Review

Minor Intron Splicing from Basic Science to Disease

Ettaib El Marabti [1], Joel Malek [1] and Ihab Younis [2,*]

[1] Weill Cornell Medicine-Qatar, Education City, Doha P.O. Box 24144, Qatar;
    ete2001@qatar-med.cornell.edu (E.E.M.); jom2042@qatar-med.cornell.edu (J.M.)
[2] Biological Sciences Program, Carnegie Mellon University in Qatar, Doha P.O. Box 24866, Qatar
* Correspondence: iyounis@andrew.cmu.edu

Abstract: Pre-mRNA splicing is an essential step in gene expression and is catalyzed by two machineries in eukaryotes: the major (U2 type) and minor (U12 type) spliceosomes. While the majority of introns in humans are U2 type, less than 0.4% are U12 type, also known as minor introns (mi-INTs), and require a specialized spliceosome composed of U11, U12, U4atac, U5, and U6atac snRNPs. The high evolutionary conservation and apparent splicing inefficiency of U12 introns have set them apart from their major counterparts and led to speculations on the purpose for their existence. However, recent studies challenged the simple concept of mi-INTs splicing inefficiency due to low abundance of their spliceosome and confirmed their regulatory role in alternative splicing, significantly impacting the expression of their host genes. Additionally, a growing list of minor spliceosome-associated diseases with tissue-specific pathologies affirmed the importance of minor splicing as a key regulatory pathway, which when deregulated could lead to tissue-specific pathologies due to specific alterations in the expression of some minor-intron-containing genes. Consequently, uncovering how mi-INTs splicing is regulated in a tissue-specific manner would allow for better understanding of disease pathogenesis and pave the way for novel therapies, which we highlight in this review.

Keywords: minor introns; U2 introns; U12 introns; minor spliceosome; RNA splicing; disease

1. RNA and Splicing

The view that RNA is simply an intermediate between DNA (the keeper of genetic information) and proteins (the executioners of all cellular functions) has long been challenged by the discoveries that gave RNA its own catalytic functions such as self-splicing introns [1] and the ribonuclease P catalyst [2]. These discoveries provided further evidence for the RNA world theory, which stipulates that life on earth started with RNA. The primary living substance on primitive earth is therefore thought to be RNA or a chemically similar structure, a process that may have started 4 billion years ago. The evidence for this 50 year old theory has grown over the years, and the importance of RNA has grown with it, giving rise to new fields in RNA biology that have expanded toward medical/clinical applications, including experiments done in the early 1990s showing protein production from an mRNA injected into mice skeletal muscle [3] and antibody generation from influenza virus mRNA injection [4], to most recently using mRNA as a vaccine against SARS-CoV-2 [5,6]. Thus, RNA-based therapies are quickly progressing to treat many ailments by targeting various pathways in RNA metabolism, including pre-mRNA splicing, which has garnered increasing attention after a handful of successful applications.

Pre-mRNA splicing was initially discovered in adenoviruses but was later described to occur in all eukaryotes [7]. More than 95% of genes in humans are transcribed as pre-mRNAs containing introns (intervening regions) that interfere with the required sequence continuity of exons (expressed regions) [7,8]. Splicing occurs co-transcriptionally to ensure the removal of introns and subsequent joining of exons, making it an essential and required step for the proper formation of a mature mRNA transcript that could be used as a template for protein synthesis [7,8]. Of note here is that a large number of intron-containing genes

are transcribed and spliced yet do not get translated into a protein but rather have key functions as noncoding RNAs. Broadly, the sequences involved in splicing include a 5′ splice site (5′ ss), a branch point sequence (BPS), and a 3′ splice site (3′ ss), and splicing is carried out as a two-step reaction, starting with 5′ splice site cleavage, generating a lariat, and ending with 3′ splice site cleavage and exon ligation [8]. This reaction is catalyzed by the spliceosome, an RNA-protein complex that requires ATP and Uridyl-rich small ribonucleoproteins (U snRNPs) [9].

2. The Other Side of the Splicing Coin—An Overview

The majority of intron splicing relies on the ubiquitous major spliceosome formed by the U1, U2, U4, U5, and U6 snRNPs [10,11]. These introns are loosely termed U2 introns and are referred to as such hereafter. Interestingly, eukaryotes carry in their genomes a less common type of introns termed minor introns (mi-INTs) or U12 introns (in this review, we use both terms interchangeably). The number of mi-INTs vary between species with humans having one of the largest number (~750) of minor-intron-containing genes (miGs) [12–15]. The mere existence of mi-INTs is remarkable for many reasons. Rather than using the highly abundant and ubiquitous major spliceosome, mi-INTs utilize a specialized spliceosome composed of a distinct set of snRNPs (U11, U12, U4atac, U5, and U6atac) [12–15]. The first mi-INTs were identified due to the unusual dinucleotide termini (AU at the 5′ end and AC at the 3′ end of the intron, compared to the conventional GU and AG, respectively) [12,16]. With the discovery of more mi-INTs, it became clear that the AU-AC dinucleotide at the intron boundaries is not what distinguishes U12 from U2 introns as most mi-INTs were shown to have a GU-AG dinucleotide but possess other unique features that set them apart from U2 introns [17,18]. These features played an important role in computationally identifying mi-INTs and include sequences within the introns as well as the conservation of their 5′ ss and BPS over 100 million years [19]. The peculiarity of mi-INTs however extends beyond their sequence features. In spite of their low abundance, the position of mi-INTs in their host genes is also highly conserved between humans, animals, and plants [14,20,21]. On the other hand, miGs tend to show commonality in their functions and have been coined by Burge et al. as "information processing genes" [19]. These functions include DNA transcription, replication and repair, RNA processing and translation, cytoskeletal organization, vesicular transport, voltage gated ion channel activity, and Ras-Raf signaling [22,23]. These very important functions seemed for a long time to be contradictory to in vitro and in vivo experiments showing slow splicing rates and hence inefficiency of the splicing of mi-INTs, making them bottlenecks for the expression of their host genes [24]. Moreover, data showing that mi-INTs can be experimentally converted to U2 introns with relative ease, as this required only a few mutations in their 5′ ss [17,19,25], adding serious doubt regarding the purpose of minor intron splicing altogether. Lastly, another puzzling feature of mi-INTs is the low abundance of their minor spliceosome (~100-fold compared to the major spliceosome), especially the catalytic component U6atac snRNP [26]. The dilemma of these seemingly contradictory features has been partially resolved as discussed later in this review.

Pinpointing the exact origin of introns in general has been elusive. However, their approximate age has been speculated based on their absence from prokaryotic genomes and presence in most eukaryotic and protist genomes [19,27,28]. Prior to the discovery of the U12 splicing pathway, attempts to explain the origin of introns followed two theories: the "exon theory of genes" and the "protosplice theory". The former stipulated that introns came into existence to allow exon assortment, giving rise to diverse genes through intron recombination, while the latter suggested that introns invaded eukaryotic genomes and were inserted at protosplice sites [29–32]. The discovery of U12 introns has challenged both theories as the presence of two distinct types of introns had not been accounted for by either theory. mi-INTs have a nonrandom distribution (clustered in information processing genes), and they occur less frequently in the genome. However, despite our lack of understanding of the origin of U12 introns, it was clear that these introns have appeared

early in eukaryotic evolution and have been conserved in many organisms, spanning several kingdoms [20,27,33]. These include plants (*Arabidopsis thaliana*), vertebrates (fish, amphibians, birds, and mammals), insects (*Drosophila melanogaster* and silkworm *Bombyx mori*), cnidarians (jellyfish), protists (*Acanthamoeba castellanii*), and fungi (*Rhizopus oryzae*). Nevertheless, it appears that mi-INTs have also been lost throughout evolution as they are absent in some fungi (*Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*), some protists, and nematode (*Caenorhabditis elegans*) [19,27]. Some evidence suggests that some U12 introns loss has occurred through conversion to U2 introns [17,19]. Again, in light of U12 introns' apparent disadvantages (slow kinetics, low abundance of minor spliceosome, etc.) and their ability to potentially easily convert to U2 introns, coupled with the extensive evidence of their conservation, and the essential functions of miGs, the conundrum of why mi-INTs even exist ensued. The working model for the why mi-INTs originated and are extremely conserved is that they play an important biological function within the genes that host them, driving their maintenance within miGs and conservation across organisms.

Several lines of evidence support this model of inefficiency for regulatory purposes and challenged the idea that mi-INTs retention is due to an inherent inefficiency in this splicing pathway. It was shown that the splicing inefficiency of mi-INTs resulting in partial or complete intron retention leads by design to specific outcomes, including degradation of their host transcripts by non-sense-mediated decay (NMD) or the exosome, alternative splicing of the transcript, or translation into a truncated protein. These outcomes are in support of mi-INTs retention being a post-transcriptional mechanism regulating miGs expression. Indeed, transcripts retaining mi-INTs were shown to constitute a pool of pre-mRNAs that is on standby to be expressed when required based on cellular needs. This is determined by the level of U6atac snRNP which is regulated through a MAPK signaling pathway. The effectiveness of this regulatory role of mi-INTs lies in the ability of a single intron to regulate the expression of a whole pre-mRNA. Therefore, the overarching function of mi-INTs in hundreds of miGs is suggested to be an expression "molecular switch" [34]. While the molecular switches model implies an all or none regulation, other data suggested a role for mi-INTs in alternative splicing as shown for the splicing factor Srsf10 [35], which is a member of the SR protein family that regulates alternative splicing. Srsf10 autoregulates its splicing through two competing minor and major 5′ splice sites, resulting in a protein-coding or a noncoding transcript, respectively. This autoregulation leads to a dynamic mechanism where the minor spliceosome activity determines *Srsf10* gene expression, and this in turns affects the levels of many other SR proteins. This work highlights a role for the minor spliceosome in regulating alternative splicing of many U2 introns indirectly through regulating the levels of SR proteins in cells [35].

The clinical relevance of splicing is well established, as 90% of disease-causing SNPs lie outside the protein-coding regions and many could have underlying mechanisms that would lead to aberrant splicing [36,37]. Indeed, some models estimate that 62% of disease mutations act by affecting splicing [38]. While it is still unclear what exact proportion of these mutations affects mi-INTs splicing as the targets of these mutations are being uncovered, several known mutations do include components of the minor splicing machinery (snRNPs), mi-INTs splice sites, and loci important for biogenesis and assembly of the minor spliceosome. Such aberrations lead to diseases ranging from neurodevelopmental and neurodegenerative to autoimmune disorders and cancer [39]. While the latter is well established in relation to mis-splicing of U2 introns and less so to that of mi-INTs [37,39], the similarity between the two splicing pathways, in addition to some evidence about alterations of mi-INTs splicing in cancer [40–42], provides further support that both U2 and U12 intron splicing may be affected in a similar fashion in cancer.

Mutations within components of the minor intron splicing machinery have been linked to the majority of diseases caused by minor intron splicing alterations. These mutations are predicted to affect hundreds of target genes containing mi-INTs, with the possibility of causing systemic defects in many body tissues. However, despite the ubiquitous nature of these mutations, most of the resulting diseases are tissue specific, indicating the possibility

that only a subset of target genes has the potential of causing disease. Fortunately, recent advances in genomics coupled with efforts to understand mi-INTs biology have shed some light on the mechanisms underlying disease tissue specificity and ways to identify target genes. In this review, we discuss this tissue specificity based on our current understanding of mi-INTs splicing and suggest experimental approaches using animal models coupled with CRISPR screens to identify target genes. We also bring to light new advances in the field of mi-INTs through discussing mi-INTs characteristic, the current debate of inefficiency vs. regulation, the fate of retained mi-INTs, and the crosstalk between the minor and major spliceosomes.

