson, and C. B. Burge. 2004. “Systematic Identification and Analysis of Exonic Splicing Silencers.” *Cell* 119 (6). https://doi.org/10.1016/j.cell.2004.11.010.

Wang, Z., X. Xiao, E. Van Nostrand, and C. B. Burge. 2006. “General and Specific Functions of Exonic Splicing Silencers in Splicing Control.” *Molecular Cell* 23 (1): 61.

Wan, R., R. Bai, X. Zhan, and Y. Shi. 2020. “How Is Precursor Messenger RNA Spliced by the Spliceosome?” *Annual Review of Biochemistry* 89 (June).

https://doi.org/10.1146/annurev-biochem-013118-111024.

Wan, R., C. Yan, R. Bai, G. Huang, and Y. Shi. 2016. “Structure of a Yeast Catalytic Step I Spliceosome at 3.4 Å Resolution.” *Science* 353 (6302). https://doi.org/10.1126/science.aag2235.

Warf, M. B., and J. A. Berglund. 2010. “Role of RNA Structure in Regulating Pre-mRNA Splicing.” *Trends in Biochemical Sciences* 35 (3). https://doi.org/10.1016/j.tibs.2009.10.004.

Watakabe, A., K. Tanaka, and Y. Shimura. 1993. “The Role of Exon Sequences in Splice Site Selection.” *Genes & Development* 7 (3). https://doi.org/10.1101/gad.7.3.407.

Wayment-Steele, Hannah K., Do Soon Kim, Christian A. Choe, John J. Nicol, Roger Wellington-Oguri, Andrew M. Watkins, R. Andres Parra Sperberg, Po-Ssu Huang, Eterna Participants, and Rhiju Das. 2021. “Theoretical Basis for Stabilizing Messenger RNA through Secondary Structure Design.” *Nucleic Acids Research* 49 (18): 10604–17.

Weber, Jeffrey, Warren Jelinek, and James E. Darnell. 1977. “The Definition of a Large Viral Transcription Unit Late in Ad2 Infection of HeLa Cells: Mapping of Nascent RNA Molecular Labeled in Isolated Nuclei.” *Cell* 10 (4): 611–16.

Wei, C. M., and B. Moss. 1974. “Methylation of Newly Synthesized Viral Messenger RNA by an Enzyme in Vaccinia Virus.” *Proceedings of the National Academy of Sciences of the United States of America* 71 (8). https://doi.org/10.1073/pnas.71.8.3014.

Wilkinson, M. E., C. Charenton, and K. Nagai. 2020. “RNA Splicing by the Spliceosome.” *Annual Review of Biochemistry* 89 (June). https://doi.org/10.1146/annurev-biochem-091719-064225.

Wirth, B., L. Brichta, B. Schrank, H. Lochmüller, S. Blick, A. Baasner, and R. Heller. 2006. “Mildly Affected Patients with Spinal Muscular Atrophy Are Partially Protected by an Increased SMN2 Copy Number.” *Human Genetics* 119 (4). https://doi.org/10.1007/s00439-006-0156-7.

Wong, M. S., J. B. Kinney, and A. R. Krainer. 2018. “Quantitative Activity Profile and Context Dependence of All Human 5’ Splice Sites.” *Molecular Cell* 71 (6). https://doi.org/10.1016/j.molcel.2018.07.033.

Wong, T. N., T. R. Sosnick, and T. Pan. 2007. “Folding of Noncoding RNAs during Transcription Facilitated by Pausing-Induced Nonnative Structures.” *Proceedings*

of the National Academy of Sciences of the United States of America 104 (46). https://doi.org/10.1073/pnas.0705038104.

Wu, Huey Nan, and Olke C. Uhlenbeck. 2002. "Role of a Bulged A Residue in a Specific RNA-Protein Interaction," May. https://doi.org/10.1021/bi00399a030.

Wu, J. Y., and T. Maniatis. 1993. "Specific Interactions between Proteins Implicated in Splice Site Selection and Regulated Alternative Splicing." *Cell* 75 (6). https://doi.org/10.1016/0092-8674(93)90316-i.

Wu, S., C. M. Romfo, T. W. Nilsen, and M. R. Green. 1999. "Functional Recognition of the 3' Splice Site AG by the Splicing Factor U2AF35." *Nature* 402 (6763). https://doi.org/10.1038/45590.

Xing, Yi, and Christopher Lee. 2006. "Alternative Splicing and RNA Selection Pressure — Evolutionary Consequences for Eukaryotic Genomes." *Nature Reviews. Genetics* 7 (7): 499–509.

