Minor spliceosome and disease

Verma, Bhupendra Kumar

2018-07

http://hdl.handle.net/10138/240373
https://doi.org/10.1016/j.semcdb.2017.09.036

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.
Review

Minor spliceosome and disease

Bhupendra Verma¹, Maureen V. Akinyi, Antto J. Norppa, Mikko J. Frilander*

Institute of Biotechnology, PL56 (Viikinkaari 5), 000014, University of Helsinki, Finland

ARTICLE INFO

Article history:
Received 4 July 2017
Received in revised form 21 September 2017
Accepted 27 September 2017
Available online xxx

Keywords:
U12-type introns
Minor spliceosome
Human diseases
Pre-mRNA splicing
Cryptic splice sites
Exon skipping
Intron retention

ABSTRACT

The U12-dependent (minor) spliceosome excises a rare group of introns that are characterized by a highly conserved 5′ splice site and branch point sequence. Several new congenital or somatic diseases have recently been associated with mutations in components of the minor spliceosome. A common theme in these diseases is the detection of elevated levels of transcripts containing U12-type introns, of which a subset is associated with other splicing defects. Here we review the present understanding of minor spliceosome diseases, particularly those associated with the specific components of the minor spliceosome. We also present a model for interpreting the molecular-level consequences of the different diseases.

© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Contents

1. Introduction .................................................................................................................. 00
2. Minor spliceosome in human disease ........................................................................ 00
   2.1. Diseases affecting the intron recognition step .................................................... 00
      2.1.1. Isolated growth hormone deficiency ......................................................... 00
      2.1.2. Cerebellar ataxia ....................................................................................... 00
      2.1.3. Myelodysplastic syndrome ...................................................................... 00
   2.2. Diseases affecting the formation of catalytically active spliceosome ................. 00
      2.2.1. MOPD1/TALS .......................................................................................... 00
      2.2.2. Roifman syndrome ................................................................................... 00
   2.3. Diseases with mutations in the U12-type splice sites ........................................... 00
   2.4. Other spliceosome mutations affecting minor spliceosome activity .................. 00
3. Minor spliceosome components as biomarkers ........................................................ 00
4. Consequences of minor spliceosome-specific mutations .......................................... 00
   4.1. The fate of the mRNA: intron retention vs cryptic splicing .............................. 00
   4.2. Fate of the mRNA and disease severity ............................................................ 00
   4.3. Defects with individual genes vs a global splicing defect .................................. 00
5. Future perspective ....................................................................................................... 00
Acknowledgements ......................................................................................................... 00
References ....................................................................................................................... 00

* Corresponding author.
E-mail address: Mikko.Frilander@Helsinki.Fi (M.J. Frilander).
¹ Present address: Department of Biotechnology, All India Institute of Medical Sciences, New Delhi, India.

https://doi.org/10.1016/j.semcdb.2017.09.036
1084-9521/© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Please cite this article in press as: B. Verma, et al., Minor spliceosome and disease, Semin Cell Dev Biol (2017),
https://doi.org/10.1016/j.semcdb.2017.09.036
1. Introduction

Pre-mRNA splicing is an essential step in the gene expression pathway of all eukaryotes. During splicing, the non-coding intron sequences are recognized and removed from precursor mRNA (pre-mRNA) by the spliceosome, a large ribonucleoprotein complex. Defects in splicing of one or more mRNA species are a major cause of human diseases and present estimates suggest that up to 60% are associated with pre-mRNA splicing [1,2]. In addition to monogenic disorders, genetic variants altering splicing are also thought to be an important contributor to complex diseases and cancer [3,4]. The majority of disease-associated splicing defects are cis-acting mutations within a single gene that disrupt either the splice site consensus sequences or splicing regulatory elements located in introns or exons [5]. Notably, exonic mutations interpreted as missense, nonsense or silent also commonly affect splicing [4,6,7]. The outcome of cis-acting splicing mutations is often either a formation of abnormal or non-functional protein species or accelerated decay of the affected individual mRNAs. In contrast, mutations in splicing factors, spliceosome components or spliceosome assembly factors often lead to widespread defects in the processing of large numbers of pre-mRNAs [8–10].

