Chapter 12

The Coupling of Alternative Splicing and Nonsense-Mediated mRNA Decay

Liana F. Lareau, Angela N. Brooks, David A.W. Soergel, Qi Meng and Steven E. Brenner*

Abstract

Most human genes exhibit alternative splicing, but not all alternatively spliced transcripts produce functional proteins. Computational and experimental results indicate that a substantial fraction of alternative splicing events in humans result in mRNA isoforms that harbor a premature termination codon (PTC). These transcripts are predicted to be degraded by the nonsense-mediated mRNA decay (NMD) pathway. One explanation for the abundance of PTC-containing isoforms is that they represent splicing errors that are identified and degraded by the NMD pathway. Another potential explanation for this startling observation is that cells may link alternative splicing and NMD to regulate the abundance of mRNA transcripts. This mechanism, which we call "Regulated Unproductive Splicing and Translation" (RUST), has been experimentally shown to regulate expression of a wide variety of genes in many organisms from yeast to human. It is frequently employed for autoregulation of proteins that affect the splicing process itself. Thus, alternative splicing and NMD act together to play an important role in regulating gene expression.

Introduction

One major result of the large-scale sequencing projects of the last decade has been an appreciation of the extent of alternative splicing of mammalian genes. Estimates vary, but most reports agree that over half of human genes are alternatively spliced.[1,2] What is the biological function of this extensive alternative splicing? Many propose that it is a major mechanism underlying proteome expansion,[3] but alternative splicing can also modulate the function or activity of a gene, for instance by adding or removing exons encoding protein domains or by altering the stability of the transcript or resulting protein.[4-6]

In the last few years, it has become clear that many alternative splice forms previously thought to encode truncated proteins are actually targets of NMD (Fig. 1). In mammals, a termination codon located more than about 50 nucleotides upstream of the final exon junction is generally recognized as premature and elicits NMD.[7-9] Understanding of this rule allowed for the identification of numerous transcripts that are predicted to be degraded rather than translated into protein. The prevalence of these predicted NMD-targeted transcripts calls for a reconsideration of the roles of alternative splicing and NMD. Since the mechanism of recognition by the NMD pathway is best understood in mammals and there are relatively few predicted or verified targets of NMD in other organisms, we will focus this review primarily on mammalian targets of NMD.

*Corresponding Author: Steven E. Brenner—Department of Molecular and Cell Biology, University of California, Berkeley. Email: brenner@compbio.berkeley.edu

Alternative Splicing in the Postgenomic Era, edited by Benjamin J. Blencowe and Brenton R. Graveley.