Xu, B., Y. Shi, Y. Wu, Y. Meng, and Y. Jin. 2019. "Role of RNA Secondary Structures in Regulating Dscam Alternative Splicing." *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1862 (11-12): 194381.

Yang, X., M. R. Bani, S. J. Lu, S. Rowan, Y. Ben-David, and B. Chabot. 1994. "The A1 and A1B Proteins of Heterogeneous Nuclear Ribonucleoparticles Modulate 5' Splice Site Selection in Vivo." *Proceedings of the National Academy of Sciences of the United States of America* 91 (15): 6924.

Yang, Yun, Leilei Zhan, Wenjing Zhang, Feng Sun, Wenfeng Wang, Nan Tian, Jingpei Bi, et al. 2011. "RNA Secondary Structure in Mutually Exclusive Splicing." *Nature Structural & Molecular Biology* 18 (2): 159–68.

Yeo, G., and C. B. Burge. 2004. "Maximum Entropy Modeling of Short Sequence Motifs with Applications to RNA Splicing Signals." *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology* 11 (2-3). https://doi.org/10.1089/1066527041410418.

Yu, A. M., P. M. Gasper, L. Cheng, L. B. Lai, S. Kaur, V. Gopalan, A. A. Chen, and J. B. Lucks. 2021. "Computationally Reconstructing Cotranscriptional RNA Folding from Experimental Data Reveals Rearrangement of Non-Native Folding Intermediates." *Molecular Cell* 81 (4). https://doi.org/10.1016/j.molcel.2020.12.017.

Yu, Yang, Patricia A. Maroney, John A. Denker, Xiang H-F Zhang, Olexandr Dybkov, Reinhard Lührmann, Eckhard Jankowsky, Lawrence A. Chasin, and

Timothy W. Nilsen. 2008. “Dynamic Regulation of Alternative Splicing by Silencers That Modulate 5′ Splice Site Competition.” *Cell* 135 (7): 1224.

Zatkova, Andrea, Ludwine Messiaen, Ina Vandenbroucke, Rotraud Wieser, Christa Fonatsch, Adrian R. Krainer, and Katharina Wimmer. 2004. “Disruption of Exonic Splicing Enhancer Elements Is the Principal Cause of Exon Skipping Associated with Seven Nonsense or Missense Alleles of NF1.” *Human Mutation* 24 (6): 491–501.

Zelin, E., Y. Wang, and S. K. Silverman. 2006. “Adenosine Is Inherently Favored as the Branch-Site RNA Nucleotide in a Structural Context That Resembles Natural RNA Splicing†,” February. https://doi.org/10.1021/bi052499l.

Zhang, H., L. Xing, W. Rossoll, H. Wichterle, R. H. Singer, and G. J. Bassell. 2006. “Multiprotein Complexes of the Survival of Motor Neuron Protein SMN with Gemins Traffic to Neuronal Processes and Growth Cones of Motor Neurons.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 26 (33). https://doi.org/10.1523/JNEUROSCI.3967-05.2006.

Zheng, Z. M. 2004. “Regulation of Alternative RNA Splicing by Exon Definition and Exon Sequences in Viral and Mammalian Gene Expression.” *Journal of Biomedical Science* 11 (3): 278.

Zheng, Z. M., M. Huynen, and C. C. Baker. 1998. “A Pyrimidine-Rich Exonic Splicing Suppressor Binds Multiple RNA Splicing Factors and Inhibits Spliceosome Assembly.” *Proceedings of the National Academy of Sciences of the United States of America* 95 (24): 14088.

Zhuang, Y., and A. M. Weiner. 1986. “A Compensatory Base Change in U1 snRNA Suppresses a 5’ Splice Site Mutation.” *Cell* 46 (6). https://doi.org/10.1016/0092-8674(86)90064-4.

Zhu, J., A. Mayeda, and A. R. Krainer. 2001. “Exon Identity Established through Differential Antagonism between Exonic Splicing Silencer-Bound hnRNP A1 and Enhancer-Bound SR Proteins.” *Molecular Cell* 8 (6). https://doi.org/10.1016/s1097-2765(01)00409-9.

Zorio, D. A., and T. Blumenthal. 1999. “Both Subunits of U2AF Recognize the 3’ Splice Site in Caenorhabditis Elegans.” *Nature* 402 (6763). https://doi.org/10.1038/45597.