Most metazoan species contain two distinct pre-mRNA splicing machineries known as the major (U2-dependent) and minor (U12-dependent) spliceosomes, which recognize and excise either the major (U2-type) or minor (U12-type) class of introns, respectively. In contrast to the major introns, U12-type introns are characterized by divergent and highly conserved 5′ splice site (5′ss) and branch point sequences (BPS) (Fig. 1A; [11]). These introns also lack the characteristic polypyrimidine tract (PPT) that is present in U2-type introns immediately upstream of the 3′ splice site (3′ss).

Minor introns constitute only ~0.35% of all human introns and have been reported to be present in 700–800 genes, each of which typically carry only a single U12-type intron and multiple U2-type introns. U12-type introns are enriched in genes that represent a rather restricted set of functional classes and pathways. Particularly, they are present in genes related to ‘information processing functions’, such as DNA replication and repair, transcription, RNA processing, and translation, but can also be found in genes related to cytoskeletal organization, vesicular transport, and voltage-gated ion channel activity, as suggested originally by Burge et al., [12] and verified later [13,14]. Both the identities of genes carrying U12-type introns and their positions within the genes are evolutionarily conserved [15].

The overall organisation and functional features of both spliceosomes are highly similar. Both are composed of five small nuclear RNA (snRNA) molecules that associate with protein factors to give rise to small nuclear ribonucleoproteins (snRNPs). Within the minor spliceosome, four of the five snRNAs are unique. Specifically, U1, U12, U4atac, and U6atac replace the major spliceosome counterparts U1, U2, U4 and U6 snRNAs, respectively. U5 snRNA is shared between the two spliceosomes. Of the 200–300 proteins associated with spliceosomes, most are thought to be shared between the two systems and only 7 proteins, associated with U11 and U12 snRNPs, are unique to the minor spliceosome [16,17].

The highly similar spliceosome composition is reflected in the conserved assembly pathway and catalytic mechanism. Both spliceosomes are assembled sequentially starting from intron recognition, followed by formation of a catalytically active spliceosome and joining of the exons flanking the excised intron. With minor introns, the 5′ss and BPS are co-operatively recognized by a pre-formed U11/U12 di-snRNP [18], contrary to the sequential recognition of these sequences by individual U1 and U2 snRNPs in major introns. Since the PPT is lacking in minor introns, the U2AF1/2 heterodimer that recognizes the PPT and 3′ss of major introns does not associate with minor introns. Instead, an integral U11/U12 di-snRNP protein component, Urp/ZRSR2 takes up the role of 3′ss recognition with minor introns [19]. Following this initial recognition, both splicing pathways proceed with association of a tri-snRNP, either U4atac/U6atac.U5 or U4/U6.U5 (Fig. 1B; [20,21]). Further rearrangements in RNA–RNA and RNA–protein interactions lead to the formation of a catalytically active spliceosome and catalytic excision of the intron [20,22,23].

Here, we review the role of the minor spliceosome in human diseases, with a specific focus on diseases caused by mutations in the integral components of the minor spliceosome. Given that most protein components and U5 snRNA are shared with the major spliceosome, there are also several diseases where mutations disrupting shared components can potentially affect the functions of both spliceosomes. Those have been discussed in detail elsewhere [24].

2. Minor spliceosome in human disease

The direct targets for human diseases specific for the minor spliceosome are the unique snRNA and protein components. This includes several components of the U11/U12 intron recognition complex and the U4atac and U6atac snRNAs in the minor tri-snRNP. The U11/U12 di-snRNP contains, in addition to the U11 and U12 snRNAs, seven integral proteins (65K, 48K, 59K, 35K, 31K, 25K and 20K) that are unique to the minor spliceosome [25,26]. Additionally, the Urp/ZRSR2 protein associated with the U11/U12 di-snRNP has been reported to function in both minor and major spliceosomes, with an essential role for U12-type intron 3′ss recognition [19].

To date five human diseases with mutations in the specific components of the minor spliceosome have been described (Table 1). Three of them affect the components of the U11/U12 di-snRNP, namely the U11/U12-65K protein (RNPC3; [27]), U12 snRNA (RN1U12; [28]) and Urp protein (ZRSR2; [29]); while two diseases are attributed to mutations in the U4atac snRNA (RNU4ATAC; [30–32]). Each of these diseases is hypomorphic, leading only to a partial loss of minor spliceosome function because correctly spliced mRNAs can be detected in the patient cells. We briefly introduce each disease and later discuss the impact of disease mutations on the assembly of the minor spliceosome and the fate of affected mRNAs.