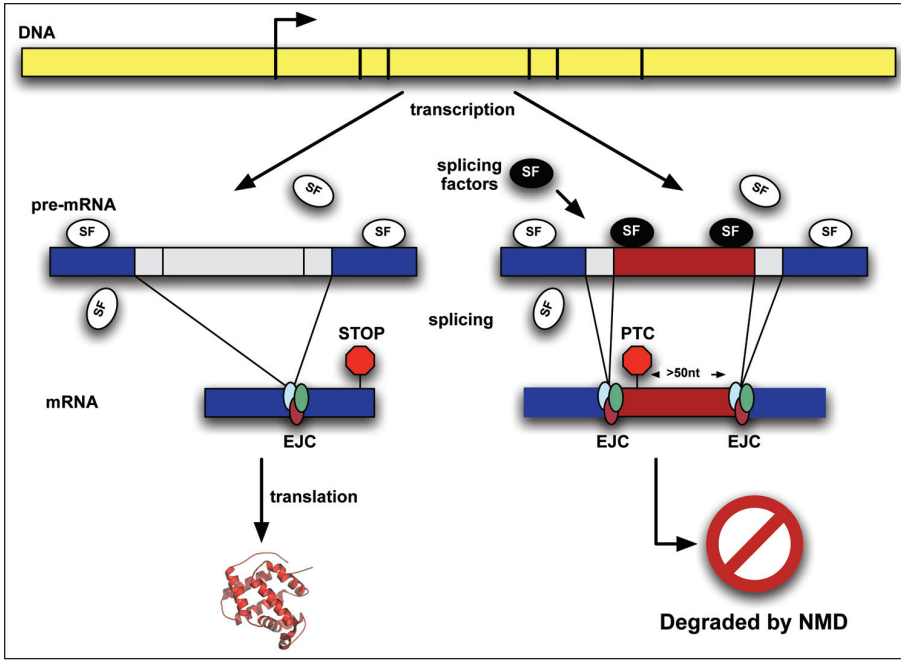

Figure 1. Some alternatively spliced transcripts are degraded by Nonsense Mediated mRNA Decay (NMD). The spliceosome deposits an Exon Junction Complex (EJC) on the mRNA ~20-24nt upstream of the splice junction, thereby marking the former location of the excised intron.[9] On the first, pioneering round of translation,[25] any in-frame stop codon found more than 50 nt upstream of the splice junction triggers NMD; such a codon is called a PTC.[8,9] Alternative splicing can lead to the inclusion of a PTC on an alternatively spliced region, or may give rise to a downstream PTC due to a frameshift. Thus, alternative splicing can give rise to unproductive transcripts. Splicing factors (labelled "SF") can alter the ratio of productive transcripts to transcripts that contain a PTC, targeting them for degradation. In this example, the dark splicing factor shown induces the inclusion of an alternative exon with a PTC, thereby decreasing the abundance of the productive isoform and downregulating protein expression. Components of the splicing machinery such as U2AF35 and PTB can similarly regulate isoform proportions.

Transcripts containing a PTC can arise through various patterns of alternative splicing (Fig. 2). For example, an exon inclusion event can introduce an in-frame PTC, thus targeting the transcript for NMD. Most alternative splicing events that induce a frameshift are predicted to give rise to a downstream PTC. Alternative splicing in noncoding regions can also give rise to NMD targets. For example, the splicing of a 3' UTR intron can create an exon-exon junction more than 50 nucleotides downstream of the original stop codon, which will consequently be recognized as premature.[10]

NMD was originally considered to be a quality control mechanism, protecting the cell from the potentially toxic effects of nonsense codons introduced by errors in replication, transcription, or splicing.[11,12] We now know that there are many targets of NMD,[13,14] including transcripts with uORFs, products of alternative splicing, byproducts of V(D)J recombination and transcripts arising from transposons and retroviruses.[15] Indeed, it now seems that a major effect of NMD is to downregulate specific transcripts, in addition to clearing the cell of aberrant transcripts.