3. mi-INTs Sequence Characteristics

Initially described as introns with AU-AC termini (Figure 1A), mi-INTs containing these termini were later found to be only a subtype of all U12 introns [12], and the majority of mi-INTs have GU-AG termini [17,18]. Furthermore, mi-INTs could also have other, albeit much rarer, termini including AU-AG, GU-AU, CU-AC, GG-AG, and GA-AG [17,43]. Consequently, mi-INTs classification had to rely on more specific characteristics to distinguish them from U2 introns that have the canonical GU-AG termini. mi-INTs key sequence features lie in the conservation of nucleotide segments in their 5′ ss and BPS immediately upstream of their 3′ ss [12]. The 5′ ss conserved segment includes 8–9 nucleotides (nts) (A/G)UAUCCUUU, and the BPS includes the conserved sequence UCCUUAAC with fewer variations compared to U2 introns [12,19,44]. The splicing of mi-INTs depends on a well-situated 3′ ss creating a functional constraint on the distance between the BPS and the 3′ ss [44]. Thus, this distance is relatively shorter (10–20 nts) in mi-INTs, with the optimal distance being 11–13 nts [44]. Furthermore, mi-INTs lack a polypyrimidine tract which is commonly found in U2 introns (Figure 1A). The splice sites sequences and their conservation in multiple species can be found at https://genome.crg.es/cgi-bin/u12db/u12db.cgi (accessed on 19 May 2021).

mi-INTs sequence features are critical for their recognition, the formation/assembly of the splicing machinery, and adequate splicing through the 5′ and 3′ ss (Figure 1B). This splicing is carried out by the less abundant minor spliceosome, with unique snRNPs and protein composition, and slower splicing kinetics (refer to the spliceosome section for more detail).
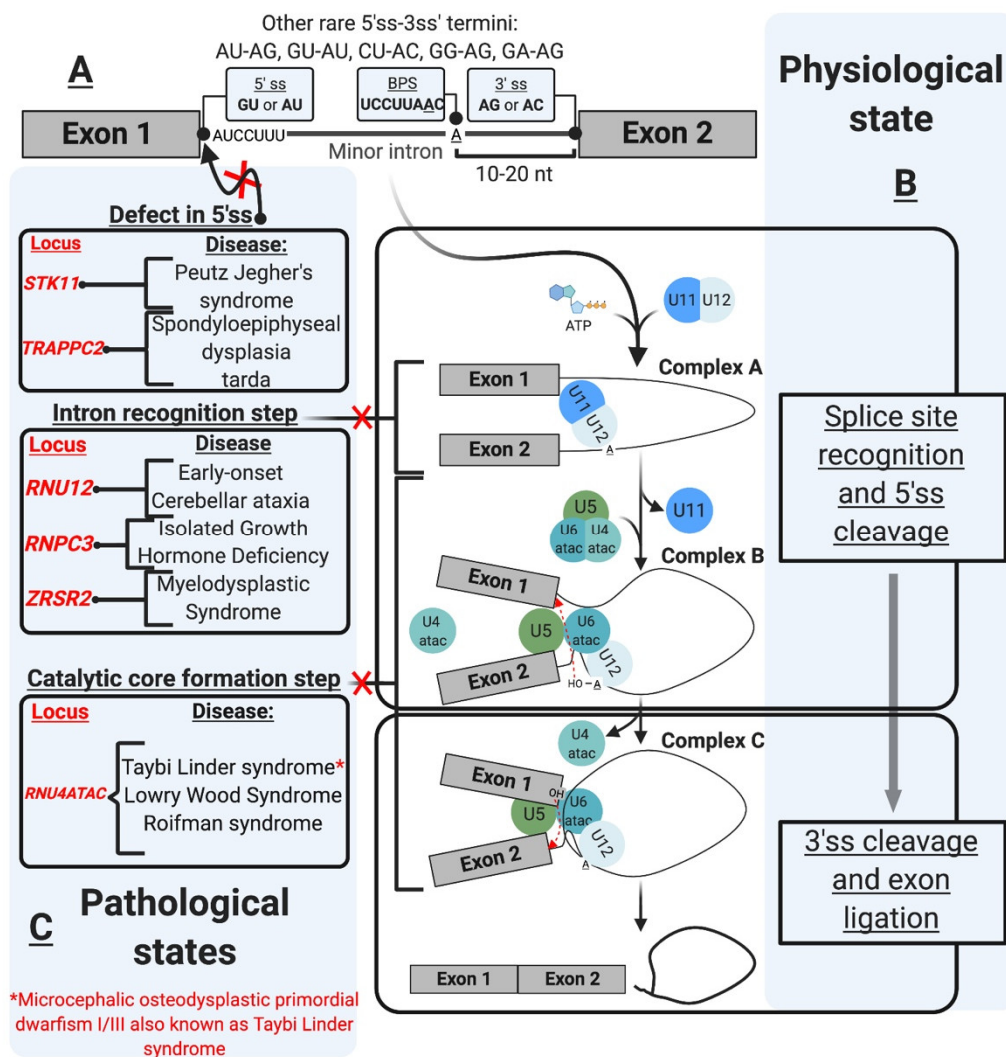
Figure 1. The minor intron splicing pathway. (**A**). U12-type intron sequence characteristics [1–6]. These sequences were initially described in the literature based on DNA nomenclature. These were changed in the figure as the transcript depicted is an RNA transcript. The commonly observed 3′ and 5′ sequences of minor introns are shown in the boxes. The conserved sequences are shown at the 5′ ss and branchpoint (**B**). Spliceosomal assembly and the splicing reaction [7–12] occur as a two-step reaction. The interaction of the preformed U11/U12 di-snRNP leads to the formation of Complex A. U4atac/U6atac.U5 tri-snRNP allows the formation of Complex B, which carries the first transesterification reaction after the release of U4atac and U11. This reaction leads to the formation of Complex C, which carries the second transesterification reaction producing ligates exons and a minor intron lariat. (**C**). U12-type intron-associated diseases classification [13–21]. The diseases are depicted in association with the affected step. Diseases not shown in the figure include amyotrophic lateral sclerosis associated with mutations in fused in sarcoma (FUS) RNA-binding protein [22], and Noonan syndrome due to U12-type intron retention in LZTR1, a regulator of RAS-related GTPases [23].

4. Origin and Evolution: The U12 Splicing Pathway Is as Old as the U2 Pathway

mi-INTs were initially exclusively described in animals [12,16] but were later identified in plants with identical sequence characteristics, dating the origin of the minor spliceo-somal pathway to at least one billion years [45], as it is thought to have appeared prior to the divergence of the animal and plant kingdoms. This is supported by the discovery of mi-INTs and the homologous component of the minor splicing pathway (U11, U12 snRNA, and minor spliceosome-specific proteins) in protists and fungi [27]. The presence of components of a minor spliceosomal pathway in protists is evidence that the minor spliceosome is possibly as old as its counterpart, the major spliceosome. The evolution of these splicing pathways could have followed three models: the group II introns parasitic

invasion model [19,46,47], the codivergence model [48], or the fission/fusion model [19]. The parasitic invasion model assumes that following endosymbiont invasion of an archaeal organism, descendants of two distinct group II introns invaded nuclear genes sequentially or simultaneously [19,28,46,47]. This may have occurred through repeated lysis of the invading organism (the developing mitochondria) causing seeding of these introns [28,47,49]. The fragmentation of these introns would have led to novel U snRNAs that could utilize proteins of the pre-existing spliceosome (supposedly U2-dependent spliceosome from a prior invasion) to splice these introns [19,47,50,51]. A protosplice site comparison of ancient introns to U2 and U12 type introns in humans and *Arabidopsis thaliana* suggests that primordial spliceosomal introns were of the U2 type [51]. Consequently, based on the parasitic invasion model, it is thought that U2 introns were the first to populate genes, followed by U12 introns [47,51].

However, this model does not explain the existence of "twintrons" in some premRNAs. A twintron is basically a major intron within a minor intron, such as that in the *Drosophila prospero* gene, which encodes a protein that plays an important role in axonal outgrowth and cell specification during the development of the nervous system [52]. Like other twintrons, that in the *prospero* gene provides a unique mechanism of alternative splicing. Remarkably, the switching between U2-dependent and U12-dependent spliceosomes is highly regulated during development and has significant consequences on the function of the gene product. More specifically, U12-dependent splicing of the twintron produces a shorter mRNA that goes on to be translated into a transcription factor with a specific homeodomain for proper DNA binding, whereas the mutually exclusive U2-dependent splicing results in a longer mRNA, that when translated, encodes a different homeodomain with potentially altered DNA-binding specificities [52]. Assuming that spliceosomal introns invaded nuclear genomes, twintrons suggest that U12 introns may not have arisen after U2 introns [48].

Furthermore, while Basu et al.'s findings were based on evidence that U12 introns positions are more conserved than those of U2 introns in humans and *Arabidopsis thaliana* [51], Moyer et al.'s data show that conservation in both pathways is similar, as 9% of U12 introns and 22% of U2 introns have been shown to have conserved position between humans and *Arabidopsis thaliana* [25]. These data argue against a parasitic invasion model in which U2 intron invasion preceded U12 invasion. On the other hand, the codivergence model assumes that following snRNA duplication in a primordial organism, introns, snRNA, and spliceosomal protein components diverged into two pathways [19,48]. Lastly, the fission/fusion model stipulates that following speciation of two different lineages from a primordial organism, the two splicing pathways diverged with each lineage containing one of the systems [19]. These organisms eventually merged, allowing the mixing of their genetic material [19]. Overall, current models support the idea that U12 and U2 introns and their splicing systems emerged around the same time.

Phylogenic reconstructions of the minor and major spliceosome's snRNAs and proteins show that both spliceosomes existed in the last eukaryotic common ancestor (LECA) [28]. The phylogenic distribution of U12 introns includes plants, vertebrates, insects, Cnidarians, protists, and fungi [19,27]. Nonetheless, they are absent in some fungi (*Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*), some protists, and nematode (*Caenorhabditis elegans*) [19]. The distribution of minor spliceosomal components shows the existence of a minor spliceosomal pathway in certain organisms that are within or below eukaryotic groups, suggesting loss of the minor splicing system during evolution [27]. One example is in fungi (Ascomycetes and Basidiomycetes) which lack minor spliceosomal components as opposed to *Rhizopus* (Zygomycota) which contains proteins components of the minor spliceosome and U12 introns. Consequently, evolutionary diversification within eukaryotes may have led to an eradication of U12 introns and the minor spliceosomal system in some organisms [27]. This loss is explained by the "class conversion hypothesis", which is believed to have occurred through the unidirectional U12 to U2 conversion. This conversion seems to preferentially occur in the GU-AG subtype of U12 introns and is

unidirectional (U12 to U2) due to the constraints in the characteristics of U12 intron splice sites sequences [20,25,27]. However, it is important to note that U12 intron loss, albeit slow in vertebrate, seems to be more common than conversion [43]. A comparison of miGs across different organisms suggests that some U12 introns have been replaced by U2 introns, potentially through splice site mutations. For example, while in humans, intron 1 of *PTEN* is a U12 intron; in *C. elegans*, intron 1 is a U2 intron. Recently, a phase 0 surplus in U2 introns and an underrepresentation of phase 0 in U12 introns [25] suggested that class conversion has occurred selectively in phase 0 U12 introns with GU-AG termini, supporting the idea that class conversion (U12 to U2) occurs through subtype switching (AU-AC → GU-AG) [19]. This conversion and loss contributed to the phase biases and low abundance of U12 introns in modern eukaryotic genomes as opposed to their ancestral counterpart.

The evolutionary loss of U12 introns poses a challenge to their intrinsic feature of conservation (phylogenetically and within gene families). This intriguing discrepancy has been partially resolved with new studies suggesting that mi-INTs function is the main driver of conservation in this class of introns.

