



Review

# Mutations of Pre-mRNA Splicing Regulatory Elements: Are Predictions Moving Forward to Clinical Diagnostics?

Lucie Grodecká<sup>1</sup>, Emanuele Buratti<sup>2</sup>  and Tomáš Freiburger<sup>1,3,\*</sup> 

<sup>1</sup> Centre for Cardiovascular Surgery and Transplantation, Brno 65691, Czech Republic; lucie.grodecka@cktch.cz

<sup>2</sup> International Centre for Genetic Engineering and Biotechnology, 34149 Trieste, Italy; buratti@icgeb.org

<sup>3</sup> Central European Institute of Technology, Masaryk University, Brno 62500, Czech Republic

\* Correspondence: tomas.freiberger@cktch.cz; Tel.: +420-543-182-548

Received: 27 June 2017; Accepted: 27 July 2017; Published: 31 July 2017

**Abstract:** For more than three decades, researchers have known that consensus splice sites alone are not sufficient regulatory elements to provide complex splicing regulation. Other regulators, so-called splicing regulatory elements (SREs) are needed. Most importantly, their sequence variants often underlie the development of various human disorders. However, due to their variable location and high degeneracy, these regulatory sequences are also very difficult to recognize and predict. Many different approaches aiming to identify SREs have been tried, often leading to the development of *in silico* prediction tools. While these tools were initially expected to be helpful to identify splicing-affecting mutations in genetic diagnostics, we are still quite far from meeting this goal. In fact, most of these tools are not able to accurately discern the SRE-affecting pathological variants from those not affecting splicing. Nonetheless, several recent evaluations have given appealing results (namely for EX-SKIP, ESRseq and Hexplorer predictors). In this review, we aim to summarize the history of the different approaches to SRE prediction, and provide additional validation of these tools based on patients' clinical data. Finally, we evaluate their usefulness for diagnostic settings and discuss the challenges that have yet to be met.

**Keywords:** splicing regulatory elements; *in silico* predictions; pre-mRNA splicing; mutation; evaluation of prediction tools; variants of unknown significance; splicing aberration

## 1. Introduction

One of the most important features distinguishing prokaryotic from eukaryotic gene expression is the process of RNA splicing. During splicing, the borders of intervening sequences (so-called introns) are recognized, cleaved, and exons are then ligated together. This process is catalyzed by a large ribonucleoprotein complex termed spliceosome. Through multiple protein–RNA and RNA–RNA interactions, this huge molecular machine recognizes pre-mRNA sequence elements indicated as “splicing signals”. These include: donor splice site, branch point site and acceptor splice site. Importantly, mammalian genomes contain a huge number of pseudo splice sites, i.e. sequences that resemble real splice sites but are never used under normal conditions. These pseudo sites outnumber the authentic splice sites by an order of magnitude. Therefore, additional regulatory sequences are necessary to discern between the two and are referred to as “splicing regulatory elements” (SREs). The function of the SREs is to bind splicing activators or repressors and influence the choice of adjacent splice sites. In general, SREs play a key role both in constitutive splicing as well as the regulation of alternative splicing (the process through which more than one mRNA isoform can be created from a single precursor transcript) [1–3].

Compared to the relatively conserved splice sites, the SRE sequences are much more degenerate [4]. SREs were originally categorized according to their localization and effect on splicing as exonic or intronic splicing enhancers (ESEs and ISEs) and silencers (ESSs and ISSs). Despite still being in use, this division now seems to be rather simplistic, as some SREs can even act adversely from various exonic or various intronic positions (e.g., SRp40 binding elements act mostly as silencers in *ADAR2* exon 8, but behave as enhancers in about 25% of positions when shifted within the same exon; similarly, neuron-specific splicing factors Nova activate alternative exon inclusion when bound to the downstream intron, but repressed its inclusion when bound to the upstream intron) [2,5]. Importantly, most regulatory contexts emerge from overlapping RNA elements. The existence of composite exonic regulatory elements of splicing (CERES) proves that some of the overlapping elements may have both silencing and enhancing properties [6]. Complicating the matter a little bit more, the elements' recognition is also influenced by its accessibility, i.e. the state of chromatin and RNA secondary structure [2].

In addition, SREs can influence not only the standard pre-mRNA splicing process but most probably also backsplicing—a special type of alternative splicing which leads to circular RNA (circRNA) formation. circRNAs are abundant, stable and evolutionary conserved noncoding RNAs often expressed in a tissue specific manner. The tissue specificity and the lack of correlation between expression levels of a circRNA and the linear transcript from which it is derived indicate that the process of circRNA biogenesis may be precisely regulated [7]. Supporting the role of SREs in the circRNA biogenesis, several splicing regulators have been found to be implicated in this process. In particular, regulation of circRNAs formation was demonstrated for splicing factor Muscleblind and several hnRNP and SR family proteins in *Drosophila* [8,9]. Similarly, splicing regulators QKI, RBM20 and FUS have been described to activate or repress biogenesis of specific circRNAs in human [10–12].

In theory, the mutation of any *cis*-acting splicing element may result in aberrant splicing. In accordance, pre-mRNA splicing defects are responsible for a substantial proportion of inherited disorders (estimated between 15% and 50%) [4]. Notably, since deregulation of specific circRNAs levels has already been shown to be associated with human diseases, this could be another way through which SRE sequence variants might lead to pathology [7]. However, further research will be needed to confirm the role of particular SRE aberrations in circRNAs genesis and disease development.

Although some exons seem to be especially prone to SRE aberrations [13–15], several systematic studies showed that SRE changes generally lead to splicing defects much less frequently than splice site disruptions [16–19]. Still, SRE-affecting mutations impose a significant burden on genetic diagnostics, as they can occur virtually anywhere in the exons or introns and are extremely difficult to be distinguished from harmless changes, which are also very abundant in human genomes [20]. This issue is even more pronounced now, as the next generation sequencing produces thousands of novel variants with every single read [21].

For all these reasons, many potential SRE-affecting variants fall into the category of so-called “variants of unknown significance” [20]. To distinguish these mutations from harmless non splicing-affecting variants, medical geneticists can either use laborious *in vitro* studies or much more feasible (yet less reliable) *in silico* predictions [22]. In this review, we summarize the current state of SRE predictions and evaluate their reliability and potential use in clinical diagnostic settings.

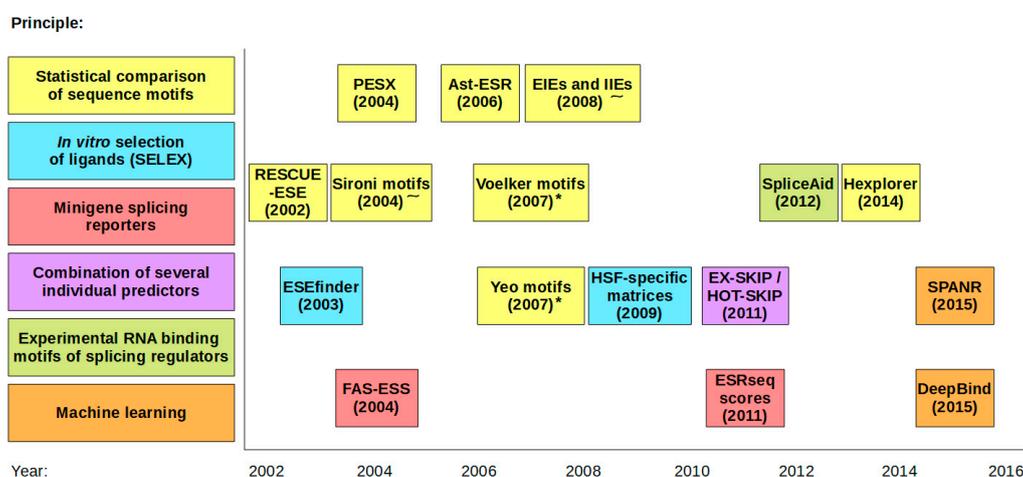
## 2. Predictions on SREs

Following the striking finding that virtually any DNA change may affect splicing by disrupting potential SREs, many attempts to locate these elements have been performed. Some systematic approaches then led to the development of several SRE-predicting tools that are listed in Table 1 and Figure 1. At first, Liu et al. attempted to define ESE motifs by using the method of “functional systematic evolution of ligands by exponential enrichment” (functional SELEX). They used this *in vitro* selection of splicing activating sequences to determine the binding preferences of four classical SR-proteins, SRSF1, SRSF2, SRSF5 and SRSF6 [23]. Based on these matrices, Cartegni et al. then