Apart from mutations in the core minor spliceosome components, we also describe the few reported human diseases caused by mutations at the splice sites of U12-type introns, though we note that this appears to be underexplored territory given the small number of cases reported so far. Finally, we mention the emerging role for the minor spliceosome in cancer and autoimmune disorders as well as neurodegenerative diseases.

2.1. Diseases affecting the intron recognition step

2.1.1. Isolated growth hormone deficiency

Recessive mutations in the RNPC3 gene, encoding U11/U12-65K, one of the seven minor spliceosome-specific proteins, have been associated with isolated growth hormone deficiency (IGHD) and associated pituitary hypoplasia. The 65K protein is part of a molecular bridge that connects the U11 and U12 snRNPs into a di-snRNP (Fig 1B; [33]). Initially, RNPC3 mutations were detected in a single family only [27], but additional cases with overlapping mutations and similar phenotypes have been described subsequently [34]. IGHD is a condition characterized by a shortage or absence of growth hormone, with the absence of associated pituitary hormone deficiencies. Genetically, IGHD is a diverse disease and can result from either recessive or dominant mutations in various genes involved in pituitary development or function [35].
All reported IGHD cases associated with RNPC3 mutations are compound heterozygous with a missense P474T mutation combined with either R502X or R205X nonsense mutation [27,34]. At the protein level the P474T and R502X mutations map to the C-terminal RNA recognition motif (RRM) of the protein, while the R205X maps to the proline-rich region between the two RRMs (Fig. 2A). Although the tissue affected in IGHD, the pituitary gland, cannot be sampled from the patients, RT-PCR and RNA-seq analyses...
Table 1

Human diseases showing defects in splicing of U12-type introns.

<table>
<thead>
<tr>
<th>Defect in</th>
<th>Locus</th>
<th>Disease</th>
<th>Abbreviation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integral components of minor snRNPs</td>
<td>RNP3C</td>
<td>Isolated Growth Hormone Deficiency</td>
<td>IGHD</td>
<td>[27]</td>
</tr>
<tr>
<td></td>
<td>ZRSR2</td>
<td>Myelodysplastic Syndrome</td>
<td>MDS</td>
<td>[29]</td>
</tr>
<tr>
<td></td>
<td>RNU12</td>
<td>Early-onset Cerbellar ataxia</td>
<td>EOCA</td>
<td>[28]</td>
</tr>
<tr>
<td></td>
<td>RN4ATAC</td>
<td>Microcephalic osteodysplastic primordial dwarfism/Taybi-Linder Syndrome</td>
<td>MOPD1/TALS</td>
<td>[31,32]</td>
</tr>
<tr>
<td></td>
<td>RN4ATAC</td>
<td>Roifman syndrome</td>
<td>RFMN</td>
<td>[30]</td>
</tr>
<tr>
<td>U12-type 5′ss</td>
<td>STK11</td>
<td>Peutz-Jegher’s syndrome</td>
<td>PJS</td>
<td>[49]</td>
</tr>
<tr>
<td>Biogenesis/Assembly</td>
<td>TRAPP2</td>
<td>Spondyloepiphysseal dysplasia tarda</td>
<td>SE1D</td>
<td>[50]</td>
</tr>
<tr>
<td></td>
<td>FUS</td>
<td>Amyotrophic lateral sclerosis</td>
<td>ALS</td>
<td>[64]</td>
</tr>
<tr>
<td></td>
<td>SMN1</td>
<td>Spinal muscular atrophy</td>
<td>SMA</td>
<td>[10,57–61]</td>
</tr>
</tbody>
</table>

**Fig. 2.** Disease-associated mutations in the specific protein and snRNA components of the minor spliceosome.

A) Domain structure of the U11/U12-65K proteins with the locations of IGHD-associated mutations indicated.

B) Secondary structure of U12 snRNA showing the location of a point mutation associated with EOCA. Binding site for the U11/U12-65K protein is also indicated.

C) Secondary structure of the U4atac/U6atac di-snRNA complex showing mutations associated with MOPD1 (red) and RFMN (cyan). A mutation shared between MOPD1/TALS and RFMN patients is indicated in red text surrounded by a cyan rectangle. Binding sites for 15.5K, PRPF31, the PRPF4/PRPF3/PHH complex proteins and the U11/U12-65K protein are also shown.