5. The Functional Relevance of Minor Introns

Minor-intron-containing genes have been described as "information processing" as opposed to "operational genes" involved in metabolism, for example [19,22,23]. mi-INTs are enriched in gene families that are involved in DNA replication/repair, transcription, translation, and RNA processing. Other enriched functional groups for miGs include voltage-gated Na and Ca channels, vesicular transport, and cytoskeletal organization [21]. Interestingly, many mi-INTs have been shown to have slow splicing kinetics (~2 fold slower than U2 introns), both in vivo [24,53] and in vitro [13,54], leading to increased mi-INT retention, while all other U2 introns in the same transcript are spliced out efficiently [24]. As the majority of miGs contain a single minor intron per gene (only a small fraction contain two or more) [55], this one intron could be rate-limiting for the expression of the miG. Thus, the nonrandom genomic distribution of mi-INTs, the fact that the majority of them exist as a single intron within miGs, and their inefficient splicing/retention led to speculations that mi-INTs play important regulatory functions. Consequently, mi-INTs were dubbed "post-transcriptional bottlenecks" which would serve to prevent gene overexpression that may be harmful to cells [24]. This is supported by the molecular switches theory [34] as shown in Figure 2A, which stipulates that mi-INTs retention forms a pool of pre-mRNA transcripts that can be rapidly expressed when needed by the cell [34]. Therefore, rather than being breaks, halting gene expression to prevent toxicity, they are considered molecular switches that can be quickly turned on or off to allow or prevent gene expression [34]. This idea was built on evidence of the high instability of U6atac snRNP, a crucial component of the minor spliceosome catalytic core [34]. This instability causes low levels of U6atac in cells leading to low levels of properly spliced mRNA for hundreds of miGs. However, once stabilized, higher U6atac rapidly enhances mi-INTs splicing and miGs expression [34]. One example of this "on-demand" miG expression is shown when U6atac stabilizes by the stress-activated protein kinase, p38MAPK, which activates these molecular switches to allow the expression of genes required to deal with cellular stress [34]. This provides an efficient mechanism through which one intron could regulate the expression of an entire transcript, a rational idea that may explain the conservation of the minor intron splicing system.
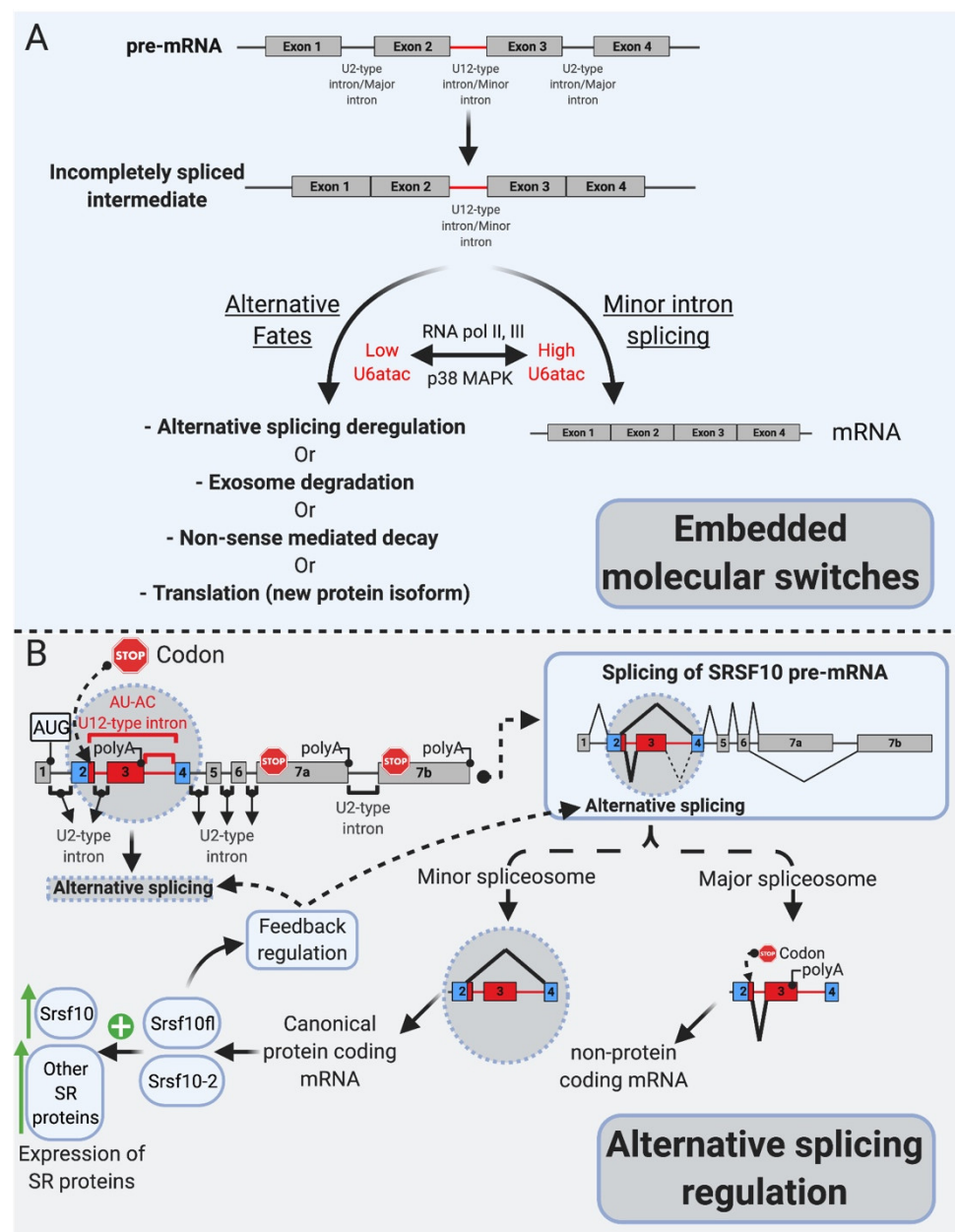
Figure 2. mi-INTs physiologic roles. (**A**). mi-INTs function as embedded molecular switches of gene expression [25]. Instability of U6atac result in low U6atac levels, which leads to pre-mRNA transcripts with a retained mi-INT whose fates are depicted in Figure 3. Decreased transcription lowers U6atac levels and the expression of minor-intron-containing genes, while activated p38MAPK stabilizes U6atac, increasing its levels and allowing the proper splicing of mi-INTs, producing a full-length mRNA. (**B**). mi-INTs regulate alternative splicing through Srsf10 [24]. Low minor spliceosome activity causes an increase in the expression of non-protein-coding Srsf10 transcript through exon 3 inclusion as a consequence of major intron splicing. High minor spliceosome activity allows exclusion of exon 3, through splicing to exon 4, allowing formation of a protein-coding transcript. The functional Srsf10 autoregulates its expression and that of other SR proteins which regulate alternative splicing of thousands of downstream targets.

Figure 3. Protein composition of minor snRNPs and autoimmune conditions associated with them. Sm proteins include B/B', D3, D2, D1, E, F, and G, whereas LSm proteins are LSm 2–8. The U4atac/U6atac.U5 tri-snRNP has two sets of Sm protein and one set of LSm proteins. Antibodies against some protein components of the minor spliceosome are found in systemic lupus erythematosus, diffuse scleroderma, and other autoimmune conditions.

More recent work has revealed another facet of mi-INTs as alternative splicing regulators, through the SR protein Srsf10, which itself functions in regulating alternative splicing [35]. One notable feature of the protein Srsf10 is that it autoregulates its expression through a cis regulatory element in its pre-mRNA [35] (Figure 2B). This autoregulation determines the use of competing major and minor splice sites leading to nonproductive or productive transcripts, respectively [35]. However, minor spliceosomal activity seems to partially overrule this autoregulation, as lower minor spliceosomal levels significantly lower the expression of productive Srsf10 transcripts [35], suggesting that minor spliceosomal abundance controls *Srsf10* gene expression by regulating its alternative splicing [35] (Figure 2B). Furthermore, in vivo evidence suggested that the expression of many other SR proteins, which themselves do not contain mi-INTs, correlates with minor spliceosomal levels in 25 mice tissues. CRISPR/Cas9 deletion of the nonproductive Srsf10 transcript led to increased expression of many other SRSF proteins [35], establishing that minor spliceosome abundance/activity controls *Srsf10* expression which in turn changes the expression of other SR proteins [35]. Taken together, these data suggest that the purpose of the minor splicing pathway is to regulate gene expression at the alternative splicing step of both miGs as well as U2 intron albeit indirectly [24,34,35].

6. mi-INTs Splicing Is Regulated

Minor-intron-containing genes show differential expression across tissues. Olthof et al. showed that both the number of miGs and their expression levels deferred across 11 mice tissues and 8 human tissues, with the heart and liver, for example expressing the least number of miGs at the lowest levels [56]. Importantly, miGs expression patterns across mice and human tissues were generally conserved [56]. Furthermore, miGs expression signatures showed functional enrichment in a tissue-specific manner [56]. For example, the cerebrum showed higher expression of miGs enriched in voltage-gated ion channel activity [56]. Intriguingly, while earlier data on a few mi-INTs have suggested an overall inefficient minor splicing pathway, recent data from 582 mi-INTs in 11 mice tissues point toward a dynamic regulatory mechanism in tissues as the inefficiency may not apply to all miGs in all tissues [56]. Therefore, mi-INTs splicing seems to be regulated in a tissue-specific manner to allow differential miGs expression. This would allow enrichment in functions that are important for specific tissues [56].

While mi-INTs retention could be an opportunity to generate a pool of unspliced transcripts standing by to be induced, it could also lead to one of the following predictable consequences (Figure 2A): (1) degradation of the host transcript by NMD or the exosome; (2) translation into a truncated protein if the transcript manages to be exported to the cytoplasm and is accessible to the ribosome, as most mi-INTs are expected to have in frame stop codons; and (3) inducing alternative splicing (AS) using cryptic or alternate splice donor, splice acceptor, or both. While the latter consequence is the least studied [21], early studies on twintrons that could switch between minor and major splicing pathways in *Drosophila* [52] and were later confirmed to occur in some human miGs [34], support the notion that mi-INTs provide a platform for AS. Recent work by Olthof et al. proposed that AS of mi-INTs occur more frequently than previously thought, and importantly it occurs in a tissue-specific manner [56]. Taken together, the idea that mi-INTs retention could be considered as a form of alternative splicing that is regulated agrees with the molecular switches mechanism previously described. Finally, the discovery of a high frequency of AS across this class of introns provides new avenues for the discovery of new RNA transcripts, protein isoforms/antigens that may be elevated in disease. These could therefore serve as biomarkers or targets for therapy.

7. The Minor Spliceosome: Composition, Assembly, and Evolution

The spliceosome is a multimegadalton ribonucleoprotein (RNP) complex that catalyzes the removal of introns in eukaryotes through two consecutive transesterification reactions [11,14]. The assembly of the spliceosome, which requires the recognition of intron consensus sequences at the 3′ and 5′ ends [11,14], is an organized process that involves a multitude of interactions between five U-rich small nuclear ribonucleoproteins (U snRNPs) and many associated proteins [11,14]. Upon assembly, RNA–RNA, RNA–protein, and protein–protein interactions allow for the catalysis of the splicing reaction; releasing the intron lariat and ligating the exons [11,14].

As described above, most eukaryotes contain two spliceosomes: the major (U2-dependent) spliceosome (reviewed in [11,57]), and the minor (U12-dependent) spliceosome which splices out mi-INTs. The minor spliceosome is made up of the minor snRNPs: U11, U12, U4atac, and U6atac, which are 100-fold less abundant [26,58] than their major spliceosome homologues U1, U2, U4, and U6, whereas U5 snRNP is shared by both spliceosomes. While the snRNPs composition in the two spliceosomes is homologous, there seems to be little sequence conservation between the minor and major snRNAs [50,58]. U11 and U12 sequences are unrelated to those of U1 and U2, respectively [26], while U4atac, U6atac have 40% sequence similarity to human U4 and U6, respectively [58]. Most of these conserved sequences are in regions important for RNA–RNA interaction within the spliceosome (between U6atac and U4atac or U12) or with the pre-mRNA [58]. However, some of them share analogous secondary structures, such as U4 and U6, and U4atac and U6atac interact through a base pairing interaction forming a di-snRNP secondary struc-

ture [58]. Furthermore, the U4atac/U6atac di-snRNP interacts with U5 to form a tri-snRNP similar to U4/U6.U5 tri-snRNP [50]. Of note, U1 and U2 exist as single particles that interact with RNA individually in the nucleus, and U11 and U12 form a U11/U12 di-snRNP, potentially through a protein-mediated interaction and interact with the pre-mRNA as a complex [59,60].

Protein composition of the minor spliceosome is yet another important aspect of the pathway that has received attention, leading to the identification of proteins that are unique to the minor spliceosome or shared with the major spliceosome. More than 300 protein have been identified in snRNP complexes, including snRNP-associated proteins (Sm proteins, Sm-like proteins, and snRNP-specific peptides) (Figure 3) and non-snRNP associated proteins (splicing factors) [14]. Interestingly, many of the snRNP-associated (i.e., Sm protein family) and non-snRNP-associated proteins (i.e., SR protein family) are shared between the minor and major spliceosomes [11,14,50,61]. For instance, U11/U12 di-snRNP has been shown to share all seven SF3b proteins with U2 snRNP, in addition to p14, hPrp43, YB-1 and Urp [14,21,62–65]. However, U11/U12 di-snRNP lacks all U1-specific proteins and instead contains unique proteins (65 K, 59 K, 48 K, 35 K, 31 K, 25 K, and 20 K), some of which are U11 specific [63]. This indicated that protein–protein and protein–RNA interactions at the 5′ ss are not conserved between the two spliceosomes [63]. Interestingly, these unique U11/U12 proteins are conserved in animals and plants [63–65]. On the other hand, the protein composition of U4atac/U6atac.U5 tri-snRNP is almost identical to that of U4/U6.U5 [50]. One notable shared protein, 15.5 K, binds the 5′ stem loop of U4 and U4atac snRNA and could play a nucleation factor role, allowing similar proteins to associate with U4/U6 and U4atac/U6atac [50,66–68]. However, unique minor tri-snRNP peptides remain to be identified [50]. The large number of proteins shared between the two spliceosome types suggest that protein–protein and protein–RNA interactions are similar or conserved to some extent.

The assembly of the minor spliceosome is very similar to that of the major spliceosome [13,58,60,68]. This is a stepwise process that involves the formation of multiple complexes (A, B, and C) through RNA–RNA, RNA–protein and protein–protein interactions, prior to the onset of catalysis [13,48,58,60,68] (Figure 1B). Early assembly involves the recognition of the 5′ ss and BPS cooperatively, through U11/5′ ss and U12/BPS interactions by the U11/U12 di-snRNP, which allows the formation of pre-spliceosomal complex A [13,58,60]. Subsequently, snRNA–snRNA base pairing allows the formation of U4atac/U6atac di-snRNP in which U4atac functions as an RNA chaperone to allow 5′ ss recognition by U6atac [48,58]. U4atac/U6atac and U5 are recruited to the spliceosome to form mature spliceosomal complex B [13,58]. The U4atac/U6atac.U5 tri-snRNP interaction disrupts the base pairing between U4atac/U6atac displacing U4atac before the first catalytic reaction [48,58]. U5 interacts with the exonic sequences at the 5′ ss and 3′ ss, and U6atac base pairs with the 5′ ss displacing U11 and interacting with U12 [48,58]. Consequently, U6atac and U12 form the activated catalytic core while U5 aligns the two exons for the second catalytic step, these form complex C that is catalytically active [13,48,58,68].