developed an online tool called ESE-finder [24]. The major advantage of this approach is that users can easily link any particular splicing-affecting sequence with its cognate binding factor. On the other hand, all sequence variants often induce ambiguous changes in the predicted scores for several individual factors, so the conclusions about the variant's effect on splicing may not be very straightforward. In addition, there is no clue what score changes are sufficient to impact splicing events, because threshold values indicated in the program may be specific for the used experimental conditions [23]. The same approach (SELEX) was later used to assign binding preferences to other splicing factors, such as hnRNPA1, Tra2beta, 9G8 and U2AF [25,26].

**Table 1.** Selected individual SRE-prediction tools.

Prediction Tool	Principle	Website	Reference	Evaluation
ESE-finder	SELEX (in vitro selection of ligands)	<a href="http://krainer01.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home">http://krainer01.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home</a>	[24]	[15,19,27–31]
ESRseq	testing of all possible k-mers for positive and negative splicing influences, based on QUEPASA method		[32]	[13,15,19]
FAS-ESS	analysis of random sequence silencing properties in the minigene settings	<a href="http://genes.mit.edu/fas-ess/">http://genes.mit.edu/fas-ess/</a>	[33]	[30]
Hexplorer	statistical comparison of hexamer sequence motifs	<a href="http://nar.oxfordjournals.org/content/early/2014/08/21/nar.gku736/suppl/DC1">http://nar.oxfordjournals.org/content/early/2014/08/21/nar.gku736/suppl/DC1</a>	[34]	[15,19]
PESX	statistical comparison of octamer sequence motifs	<a href="http://cubio.biology.columbia.edu/pesx/pesx/">http://cubio.biology.columbia.edu/pesx/pesx/</a>	[35]	[19,27]
RESCUE-ESE	statistical comparison of hexamer sequence motifs	<a href="http://genes.mit.edu/burgelab/rescue-ese/">http://genes.mit.edu/burgelab/rescue-ese/</a>	[36]	[19,27,30,31]
SPANR	splicing code, machine learning	<a href="http://tools.genes.toronto.edu/">http://tools.genes.toronto.edu/</a>	[37]	[15]
SpliceAid2	database of in vitro proved splicing factors binding sites	<a href="http://www.introni.it/spliceaid.html">www.introni.it/spliceaid.html</a>	[38]	



**Figure 1.** Individual approaches to SRE predictions showing the first release of the online tools (\* denotes the tools that are accessible via Sroogle engine; ~ denotes the tools that are accessible via Human Splicing Finder (HSF) engine).

Next, Fairbrother et al. presented another pioneering approach to identifying ESEs: Relative Enhancer and Silencer Classification by Unanimous Enrichment (RESCUE) [39]. They compared the occurrence of all possible hexanucleotide motifs in exonic and intronic sequences and both in exons with strong and weak splice sites. The rationale behind such an experiment is that ESEs are more prevalent in exons than in introns and that exons with weak splice sites are more dependent on the presence of ESEs. The authors made their results easily accessible with an online tool named RESCUE-ESE [36]. This tool is easy to use, simply showing the number and identity of predicted ESEs. However, there is no correction for possible elements' overlap and for any related context dependence.

Other scientists later adapted the same principle for further SRE identification, only using different rationales. From those, we would like to point out the more recent Hexplorer approach by the Schaal laboratory which showed improved efficiency in current evaluation studies (see below) [34]. In analogy with RESCUE-ESE, this approach relies on the computational comparison of hexamer frequencies in exonic and intronic sequences around weak and strong 5' splice sites. Its advantage comes from counting each nucleotide's probability of being part of an enhancer or a silencer from all the six hexamers overlapping it. Therefore, the program is able to better cover context dependence compared to methods based on individual predicted SRE motifs [34].

In parallel to analyzing existing genomic data, another way to define functional SREs is to use splicing reporter systems to test many different sequence combinations for enhancer or silencer properties. For example, Wang et al. adopted a systematic minigene analysis to test random decanucleotides for their ESS properties [33]. More specifically, they inserted random sequences into the middle exon of a three-exon minigene. The sequences that led to exon skipping of this otherwise constitutive exon were assigned as ESS. They can easily be searched using an online tool, FAS-ESS. Notably, one of the major drawbacks to this approach is that the discovered motifs might be functional only in the particular cellular and genomic (or minigene) context [2]. Later, Ke et al. used an analogous approach to assign enhancing or silencing properties to all hexanucleotide combinations (ESRseq scores) [32]. In order to make provision for the context dependence of the hexamer activities, these authors tested all the sequence motifs in five different exonic locations. Similar to the Hexplorer method, the ESRseq score for each hexamer was counted based on all the overlapping hexamers' individual scores. Unfortunately, these scores are accessible only as supplementary tables to the article and no online program based on this approach has yet been made available to the public.

Recently, Xiong et al. [37] presented another approach based on splicing code modeling, which they used to design the SPANR tool (Splicing-based Analysis of Variants) [37]. The tool predicts the effects that a nucleotide variant exerts on cassette exon alternative splicing. This method extracts a huge number of DNA sequence features (*cis*-elements) from exon triplets (three exons and their intervening introns) and uses them to predict the percentage of middle exon inclusion. Owing to its non-linear nature, the model incorporates context dependent effects. In addition, the tool was trained on data from 16 different human tissues, which should enable it to make provision for some tissue specific features.

In addition to genomics and minigene systems, other possible approaches to identify splicing regulatory evidence have also been followed. In particular, Piva et al. have presented their database (called SpliceAid2) of human splicing factor expression data and RNA target motifs [38]. Here, users can get predictions on potential SREs based strictly on comparison with experimentally proved splicing regulators' binding sites. The advantage of this approach is that it allows the identification of both the potential SRE element and the protein which may bind to it. Of course, the limitation is that we still largely ignore the binding site specificities of many RNA binding proteins. In addition, recent works on identifying RNA binding proteins in HeLa cells have uncovered a huge number of RNA binding proteins for which we completely ignore their potential to alter the splicing process [40].

Up until this moment, the best way to use these programs to obtain an accurate prediction of SRE elements in a given experimental system has been to use them in combination. For example, Raponi et al. adopted an integrated approach: they got the best predictions on mutation-induced exon skipping when they combined multiple individual SRE predictions together (see Table 2 for details) [41]. The resulting tools, EX-SKIP and HOT-SKIP, count the sum of ESEs and ESSs derived from individual predictions and then calculate the ESS/ESE ratio, either for the specific variant (in EX-SKIP) or for each possible single nucleotide substitution in a selected exon (in HOT-SKIP). An advantage of these tools is that the individual predictions are collectively shown on the results page, so the user does not have to approach each program individually. Such an advantage is also held by several online engines/web pages that include several individual prediction tools at one web location, e.g., Sroogle and Human Splicing Finder (Table 2) [26,42].