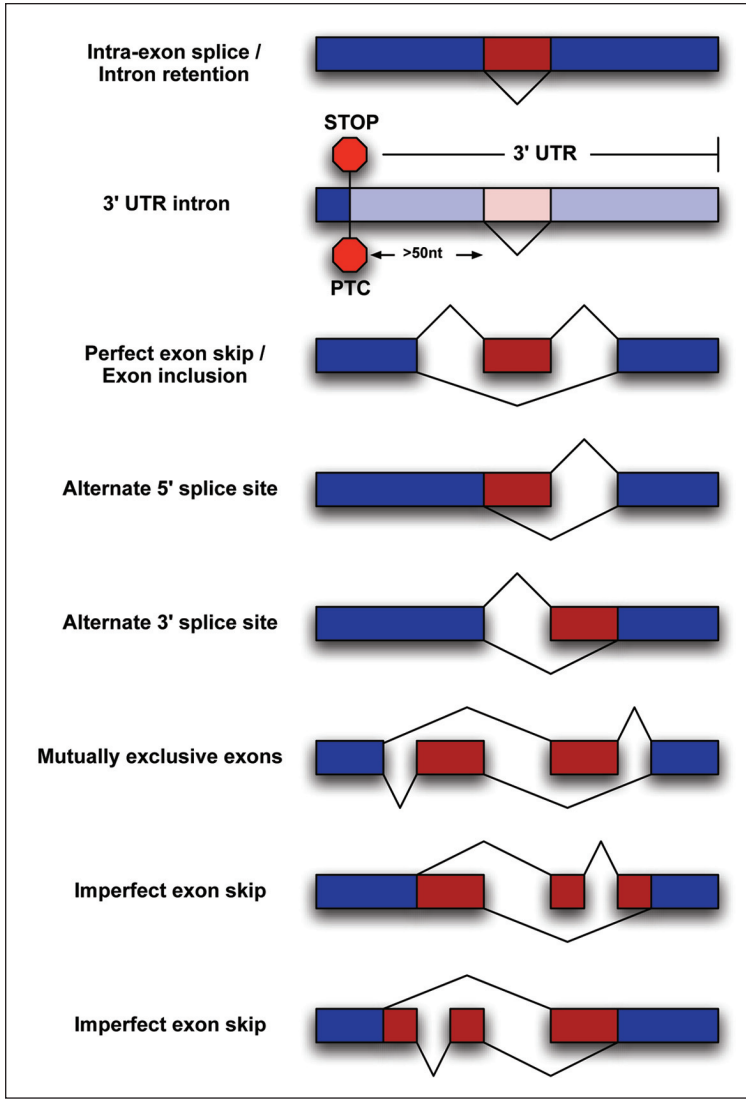

Figure 2. Patterns of alternative splicing. Alternative selection of 5′ and 3′ splice sites can lead to various patterns of included exons. Any exon that is included in an alternative form may harbor a PTC. Also, whenever an exon whose length is not a multiple of 3 is included or removed, the concomitant frame shift may result in a downstream PTC. Finally, splicing out an intron in the 3′ untranslated region (UTR) can cause the normal stop codon to be considered premature.

Many Alternative Splice Forms Are Targets of NMD

While it was long known that alternative splicing may produce isoforms that are degraded by NMD, this was not appreciated as a pervasive phenomenon until genome-wide studies indicated that a substantial fraction of human genes are spliced to produce isoforms that may be targeted for NMD.

The first study to predict widespread NMD of alternative splice forms used human mRNA and EST sequences from public databases to infer alternative splice forms and identify PTCs.[16] Lewis et al considered 16,780 human mRNA sequences from the reviewed category of RefSeq, a set of well-characterized, experimentally confirmed transcript sequences.[17] Alignment of the RefSeq mRNAs to their genomic loci showed that 617 of these curated mRNA sequences, or 3.7%, contained PTCs. However, the alternative splice forms inferred by aligning EST sequences from dbEST[18] to the RefSeq-defined genomic loci substantially increased the estimated fraction of genes with PTC^+ isoforms (Fig. 3). Based on the EST data, over 3000 of the RefSeq genes had alternative splice forms and of these alternatively spliced genes, 45% were predicted to encode at least one splice isoform that is a target of NMD.[16] Therefore, the study found that at least 12% of human genes have a PTC^+ isoform.

These results have been confirmed and strengthened by more recent studies. An analysis of the isoforms contained in SWISS-PROT[19] showed that even this reliable, curated database contained presumed translation products of mRNA sequences that are likely to be degraded by NMD. Alignment of the mRNA sequence of each protein isoform reported in SWISS-PROT to the human genome identified reliable exon-intron structures for 2483 isoforms from 1363 genes. The 50-nucleotide rule predicted that 144 isoforms (5.8% of 2483) from 107 genes (7.9% of 1363) contain a PTC and are likely targets of NMD.[20]

An elegant study by Baek and Green extended the analysis of PTC^+ alternative splicing to consider conservation of splice forms between human and mouse.[21] This approach helps distinguish aberrant splicing events from rare but functional variants. Starting from a large set of cDNA and EST sequences, Baek and Green identified about 1500 pairs of exon inclusion/exclusion splice forms found in both human and mouse. A quarter of the conserved alternative forms contain a conserved PTC,[21] which is consistent with subsequent findings,[23] suggesting that these isoforms play a functional role and that the PTC is important to their function.