The identification of the snRNA and protein components in the minor spliceosome has helped shed light on the evolutionary relationship between the minor and major splicing systems. This relationship was described by the three models (previously discussed), which attempt to answer the question of whether these two systems originated from a common ancestor (homologous origin) [19]. The fact that most proteins in U4/U6.U5 are shared with U4atac/U6atac.U5, while many other proteins are shared between U2 and U11/U12 [50,61,65], supports a model of common ancestry [19]. However, the identification of novel unique U11/U12 proteins that have 5′ ss interactions, distinct from U1 [63], challenges the homologous origin model. This is further exacerbated by evidence of low percentage/lack of sequence conservation between minor and major spliceosomal snRNAs [26,58]. Therefore, a nonhomologous model such as a parasitic group II intron invasion model could explain the highly diverged snRNA and the high protein similarity in the minor and major spliceosomes [19,50,63]. However, this model does not explain

the novel proteins identified in the minor spliceosome [63]. Instead, a homologous model with high divergence such as the fission fusion model could account for the current observations [19,63]. Speciation may have led to two separate lineages, each containing one splicing pathway in which snRNAs and spliceosomal proteins diverged separately [19,63]. Eventually, endosymbiosis allowed the fusion of the two lineages, forming an ancestral eukaryote with both splicing systems [19,63].

8. Splicing and Disease

The ability of splicing mutations to cause disease is well recognized through the myriad of described diseases that result from altered major or minor and/or alternative splicing [37,39,69]. Generally, these alterations in splicing patterns are a consequence of mutations to the cis-acting elements or the trans-acting factors. The cis-acting elements include the core splicing consensus sequences (5′ ss, 3′ ss, and BPS). In fact, 15% of disease-causing SNPs are located within the core splice sites. The other key set of cis-acting elements are splicing enhancers and silencers sequences also known as splicing regulatory elements (SREs). These could be within exons, exonic splicing enhancers (ESE), and exonic splicing silencers (ESS), or within introns, intronic splicing enhancers (ISE), and intronic splicing silencers (ISS) [70]. On the other hand, trans-acting factors include RNA-binding proteins, classified as spliceosomal and non-spliceosomal components that facilitate spliceosomal recruitment or suppression. Several disease-causing mutations affecting splicing through SREs or trans-acting elements have been identified [39,70–72], but many more are yet to be characterized, which undoubtedly would uncover a large set of disease causing mutations in the splicing realm. In addition, while most research attention has been focused on diseases affecting U2 intron splicing (review in [37]), a growing literature suggests that mi-INTs splicing alterations may be as important in disease [39,71–75].

9. Minor Intron Splicing and Diseases

The minor splicing pathway has been linked to three classes of disease, autoimmune, cancer/cancer predisposing conditions, and neurological (congenital and degenerative) [39,72–75] (Figures 1C and 3 and Table 1). A testament to the essential role of mi-INTs, mutations that completely abolish their splicing are lethal [76,77]. For instance, complete loss of *Rnpc3*, which codes for 65 K, a protein unique to the minor spliceosome, causes death in mice embryos prior to blastocyst implantation [78]. Constitutive biallelic loss of *Rnu11* in mice is embryonically lethal [79]. Furthermore, Microcephalic Osteodysplastic Primordial Dwarfism type I, a disorder caused by homozygous or compound heterozygous *RNU4ATAC* mutations [80,81], albeit hypomorphic, causes early death in homozygous patients [82,83]. Another example is complete loss of SMN protein, which functions in snRNP maturation affecting miGs to a great extent [84], is lethal in mice, prior to implantation [76]. However, many of the mutations related to minor intron splicing cause partial loss, rather than complete loss of the minor spliceosome, or the function of miGs. This allows organismal survival and phenotypic manifestation of pathologies. Below is a brief description of the role of mi-INTs splicing in diseases.

Table 1. Human diseases associated with defects in minor intron splicing.

Splicing Defect	Gene Affected	Disease	References
Aberrant 5′ splice site	*STK11*	Peutz Jegher's syndrome (PJS)	[85]
	TRAPPC2	Spondyloepiphyseal dysplasia tarda (SEDT)	[86]
Intron recognition/mutant minor snRNP component	*RNU12*	Early-onset cerebellar ataxia (EOCA)	[87]
	RNPC3	Isolated growth hormone deficiency (IGHD)	[74]
	ZRSR2	Myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML)	[88]
	RNU4ATAC	Taybi-Linder syndrome (TLS)	[89]
	RNU4ATAC	Lowry Wood syndrome (LWS)	[90]
	RNU4ATAC	Roifman syndrome	[91]
Branch point mutation	*LZTR1*	Noonan syndrome (NS)	[88]
	LZTR1	Schwannomatosis	[92]
Biogenesis/Assembly	*FUS*	Amyotrophic lateral sclerosis (ALS)	[93]
	SMN1	Spinal muscular atrophy (SMA)	[94–100]

A strong support for a role of splicing in autoimmune diseases comes from the observation that spliceosomal proteins (major and minor) have been shown to have important roles in systemic lupus erythematosus, polymyositis and diffuse scleroderma [101–103]. Interestingly, Ng et al. have shown a significantly higher rate of minor intron splicing in transcripts generating autoantigens as opposed to randomly selected genes (80% vs. 1%) [75].

The congenital disorders related to minor intron splicing include: spondyloepiphyseal dysplasia tarda (SEDT) [86], Microcephalic Osteodysplastic Primordial Dwarfism type I/III or Taybi-Linder syndrome (MOPDI/III or TALS) [89,104], Roifman syndrome (RFMN) [91], Lowry Wood syndrome (LWS) [90], early-onset cerebellar ataxia (EOCA) [87], isolated growth hormone deficiency (IGHD) [74] (reviewed in [39.72]), and Noonan syndrome [88]. On the other hand, the neurodegenerative diseases include spinal muscular atrophy (SMA) and amyotrophic lateral sclerosis (ALS) (reviewed in [39,72]). These diseases can be further classified by the mutation type that would lead to loss of function of the minor splicing system [39]. These involve defects in components of the minor snRNPs, mutation of U12-type 5′ ss, and defect in biogenesis and/or assembly of the minor spliceosome [39,72] (Figures 1C and 2).

While deregulated alternative splicing in cancer is well established [10,105], most studies have focused on major intron splicing. However, the high similarity of the two pathways in their physiological states and the prevalence of alternative splicing in the minor splicing pathway [55] suggest that the two pathways are affected similarly in pathological states. Furthermore, the minor splicing system has been linked to myelodysplastic syndrome [71,88] and Peutz-Jegher's syndrome [85], which increase the risk of acute myeloid leukemia [71] and gastrointestinal/extra gastrointestinal cancers [105], respectively. The cause of six congenital diseases and two neurodegenerative disorders have been attributed completely or in part to the minor splicing pathway.

One property of mi-INTs is their enrichment in information-processing genes. This creates a functional exclusiveness, crucial for cell survival and signaling (i.e., DNA transcription, replication and repair, RNA processing and translation, cytoskeletal organization, vesicular transport, voltage-gated ion channels, and Ras-Raf signaling) [22,23]. Therefore, based on the exclusivity of minor introns to certain gene families and their incomplete functional loss in diseases, it is highly likely that they are implicated in diseases such as cancer, where alterations in these gene families contribute to the hallmarks of cancer. Indeed, mi-INTs are enriched in cancer-relevant oncogenes such as *BRAF*, *ERK2*, *MAPK11/P38beta*,

and *JNK1* [78,106]. Furthermore, p38MAPK, which has been implicated in mi-INTs splicing regulation [34], itself harbors an mi-INT. Other examples of cancer-relevant miGs include *PARP1*, which has many functions including DNA damage/repair [107] and the tumor suppressor *PTEN*. As shown in Figure 4, mi-INTs splicing deregulation in such genes could lead to consequences that may cause functional loss of the gene product or the production of an alternative isoform that could possibly act as dominant negative. In the absence of gene redundancy that could compensate partially for functional loss, the consequences could be genomic instability and chromosomal abnormalities [107] in addition to abnormal cell proliferation. We therefore propose that mutations affecting the minor splicing pathway would impact many pathways, and the culmination of effects from all miGs affected is thus expected to contribute to cancer pathogenesis.
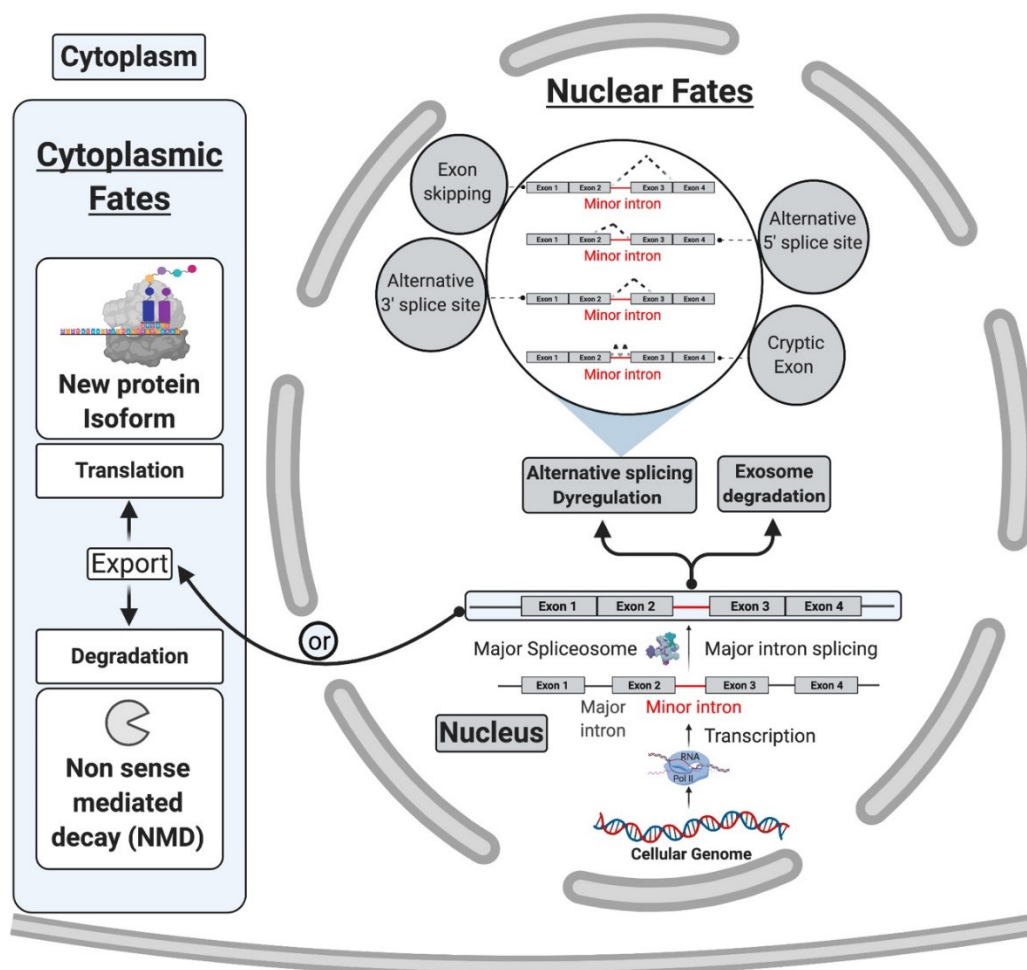
Figure 4. Cytoplasmic and nuclear fates of pre-mRNA transcripts with retained mi-INTs [20,26,27]. mi-INTs retention may lead to trapping of the transcript within the nucleus. The trapped transcript may undergo exosomal degradation, activation of cryptic U2-type intron cryptic sites or exon skipping. Transcripts that escape degradation in the nucleus are exported to the cytoplasm they may be degraded by NMD or translated to a truncated or novel protein isoform.

10. Systemic vs. Tissue-Specific Pathologies

Theoretically, perturbation of mi-INTs splicing or ubiquitously expressed proteins involved in this splicing pathway should lead to multisystem pathologies, reflecting the ubiquitous nature of the etiology. However, in most diseases related to defects in mi-INTs splicing, only a single or a handful of tissues are affected. For instance, MOPD I/III, RFMN, and LWS which are caused by hypomorphic mutations in U4atac snRNA [81,90,91,104] lead to intrauterine growth restriction, poor growth, developmental abnormalities spanning the skeletal and nervous systems, and retinal abnormalities [83,90]. A clear mechanism

behind these phenotypes has not been identified; however, it is suggested that they are certainly tied to abnormalities in mi-INTs splicing [79,91]. Another example is SMA, which manifests primarily in the nervous system and neuromuscular junctions, through motor neuron loss and muscle atrophy (Figure 5) [108,109]. This disease is caused by a mutation or homozygous deletion in *SMN1* with preservation of *SMN2* allowing the production of the SMN protein at low levels (Figure 5). SMN is essential for major and minor snRNPs biogenesis and assembly (Figure 6), a function that is essential in all cells of any tissue type. However, it has been shown that low SMN causes a biased decrease in minor snRNAs in SMA mice [94–96], causing splicing defects in some miGs [97–100]. In fact, selective enhancement of mi-INTs splicing in SMN-deficient mice, improved SMA disease phenotype, motor function, rescue of the loss of proprioceptive sensory synapses, and prolonged survival [84]. Thus, the ubiquitous nature of SMN and minor snRNPs implies that a deficiency would affect all tissues. Nevertheless, similar to MOPD I/III, RFMN, and LWS, SMA is also tissue specific. This dilemma has remained largely unresolved, but recent studies provide some possible explanations. Olthof et al. showed that miGs are differentially expressed in mice and human tissues, with tissues having the highest number of expressed miGs such as the cerebrum showing high expression levels [55]. Furthermore, the expressed miGs in all tissues are enriched in functions important for that specific tissue [55]. This suggests that tissues with highest number of expressed miGs, especially those with tissue-specific functions, are more sensitive to mi-INTs splicing deregulation. Indeed, 7 out of the 10 currently described mi-INTs-related diseases involve the nervous system; IGHD, EOCA, MOPD I/III, RFMN, LWS, ALS, and SMA. Furthermore, transcriptome analysis from MOPD I patients shows differential mi-INTs retention across tissues and miGs, again indicating that certain tissues and miGs are more susceptible/sensitive to mi-INTs retention. Consequently, the same mutation (i.e., *RNU4ATAC*) could cause a different mi-INTs retention levels in different tissues with certain miGs more affected than others. The biological determinants of mi-INTs retention susceptibility in tissues are not fully understood, but work by Doggett et al. shows that minor spliceosomal protein 65 K is localized to the stem/progenitor cell compartment, which is responsible for tissue maintenance in the small intestine [78]. Loss of this protein impaired minor intron splicing leading to apoptosis and decreased cell proliferation in the stem/progenitor cells population, with negligible effects on other small-intestine cell populations [78]. Similar findings were demonstrated by Baumgartner et al. in the context of microcephaly due to MOPD I, RFMN, and LWS. U11 loss in developing mice showed elevated mi-INTs retention contributing to preferential loss/apoptosis of a specific population of neural progenitor cells; the radial glial cell leading to microcephaly [79]. Most of mi-INTs retention led to NMD of the transcripts, the majority of which were enriched in functions such as nucleotide binding (e.g., *Parp1*) and protein transport and cell-cycle functions (e.g., *Pten*) [79]. Taken together, these findings demonstrate that within tissues, mi-INTs splicing and minor spliceosome components seem to be more important to certain cell populations than others. These cells are mostly those responsible for tissue development and maintenance, which could explain the high prevalence of developmental and cancer syndromes in mi-INTs spliceopathies.