**Table 2.** Selected tools and online engines combining multiple SRE-prediction tools.

Prediction Tool	Included Tools	Website	Reference	Evaluation
EX-SKIP	PESE and PESS [35], FAS-ESS [33], RESCUE-ESE [36], EIEs and IIEs [43], NI-ESE and NI-ESS [44]	<a href="http://ex-skip.img.cas.cz/">http://ex-skip.img.cas.cz/</a>	[41]	[13,15,19,31,45]
HOT-SKIP	PESE and PESS, FAS-ESS, RESCUE-ESE, EIEs and IIEs, NI-ESE and NI-ESS	<a href="http://hot-skip.img.cas.cz/">http://hot-skip.img.cas.cz/</a>	[41]	
Human Splicing Finder (HSF)	ESE-finder [24], RESCUE-ESE, PESE and PESS, EIEs and IIEs, FAS-ESS and ESS decamers [33], Exonic splicing regulatory sequences [5], HSF-specific matrices for Tra2- $\beta$ , 9G8 and hnRNP A1 [26]	<a href="http://www.umd.be/HSF3/index.html">http://www.umd.be/HSF3/index.html</a>	[26]	[15]
Sroogle	ESE-finder, RESCUE-ESE, FAS-ESS, PESE and PESS, other SRE predictions according to Voelker [46], Yeo [47], Goren [5]	<a href="http://sroogle.tau.ac.il/">http://sroogle.tau.ac.il/</a>	[42]	

### 3. Efficiency of SRE Predictions

An important point about SRE predictors is their reliability in terms of concordance between predicted events and the real situation in cells, tissues and organisms. However, due to methodological constraints and lack of primary samples, most researchers often have had to rely only on in vitro results. Yet profound analyses showed rather good agreement between in vitro and in vivo splicing affection [48], suggesting that this drawback does not compromise substantially the validity of SRE predictions.

Despite the fact that developers of SRE-predictors mostly proved their functionality using independent in vitro analyses, additional studies often put the reliability of these tools in question. On one hand, SRE-prediction tools have been shown to recognize motifs with general splicing-regulatory properties. Particularly, the tools were demonstrated to statistically distinguish sequences with different propensity to activate cryptic splice sites or the splicing-affecting variants from harmless SNPs [30,49]. In parallel, Raponi et al. indicated several statistically significant correlations between SRE predictions and the level of *BRCA1* exon 6 inclusion [41]. On the other hand, difficulties arose when these predictors were tested for discerning individual splicing-affecting variants from harmless sequence changes. Many studies have therefore indicated SRE-predicting tools as less efficient, often inconclusive and difficult to interpret, possibly because most of these programs were not designed for this purpose [13,27–31]. For this reason, these programs have not been regarded as useful in clinical investigations [4].

Interestingly, several recent evaluations have pinpointed some promising achievements of newly developed algorithms (EX-SKIP, ESRseq scores and Hexplorer) to recognize SRE-affecting variants [13,15,19,45]. In particular, testing EX-SKIP with 29 variants found in five *CFTR* gene exons showed, on average, a 72.5% success rate in predicting the direction of exon inclusion change [45]. When assessing its capacity to distinguish variants capable of increasing exon skipping, the predictions suffered more from a lower sensitivity (71%) than specificity issues (75%). However, these numbers could be biased due to the higher number of mutations leading to exon skipping in the dataset compared to the silent variants and variants promoting exon inclusion. In another evaluation using 35 exonic variants in six different immunity-related genes, EX-SKIP showed reasonable sensitivity (75%) but poor specificity, possibly due to a low representation of splicing-affecting mutations in the testing data [19]. Likewise, Soukarieh et al. detected a similar sensitivity (75%) but again low specificity of EX-SKIP predictions (46%) [15].

For ESRseq scores, Di Giacomo et al. was the first study to independently show its promising potential in discerning splicing-affecting from non-affecting changes [13]. Using a tentative threshold for ESRseq score difference on 32 variants from *BRCA2* exon 7, they obtained no false negatives and just two false positive predictions on exon skipping induction. Later, Soukarieh et al. extended this analysis with four other sets of variants (from *MLH1* exon 10, *BRCA1* exon 6, *CFTR* exon 12 and *NF1*

exon 37) including 154 individual point mutations in total [15]. The predictions on exon skipping variants in individual datasets showed sensitivity to be between 67% and 100% (weighed mean: 85%) and specificity between 66% and 94% (weighed mean: 83%). The high sensitivity of these predictions was later corroborated by Grodecká et al., although the specificity remained poorer [19].

Finally, the Hexplorer tool has been shown to perform comparably well with respect to the ESRseq scores. With the five extensive variant sets tested by Soukariéh et al., it provided a sensitivity between 57% and 100% (weighed mean: 79%) and specificity between 63% and 89% (weighed mean: 74%) [15]. In another study, this tool showed 100% sensitivity but again quite a poor specificity, possibly due to the chosen dataset [19].

In fact, all the above-described evaluations of EX-SKIP, ESRseq and Hexplorer have been done using data derived from minigene splicing analysis. In most cases, this analysis reliably mirrors the splicing effects, at least in terms of induction and the direction of a splicing defect [15,19,31,48].

Therefore, to extend these observations, we have decided to further evaluate these three predictors using 20 gene variants retrieved from the literature, for which the results on splicing affection were ascertained directly from the patients' tissues. As shown in Table 3, even with this smaller set of mutations we have reached sensitivity and specificity in discerning skipping-inducing mutations comparable to the previous results [15,45], differing only in a somewhat higher specificity for the EX-SKIP tool compared to [15].

**Table 3.** Evaluation of ESRseq, Hexplorer and EX-SKIP predictors on patients' RNA-based results.

Gene	cDNA Variant	Exon	Effect on Exon Skipping	$\Delta$ ESRseq (−0.5)	Hexplorer: $\Delta$ HZ <sub>EI</sub> (−0.5)	EX-SKIP: ESS/ESE mut/wt (1)
<i>BRCA1</i>	c.5123C > A	18	increased	−2.574	−10.85	1.05
	c.5434C > G	23	increased	0.558	−1.28	1.09
	c.5453A > G	23	increased	−2.176	−15.02	1.43
	c.5096G > A	18	none	−1.731	−0.22	0.93
	c.5116G > A	18	none	1.582	2.67	0.86
<i>BRCA2</i>	c.5411T > A	23	none	0.66	4.33	0.83
	c.231T > G	3	increased	1.65	4.76	0.97
	c.439C > T	5	increased	−2.69	−13.06	1.77
	c.7992T > A	18	increased	−1.11	0.00	1.00
	c.8257_8259delCTT	18	increased	−0.51	0.52	0.98
	c.9234C > T	24	increased	−1.24	−12.07	1.12
	c.223 > C	3	none	0.37	−11.19	0.97
	c.433_435delGTT	5	none	−0.23	6.46	0.51
	c.7994A > G	18	none	−1.21	0.24	1.02
	c.8182G > A	18	none	−1.65	0.00	0.98
	c.9216G > T	24	none	−2.88	−10.94	1.04
	<i>NF1</i>	c.557A > T	5	increased	−2.43	−12.21
c.528T > A		5	none	1.17	0.70	0.90
<i>DMD</i>	c.5287C > T	37	increased	−0.70	−16.84	1.16
	c.5308A > T	37	none	−0.05	−2.85	1.05
True calls				70.0%	70.0%	70.0%
Sensitivity				80.0%	70.0%	70.0%
Specificity				60.0%	70.0%	70.0%

To further evaluate the three prediction tools, we have retrieved 20 gene variants detected in genes *BRCA1*, *BRCA2*, *NF1* and *DMD* (10 inducing aberrant splicing and 10 harmless at the level of splicing) from the literature. In all these cases, nonsense mediated decay was either prevented or not expected. For an easy comparison, we used the same thresholds as described in Soukariéh et al. (shown in table headings) [15], except from the original Hexplorer threshold which was not applicable to our data, so we used a threshold −0.5 instead. The true calls are shown in bold.  $\Delta$ ESRseq: score difference between predicted mutant and wild type ESRseq score.  $\Delta$ HZ<sub>EI</sub>: score difference between predicted mutant and wild type HZ<sub>EI</sub> score.