Several microarray experiments have provided direct evidence to support these computational results.[15,22] In one example of these experiments, Mendell and coworkers depleted HeLa cells of Upf1, an essential component of the NMD pathway, and used microarrays to compare mRNA abundances in these cells to mRNA abundances in mock-treated cells.[15] They found that 4.9% of the ~4000 transcripts tested showed significantly higher abundances in cells deficient in NMD, suggesting that NMD normally downregulates those transcripts. Evidence that their observations were largely due to the direct action of NMD, rather than being a downstream regulatory consequence, was provided by showing that several of the putative NMD-targeted transcripts they identified decayed faster in normal cells than in cells depleted of Upf1. They also provided evidence that the effect they observed was due to NMD by showing that the PTC^+ transcript abundances responded similarly to depletion of Upf2, another protein that is essential for NMD. Finally, Mendell et al also observed that 4.3% of the transcripts decreased in abundance in NMD-deficient cells. The stability of those transcripts was not altered by NMD deficiency, showing that the change in their abundance was an indirect effect. Because this microarray experiment detected changes in total transcript levels across all isoforms of a particular gene, it may not have detected changes in transcript levels of a specific PTC^+ isoform. Therefore, many true NMD targets would not have been identified.

To specifically detect changes in specific isoform abundances due to NMD inhibition, Pan et al used an alternative splicing microarray platform.[23] By distinguishing relative levels of PTC^+ versus PTC^- isoforms, they found that approximately 10% of the PTC-containing isoforms increased in abundance by at least 15 percentage points upon inhibition of NMD. Although Pan et al were able to detect relatively few targets of NMD, they reported that a majority of the PTC^+ isoforms are present at relatively low abundances, even when NMD is inhibited. They concluded that many of these may represent nonfunctional transcripts or transcripts that are not under strong selection pressure. This conclusion is consistent with the observation that the majority of PTC-containing splice variants identified in sequenced transcripts are not conserved between human and mouse.[21,23]

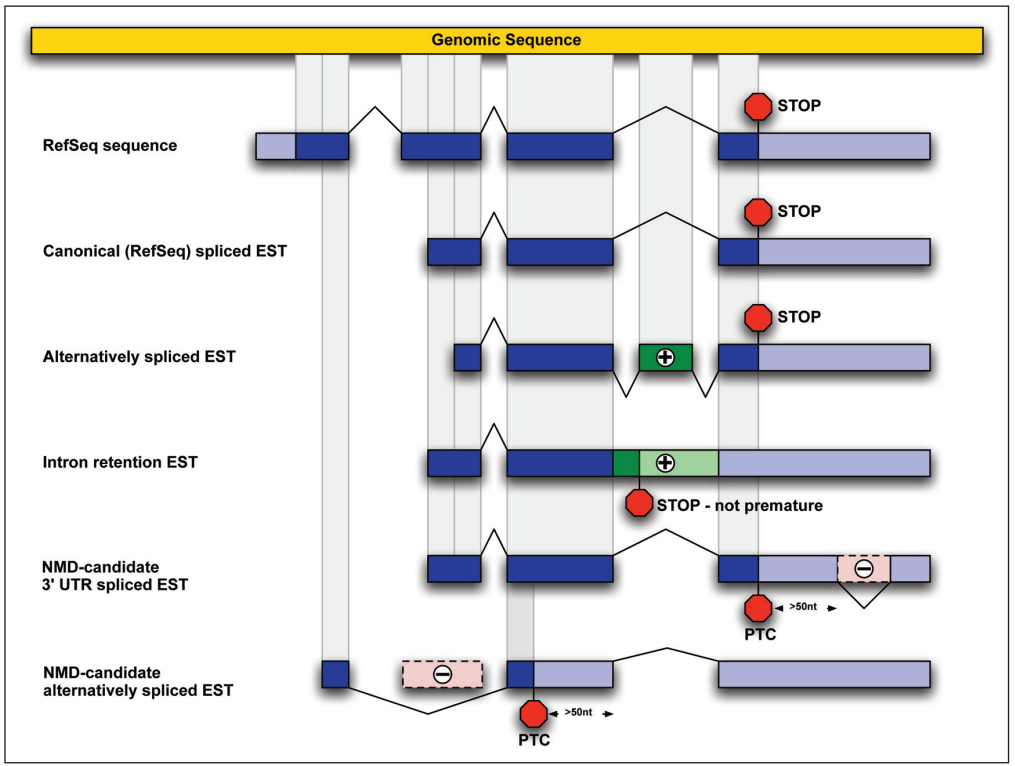

Figure 3. Inference of alternative splice forms and PTCs from RefSeq and EST data. Lewis et al aligned coding regions of RefSeq mRNAs to the genomic sequence to determine canonical splicing patterns.[16] EST alignments to the genomic sequence confirmed the canonical splices and indicated alternative splices. Canonical (RefSeq) splices are indicated above the exons, whereas alternative splices are indicated below the exons. When an alternative splice introduced a stop codon >50 nucleotides upstream of the final exon-exon junction of an inferred mRNA isoform, the stop codon was classified as a PTC and the corresponding mRNA isoform was labeled an NMD candidate.