D) Domain structure of the ZRSR2 protein showing mutations associated with MDS. Mutations are from the COSMIC database [76]. Mutations studied by Madan et al. [29] are shown in blue.

was carried out on P474T/R502X patient lymphocytes [27]. Consistent with the function of the 65K protein in the intron recognition complex, several types of aberrant splicing events associated with a subset of U12-type introns were detected, including increased retention of U12-type introns, activation of nearby cryptic U2-type splice sites and exon skipping events. Biochemical analysis of patient cells revealed a reduction in levels of the 65K protein, as well as reduced stability of the U11/U12 di-snRNP complex. This result is consistent with the function of the 65K protein as a component of the molecular bridge between U11 and U12 snRNPs through interactions with the U11-59K protein and U12 snRNA [33]. The 65K-U12 snRNA interaction is mediated by the C-terminal RRM domain and, as predicted, the IGHD mutations were shown to either reduce (P474T) or eliminate (R502X) the binding of 65K
to the stem–loop III of U12 snRNA [Fig. 2B; [36]]. The picture that is emerging from these studies is that the IGHD mutations lead to impaired interaction between the 65K protein and U12 snRNA, which in turn causes destabilization or reduced formation of the U11/U12 di-snRNP. Additionally, it is likely that both RNPC3 nonsense mutations described for IGHD lead to allele–specific decay of the 65K mRNA via nonsense mediated decay (NMD) pathway, as shown recently for the R502X allele [36]. This is expected to significantly reduce the levels of 65K proteins encoded by the nonsense alleles.

However, we note that additional mechanisms may be at play, as it is not currently known whether the 65K protein is purely a structural component of the U11/U12 di-snRNP, or whether it has additional functions in the splicing process. A zebrafish study indicated that the 65K protein may have other functions later in the spliceosome assembly [37]. Consistently, recent work reported an interaction between 65K and the 3′ terminal stem–loop of U6atac snRNA (Fig. 2C; [38]), which is similarly impaired by the IGHD mutations [36].

2.1.2. Cerebellar ataxia

A mutation in the noncoding U12 snRNA gene (RNU12) has recently been associated with an early-onset cerebellar ataxia (EOCA) [28]. In general, ataxias are a diverse group of disorders characterized by defects in muscle coordination which, in the case of cerebellar ataxias, results from abnormal development and/or degeneration of the cerebellum. The patients in this single study are homozygous for a point mutation (84C>T; see Fig. 2B) with symptoms of hypotonia at infancy, delayed motor development, abnormal gait, speech and learning difficulties [28]. RNAseq and RT-PCR analyses of patient mononuclear cells showed increased U12-type intron retention and upregulation of the U12 snRNA. The 84C>T mutation is located at the base of U12 snRNA stem–loop III, and is predicted to weaken the stem–closing G–C base-pair by converting it to a weaker G–U base-pair (Fig. 2B). Mechanistic understanding of how this mutation leads to a defect in minor spliceosome function is currently lacking. The mutation does not directly overlap with any known functional elements, such as sequences involved in known RNA–RNA or RNA–protein interactions; however, the apical part of this long stem-loop acts as the binding site for the 65K protein (Fig 2B; [33]) and the assembly site for the Sm protein ring, an integral part of all snRNPs, is located only two nucleotides away. Furthermore, as the U12 snRNA not only recognizes the branch point sequence during intron recognition, but is also part of the catalytic core of the minor spliceosome, several additional mechanistic scenarios are conceivable. Further functional studies are needed to uncover the mechanistic basis of this disease.

2.1.3. Myelodysplastic syndrome

Myelodysplastic syndrome (MDS; myelodysplasia) is a term for a group of disorders characterized by inefficient hematopoiesis, with a risk of progression to acute myeloid leukemia (AML). In recent years, several whole-exome sequencing studies have revealed frequent somatic mutations in genes encoding splicing factors, including U2AF1, ZRSR2, SRSE2 and SF3B1 in patients with MDS reviewed by [9,39]. Although patients with a combination of different splicing factor mutations have been reported, typically only one splicing factor gene in each patient has been mutated. Furthermore, the splicing factor mutations frequently co-occur with mutations in genes encoding epigenetic factors, cell signaling regulators and transcriptional regulators [39,40]. Here, we focus on the U1p/ZRSR2 protein involved in U12-type 3′ss recognition.