#### 4. Exons Susceptibility to the Splicing Defects Due to SREs Changes

As mentioned in the introduction, some exons are being described as more susceptible to SRE disruptions than others. Typical signs for these susceptible exons are: (i) above- or below-average length (exon lengths below 50 nts and above 300 nts have been connected to lower splicing efficiency, probably due to affected recognition through the exon-definition process); (ii) weak splice sites; and (iii)

alternative splicing [27,45,50,51]. However, these potentially important variables that affect the variants' effects on splicing are not considered by most of the SRE-prediction algorithms. In addition, there are several other related factors that influence the basic level of exon recognition and might thus influence exons susceptibility to SRE-affecting variants, which might be more difficult to predict *in silico*. These are RNA secondary structure, nucleosome density and transcription rate [45]. Hand in hand with that, even distant regulatory elements in promoters and transcription enhancers have been shown to frequently influence alternative splicing. Besides impacting the basic level of exons' recognition, this also represents another level of (hardly predictable) interindividual variability, as variations in these functional elements (e.g., SNPs or differences in methylation) are substantially associated with alternative splicing [52].

All these reasons may explain the discrepancies regarding results of SRE-predictor tools, that have been much more successful for some exons than for others. For example, compare the 57% vs. 83% sensitivity of Hexplorer predictions on variants in *MLH1* exon 10 and *NF1* exon 37, respectively, in Soukariéh et al. [15], or 87.5% vs. 50% success rate of EX-SKIP predictions on variants in *CFTR* exon 3 and exon 10, respectively, in Aissat et al. [45]. Hand in hand with that, while some predicted score differences showed significant correlations with the level of the induced exon skipping within individual exons, the overall extent of exon skipping between individual exons seems to be better predicted by the strength of splicing signals [15,45]. All these facts once again pinpoint the high context-dependence of the splicing regulation and the need for more complex prediction approaches to be developed.

## 5. Future Approaches at Evaluating the Effects of Splicing Disruption

The first issue, which still needs to be addressed by splicing research, is the prediction of the aberrant splicing pattern that emerges when splicing is affected. For obvious reasons, the particular outcome of a splicing-affecting mutation is crucial for diagnostic decisions. In general, the SRE-affecting mutations most usually increase or decrease inclusion of the exon they are located in [6,15,17,27,31]. However, they can also result in multiple exon skipping or activate a cryptic splice site [53–55]. In specific cases, intronic SRE mutations may lead to the activation of a pseudoexon (an intronic sequence that is neglected by the spliceosome under standard conditions) [56]. Once again, also in these cases the resulting splicing defects are extremely context-dependent.

Thus far, to the best of our knowledge, there were only a few pioneering attempts to use *in silico* prediction tools to assess whether a mutation will lead to exon skipping or to the cryptic splice site activation [57,58]. However, none of these approaches were directed at looking for potential SRE-affecting variants and it would be interesting if future research could specifically be targeted to this specific aim.

Finally, another issue that still needs to be addressed will be to predict the extent of aberrant splicing. This particular area is still in its infancy, with limited correlations between predicted variables and the level of exon skipping or inclusion [15,45]. Such correlations might be useful for prediction tool developers, but they are rather useless for the diagnosticians who have to cope with individual sequence changes. In particular, all these approaches often disregard the fact that for many genes, the level of exon inclusion is often tissue specific [6,27]. This problem will eventually have to be included in predictions, if they are ever going to be useful in diagnostics.

## 6. Clinical Significance of Splicing Aberrations

As outlined above, if we consider the huge complexity of causes and consequences of splicing aberrations, one can easily understand the challenges that a genetic diagnostician can meet when working with novel gene variants. As a result, extensive effort has been put into defining classification criteria for assessment of variants' pathogenicity, as this issue markedly impacts patients' clinical management. Some publications proposed criteria for overall pathogenicity assessment, either for disease-causing genes in general or for a specific set of genes (e.g., mismatch repair genes) [52,53].

Others aimed more specifically at coping with splicing-affecting mutations [54,55]. All these works based their classification on a five-tier scheme where the most probably pathogenic variants fall into the class 5 while those most probably benign are included in class 1. For clarity, we have summarized the classification criteria that apply to SRE-affecting variants in Table 4.

**Table 4.** Summary of clinical classification guidelines that apply to SRE-affecting variants.

Class	Observation	Reference
5: pathogenic	• assay on mRNA from patients tissue samples	[59] <sup>1</sup> [60]
	AND no wt transcript detected from variant allele	
	AND aberrant transcripts introduce PTC or deletion disrupting functional domain	only in [60]
	OR deletion disrupting protein conformation	
	OR damaging effect on the gene or gene product (extent not specified)	[61]
	AND other lines of evidence supporting variant pathogenicity <sup>2</sup> (stronger than for class 4)	
	• lab assays based on mRNA (e.g., minigenes)	[60,61]
	AND variant-specific abrogated function (extent not specified)	
AND additional frequency/co-segregation/clinical data, additional molecular/mechanistic evidences from other sources, supporting variant pathogenicity (stronger than for class 4)		
4: probably pathogenic	• assay on mRNA from patients tissue samples	[61]
	AND damaging effect on the gene or gene product (extent not specified)	
	AND other lines of evidence supporting variant pathogenicity (milder than for class 5)	
	• lab assays based on mRNA (e.g., minigenes)	[60,61]
	AND variant-specific abrogated function (extent not specified)	
	AND additional frequency/co-segregation/clinical data, additional molecular/mechanistic evidences from other sources, supporting variant pathogenicity (milder than for class 5)	
3: uncertain pathogenicity	• minigene assays	[48]
	AND complete aberrant and frameshifting effect/deletion of a functional domain effect	
	e.g., all variants that do not fall into other classes	[59]
	e.g., aberrant transcripts produce deletion not disrupting known functional domains	
e.g., change in the level of alternative transcripts, at least some of which do not introduce PTC or protein-disrupting deletion		
e.g., leaky aberrant splicing	[61]	
e.g., contradictory benign and pathogenic criteria		

Table 4. Cont.