There is limited information about the prevalence of alternative splicing coupled to NMD in *Drosophila*. Rehwinkel et al used a gene expression array and found that 3.4% of genes had a significant increase in overall transcript abundances when NMD was inhibited; intriguingly, the NMD protein SMG5 was among these.[24] More recently, an alternative splicing array platform capable of distinguishing distinct isoforms found an order of magnitude more isoforms that are targets of NMD in fly (manuscript in preparation). However, even this finding represents a modest level of coupling relative to the amount predicted based on the analysis of EST and cDNA transcripts from human and mouse tissues.

The striking number of predicted PTC$^+$ alternative splice forms demands more detailed explanation. Are some of these isoforms translated at levels sufficient to impact physiology? Are they an unavoidable side effect of productive alternative splicing? How many of the observed PTC$^+$ isoforms are due to transcriptional or splicing noise? To what extent do PTC$^+$ isoforms represent the coupling of alternative splicing and NMD in order to regulate gene expression? We shall consider each of these potential explanations in turn.

Do the Observed PTC$^+$ mRNA Isoforms Evade NMD to Produce Functional Protein?

The existence of numerous PTC$^+$ isoforms was first inferred from EST data.[16] One may wonder why EST evidence exists at all for isoforms that are expected to be degraded by NMD. As observed in numerous examples (Table 1), NMD substantially reduces the abundance of PTC$^+$ transcripts, but does not eliminate them entirely. One explanation for the presence of these ESTs is that NMD surveillance may not be completely effective. Furthermore, PTC$^+$ isoforms are not degraded instantly upon being spliced; rather, their degradation occurs only after a pioneer round of translation,[25] which might occur near the nuclear pore during or soon after export of the message from the nucleus (reviewed in ref. 26). Thus, we expect there to be some steady-state abundance of PTC$^+$ isoforms that have not yet been degraded, especially inside the nucleus. A series of elegant experiments and computational modeling in yeast suggest that the dominant reason for the presence of PTC$^+$ mRNAs in the cell is the temporal lag between splicing and degradation, rather than incomplete surveillance.[27] Evidently, the resulting abundance of PTC$^+$ isoforms is in many cases high enough for ESTs derived from those isoforms to be observed and deposited in dbEST. Indeed, many of the alternative splice junctions that generate a PTC are supported by multiple ESTs.

Nevertheless, less stable isoforms will be underrepresented in EST libraries. Using sequence features such as splice site strength, Baek and Green modeled the predicted inclusion rates of alternative exons.[21] They showed that PTC$^+$ isoforms are probably produced at a higher rate than they are observed in EST data and that many are degraded before they can be sequenced. Thus, the EST data underestimate the fraction of a given gene's mRNA that is PTC$^+$ and also underestimate the number of genes with PTC$^+$ alternative splicing. For this reason and also because the quality filters used in the above studies excluded many genes and isoforms, EST-based reports offer a lower bound on the number of PTC$^+$ isoforms; the true prevalence of alternative splicing and of PTC$^+$ isoforms may be substantially higher.

Some PTC$^+$ transcripts may evade NMD, increasing their likelihood of being observed and deposited in sequence databases. This evasion can happen in two ways—by the incomplete action of NMD to degrade the PTC$^+$ transcript, or by a specific mechanism that allows the transcript to evade NMD to ensure protein production. There are a few known examples in which a transcript which should be degraded according to the 50-nucleotide rule is in fact stable and is translated to produce protein. These include polycistronic transcripts on which translation is reinitiated downstream of a PTC[28-30]; apolipoprotein B, which is protected from NMD by an RNA editing complex[31]; some transcripts with a PTC near the initiation codon[32]; cytokine thrombopoietin (TPO) mRNA with several uORFs[33]; and an aberrant beta-globin transcript which is protected from NMD by an unknown mechanism.[34] Although NMD does not prevent protein production entirely in such cases, it may nonetheless limit expression from PTC$^+$ transcripts substantially, as was shown for an alternative transcript of FAH[35] and for ARD-1.[28] Nonetheless, documented