Whether the cell population affected is a function of susceptible miGs/biological functions or vice versa is still unclear. On the one hand, the susceptibility of miGs to mi-INTs retention seems to be related to their splicing efficiency and expression levels in normal tissues, with those having low splicing efficiency and expression levels being more susceptible [82]. Another determinant of tissue specificity could be drawn from SMA pathogenesis. In this disease, minor snRNAs are decreased in brain, spinal cord, and heart tissue but remain unchanged in kidney and skeletal muscle tissues [95]. This differential expression of minor snRNAs could exist to some extent in normal tissues, which could be regulated in a tissue-specific manner and could potentially render tissues with the lowest snRNA expression at risk of higher rates of mi-INTs retention due to low minor snRNPs. This regulation could be at the level of transcription, biogenesis/assembly, or in the fully formed snRNP. Evidence for the latter has been documented in the fully formed U6atac,

whose levels increase following stabilization by p38MAPK in cells [34]. Therefore, the idea of a possible regulatory mechanism that determines minor snRNP levels in healthy tissues and tissue susceptibility to mi-INTs retention is probable.
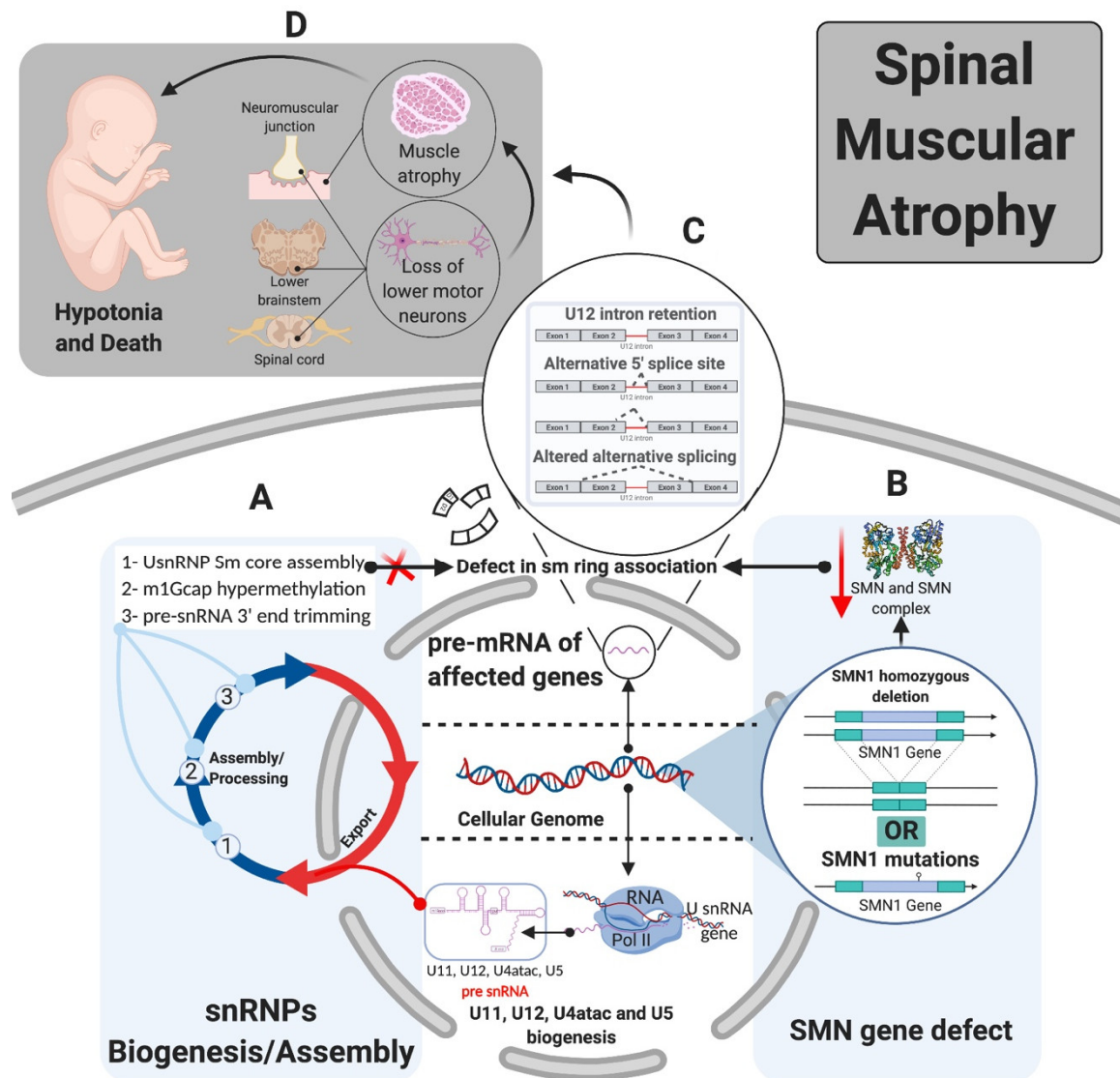

Figure 5. Tissue specificity of mi-INTs-related diseases (SMA). (**A**). snRNP biogenesis and assembly requires SMN along with other proteins (SMN complex) for Sm core assembly, m1Gcap hypermethylation, and pre-snRNA 3'end trimming. (**B**). Loss or mutation in SMN1 with retention of SMN2 product allows the production of SMN at lower levels leading to lower SMN complex level and lower ability to interact with Sm proteins, preventing the formation of the Sm ring and mature snRNPs (major and minor) which preferentially affects minor snRNPs [28–30] leading to splicing defects. (**C**). Splicing defect in U2 and U12-type intron splicing with bias toward U12-type intron retention [29,31] leads to aberrant mRNA transcripts that may be degraded or produce aberrant proteins. (**D**). Phenotype of SMA depends on the severity of the illness, which is caused by motor neuron degeneration in the brain stem, anterior horn of the spinal cord, and muscle atrophy. This leads to hypotonia and eventual death [32,33].
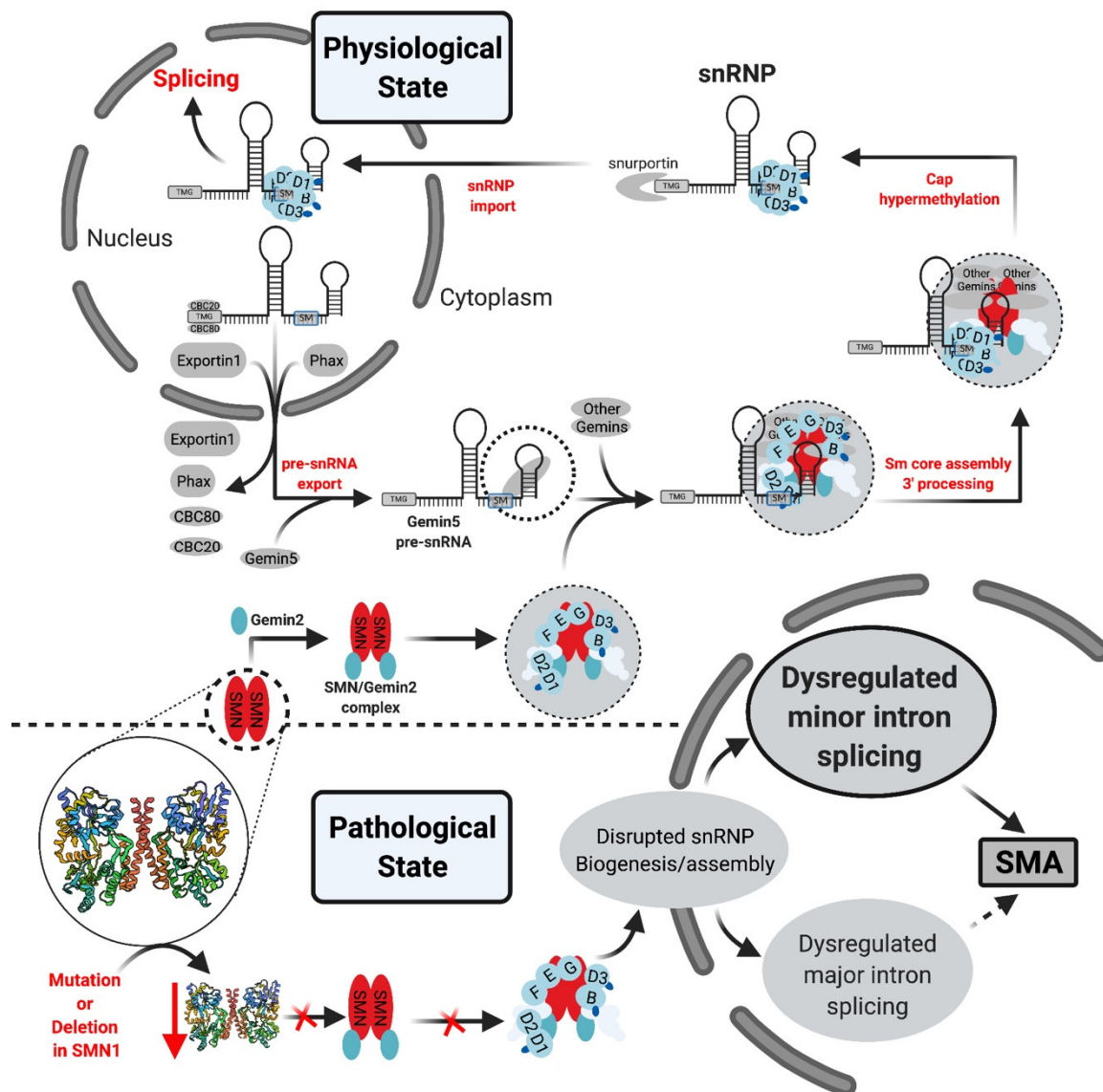
Figure 6. U snRNP biogenesis pathway. In the physiological state, once the U snRNA is transcribed and capped it associates with multiple factors which form the export complex and allow U snRNA export into the cytoplasm. Subsequently, in a process mediated by SMN complex and Gemin proteins, the Sm proteins interact with the U snRNA at the Sm site. The SMN complex dissociates, followed by 3′ processing/trimming and cap hypermethylation. The final product is imported into the nucleus. Further modifications occur on the U snRNA, and snRNP (U1, U2, U4, U5 or U11/U12, U4atac)-specific proteins associate with U snRNP Sm core [34].

In the pathological state disruption of the SMN1 gene via a mutation or deletion, with preservation of SMN2 gene, lowers the levels of the SMN protein, decreasing the levels of the SMN complex and its ability to interact with Sm proteins. This leads to a preferential decrease in minor snRNPs. Ultimately, both the major and minor splicing pathways are affected with minor splicing defects being more frequent as a result of U12-type intron retention [95,99].

11. Minor Splicing and Disease Onset

The minor splicing system is important for tissue maintenance and renewal during development and adult life. In other words, this is a system that is indispensable for a normal physiology and life (as described above). The crucial role of this splicing system in development and adulthood bares the question of whether the same mutation affecting mi-INTs splicing could manifest differently depending on the age of onset. For instance, *Rnpc3*

loss during the embryonic stage prevented implantation in mice, while in adult tissue, it affected the integrity of many tissues, especially the gastrointestinal tract [78]. On the other hand, the biallelic loss of *RNPC3* has been linked to IGHD in human [74]. Similarly, constitutive loss of U11 was embryonically lethal in mice but led to microcephaly when the loss occurred later in development [79]. Furthermore, U12-type intron splicing seems to be important for specific populations of cells namely stem/progenitor cells [78,79]. These findings allow us to speculate that age of onset may play an important role in determining the MiGs affected and consequently the phenotype observed. Under this premise, the loss of U11 later into adulthood may contribute to a neurodegenerative condition as opposed to microcephaly. Understanding the relationship between mutations that cause defects in mi-INTs splicing and age of onset may uncover more diseases that may be related to this splicing pathway. It would allow for a better appreciation of the differential regulatory roles that this pathway plays in developing and adult tissues. Furthermore, it would enhance our grasp of disease mechanisms related to this class of introns, which are still lacking.