The ZRSR2 protein shares both primary structure and functional features with the splicing factor U2AF1, that facilitates 3′ss recognition with the major introns. Both contain a central U2AF-homology domain (UHM) flanked on each side by a CCCH-type zinc finger, and an RS domain near the C-terminus (Fig. 2D). ZRSR2 is an integral component of the U11/U12 di-snRNP and functions in the recognition of the U12-type 3′ splice site, but it has also been reported to function during the second catalytic step of U2-type intron splicing [19,25]. MDS-associated mutations in ZRSR2 include a wide spectrum of nonsense, missense, frameshift and splice site mutations, which are uniformly scattered across the gene (Fig. 2D). This distribution is in striking contrast to MDS cases with mutations in U2AF1, SRSE2 and SF3B1 genes that show a narrower mutation spectrum, with particular preference for missense mutations that are located mostly within a small number of hotspot positions.

Of the numerous somatic mutations reported in the ZRSR2 (Fig. 2D), only few have been studied in detail [29]. In the Madan et al. [29] RNAseq study, bone marrow samples from eight ZRSR2-mutant MDS patients, each carrying a different mutation, were compared to both ZRSR2 wild-type MDS patients and healthy individuals. The analysis revealed a global increase in U12-type intron retention, as well as activation of cryptic U2-type splice sites only in MDS patients carrying the mutated ZRSR2 allele. Importantly, the effect of ZRSR2 mutations and knockdown of ZRSR2 on U2-type intron splicing were comparatively small. The affected U2-type introns were mostly in genes containing U12-type introns. Similar conclusions have also been reported in plants [41]. These findings are in stark contrast to the earlier biochemical experiments which suggested that an essential role for ZRSR2 was in splicing of both U2– and U12-type introns [19].

2.2. Diseases affecting the formation of catalytically active spliceosome

2.2.1. MOPD1/TALS

Recessive mutations in the U4atac snRNA gene (RNU4atac) associated with MOPD1/TALS, a rare autosomal recessive disease, represent the first human disease with mutations in a specific component of the minor spliceosome [31,32,42]. U4atac snRNA is an essential component of the U4atac/U6atac.U5 tri-snRNP (minor tri-snRNP), which joins the nascent spliceosome after the intron boundaries have been recognized by the U11/U12 di-snRNP (Fig. 1B). During the assembly of the minor tri-snRNP, U4atac forms an extensively base-paired duplex with the U6atac snRNA (Figs. 1B and 2C). This U4atac/U6atac di-snRNA complex forms a platform for the binding of tri-snRNP-specific proteins that are needed for the association of the U5 snRNP to the complex.

MOPD1/TALS in its severe form is characterized by intrauterine and postnatal growth retardation, developmental defects in several organs, including microcephaly, and death typically within 3 years after birth. However, milder forms of the disease have been reported displaying more subtle growth retardation, less severe developmental defects and/or survival to at least 12 years of age or even adulthood [43,44]. Presently, individual single-nucleotide point mutations at ten different positions within the U4atac snRNA have been associated with MOPD1/TALS (Fig. 2C); in addition, one patient with a duplication of U4atac nucleotides 16–100 has been reported [44]. All patients are either homozygous or compound heterozygous for the disease mutations. In contrast, heterozygous parents carrying only a single disease-causing mutation are phenotypically normal, suggesting that a loss of a single allele is well-tolerated at the cellular level.

The majority of disease-causing mutations, particularly those leading to severe forms of the disease, are found in the U4atac 5′ stemloop, with only few pathogenic mutations located elsewhere in the U4atac snRNA (Fig. 2C). Earlier studies from the major spliceosome have indicated that this region is recognized by a group of proteins (Fig. 2C), particularly the 15.5K (Snu13) and PRPF31 proteins, that are both necessary for the association of U5 snRNP to
the tri-snRNP [45]. Consistently, the MOPD1/TALS mutations in this region reduce the 15.5K binding, which in turn decreases the cellular levels of U4atac/U6atac.U5 tri-snRNP [46]. Importantly, there is a correlation between disease severity and the affinity of the 15.5K protein. Mutations on the tip of the 5′ stemloop that reduce most the 15.5K binding, also lead to severe forms of the disease. In contrast, the more distal mutations (such as U4atac positions 30, 53 and 55—see Fig. 2C) have little effect on the binding of the 15.5K protein and are consequently associated with milder symptoms, such as longer life expectancy [43,46]. Additionally, at least one of the mutations (124G>A) reduces the U4atac snRNA expression levels [46].