Table 1. *Experimentally confirmed examples of unproductive splicing in natural wild-type isoforms*

Name	Organism	AS → PTC	Regulated AS	PTC → Low Abundance	NMD	Notes	References
a) Unproductive splicing							
Calpain-10	Human	•		•	•		20, 45, 46
HPRT	Human	•		•	•	Unproductive transcripts are likely noise	84
POLB	Human	•		•	•	It is not clear whether AS is regulated or if splice variants are noise	84
TCR-beta	Human	•		•	•	V(D)J cleanup. NMD strength boosted by sequence elements	85, 86
ABCC4	Human, Monkey, Mouse	•		•	•	High conservation of PTC-producing exons from mouse to human suggests that they are under ESE control, or that translation is reinitiated downstream of the PTC	78
FGFR2	Rat	•		•	•	Side effect. Productive forms are tissue-specific	37, 87
cbp-2ps	Worm	•		•	•	Pseudogene of *cbp-1*, a gene that encodes a cyclic AMP response element binding protein (CREB) binding protein	88
F45D11.2, F45D11.3, F45D11.4	Worm	•		•	•	Pseudogenes from triplication of gene/pseudogene pairs from the F45D11 family of genes	88
rpl-7aps	Worm	•		•	•	Pseudogene of *rpl-7a*, which encodes ribosomal protein L7a	88
ubq-3ps, ubq-4ps	Worm	•		•	•	Pseudogenes of *ubq-1*	88
b) Regulated unproductive splicing							
FAH	Human	•	•	•	•	PTC^+ isoform is an NMD target and produces a short protein	35

continued on next page

Table 1. Continued

Name	Organism	AS → PTC	Regulated AS	PTC → Low Abundance	NMD	Notes	References
U2AF35	Human	•	•	•	•	Mutually exclusive exons; PTC$^+$ isoforms are an apparent side effect	89-93
MID-1	Human, Mouse, Fugu	•	•	•	•	Tissue specific alternative splicing	52
ClC-1	Human, Mouse	•	•	•	•	Misregulation of PTC-containing isoform leads to myotonic dystrophy	57, 58
SSAT	Mouse	•	•	•	•	Relative abundance of PTC$^+$ isoform is regulated by substrates that SSAT acts upon (i.e., depletion of polyamine substrates causes an increase in relative levels of the NMD targeted isoform)	53
Nicastrin	Rat	•	•	•	•	PTC$^+$ isoform expressed in neurons	94
MER2	Yeast	•	•	•	•	Splicing is regulated by MER1, which is produced only in meiotic cells. As a result, MER2 transcripts are productively spliced only during meiosis. In mitotic cells, a PTC$^+$ form is produced and degraded	63, 95
ITSN-1	Human, Mouse	•	•	•		Productive and unproductive isoforms are tissue-specific	96
NDUFS4	Human	•		•	•		97
LARD	Human	•	•			PTC$^+$ isoforms are abundant	98
ARD-1, NIPP-1	Human	•		•		Translation reinitiation. ARD1 is downregulated by NMD, but is nonetheless expressed. Also, ARD1 and NIPP1 may influence splicing via PP1	28

continued on next page

Table 1. Continued

Name	Organism	AS → PTC	Regulated AS	PTC → Low Abundance	NMD	Notes	References
c) Autoregulatory unproductive splicing not affecting splicing factors							
RPL3	Human	•	•	•	•		99
RPL12	Worm, Human	•	•	•	•		64, 99
RPL30	Yeast	•	•	•	•		62
AtGRP7	Arabidopsis	•	•	•			73
d) Autoregulatory unproductive splicing affecting splicing factors							
PTB	Human	•	•	•	•		67, 69, 70
Sc35	Human	•	•	•	•		10
TIAR / TIA-1	Human	•	•	•	•		100
CLKs	Human, Mouse, Ciona	•	•	•	•		20, 71, 72, 101, 102
AUF1	Human	•	•	•			103, 104
nPTB	Human	•	•	•		Splicing is regulated by PTB. Direct evidence of additional autoregulation has not been shown	69
SRp20, SRp30b	Worm	•		•	•		105
TRA2-beta	Human	•	•			PTC^+ forms are abundant, but are not translated, perhaps due to sequestration	106