12. Identification of Key Disease-Associated miGs

The current understanding of mi-INTs splicing-related diseases is limited. Importantly, whether the phenotypes observed are a consequence of a systemic splicing disruption or a subset of disrupted miGs remains unanswered. The tissue specificity of pathological conditions associated with mi-INTs splicing and its physiologic role in regulating gene expression and alternative splicing [34,35,39,55,56,78,79] suggest that only a subset of miGs is responsible for a given disease. Based on this, a number of in vitro and in vivo experiments could be designed to identify disease-causing miGs using CRISPR/Cas system and antisense oligonucleotides (ASO). An overall scheme of the approaches that are needed to move our understanding of mi-INTs splicing from the laboratory to the clinic is presented in Figure 7. This scheme relies initially on comprehensive methods such as RNA-seq to identify splicing changes, especially those in miGs, in a given disease. Splicing aberrations that are potentially relevant to the disease are then shortlisted for targeting. We propose two parallel approaches moving forward. (1) A CRISPR approach in which the miG or the specific intron is modified in the genome of cells extracted from patients or controls followed by testing the effects of this genome editing on disease progression. (2) ASOs on the other hand can modulate splicing at the RNA level without modifying the genome. A set of ASOs can be used to modulate several splicing events and studied in vitro to narrow it down to a handful that can be tested in an animal model. It is worth noting here that these ambitious approaches come with their own hurdles. These include but are not limited to issues with delivery to specific cells, in addition to off-target genome editing in the case of CRISPR or off-target splicing modulation in the case of ASOs. These hurdles necessitate in-depth follow-up studies before these approaches can be adapted as therapies in humans. One way to help mitigate these hurdles is creating animal models with aberrant splicing in miGs that mimic disease progression. These models can be used to further understand the relationship between altered mi-INTS splicing and disease, optimizing some of the limitations, and eventually drug testing and further preclinical trials testing. The identified genes along with the targeting methods used to identify them could have tremendous value in developing targeted therapies.
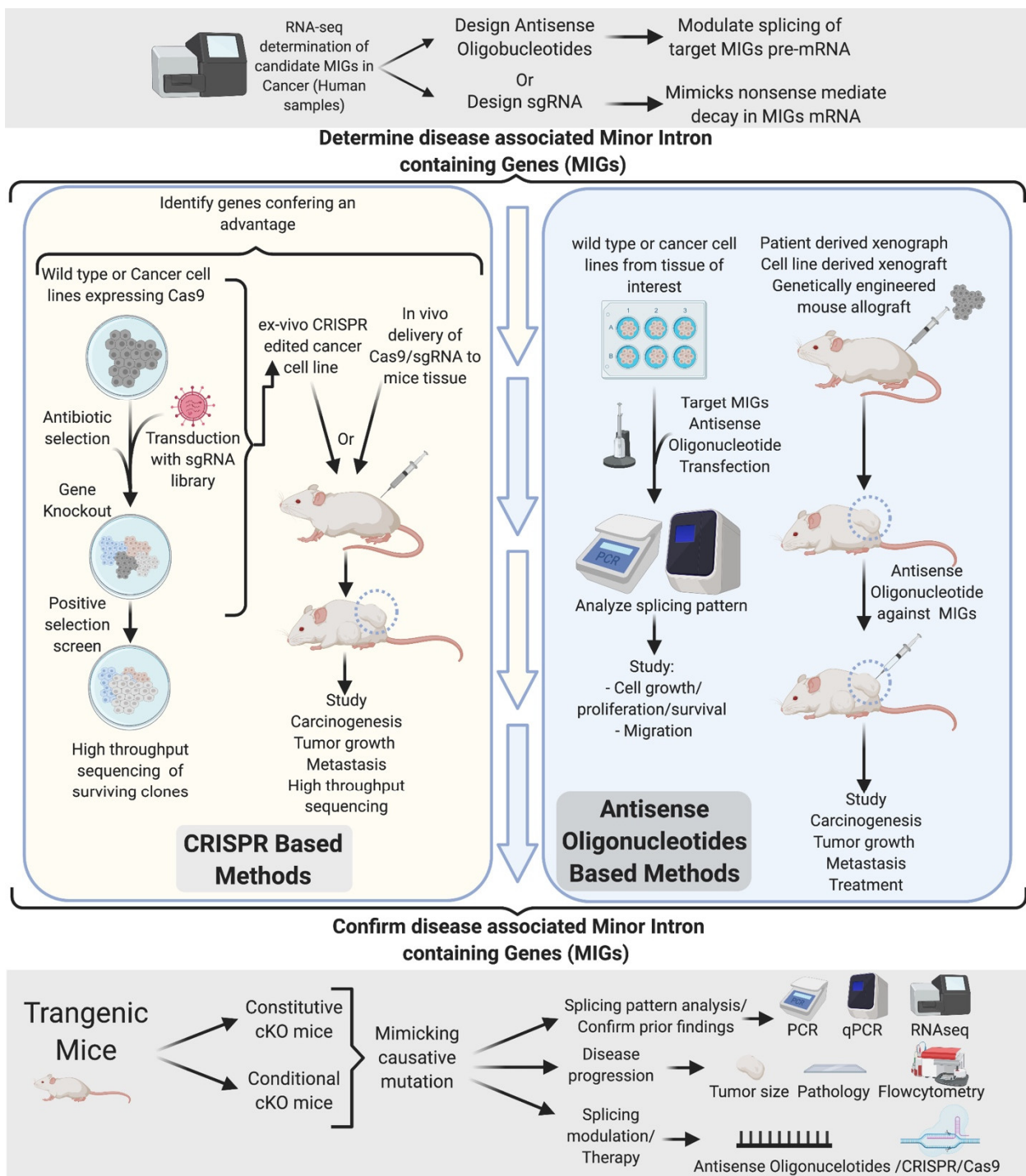
Figure 7. Experimental determination of disease-causing/-associated MIGs. Initial determination of disease-related MIGs through RNA-seq generates a list of potential candidates. Targeting of these candidates through CRISPR and ASOs in vivo and vitro narrows down the gene targets to a few. The ultimate confirmation comes from a genetically engineered mouse, mimicking the disease mutation found in human. This allows confirmation of mi-INTs splicing defects observed and determination of their contribution to the disease.

Targeted therapy has been successfully used in the treatment of many diseases. In cancer, targeted therapy has been used after identifying the gene or protein that are crucial to cancer progression and modulating them. However, with splicing being a targetable process, the list of targetable elements should include genes and mRNA transcripts that are underexpressed. The monogenic nature of mutations contributing to mi-INTs associated

diseases makes these mutations good targets (i.e., SMA) [110]. The majority of mi-INTs defects result in mRNA transcripts retaining mi-INTs, which, if they escape NMD, may cause the production of a novel protein isoform or nonfunctional protein. The occurrence of such an event in the tumor suppressor *PTEN*, for instance, could be detrimental. However, the use of ASOs or splice switching oligonucleotides (SSOs) [111] have the potential to promote adequate splicing in the setting of mi-INTs retention. A successful application of splice site targeting was done using CRISPR-guided cytidine deaminase (TAM) to restore the open reading frame of a mutant *DMD* gene [112]. The ability of the TAM system to modify splice sites could potentially have a role in the treatment of mi-INTs splice site mutations such as those in *STK11* and *TRAPPC2*. These methods provide a broad spectrum of approaches by which mi-INTs splicing could be targeted through correcting the main causative mutation or the splicing of affected MiGs.

Author Contributions: Conceptualization, E.E.M., J.M. and I.Y.; writing—original draft preparation, E.E.M.; writing—review and editing, E.E.M., J.M. and I.Y.; visualization, E.E.M.; supervision, J.M. and I.Y.; project administration, J.M. and I.Y.; funding acquisition, I.Y. All authors have read and agreed to the published version of the manuscript.

References

1. Kruger, K.; Grabowski, P.J.; Zaug, A.J.; Sands, J.; Gottschling, D.E.; Cech, T.R. Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* **1982**, *31*, 147–157. [CrossRef]
2. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **1983**, *35*, 849–857. [CrossRef]
3. Wolff, J.A.; Malone, R.W.; Williams, P.; Chong, W.; Acsadi, G.; Jani, A.; Felgner, P.L. Direct gene transfer into mouse muscle in vivo. *Science* **1990**, *247*, 1465–1468. [CrossRef] [PubMed]
4. Martinon, F.; Krishnan, S.; Lenzen, G.; Magné, R.; Gomard, E.; Guillet, J.-G.; Lévy, J.-P.; Meulien, P. Induction of virus-specific cytotoxic T lymphocytesin vivo by liposome-entrapped mRNA. *Eur. J. Immunol.* **1993**, *23*, 1719–1722. [CrossRef] [PubMed]
5. Polack, F.P.; Thomas, S.J.; Kitchin, N.; Absalon, J.; Gurtman, A.; Lockhart, S.; Perez, J.L.; Marc, G.P.; Moreira, E.D.; Zerbini, C.; et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N. Engl. J. Med.* **2020**, *383*, 2603–2615. [CrossRef]
6. Jackson, L.A.; Anderson, E.J.; Rouphael, N.G.; Roberts, P.C.; Makhene, M.; Coler, R.N.; McCullough, M.P.; Chappell, J.D.; Denison, M.R.; Stevens, L.J.; et al. An mRNA Vaccine against SARS-CoV-2—Preliminary Report. *N. Engl. J. Med.* **2020**, *383*, 1920–1931. [CrossRef] [PubMed]
7. Abelson, J. RNA Processing and the Intervening Sequence Problem. *Annu. Rev. Biochem.* **1979**, *48*, 1035–1069. [CrossRef]
8. Green, M.R. PRE-mRNA Splicing. *Ann. Rev. Genet.* **1986**, *20*, 671–708. [CrossRef]
9. Padgett, R.A.; Grabowski, P.J.; Konarska, M.M.; Seiler, S.; Sharp, P.A. Splicing of Messenger RNA Precursors. *Annu. Rev. Biochem.* **1986**, *55*, 1119–1150. [CrossRef]
10. El Marabti, E.; Younis, I. The Cancer Spliceome: Reprograming of Alternative Splicing in Cancer. *Front. Mol. Biosci.* **2018**, *5*, 80. [CrossRef]
11. Will, C.L.; Lührmann, R. Spliceosome structure and function. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, 1–2. [CrossRef] [PubMed]
12. Hall, S.L.; Padgett, R. Conserved Sequences in a Class of Rare Eukaryotic Nuclear Introns with Non-consensus Splice Sites. *J. Mol. Biol.* **1994**, *239*, 357–365. [CrossRef]
13. Tarn, W.-Y.; Steitz, J.A. A Novel Spliceosome Containing U11, U12, and U5 snRNPs Excises a Minor Class (AT–AC) Intron In Vitro. *Cell* **1996**, *84*, 801–811. [CrossRef]

14. Patel, A.A.; Steitz, J.A. Splicing double: Insights from the second spliceosome. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 960–970. [CrossRef] [PubMed]

15. Hall, S.L.; Padgett, R.A. Requirement of U12 snRNA for in Vivo Splicing of a Minor Class of Eukaryotic Nuclear Pre-mRNA Introns. *Science* **1996**, *271*, 1716–1718. [CrossRef]

16. Jackson, I.J. A reappraisal of non-consensus mRNA splice sites. *Nucleic Acids Res.* **1991**, *19*, 3795–3798. [CrossRef]

17. Dietrich, R.C.; Incorvaia, R.A.; Padgett, R. Terminal Intron Dinucleotide Sequences Do Not Distinguish between U2- and U12-Dependent Introns. *Mol. Cell* **1997**, *1*, 151–160. [CrossRef]

18. Sharp, P.A.; Burge, C.B. Classification of Introns: U2-Type or U12-Type. *Cell* **1997**, *91*, 875–879. [CrossRef]

19. Burge, C.B.; Padgett, R.; Sharp, P.A. Evolutionary Fates and Origins of U12-Type Introns. *Mol. Cell* **1998**, *2*, 773–785. [CrossRef]

20. Basu, M.K.; Makalowski, W.; Rogozin, I.B.; Koonin, E.V. U12 intron positions are more strongly conserved between animals and plants than U2 intron positions. *Biol. Direct* **2008**, *3*, 19. [CrossRef] [PubMed]

21. Turunen, J.J.; Niemelä, E.H.; Verma, B.; Frilander, M.J. The significant other: Splicing by the minor spliceosome. *Wiley Interdiscip. Rev. RNA* **2012**, *4*, 61–76. [CrossRef]

22. Yeo, G.W.; Van Nostrand, E.L.; Liang, T.Y. Discovery and Analysis of Evolutionarily Conserved Intronic Splicing Regulatory Elements. *PLoS Genet.* **2007**, *3*, e85. [CrossRef]