No transcriptome studies have yet been conducted for MOPD1/TALS, but qPCR analyses have shown increased U12-type intron retention in patient cells that can be rescued by expression of wild type U4atac snRNA [31,32]. Significantly, an analysis with a reporter system indicates that >90% of the minor spliceosome activity may have been compromised [31]. However, it is possible that the reporter system used may have exacerbated the splicing defect as the endogenous transcripts in patient cell lines and iPSC cells carrying the mutation show much milder splicing defects [31,32,46].

2.2.2. Roifman syndrome

In addition to MOPD1/TALS, mutations in the RNU4atac locus have been associated with a phenotypically different Roifman syndrome (RFMN) [30]. Patients suffering from RFMN are characterized by poor pre- and post-natal growth, distinct facial dysmorphism, cognitive delay and immunological defects which, with the exception of growth defects, are distinct to Roifman patients [30]. Interestingly, all RFMN patients are compound heterozygotes, with one mutation shared with MOPD1/TALS patients with severe forms of the diseases (Fig. 2C). The other mutation is predominantly located among the highly conserved positions of the RNU4atac stem II, except for a single case with a mutation within the Sm site.

No functional studies have been performed with the RFMN mutations, but it is safe to assume that the mutations shared with the MOPD1/TALS lead to a similar defect in minor tri-snRNP assembly [46]. In contrast, the stem II mutations presumably lead to a milder functional defect in minor splicing, given the overall less severe disease phenotype. While the molecular defects resulting from the stem II mutations remain to be determined, the investigation from the major spliceosome assembly tentatively suggest that the association of the PRPF4/PRPF3/PPIH complex (Fig. 2C) with the stem II may be compromised [30,47,48].

In addition, RNAseq analyses of the patient cells revealed a widespread retention of U12-type introns. Significantly, neither increased retention of U2-type introns, nor other types of alternative splicing changes were detected in the patient cells [30], suggesting a very specific splicing defect affecting U12-type introns only.

2.3. Diseases with mutations in the U12-type splice sites

To our knowledge there are only two reported cases in literature that specifically attribute U12-type splice site mutations with human disease. These include the autosomal dominant disorder Peutz-Jeghers syndrome (PJS), characterized by the presence of multiple intestinal polyps and a high risk for cancer [49], and Spondyloepiphyseal dysplasia tarda (SEDT), an X-linked recessive disorder [50]. Both are caused by mutations in a U12-type 5′ss that lead to incorrect splicing of the respective mRNAs. Additionally, the Motor endplate disease in mouse reports a mutation of a U2-type 5′ss in a downstream intron, leading to skipping of the upstream U12-type intron [51].

These three cases do not only demonstrate that a single point mutation in the highly conserved U12-type 5'ss sequence can efficiently disable the U12-type intron recognition, but also that splicing of a U12-type intron is linked to other splicing events in the same transcript and that in some cases the outcome may be difficult to predict. PJS, in which a mutation at the +1 position of an AT–AC-type U12-type intron is changed to G, is a particularly illuminating case [49]. At the sequence level, it appears to be a simple U12-type 5′ss subtype change, but the experimentally determined outcome is in fact complex, showing activation of several non-canonical 3′ splice sites that illuminate the flexibility of the U12-dependent spliceosome in 3′ss recognition. Even though the outcome in each case is the same, i.e. introduction of a premature STOP-codon and mRNA decay via the NMD pathway, this particular mutation demonstrates that simple point mutations can lead to unexpected consequences.

2.4. Other spliceosome mutations affecting minor spliceosome activity

In addition to diseases caused by mutations in the specific components of the minus spliceosome, there are several additional human diseases caused by congenital or somatic mutations in either the shared components of the two spliceosomes or the assembly factors necessary for both major and minor snRNPs. Examples of the former include Retinitis pigmentosa (RP) mutations in genes encoding the tri-snRNP specific proteins PRP5, PRP4, PRP6, PRP8, PRPF31 and BRR2 [Reviewed in 8], and MDS mutations in PRPF8 [52], in SF3b1, a component of both U2 snRNP and U11/U12 di-snRNP [26], and in SRSF2, a splicing activator potentially affecting both spliceosomes [53]. Spinal muscular atrophy (SMA) represents the latter group and is caused by mutations in the SMN1 gene, coding for an essential factor for the assembly of all Sm-class snRNPs, including components of both major and minor spliceosomes [54].