Dots indicate direct experimental confirmation; lack of a dot means only that the experiment has not been performed to our knowledge. For instance, while it seems certain that ITSN-1 PTC^+ isoforms are degraded by NMD, this has not been directly observed using Upf1 knockdown or another NMD assay. The few cases where an experiment was performed but yielded a negative result are noted. AS, alternative splicing

exceptions to the 50-nt rule are rare and there are many more known cases in which the 50-nt rule is obeyed.

Even for PTC$^+$ transcripts that do not evade NMD, the possibility remains that the single truncated protein product of the pioneer round of translation is functionally significant, since some regulatory proteins can have an effect even at a very low copy number.[36] Also, to the extent that NMD is not completely effective at detecting and degrading typical PTC$^+$ transcripts, the transcripts that escape this process may be translated to produce truncated proteins. However, these proteins will frequently lack critical domains, rendering them inactive or even harmful. In any case, it is hard to imagine that functional roles of truncated proteins could explain the high prevalence of genes with PTC$^+$ isoforms, especially given the wide functional diversity of these genes, and no data exist to support such a view.

While there may be exceptions, it seems unlikely that many PTC$^+$ isoforms produce functional protein, either during the pioneer round of translation, due to incomplete surveillance, or by evading NMD altogether.

Are the Observed PTC$^+$ mRNA Isoforms a Side Effect of Productive Alternative Splicing?

In the particular situation of mutually-exclusive exon usage, NMD may be a mechanism for removing transcripts that erroneously include both exons or neither exon. For isoforms of *FGFR2*, including both exons or neither exon introduces a frameshift and PTC into the mRNA, targeting it for degradation.[37] In this circumstance, an isoform including exon IIIb while skipping exon IIIc is productive; similarly, the isoform including exon IIIc but excluding exon IIIb is productive. However, the spliceosome may also pair the same splice sites differently such that both exons are included, or such that neither are included. Both of these latter possibilities introduce a PTC (Fig. 4).

Each splice site involved in the removal of exons IIIb and IIIc is required for the production of at least one productive isoform; the unproductive isoforms arise simply from alternate pairings of these otherwise productive splice sites. Given that the spliceosome is prone to such alternate pairings, there may be evolutionary pressure to ensure that the undesired isoforms include a PTC. This results in an inevitable side effect of the mechanism for productive alternative splicing. NMD can be used as a filter to remove these "side effect" isoforms, which may comprise a substantial fraction of the transcripts produced (up to 50% in the case of *FGFR2*).[37]

We examined the alternative isoforms inferred from human dbEST data (see above) and found that PTC$^+$ isoforms could be explained as a side effect for 34% of the genes that produce them. That is, 66% of the genes with a PTC$^+$ isoform have a splice site that is specific to PTC$^+$ isoforms and that is responsible for introducing the PTC (Soergel 2005, unpublished data). If these unproductive isoforms were on the whole detrimental to the cell, then we would expect evolutionary pressure to have selected against PTC$^+$ specific splice sites, but in fact many of them are strikingly conserved, as we discuss below. Thus, while the contribution of "side effect" isoforms may be significant, they alone cannot explain the high prevalence of PTC$^+$ isoforms.

Do the Observed PTC$^+$ mRNA Isoforms Represent Missplicing or Cellular Noise?

NMD was originally described as a means of clearing erroneous transcripts from the cell.[11,12] In keeping with this role, some alternative splice forms that are degraded by NMD could represent splicing errors. Such errors could arise from mutations disrupting splice sites or regulatory sequences, including mutations in intronic regions that are invisible after intron removal. Also, the splicing machinery itself could recognize incorrect splice sites. The spliceosome distinguishes true splice sites from nearby cryptic sites with impressive fidelity, but splice site recognition is a complex process and errors occur at some rate. Although there are at present no clear data on the extent of missplicing, EST libraries contain millions of transcript sequences and even extremely rare events, such as those arising from missplicing, may be represented.