23. Wu, Q.; Krainer, A.R. AT-AC Pre-mRNA Splicing Mechanisms and Conservation of Minor Introns in Voltage-Gated Ion Channel Genes. *Mol. Cell. Biol.* **1999**, *19*, 3225–3236. [CrossRef] [PubMed]

24. Patel, A.A.; McCarthy, M.; Steitz, J.A. The splicing of U12-type introns can be a rate-limiting step in gene expression. *EMBO J.* **2002**, *21*, 3804–3815. [CrossRef] [PubMed]

25. Moyer, D.C.; LaRue, E.G.; Hershberger, C.E.; Roy, S.W.; Padgett, R.A. Comprehensive database and evolutionary dynamics of U12-type introns. *Nucleic Acids Res.* **2020**. [CrossRef] [PubMed]

26. Montzka, K.A.; Steitz, J.A. Additional low-abundance human small nuclear ribonucleoproteins: U11, U12, etc. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 8885–8889. [CrossRef]

27. Russell, A.G.; Charette, J.M.; Spencer, D.F.; Gray, M.W. An early evolutionary origin for the minor spliceosome. *Nat. Cell Biol.* **2006**, *443*, 863–866. [CrossRef] [PubMed]

28. Rogozin, I.B.; Carmel, L.; Csuros, M.; Koonin, E.V. Origin and evolution of spliceosomal introns. *Biol. Direct* **2012**, *7*, 11. [CrossRef] [PubMed]

29. Gilbert, W. The Exon Theory of Genes. *Cold Spring Harb. Symp. Quant. Biol.* **1987**, *52*, 901–905. [CrossRef]

30. Dibb, N.J.; Newman, A.J. Evidence that introns arose at proto-splice sites. *EMBO J.* **1989**, *8*, 2015–2021. [CrossRef]

31. Dibb, N. Proto-splice site model of intron origin. *J. Theor. Biol.* **1991**, *151*, 405–416. [CrossRef]

32. Sverdlov, A.V.; Rogozin, I.B.; Babenko, V.N.; Koonin, E.V. Reconstruction of Ancestral Protosplice Sites. *Curr. Biol.* **2004**, *14*, 1505–1508. [CrossRef] [PubMed]

33. Shukla, G.C.; Padgett, R. Conservation of functional features of U6atac and U12 snRNAs between vertebrates and higher plants. *RNA* **1999**, *5*, 525–538. [CrossRef] [PubMed]

34. Younis, I.; Dittmar, K.; Wang, W.; Foley, S.W.; Berg, M.G.; Hu, K.Y.; Wei, Z.; Wan, L.; Dreyfuss, G. Minor introns are embedded molecular switches regulated by highly unstable U6atac snRNA. *eLife* **2013**, *2*, e00780. [CrossRef]

35. Meinke, S.; Goldammer, G.; Weber, A.I.; Tarabykin, V.; Neumann, A.; Preussner, M.; Heyd, F. Srsf10 and the minor spliceosome control tissue-specific and dynamic SR protein expression. *eLife* **2020**, *9*, e56075. [CrossRef]

36. Hindorff, L.A.; Sethupathy, P.; Junkins, H.A.; Ramos, E.M.; Mehta, J.P.; Collins, F.S.; Manolio, T.A. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9362–9367. [CrossRef]

37. Scotti, M.M.; Swanson, M.S. RNA mis-splicing in disease. *Nat. Rev. Genet.* **2016**, *17*, 19–32. [CrossRef]

38. López-Bigas, N.; Audit, B.; Ouzounis, C.; Parra, G.; Guigó, R. Are splicing mutations the most frequent cause of hereditary disease? *FEBS Lett.* **2005**, *579*, 1900–1903. [CrossRef]

39. Verma, B.; Akinyi, M.; Norppa, A.J.; Frilander, M.J. Minor spliceosome and disease. *Semin. Cell Dev. Biol.* **2018**, *79*, 103–112. [CrossRef]

40. Shah, A.A.; Xu, G.; Rosen, A.; Hummers, L.K.; Wigley, F.M.; Elledge, S.J.; Casciola-Rosen, L. Brief Report: Anti–RNPC-3 Antibodies As a Marker of Cancer-Associated Scleroderma. *Arthritis Rheumatol.* **2017**, *69*, 1306–1312. [CrossRef]

41. Fischer, D.; Wahlfors, T.; Mattila, H.; Oja, H.; Tammela, T.L.J.; Schleutker, J. MiRNA Profiles in Lymphoblastoid Cell Lines of Finnish Prostate Cancer Families. *PLoS ONE* **2015**, *10*, e0127427. [CrossRef] [PubMed]

42. Tian, Y.; Huang, Z.; Wang, Z.; Yin, C.; Zhou, L.; Zhang, L.; Huang, K.; Zhou, H.; Jiang, X.; Li, J.; et al. Identification of Novel Molecular Markers for Prognosis Estimation of Acute Myeloid Leukemia: Over-Expression of PDCD7, FIS1 and Ang2 May Indicate Poor Prognosis in Pretreatment Patients with Acute Myeloid Leukemia. *PLoS ONE* **2014**, *9*, e84150. [CrossRef] [PubMed]

43. Lin, C.-F.; Mount, S.M.; Jarmołowski, A.; Makałowski, W. Evolutionary dynamics of U12-type spliceosomal introns. *BMC Evol. Biol.* **2010**, *10*, 47. [CrossRef] [PubMed]

44. Dietrich, R.C.; Peris, M.J.; Seyboldt, A.S.; Padgett, R.A. Role of the 3′ Splice Site in U12-Dependent Intron Splicing. *Mol. Cell. Biol.* **2001**, *21*, 1942–1952. [CrossRef]

45. Wu, H.-J.; Gaubiercomella, P.; Delseny, M.; Grellet, F.; Van Montagu, M.; Rouze, P. Non–canonical introns are at least 109 years old. *Nat. Genet.* **1996**, *14*, 383–384. [CrossRef]

46.   Sharp, P.A. Five easy pieces. *Science* **1991**, *254*, 663–664. [CrossRef]
47.   Lynch, M. The evolution of spliceosomal introns. *Curr. Opin. Genet. Dev.* **2002**, *12*, 701–710. [CrossRef]
48.   Tarn, W.-Y.; Steitz, J.A. Pre-mRNA splicing: The discovery of a new spliceosome doubles the challenge. *Trends Biochem. Sci.* **1997**, *22*, 132–137. [CrossRef]
49.   Martin, W.; Koonin, E.V. Introns and the origin of nucleus–cytosol compartmentalization. *Nat. Cell Biol.* **2006**, *440*, 41–45. [CrossRef]
50.   Schneider, C.; Will, C.L.; Makarova, O.V.; Makarov, E.M.; Lührmann, R. Human U4/U6.U5 and U4atac/U6atac.U5 Tri-snRNPs Exhibit Similar Protein Compositions. *Mol. Cell. Biol.* **2002**, *22*, 3219–3229. [CrossRef]
51.   Basu, M.K.; Rogozin, I.B.; Koonin, E.V. Primordial spliceosomal introns were probably U2-type. *Trends Genet.* **2008**, *24*, 525–528. [CrossRef]
52.   Scamborova, P.; Wong, A.; Steitz, J.A. An Intronic Enhancer Regulates Splicing of the Twintron of Drosophila melanogaster prospero Pre-mRNA by Two Different Spliceosomes. *Mol. Cell. Biol.* **2004**, *24*, 1855–1869. [CrossRef]
53.   Singh, J.; Padgett, R. Rates of in situ transcription and splicing in large human genes. *Nat. Struct. Mol. Biol.* **2009**, *16*, 1128–1133. [CrossRef] [PubMed]
54.   Wu, Q.; Krainer, A.R. U1-Mediated Exon Definition Interactions between AT-AC and GT-AG Introns. *Science* **1996**, *274*, 1005–1008. [CrossRef] [PubMed]
55.   Levine, A.; Durbin, R. A computational scan for U12-dependent introns in the human genome sequence. *Nucleic Acids Res.* **2001**, *29*, 4006–4013. [CrossRef]
56.   Olthof, A.M.; Hyatt, K.C.; Kanadia, R.N. Minor intron splicing revisited: Identification of new minor intron-containing genes and tissue-dependent retention and alternative splicing of minor introns. *BMC Genom.* **2019**, *20*, 686. [CrossRef]
57.   Wilkinson, M.E.; Charenton, C.; Nagai, K. RNA Splicing by the Spliceosome. *Annu. Rev. Biochem.* **2020**, *89*, 359–388. [CrossRef] [PubMed]
58.   Tarn, W.-Y.; Steitz, J.A. Highly Diverged U4 and U6 Small Nuclear RNAs Required for Splicing Rare AT-AC Introns. *Science* **1996**, *273*, 1824–1832. [CrossRef]
59.   Wassarman, K.M.; Steitz, J.A. The low-abundance U11 and U12 small nuclear ribonucleoproteins (snRNPs) interact to form a two-snRNP complex. *Mol. Cell. Biol.* **1992**, *12*, 1276–1285. [CrossRef]
60.   Frilander, M.J.; Steitz, J.A. Initial recognition of U12-dependent introns requires both U11/5′ splice-site and U12/branchpoint interactions. *Genes Dev.* **1999**, *13*, 851–863. [CrossRef]
61.   Will, C.L.; Schneider, C.; Reed, R.; Lührmann, R. Identification of Both Shared and Distinct Proteins in the Major and Minor Spliceosomes. *Science* **1999**, *284*, 2003–2005. [CrossRef] [PubMed]
62.   Golas, M.M.; Sander, B.; Will, C.L.; Lührmann, R.; Stark, H. Molecular Architecture of the Multiprotein Splicing Factor SF3b. *Science* **2003**, *300*, 980–984. [CrossRef] [PubMed]
63.   Will, C.L.; Schneider, C.; Hossbach, M.; Urlaub, H.; Rauhut, R.; Elbashir, S.; Tuschl, T.; Lührmann, R. The human 18S U11/U12 snRNP contains a set of novel proteins not found in the U2-dependent spliceosome. *RNA* **2004**, *10*, 929–941. [CrossRef] [PubMed]
64.   Will, C.L.; Schneider, C.; Macmillan, A.M.; Katopodis, N.F.; Neubauer, C.; Wilms, M.; Lührmann, R.; Query, C.C. A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site. *EMBO J.* **2001**, *20*, 4536–4546. [CrossRef] [PubMed]
65.   Lorkovic, Z.J.; Lehner, R.; Forstner, C.; Barta, A. Evolutionary conservation of minor U12-type spliceosome between plants and humans. *RNA* **2005**, *11*, 1095–1107. [CrossRef] [PubMed]
66.   Park, S.J.; Jung, H.J.; Dinh, S.N.; Kang, H. Structural features important for the U12 snRNA binding and minor spliceosome assembly of Arabidopsis U11/U12-small nuclear ribonucleoproteins. *RNA Biol.* **2016**, *13*, 670–679. [CrossRef]
67.   Nottrott, S.; Hartmuth, K.; Fabrizio, P.; Urlaub, H.; Vidovic, I.; Ficner, R.; Luhrmann, R. Functional interaction of a novel 15.5 kD [U4/U6·U5] tri-snRNP protein with the 5′ stem-loop of U4 snRNA. *EMBO J.* **1999**, *18*, 6119–6133. [CrossRef]
68.   Frilander, M.J.; Steitz, J.A. Dynamic Exchanges of RNA Interactions Leading to Catalytic Core Formation in the U12-Dependent Spliceosome. *Mol. Cell* **2001**, *7*, 217–226. [CrossRef]
69.   Montes, M.; Sanford, B.L.; Comiskey, D.F.; Chandler, D.S. RNA Splicing and Disease: Animal Models to Therapies. *Trends Genet.* **2019**, *35*, 68–87. [CrossRef]
70.   Singh, R.; Cooper, T.A. Pre-mRNA splicing in disease and therapeutics. *Trends Mol. Med.* **2012**, *18*, 472–482. [CrossRef]
71.   Madan, V.; Kanojia, D.; Jia, L.; Okamoto, R.; Sato-Otsubo, A.; Kohlmann, A. ZRSR2 mutations cause dysregulated RNA splicing in MDS. *Blood* **2014**, *124*, 4609. [CrossRef]
72.   Jutzi, D.; Akinyi, M.; Mechtersheimer, J.; Frilander, M.J.; Ruepp, M.-D. The emerging role of minor intron splicing in neurological disorders. *Cell Stress* **2018**, *2*, 40–54. [CrossRef] [PubMed]
73.   Drake, K.D.; Lemoine, C.; Aquino, G.S.; Vaeth, A.M.; Kanadia, R.N. Minor spliceosome disruption causes limb growth defects without altering patterning. *BioRxiv* **2020**. [CrossRef]
74.   Argente, J.; Flores, R.Z.; Gutiérrez-Arumí, A.; Verma, B.; Moreno, G.; Ángel, M.; Cuscó, I.; Oghabian, A.; Chowen, J.A.; Frilander, M.J.; et al. Defective minor spliceosome mRNA processing results in isolated familial growth hormone deficiency. *EMBO Mol. Med.* **2014**, *6*, 299–306. [CrossRef]
75.   Ng, B.; Yang, F.; Huston, D.P.; Yan, Y.; Yang, Y.; Xiong, Z.; Peterson, L.E.; Wang, H.; Yang, X.-F. Increased noncanonical splicing of autoantigen transcripts provides the structural basis for expression of untolerized epitopes. *J. Allergy Clin. Immunol.* **2004**, *114*, 1463–1470. [CrossRef] [PubMed]