With RP, the shared protein composition between the minor and major tri-snRNPs [17] and functional studies [55,56] predict that the mutations should have an equal effect on both the major and minor tri-snRNPs. With SMA, however, several investigations using different model systems have reported a preferential reduction in the cellular levels of minor snRNPs [10,57–61]. In a subset of the studies this has been reported to lead to an increased retention U12-type introns, suggesting that defects in the splicing of minor introns may contribute to the SMA symptoms. In a Drosophila model of SMA, the effects on neuronal cells were attributed to a single U12-type intron containing gene, Stasimon, encoding a transmembrane protein required for axonal growth and regulation of synaptic transmission of motor neurons [60]. However, some aspects of this study have been challenged subsequently [62].

Finally, two independent studies have linked minor spliceosome to amyotrophic lateral sclerosis (ALS). Ishihara et al. [63] reported that ALS-associated depletion of TDP-43 protein leads to a co-depletion of U12 snRNA both in cultured cells and in patient samples, presumably due to defects in snRNP assembly. In another study [64], ALS-associated mutations in the Fused in sarcoma (FUS) protein lead to reduced nuclear import of FUS and its accumulation in the cytoplasm. FUS is a multifunctional RNA-binding protein that interacts with several splicing factors, including U11 snRNP and is needed for the splicing of at least a subset of minor introns [64]. Significantly, the cytoplasmic accumulation of a particular FUS mutant leads to a concomitant mislocalization of U11 and U12 snRNPs in the cytoplasm, which in turn leads to defects in the splicing of U12-type introns.
3. Minor spliceosome components as biomarkers

In addition to being a target for disease-causing mutations, several components of the minor spliceosome have been reported as cancer biomarkers. A recent study by Xu et al. [65] using serum from patients with scleroderma and coincident cancer reported autoantibodies targeted to components of the U11/U12 di-snRNP, particularly the U11/U12-65K protein, but also the 25K, 35K and 59K proteins. Of these, the U11/U12-65K autoantibodies appear to show a reliable association with an increased risk of cancer occurring within 2 years of the onset of scleroderma [66]. Similar associations have been described for U11 snRNA in familial prostate cancers [67] and U11-59K protein in Acute Myeloid Leukemia [68]. However, the role of minor spliceosome-specific autoantibodies in the pathogenesis of scleroderma and/or coincident cancer, or the expression level changes of other minor spliceosome components in other diseases is presently unclear.

4. Consequences of minor spliceosome-specific mutations

Presently, mutations in four specific components of the minor spliceosome have been associated with human diseases showing very different disease phenotypes. RT-PCR and RNAseq analyses have demonstrated that each of the five human diseases show the expected molecular level phenotype, i.e. the increased levels of unspliced U12-type introns in the patient cells. With a subset of diseases, additional splicing defects, particularly activation of cryptic U2-type splice sites and exon skipping events have also been described. Presently, there has not been clear consensus on how to interpret the molecular consequences of minor spliceosome mutations. Here, we propose a simple classification of disease-causing mutations according to their impact on the recognition of the U12-type introns. This provides a simple framework for predicting the molecular consequences of the mutations, but may also help to interpret the subsequent disease phenotypes.

4.1. The fate of the mRNA: intron retention vs cryptic splicing

According to this classification the mutations associated with both the MOPD1/TALS and RFMN affect the formation of the U4atac/U6atac.U5 tri-snRNP (see 2.2.1 and 2.2.2), but not the initial intron recognition step. Therefore, in these diseases the spliceosome assembly is arrested at the pre-spliceosome stage where the U11/U12 di-snRNP is bound to the intron but the subsequent assembly steps are inhibited. Consequently, the bound U11/U12 di-snRNP can be expected to maintain exon-definition interactions with the surrounding (U2-type) introns via the RS domain containing U11-35K [69] and possibly Urp2/R5SR2 proteins. Furthermore, the U11/U12 di-snRNP bound to the partially spliced mRNA can trap such transcripts in the nucleus [70], where they may be targeted by the nuclear quality control mechanisms [71,72].

In contrast, the loss of U11/U12 binding and the recognition of U12-type introns in IGHD and MDS (and possibly ECOA) leads to a different outcome. Due to the loss of U11/U12 binding, transcripts containing unspliced U12-type introns are free to be exported to the cytoplasm. Concomitantly, the loss of local exon definition interactions may lead to increased activation of cryptic splice sites or exon skipping events near the U12-type introns [73]. The outcome can be either truncated or altered protein, or mRNA decay via the NMD quality control pathway (Fig. 3).