2: likely not pathogenic	•	assay on mRNA from patients tissue samples	[59–61]
	AND	no associated mRNA aberration detected	
	AND	analysis including NMD inhibition	only [60]
	AND	other lines of evidence disproving variant pathogenicity	only [61]
	•	lab assays based on mRNA (e.g., minigenes)	[60,61]
	AND	variant-specific proficient function	
1: not pathogenic	AND	additional frequency/co-segregation/clinical data, additional molecular/mechanistic evidences from other sources, disproving variant pathogenicity (milder than for class 1)	[60,61]
	•	lab assays based on mRNA (e.g., minigenes)	
	AND	variant-specific proficient function	[60,61]
	AND	additional frequency/co-segregation/clinical data, additional molecular/mechanistic evidences from other sources, disproving variant pathogenicity (stronger than for class 2)	

<sup>1</sup> Richards et al. combines multiple individual pathogenic and benign criteria to reach the classification [61];

<sup>2</sup> Walker et al. recommends predicted/expected SRE aberrations only for research testing [59]. • denotes a single set of criteria for each class. A variant has to meet all the criteria in at least one set to be included into the respective class. The criteria listed in the table were extracted from several publications that were written, in fact, for somewhat differing purposes. Walker et al. published an evaluation and update of clinical classification guidelines previously designed by Spurdle et al. that were specifically aimed at potential splicing affection [59,62]. On the other hand, Thompson et al. and Richards et al. dealt with overall pathogenic potential of novel variants including direct effects on protein coding [60,61]. While Thompson et al. limited their criteria on mismatch repair genes [60], Richards et al. designed criteria for classification of all variants identified in genes that cause Mendelian disorders [61]. Finally, we have included one classification proposal derived from a publication that thoroughly evaluated the reliability of minigenes [48]. Of note, we did not include in the table the splicing-specific classification designed by Houdayer et al. [17], as this recommended a particular (specific) re-classification of class 3 variants (or variants of unknown significance) into three other classes (1S, 2S, 3S). Concerning the in silico predictions, Richards et al. propose using them as a supportive criterion, if all the in silico programs tested agree on the prediction [61]. However, they do not directly mention predictions of SREs in their publication, describing only the splice site predictors. In comparison, Walker et al. allows usage of SRE predictors in general. If a variant causes a loss or creation of the same SRE predicted by at least two of three used programs, then these guidelines propose that it should be experimentally tested (as a research testing) [59]. Other guidelines mentioned in this table do not propose the usage of SRE predictions. Please note that only an effect on splicing is taken into account in this table. The effect of a variant on protein coding per se should always be considered as well.

Regarding splicing alterations, variants leading to aberrant splicing of 100% transcripts, with clear effect on protein function (i.e., introducing PTC or in-frame deletion disrupting a functional domain) fall into class 5 (classified as pathogenic) [59,60]. According to some guidelines, aberrant splicing has to be shown on patient's mRNA samples [59], while others allow this classification to also be based on in vitro assays when the results are supported by additional clinical evidence (family co-segregation, number of tumors, etc.) [60,61].

According to Thompson et al. [60], variants shown to abrogate mRNA splicing using in vitro assays only (without patient's RNA-derived data) can fall into class 4 (likely pathogenic), if the additional clinical evidence does not allow class 5 ranking. On the other hand, after careful comparison of results obtained from patients' RNA and the minigenes, van der Klift et al. proposed that minigene assay alone should be sufficient for class 4 of clinical classification, if a splicing affecting variant demonstrates complete aberrant and frameshifting effect [48]. Such reclassification would have important consequences for patients, since class 4 and class 5 variants are recommended to have the same clinical management.

However, after analyzing data from several systemic splicing studies, we found that only a tiny minority of SRE-affecting mutations led to a complete splicing defect in minigenes (in particular, 5 of 136, i.e. 3.7%, variants increasing exon skipping showed 100% splicing aberration) [14,15,17,27,63].

From this point of view, the guidelines proposed by Richards et al. seem to be more pertinent for the SRE-affecting variants, since these always use a combination of several criteria to classify a variant and do not stick to 100% aberrant effect at the same time [61].

This pinpoints another difficulty that arises from the uncertainty in the minimal extent of aberrant splicing that can lead to a particular disease. In other words, the minimal amount of the residual full-length transcript sufficient to cover the biological function (in the relevant tissue) and prevent the disease development needs to be determined for individual conditions. Importantly, we cannot forget any possible adverse effects that the products of aberrantly spliced transcripts may exert on the wild type ones, e.g., dominant negativity in protein dimers [55,64].

At present, it follows that many variants fall into class 3—uncertain pathogenicity. Besides the mutations resulting in leaky aberrant splicing (variant allele produces full-length transcripts to some extent), this pertains, e.g., for exon-skipping mutations leading to in frame deletions not disrupting known functional domains or to upregulation of physiological alternative transcripts, at least some of which do not clearly disrupt protein function [59,60].

While the splice site related predictions are beginning to be included in classification guidelines, the SRE-related predictions stand aside [59,61]. In fact, despite the significant progress of these programs achieved in the past few years, they still await proper validation [4,15,17,48]. Currently, we can propose using the best performing tools (ESRseq, Hexplorer, and possibly EX-SKIP), ideally in combination with some a priori knowledge about the exons alternative splicing, as a rough preliminary filter to select the variants for further in vitro testing. Naturally, such a selection should be based on an overall estimation of each case (careful assessment of the patient's condition, variant co-segregation, etc.). Recognition of the splicing affecting variants is beneficial not only for the diagnostics, but it may eventually lead to the development of effective RNA-targeting therapy. This issue has been addressed by several recent reviews [65–68].

## 7. Conclusions

In the past two decades, many prediction tools have assisted researchers in defining potential SREs. On the other hand, only the approaches that either consider a nucleotide change in its larger sequence environment or combine several individual prediction methods have been successful in accurately recognizing splicing-affecting SRE changes. Repeatedly, the best outcomes have been shown for ESRseq scores, followed by Hexplorer and EX-SKIP tools. What still prevents these tools from being firmly included in diagnostic settings is their somewhat variable reliability for different exons, resulting defect ambiguity, and the limited number of evaluating studies.

In general, the future aim for SRE predictors is to make maximal provision for the context dependence of splicing regulation to improve both the recognition of splicing-affecting variants and the assessment of the resulting defects. To this point, combining the SRE predictions with an estimate of exon susceptibility to be skipped or to activate nearby cryptic splice sites may be helpful. Other approaches could include data gained from functional (high throughput) SRE mapping or from CLIP-detected binding sites of splicing regulators, accompanied with the knowledge of splicing effects upon their knock down.

**Acknowledgments:** This project was supported by Ministry of Health of the Czech Republic, grant No. 16-34414A. All rights reserved.