76. Schrank, B.; Götz, R.; Gunnersen, J.M.; Ure, J.M.; Toyka, K.V.; Smith, A.G.; Sendtner, M. Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 9920–9925. [CrossRef]

77. Otake, L.R.; Scamborova, P.; Hashimoto, C.; Steitz, J.A. The divergent U12-type spliceosome is required for pre-mRNA splicing and is essential for development in Drosophila. *Mol. Cell* **2002**, *9*, 439–446. [CrossRef]

78. Doggett, K.; Williams, B.B.; Markmiller, S.; Geng, F.-S.; Coates, J.; Mieruszynski, S.; Ernst, M.; Thomas, T.; Heath, J.K. Early developmental arrest and impaired gastrointestinal homeostasis in U12-dependent splicing-defective Rnpc3-deficient mice. *RNA* **2018**, *24*, 1856–1870. [CrossRef]

79. Baumgartner, M.; Olthof, A.M.; Aquino, G.S.; Hyatt, K.C.; Lemoine, C.; Drake, K.; Sturrock, N.; Nguyen, N.; al Seesi, S.; Kanadia, R.N. Minor spliceosome inactivation causes microcephaly, owing to cell cycle defects and death of self-amplifying radial glial cells. *Development* **2018**, *145*, 166322. [CrossRef]

80. Abdel-Salam, G.M.; Abdel-Hamid, M.S.; Issa, M.; Magdy, A.; El-Kotoury, A.; Amr, K. Expanding the phenotypic and mutational spectrum in microcephalic osteodysplastic primordial dwarfism type I. *Am. J. Med. Genet. Part A* **2012**, *158A*, 1455–1461. [CrossRef]

81. Abdel-Salam, G.M.; Miyake, N.; Eid, M.M.; Abdel-Hamid, M.S.; Hassan, N.A.; Eid, O.M.; Effat, L.K.; El-Badry, T.H.; El-Kamah, G.Y.; El-Darouti, M.; et al. A homozygous mutation in RNU4ATAC as a cause of microcephalic osteodysplastic primordial dwarfism type I (MOPD I) with associated pigmentary disorder. *Am. J. Med. Genet. Part A* **2011**, *155*, 2885–2896. [CrossRef]

82. Cologne, A.; Benoit-Pilven, C.; Besson, A.; Putoux, A.; Campan-Fournier, A.; Bober, M.B.; De Die-Smulders, C.E.; Paulussen, A.D.; Pinson, L.; Toutain, A.; et al. New insights into minor splicing-a transcriptomic analysis of cells derived from TALS patients. *RNA* **2019**, *25*, 1130–1149. [CrossRef] [PubMed]

83. Pierce, M.J.; Morse, R.P. The neurologic findings in Taybi-Linder syndrome (MOPD I/III): Case report and review of the literature. *Am. J. Med. Genet. Part A* **2012**, *158*, 606–610. [CrossRef]

84. Osman, E.Y.; Van Alstyne, M.; Yen, P.-F.; Lotti, F.; Feng, Z.; Ling, K.K.; Ko, C.-P.; Pellizzoni, L.; Lorson, C.L. Minor snRNA gene delivery improves the loss of proprioceptive synapses on SMA motor neurons. *JCI Insight* **2020**, *5*. [CrossRef]

85. Hastings, M.L.; Resta, N.; Traum, D.; Stella, A.; Guanti, G.; Krainer, A.R. An LKB1 AT-AC intron mutation causes Peutz-Jeghers syndrome via splicing at noncanonical cryptic splice sites. *Nat. Struct. Mol. Biol.* **2004**, *12*, 54–59. [CrossRef] [PubMed]

86. Shaw, M.A.; Brunetti-Pierre, N.; Kadasi, L.; Kovacova, V.; Van Maldergem, L.; De Brasi, D.; Salerno, M.; Gecz, J. Identification of three novel SEDL mutations, including mutation in the rare, non-canonical splice site of exon 4. *Clin. Genet.* **2003**, *64*, 235–242. [CrossRef] [PubMed]

87. Elsaid, M.F.; Chalhoub, N.; Ben-Omran, T.; Kumar, P.; Kamel, H.; Ibrahim, K.; Mohamoud, Y.; Al-Dous, E.; Al-Azwani, I.; Malek, J.A.; et al. Mutation in noncoding RNA RNU12 causes early onset cerebellar ataxia. *Ann. Neurol.* **2017**, *81*, 68–78. [CrossRef] [PubMed]

88. Inoue, D.; Polaski, J.T.; Taylor, J.; Castel, P.; Chen, S.; Kobayashi, S.; Hogg, S.J.; Hayashi, Y.; Pineda, J.M.B.; El Marabti, E.; et al. Minor intron retention drives clonal hematopoietic disorders and diverse cancer predisposition. *Nat. Genet.* **2021**, *53*, 707–718. [CrossRef] [PubMed]

89. Edery, P.; Marcaillou, C.; Sahbatou, M.; Labalme, A.; Chastang, J.; Touraine, R.; Tubacher, E.; Senni, F.; Bober, M.B.; Nampoothiri, S.; et al. Association of TALS Developmental Disorder with Defect in Minor Splicing Component U4atac snRNA. *Science* **2011**, *332*, 240–243. [CrossRef]

90. Farach, L.S.; Little, M.E.; Duker, A.L.; Logan, C.V.; Jackson, A.; Hecht, J.T.; Bober, M. The expanding phenotype of RNU4ATAC pathogenic variants to Lowry Wood syndrome. *Am. J. Med. Genet. Part A* **2018**, *176*, 465–469. [CrossRef]

91. Merico, D.; Roifman, M.; Braunschweig, U.; Yuen, R.K.C.; Alexandrova, R.; Bates, A.; Reid, B.; Nalpathamkalam, T.; Wang, Z.; Thiruvahindrapuram, B.; et al. Compound heterozygous mutations in the noncoding RNU4ATAC cause Roifman Syndrome by disrupting minor intron splicing. *Nat. Commun.* **2015**, *6*, 8718. [CrossRef] [PubMed]

92. Piotrowski, A.; Xie, J.; Liu, Y.F.; Poplawski, A.B.; Gomes, A.R.; Madanecki, P.; Fu, C.; Crowley, M.R.; Crossman, D.K.; Armstrong, L.; et al. Germline loss-of-function mutations in LZTR1 predispose to an inherited disorder of multiple schwannomas. *Nat. Genet.* **2014**, *46*, 182–187. [CrossRef] [PubMed]

93. Reber, S.; Stettler, J.; Filosa, G.; Colombo, M.; Jutzi, D.; Lenzken, S.C.; Schweingruber, C.; Bruggmann, R.; Bachi, A.; Barabino, S.M.; et al. Minor intron splicing is regulated by FUS and affected by ALS-associated FUS mutants. *EMBO J.* **2016**, *35*, 1504–1521. [CrossRef]

94. Gabanella, F.; Butchbach, M.E.R.; Saieva, L.; Carissimi, C.; Burghes, A.H.M.; Pellizzoni, L. Ribonucleoprotein Assembly Defects Correlate with Spinal Muscular Atrophy Severity and Preferentially Affect a Subset of Spliceosomal snRNPs. *PLoS ONE* **2007**, *2*, e921. [CrossRef] [PubMed]

95. Zhang, Z.; Lotti, F.; Dittmar, K.; Younis, I.; Wan, L.; Kasim, M.; Dreyfuss, G. SMN Deficiency Causes Tissue-Specific Perturbations in the Repertoire of snRNAs and Widespread Defects in Splicing. *Cell* **2008**, *133*, 585–600. [CrossRef] [PubMed]

96. Workman, E.; Saieva, L.; Carrel, T.L.; Crawford, T.O.; Liu, D.; Lutz, C.; Beattie, C.E.; Pellizzoni, L.; Burghes, A.H. A SMN missense mutation complements SMN2 restoring snRNPs and rescuing SMA mice. *Hum. Mol. Genet.* **2009**, *18*, 2215–2229. [CrossRef]

97. Boulisfane, N.; Choleza, M.; Rage, F.; Neel, H.; Soret, J.; Bordonné, R. Impaired minor tri-snRNP assembly generates differential splicing defects of U12-type introns in lymphoblasts derived from a type I SMA patient. *Hum. Mol. Genet.* **2010**, *20*, 641–648. [CrossRef]

98. Lotti, F.; Imlach, W.L.; Saieva, L.; Beck, E.S.; Hao, L.T.; Li, D.K.; Jiao, W.; Mentis, G.Z.; Beattie, C.E.; McCabe, B.D.; et al. An SMN-Dependent U12 Splicing Event Essential for Motor Circuit Function. *Cell* **2012**, *151*, 440–454. [CrossRef]

99. Jangi, M.; Fleet, C.; Cullen, P.; Gupta, S.V.; Mekhoubad, S.; Chiao, E.; Allaire, N.; Bennett, C.F.; Rigo, F.; Krainer, A.R.; et al. SMN deficiency in severe models of spinal muscular atrophy causes widespread intron retention and DNA damage. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E2347–E2356. [CrossRef]

100. Doktor, T.K.; Hua, Y.; Andersen, H.S.; Brøner, S.; Liu, Y.H.; Wieckowska, A.; Dembic, M.; Bruun, G.H.; Krainer, A.R.; Andresen, B.S. RNA-sequencing of a mouse-model of spinal muscular atrophy reveals tissue-wide changes in splicing of U12-dependent introns. *Nucleic Acids Res.* **2017**, *45*, 395–416. [CrossRef]

101. Gilliam, A.C.; Steitz, J.A. Rare scleroderma autoantibodies to the U11 small nuclear ribonucleoprotein and to the trimethylguanosine cap of U small nuclear RNAs. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 6781–6785. [CrossRef]

102. Zieve, G.W.; Sauterer, R.A.; Margolis, R.L. Cell Biology of the snRNP Particle. *Crit. Rev. Biochem. Mol. Biol.* **1990**, *25*, 1–46. [CrossRef] [PubMed]

103. Lehmeier, T.; Foulaki, K.; Lührmann, R. Evidence for three distinct D proteins, which react differentially with anti-Sm autoantibodies, in the cores of the major snRNPs U1, U2, U4/U6 and U5. *Nucleic Acids Res.* **1990**, *18*, 6475–6484. [CrossRef] [PubMed]

104. He, H.; Liyanarachchi, S.; Akagi, K.; Nagy, R.; Li, J.; Dietrich, R.C.; Li, W.; Sebastian, N.; Wen, B.; Xin, B.; et al. Mutations in U4atac snRNA, a component of the minor spliceosome, in the developmental disorder MOPD I. *Science* **2011**, *332*, 238–240. [CrossRef] [PubMed]

105. Wang, Z.; Lo, H.S.; Yang, H.; Gere, S.; Hu, Y.; Buetow, K.H.; Lee, M.P. Computational analysis and experimental validation of tumor-associated alternative RNA splicing in human cancer. *Cancer Res.* **2003**, *63*, 655–657. [PubMed]

106. Markmiller, S.; Cloonan, N.; Lardelli, R.M.; Doggett, K.; Keightley, M.-C.; Boglev, Y.; Trotter, A.J.; Ng, A.Y.; Wilkins, S.J.; Verkade, H.; et al. Minor class splicing shapes the zebrafish transcriptome during development. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 3062–3067. [CrossRef]

107. Krishnakumar, R.; Kraus, W.L. The PARP Side of the Nucleus: Molecular Actions, Physiological Outcomes, and Clinical Targets. *Mol. Cell* **2010**, *39*, 8–24. [CrossRef]

108. Burghes, A.H.M.; Beattie, C.E. Spinal muscular atrophy: Why do low levels of survival motor neuron protein make motor neurons sick? *Nat. Rev. Neurosci.* **2009**, *10*, 597–609. [CrossRef]

109. Tisdale, S.; Pellizzoni, L. Disease mechanisms and therapeutic approaches in spinal muscular atrophy. *J. Neurosci.* **2015**, *35*, 8691–8700. [CrossRef]

110. Rigo, F.; Hua, Y.; Krainer, A.R.; Bennett, C.F. Antisense-based therapy for the treatment of spinal muscular atrophy. *J. Cell Biol.* **2012**, *199*, 21–25. [CrossRef]

111. Martinez-Montiel, N.; Rosas-Murrieta, N.H.; Ruiz, M.A.; Monjaraz-Guzman, E.; Martinez-Contreras, R. Alternative Splicing as a Target for Cancer Treatment. *Int. J. Mol. Sci.* **2018**, *19*, 545. [CrossRef] [PubMed]

112. Yuan, J.; Ma, Y.; Huang, T.; Chen, Y.; Peng, Y.; Li, B.; Li, J.; Zhang, Y.; Song, B.; Sun, X.; et al. Genetic Modulation of RNA Splicing with a CRISPR-Guided Cytidine Deaminase. *Mol. Cell* **2018**, *72*, 380–394. [CrossRef] [PubMed]