Is there evidence supporting this model? RNAseq analyses from both IGHD and MDS patient cells provide extensive evidence of activation of cryptic U2-type splice sites and increased levels of exon skipping near U12-type introns in addition to the expected retention of U12-type introns [27,29]. In contrast, besides the increased intron retention of U12-type introns seen with both MOPD1/TALS and RFMN, there is no evidence of the associated cryptic U2-type
splicing events [30–32,46] as stated by Merico et al. [30] in their detailed RNAseq analysis of RFMN patient cells. Together, present evidence supports the proposal that the consequences of the minor spliceosome mutations are linked to the type of assembly defect introduced. However, with both MOPD1/TALS and RFMN it is possible that some arrested transcripts trapped in the nucleus may either lose the bound U11/U12 di-snRNP or undergo the splicing process post-transcriptionally (see Fig. 3).

4.2. Fate of the mRNA and disease severity

It is somewhat puzzling to note that the severity of congenital minor spliceosome diseases and the level of mRNA splicing defect follow a counterintuitive association as described above. The two diseases showing only relatively mild intron retention defects (MOPD1/TALS and RFMN) show pleiotropic, and often severe, defects in many tissues [30–32]. In contrast, IGHD and EOCA show large splicing defects that were detected presumably ubiquitously – also from cells/tissues not associated with the disease [27,28]. Even though the present small dataset does not allow generalization, it is tempting to speculate that the difference may be linked to the fate of the unspliced mRNAs. Specifically, it is possible that the unspliced U12-type introns or transcripts containing cryptic splicing events are better tolerated when exported to cytoplasm where they can be targeted by the cytoplasmic quality control mechanisms. Transcripts trapped in the nucleus may, in contrast, accumulate in some tissues or lead to other detrimental downstream effects. This hypothesis receives tentative support from a zebrafish study where a specific 65K mutation arrests the spliceosome assembly at a late stage and presumably traps such mRNAs in the nucleus. Unlike the more restricted phenotype seen with IGHD 65K mutations, the zebrafish phenotype is severe, with developmental defects in multiple organs [37].

4.3. Defects with individual genes vs a global splicing defect

One of the outstanding questions is whether the pathological consequences of the minor spliceosome diseases are caused by a global defect in the splicing of U12-type introns or whether the various effects seen in different tissues are caused by missplicing of a small subset of genes containing U12-type introns. Gene expression analyses have suggested various candidate genes responsible for the observed disease phenotypes. These must, however, be interpreted with caution since in all cases except MDS, the gene expression analyses have been done using cells other than the actual affected tissue.

Furthermore, it is not yet possible to conclude whether a particular tissue or cell type would be more susceptible to defects in the U12-dependent spliceosome. Of the present congenital minor spliceosome diseases, the only common target appears to be the nervous system. Specifically, microcephaly has been reported for three of the diseases (severe for MOPD1/TALS, mild for RFMN and IGHD), which with MOPD1/TALS is associated with structural abnormalities of brain. Additionally, RFMN and EOCA patients display cognitive delay and cerebellar ataxia, respectively. The observed tissue-specific effects may be linked to the associations of the minor spliceosome activity with cell proliferation [37,74] or cellular response to stress [75]. Consistently, recent work has reported that components of the U11/U12 di-snRNP, including the 65K, 48K and Urp/2RSR2 proteins are downregulated during neuron terminal differentiation [70].

5. Future perspective

The recent discoveries of human diseases affecting the minor spliceosome have significantly advanced our understanding of the significance and function of the U12-dependent spliceosome. However, many aspects of the disease process, including the basis of the tissue specificity and the detailed mechanism of the disease process are yet to be determined for most diseases, requiring animal models of each disease. These will also help to address the more fundamental question of the existence of two separate spliceosomes.

Acknowledgements

This work was supported by the Academy of Finland (Grant 140087 to MJF and Grant 278798 to BV) and Sigrid Jusélius Foundation (MJF). AJN was supported by the Integrative Life Science doctoral program at the University of Helsinki.

References


Please cite this article in press as: B. Verma et al., Minor spliceosome deficiency, and disease. Semin Cell Dev Biol (2017), https://doi.org/10.1016/j.semcdb.2017.09.036