**Author Contributions:** Lucie Grodecká, Emanuele Buratti and Tomáš Freiberger substantially contributed to the work by writing and proofreading the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

CERES	Composite exonic regulatory element of splicing
circRNA	Circular RNA
CLIP	Cross-linking immunoprecipitation
ESE	Exonic splicing enhancer
ESS	Exonic splicing silencer
HSF	Human Splicing Finder
ISE	Intronic splicing enhancer
ISS	Intronic splicing silencer
PTC	Premature termination codon
RESCUE	Relative Enhancer and Silencer Classification by Unanimous Enrichment
SELEX	Functional systematic evolution of ligands by exponential enrichment
SNP	Single nucleotide polymorphism
SPANR	Splicing-based Analysis of Variants
SRE	Splicing regulatory element

## References

1. Sun, H.; Chasin, L.A. Multiple splicing defects in an intronic false exon. *Mol. Cell. Biol.* **2000**, *20*, 6414–6425. [[CrossRef](#)] [[PubMed](#)]
2. Fu, X.-D.; Ares, M. Context-dependent control of alternative splicing by RNA-binding proteins. *Nat. Rev. Genet.* **2014**, *15*, 689–701. [[CrossRef](#)] [[PubMed](#)]
3. Lee, Y.; Rio, D.C. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu. Rev. Biochem.* **2015**, *84*, 291–323. [[CrossRef](#)] [[PubMed](#)]
4. Baralle, D.; Buratti, E. RNA splicing in human disease and in the clinic. *Clin. Sci.* **2017**, *131*, 355–368. [[CrossRef](#)] [[PubMed](#)]
5. Goren, A.; Ram, O.; Amit, M.; Keren, H.; Lev-Maor, G.; Vig, I.; Pupko, T.; Ast, G. Comparative analysis identifies exonic splicing regulatory sequences—The complex definition of enhancers and silencers. *Mol. Cell* **2006**, *22*, 769–781. [[CrossRef](#)] [[PubMed](#)]
6. Pagani, F.; Stuani, C.; Tzetis, M.; Kanavakis, E.; Efthymiadou, A.; Doudounakis, S.; Casals, T.; Baralle, F.E. New type of disease causing mutations: The example of the composite exonic regulatory elements of splicing in CFTR exon 12. *Hum. Mol. Genet.* **2003**, *12*, 1111–1120. [[CrossRef](#)] [[PubMed](#)]
7. Greene, J.; Baird, A.-M.; Brady, L.; Lim, M.; Gray, S.G.; McDermott, R.; Finn, S.P. Circular RNAs: Biogenesis, Function and Role in Human Diseases. *Front. Mol. Biosci.* **2017**, *4*. [[CrossRef](#)] [[PubMed](#)]
8. Ashwal-Fluss, R.; Meyer, M.; Pamudurti, N.R.; Ivanov, A.; Bartok, O.; Hanan, M.; Evtantal, N.; Memczak, S.; Rajewsky, N.; Kadener, S. circRNA Biogenesis Competes with Pre-mRNA Splicing. *Mol. Cell* **2014**, *56*, 55–66. [[CrossRef](#)] [[PubMed](#)]
9. Kramer, M.C.; Liang, D.; Tatomer, D.C.; Gold, B.; March, Z.M.; Cherry, S.; Wilusz, J.E. Combinatorial control of *Drosophila* circular RNA expression by intronic repeats, hnRNPs, and SR proteins. *Genes Dev.* **2015**, *29*, 2168–2182. [[CrossRef](#)] [[PubMed](#)]
10. Conn, S.J.; Pillman, K.A.; Toubia, J.; Conn, V.M.; Salmandis, M.; Phillips, C.A.; Roslan, S.; Schreiber, A.W.; Gregory, P.A.; Goodall, G.J. The RNA Binding Protein Quaking Regulates Formation of circRNAs. *Cell* **2015**, *160*, 1125–1134. [[CrossRef](#)] [[PubMed](#)]
11. Khan, M.A.; Reckman, Y.J.; Aufiero, S.; van den Hoogenhof, M.M.; van der Made, I.; Beqqali, A.; Koolbergen, D.R.; Rasmussen, T.B.; Van Der Velden, J.; Creemers, E.E.; et al. RBM20 Regulates Circular RNA Production From the Titin Gene. *Circ. Res.* **2016**, *119*, 996–1003. [[CrossRef](#)] [[PubMed](#)]
12. Errichelli, L.; Dini Modigliani, S.; Laneve, P.; Colantoni, A.; Legnini, I.; Capauto, D.; Rosa, A.; De Santis, R.; Scarfò, R.; Peruzzi, G.; et al. FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons. *Nat. Commun.* **2017**, *8*, 14741. [[CrossRef](#)] [[PubMed](#)]
13. Di Giacomo, D.; Gaildrat, P.; Abuli, A.; Abdat, J.; Frébourg, T.; Tosi, M.; Martins, A. Functional Analysis of a Large set of BRCA2 exon 7 Variants Highlights the Predictive Value of Hexamer Scores in Detecting Alterations of Exonic Splicing Regulatory Elements: HUMAN MUTATION. *Hum. Mutat.* **2013**, *34*, 1547–1557. [[CrossRef](#)] [[PubMed](#)]

14. Julien, P.; Miñana, B.; Baeza-Centurion, P.; Valcárcel, J.; Lehner, B. The complete local genotype–phenotype landscape for the alternative splicing of a human exon. *Nat. Commun.* **2016**, *7*, 11558. [[CrossRef](#)] [[PubMed](#)]
15. Soukariéh, O.; Gaildrat, P.; Hamieh, M.; Drouet, A.; Baert-Desurmont, S.; Frébourg, T.; Tosi, M.; Martins, A. Exonic Splicing Mutations Are More Prevalent than Currently Estimated and Can Be Predicted by Using in Silico Tools. *PLoS Genet.* **2016**, *12*, e1005756. [[CrossRef](#)] [[PubMed](#)]
16. Houdayer, C.; Dehainault, C.; Mattler, C.; Michaux, D.; Caux-Moncoutier, V.; Pagès-Berhouet, S.; d’Enghien, C.D.; Laugé, A.; Castera, L.; Gauthier-Villars, M.; et al. Evaluation of in silico splice tools for decision-making in molecular diagnosis. *Hum. Mutat.* **2008**, *29*, 975–982. [[CrossRef](#)] [[PubMed](#)]
17. Houdayer, C.; Caux-Moncoutier, V.; Krieger, S.; Barrois, M.; Bonnet, F.; Bourdon, V.; Bronner, M.; Buisson, M.; Coulet, F.; Gaildrat, P.; et al. Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined in silico/in vitro studies on BRCA1 and BRCA2 variants. *Hum. Mutat.* **2012**, *33*, 1228–1238. [[CrossRef](#)] [[PubMed](#)]
18. Kergourlay, V.; Raï, G.; Blandin, G.; Salgado, D.; Bérout, C.; Lévy, N.; Krahn, M.; Bartoli, M. Identification of Splicing Defects Caused by Mutations in the Dysferlin Gene. *Hum. Mutat.* **2014**, *35*, 1532–1541. [[CrossRef](#)] [[PubMed](#)]
19. Grodecká, L.; Hujová, P.; Kramárek, M.; Kršjaková, T.; Kováčová, T.; Vondrášková, K.; Ravčuková, B.; Hrnčířová, K.; Souček, P.; Freiberger, T. Systematic analysis of splicing defects in selected primary immunodeficiencies-related genes. *Clin. Immunol.* **2017**, *180*, 33–44. [[CrossRef](#)] [[PubMed](#)]
20. Baralle, D.; Lucassen, A.; Buratti, E. Missed threads. The impact of pre-mRNA splicing defects on clinical practice. *EMBO Rep.* **2009**, *10*, 810–816. [[CrossRef](#)] [[PubMed](#)]
21. DePristo, M.A.; Banks, E.; Poplin, R.; Garimella, K.V.; Maguire, J.R.; Hartl, C.; Philippakis, A.A.; del Angel, G.; Rivas, M.A.; Hanna, M.; et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **2011**, *43*, 491–498. [[CrossRef](#)] [[PubMed](#)]
22. Buratti, E.; Baralle, M.; Baralle, F.E. From single splicing events to thousands: The ambiguous step forward in splicing research. *Brief Funct. Genom.* **2013**, *12*, 3–12. [[CrossRef](#)] [[PubMed](#)]
23. Liu, H.X.; Zhang, M.; Krainer, A.R. Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. *Genes Dev.* **1998**, *12*, 1998–2012. [[CrossRef](#)] [[PubMed](#)]
24. Cartegni, L. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res.* **2003**, *31*, 3568–3571. [[CrossRef](#)] [[PubMed](#)]
25. Wu, S.; Romfo, C.M.; Nilsen, T.W.; Green, M.R. Functional recognition of the 3′ splice site AG by the splicing factor U2AF35. *Nature* **1999**, *402*, 832–835. [[CrossRef](#)] [[PubMed](#)]
26. Desmet, F.-O.; Hamroun, D.; Lalande, M.; Collod-Beroud, G.; Claustres, M.; Beroud, C. Human Splicing Finder: An online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* **2009**, *37*, e67. [[CrossRef](#)] [[PubMed](#)]
27. Lastella, P.; Surdo, N.C.; Resta, N.; Guanti, G.; Stella, A. In silico and in vivo splicing analysis of MLH1 and MSH2 missense mutations shows exon- and tissue-specific effects. *BMC Genom.* **2006**, *7*, 243. [[CrossRef](#)] [[PubMed](#)]
28. ElSharawy, A.; Hundrieser, B.; Brosch, M.; Wittig, M.; Huse, K.; Platzer, M.; Becker, A.; Simon, M.; Rosenstiel, P.; Schreiber, S.; et al. Systematic evaluation of the effect of common SNPs on pre-mRNA splicing. *Hum. Mutat.* **2009**, *30*, 625–632. [[CrossRef](#)] [[PubMed](#)]
29. kConFab Investigators; Whiley, P.J.; Pettigrew, C.A.; Brewster, B.L.; Walker, L.C.; Spurdle, A.B.; Brown, M.A. Effect of BRCA2 sequence variants predicted to disrupt exonic splice enhancers on BRCA2 transcripts. *BMC Med. Genet.* **2010**, *11*, 80. [[CrossRef](#)] [[PubMed](#)]
30. Woolfe, A.; Mullikin, J.C.; Elnitski, L. Genomic features defining exonic variants that modulate splicing. *Genome Biol.* **2010**, *11*, R20. [[CrossRef](#)] [[PubMed](#)]
31. Gaildrat, P.; Krieger, S.; Di Giacomo, D.; Abdat, J.; Révillion, F.; Caputo, S.; Vaur, D.; Jamard, E.; Bohers, E.; Ledemeny, D.; et al. Multiple sequence variants of BRCA2 exon 7 alter splicing regulation. *J. Med. Genet.* **2012**, *49*, 609–617. [[CrossRef](#)] [[PubMed](#)]
32. Ke, S.; Shang, S.; Kalachikov, S.M.; Morozova, I.; Yu, L.; Russo, J.J.; Ju, J.; Chasin, L.A. Quantitative evaluation of all hexamers as exonic splicing elements. *Genome Res.* **2011**, *21*, 1360–1374. [[CrossRef](#)] [[PubMed](#)]
33. Wang, Z.; Rolish, M.E.; Yeo, G.; Tung, V.; Mawson, M.; Burge, C.B. Systematic identification and analysis of exonic splicing silencers. *Cell* **2004**, *119*, 831–845. [[CrossRef](#)] [[PubMed](#)]

34. Erkelenz, S.; Theiss, S.; Otte, M.; Widera, M.; Peter, J.O.; Schaal, H. Genomic HEXploring allows landscaping of novel potential splicing regulatory elements. *Nucleic Acids Res.* **2014**, *42*, 10681–10697. [[CrossRef](#)] [[PubMed](#)]
35. Zhang, X.H.-F. Computational definition of sequence motifs governing constitutive exon splicing. *Genes Dev.* **2004**, *18*, 1241–1250. [[CrossRef](#)] [[PubMed](#)]
36. Fairbrother, W.G.; Yeo, G.W.; Yeh, R.; Goldstein, P.; Mawson, M.; Sharp, P.A.; Burge, C.B. RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. *Nucleic Acids Res.* **2004**, *32*, W187–W190. [[CrossRef](#)] [[PubMed](#)]
37. Xiong, H.Y.; Alipanahi, B.; Lee, L.J.; Bretschneider, H.; Merico, D.; Yuen, R.K.C.; Hua, Y.; Gueroussov, S.; Najafabadi, H.S.; Hughes, T.R.; et al. The human splicing code reveals new insights into the genetic determinants of disease. *Science* **2015**, *347*, 1254806. [[CrossRef](#)] [[PubMed](#)]
38. Piva, F.; Giulietti, M.; Burini, A.B.; Principato, G. SpliceAid 2: A database of human splicing factors expression data and RNA target motifs. *Hum. Mutat.* **2012**, *33*, 81–85. [[CrossRef](#)] [[PubMed](#)]
39. Fairbrother, W.G. Predictive Identification of Exonic Splicing Enhancers in Human Genes. *Science* **2002**, *297*, 1007–1013. [[CrossRef](#)] [[PubMed](#)]
40. Castello, A.; Fischer, B.; Frese, C.; Horos, R.; Alleaume, A.-M.; Foehr, S.; Curk, T.; Krijgsveld, J.; Hentze, M. Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol. Cell* **2016**, *63*, 696–710. [[CrossRef](#)] [[PubMed](#)]
41. Raponi, M.; Kralovicova, J.; Copson, E.; Divina, P.; Eccles, D.; Johnson, P.; Baralle, D.; Vorechovsky, I. Prediction of single-nucleotide substitutions that result in exon skipping: Identification of a splicing silencer in BRCA1 exon 6. *Hum. Mutat.* **2011**, *32*, 436–444. [[CrossRef](#)] [[PubMed](#)]
42. Schwartz, S.; Hall, E.; Ast, G. SROOGLE: Webserver for integrative, user-friendly visualization of splicing signals. *Nucleic Acids Res.* **2009**, *37*, W189–W192. [[CrossRef](#)] [[PubMed](#)]
43. Zhang, C.; Li, W.-H.; Krainer, A.R.; Zhang, M.Q. RNA landscape of evolution for optimal exon and intron discrimination. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5797–5802. [[CrossRef](#)] [[PubMed](#)]
44. Stadler, M.B.; Shomron, N.; Yeo, G.W.; Schneider, A.; Xiao, X.; Burge, C.B. Inference of Splicing Regulatory Activities by Sequence Neighborhood Analysis. *PLoS Genet.* **2006**, *2*, e191. [[CrossRef](#)] [[PubMed](#)]
45. Aissat, A.; de Becdelièvre, A.; Golmard, L.; Vasseur, C.; Costa, C.; Chaoui, A.; Martin, N.; Costes, B.; Goossens, M.; Girodon, E.; et al. Combined Computational-Experimental Analyses of CFTR Exon Strength Uncover Predictability of Exon-Skipping Level. *Hum. Mutat.* **2013**, *34*, 873–881. [[CrossRef](#)] [[PubMed](#)]
46. Voelker, R.B.; Berglund, J.A. A comprehensive computational characterization of conserved mammalian intronic sequences reveals conserved motifs associated with constitutive and alternative splicing. *Genome Res.* **2007**, *17*, 1023–1033. [[CrossRef](#)] [[PubMed](#)]
47. Yeo, G.W.; Van Nostrand, E.L.; Nostrand, E.L.V.; Liang, T.Y. Discovery and analysis of evolutionarily conserved intronic splicing regulatory elements. *PLoS Genet.* **2007**, *3*, e85.
48. Van der Klift, H.M.; Jansen, A.M.L.; van der Steenstraten, N.; Bik, E.C.; Tops, C.M.J.; Devilee, P.; Wijnen, J.T. Splicing analysis for exonic and intronic mismatch repair gene variants associated with Lynch syndrome confirms high concordance between minigene assays and patient RNA analyses. *Mol. Genet. Genom. Med.* **2015**, *3*, 327–345. [[CrossRef](#)] [[PubMed](#)]
49. Královicová, J.; Vorechovsky, I. Global control of aberrant splice-site activation by auxiliary splicing sequences: Evidence for a gradient in exon and intron definition. *Nucleic Acids Res.* **2007**, *35*, 6399–6413. [[CrossRef](#)] [[PubMed](#)]
50. Flanigan, K.M.; Dunn, D.M.; von Niederhausern, A.; Soltanzadeh, P.; Howard, M.T.; Sampson, J.B.; Swoboda, K.J.; Bromberg, M.B.; Mendell, J.R.; Taylor, L.E.; et al. Nonsense mutation-associated Becker muscular dystrophy: Interplay between exon definition and splicing regulatory elements within the DMD gene. *Hum. Mutat.* **2011**, *32*, 299–308. [[CrossRef](#)] [[PubMed](#)]
51. De Conti, L.; Baralle, M.; Buratti, E. Exon and intron definition in pre-mRNA splicing. *Wiley Interdiscip. Rev. RNA* **2013**, *4*, 49–60. [[PubMed](#)]
52. Gutierrez-Arcelus, M.; Ongen, H.; Lappalainen, T.; Montgomery, S.B.; Buil, A.; Yurovsky, A.; Bryois, J.; Padioleau, I.; Romano, L.; Planchon, A.; et al. Tissue-Specific Effects of Genetic and Epigenetic Variation on Gene Regulation and Splicing. *PLoS Genet.* **2015**, *11*, e1004958.

53. Baralle, M.; Skoko, N.; Knezevich, A.; De Conti, L.; Motti, D.; Bhuvanagiri, M.; Baralle, D.; Buratti, E.; Baralle, F.E. *NF1* mRNA biogenesis: Effect of the genomic milieu in splicing regulation of the *NF1* exon 37 region. *FEBS Lett.* **2006**, *580*, 4449–4456. [[CrossRef](#)] [[PubMed](#)]
54. Mine, M.; Brivet, M.; Touati, G.; Grabowski, P.; Abitbol, M.; Marsac, C. Splicing Error in E1 $\alpha$  Pyruvate Dehydrogenase mRNA Caused by Novel Intronic Mutation Responsible for Lactic Acidosis and Mental Retardation. *J. Biol. Chem.* **2003**, *278*, 11768–11772. [[PubMed](#)]
55. Moseley, C.T.; Mullis, P.E.; Prince, M.A.; Phillips, J.A. An Exon Splice Enhancer Mutation Causes Autosomal Dominant GH Deficiency. *J. Clin. Endocrinol. Metab.* **2002**, *87*, 847–852. [[CrossRef](#)] [[PubMed](#)]
56. Dhir, A.; Buratti, E. Alternative splicing: Role of pseudoexons in human disease and potential therapeutic strategies: Pseudoexons in human disease. *FEBS J.* **2010**, *277*, 841–855. [[CrossRef](#)] [[PubMed](#)]
57. Divina, P.; Kvitkovicova, A.; Buratti, E.; Vorechovsky, I. Ab initio prediction of mutation-induced cryptic splice-site activation and exon skipping. *Eur. J. Hum. Genet.* **2009**, *17*, 759–765. [[PubMed](#)]
58. Sharma, N.; Sosnay, P.R.; Ramalho, A.S.; Douville, C.; Franca, A.; Gottschalk, L.B.; Park, J.; Lee, M.; Vecchio-Pagan, B.; Raraigh, K.S.; et al. Experimental Assessment of Splicing Variants Using Expression Minigenes and Comparison with In Silico Predictions. *Hum. Mutat.* **2014**, *35*, 1249–1259. [[PubMed](#)]
59. Walker, L.C.; Whiley, P.J.; Houdayer, C.; Hansen, T.V.O.; Vega, A.; Santamarina, M.; Blanco, A.; Fachal, L.; Southey, M.C.; Lafferty, A.; et al. Evaluation of a 5-Tier Scheme Proposed for Classification of Sequence Variants Using Bioinformatic and Splicing Assay Data: Inter-Reviewer Variability and Promotion of Minimum Reporting Guidelines. *Hum. Mutat.* **2013**, *34*, 1424–1431. [[PubMed](#)]
60. Thompson, B.A.; Spurdle, A.B.; Plazzer, J.-P.; Greenblatt, M.S.; Akagi, K.; Al-Mulla, F.; Bapat, B.; Bernstein, I.; Capellá, G.; den Dunnen, J.T.; et al. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. *Nat. Genet.* **2014**, *46*, 107–115. [[PubMed](#)]
61. Richards, S.; Aziz, N.; Bale, S.; Bick, D.; Das, S.; Gastier-Foster, J.; Grody, W.W.; Hegde, M.; Lyon, E.; Spector, E.; et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **2015**, *17*, 405–423. [[CrossRef](#)] [[PubMed](#)]
62. Spurdle, A.B.; Couch, F.J.; Hogervorst, F.B.L.; Radice, P.; Sinilnikova, O.M. Prediction and assessment of splicing alterations: Implications for clinical testing. *Hum. Mutat.* **2008**, *29*, 1304–1313. [[CrossRef](#)] [[PubMed](#)]
63. Fraile-Bethencourt, E.; Díez-Gómez, B.; Velásquez-Zapata, V.; Acedo, A.; Sanz, D.J.; Velasco, E.A. Functional classification of DNA variants by hybrid minigenes: Identification of 30 spliceogenic variants of *BRCA2* exons 17 and 18. *PLoS Genet.* **2017**, *13*, e1006691. [[CrossRef](#)] [[PubMed](#)]
64. Petkovic, V.; Godi, M.; Lochmatter, D.; Eblé, A.; Flück, C.E.; Robinson, I.C.; Mullis, P.E. Growth Hormone (GH)-Releasing Hormone Increases the Expression of the Dominant-Negative GH Isoform in Cases of Isolated GH Deficiency due to GH Splice-Site Mutations. *Endocrinology* **2010**, *151*, 2650–2658. [[CrossRef](#)] [[PubMed](#)]
65. Suñé-Pou, M.; Prieto-Sánchez, S.; Boyero-Corral, S.; Moreno-Castro, C.; El Yousfi, Y.; Suñé-Negre, J.; Hernández-Munain, C.; Suñé, C. Targeting Splicing in the Treatment of Human Disease. *Genes* **2017**, *8*, 87. [[CrossRef](#)] [[PubMed](#)]
66. Sardone, V.; Zhou, H.; Muntoni, F.; Ferlini, A.; Falzarano, M. Antisense Oligonucleotide-Based Therapy for Neuromuscular Disease. *Molecules* **2017**, *22*, 563. [[CrossRef](#)] [[PubMed](#)]
67. Tian, N.; Li, J.; Shi, J.; Sui, G. From General Aberrant Alternative Splicing in Cancers and Its Therapeutic Application to the Discovery of an Oncogenic DMTF1 Isoform. *Int. J. Mol. Sci.* **2017**, *18*, 191. [[CrossRef](#)] [[PubMed](#)]
68. Havens, M.A.; Hastings, M.L. Splice-switching antisense oligonucleotides as therapeutic drugs. *Nucleic Acids Res.* **2016**, *44*, 6549–6563. [[CrossRef](#)] [[PubMed](#)]