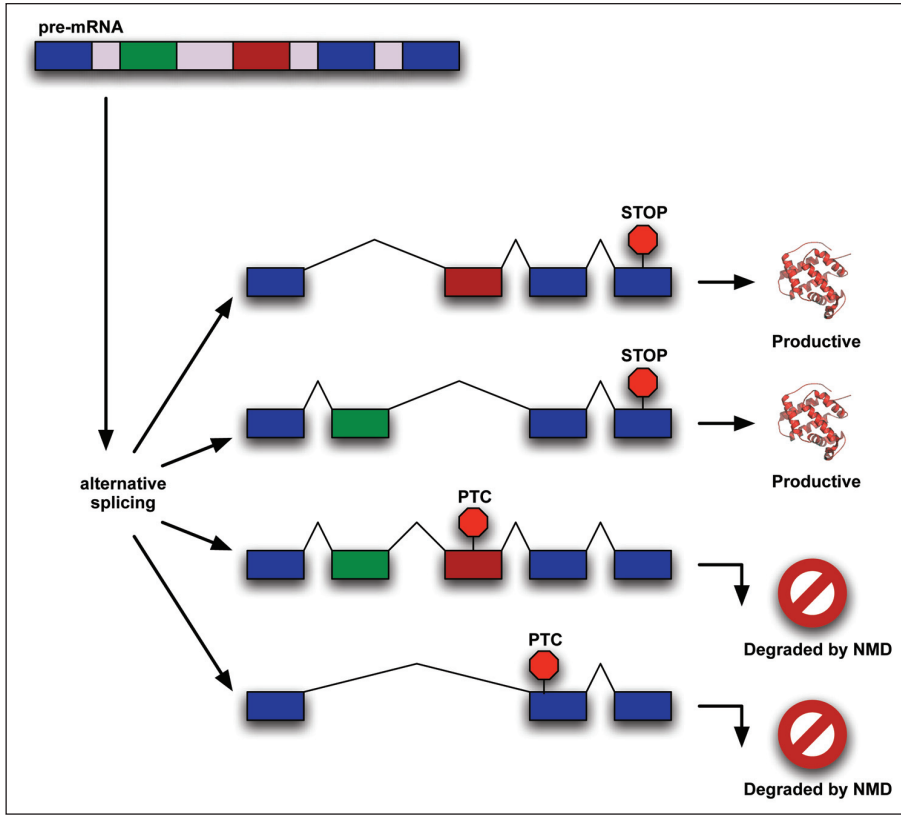

Figure 4. NMD can be employed to remove "side effect" isoforms in the case of mutually exclusive exons. Alternative splicing may generate two productive isoforms including one or the other of a pair of mutually exclusive exons. By choosing different pairings from the same set of 5′ and 3′ splice sites, the spliceosome may also generate isoforms including both exons, or neither exon. Frameshifts can give rise to PTCs on these undesired isoforms so that they will be degraded by NMD.

In EST-based computational analyses, splicing errors can be identified to some extent by filtering out splicing events that are seen only in a few ESTs, but this method cannot distinguish errors from legitimate rare splice forms. With multiple mammalian genomes available, recent work has focused on evolutionary conservation to suggest negative selection and, perhaps, functional roles for conserved alternative forms,[21,38] (also reviewed in refs. 39,40). Minor isoforms, those that occur only a fraction of the time, are less often conserved than major isoforms[41] and may sometimes represent recent mutations or splicing errors. The minor isoforms that are conserved, including PTC-containing isoforms, are more likely to be functional than minor isoforms that are seen only in one species, although species-specific isoforms may also be functional.[42]

As described above, Baek and Green identified PTC$^+$ isoforms that were conserved between human and mouse to filter out aberrant splicing. They note that the inclusion of the same "accidental" alternative exon is unlikely to happen by chance in both species, but that occasional accidental skipping of the same exon could more readily happen by chance in both human and mouse. To reduce the influence of these conserved but aberrant splicing events on their data set, Baek and Green designed a statistical method to discriminate between splicing errors and functional alternative splicing. Using this method, they inferred that 80% of the conserved PTC-producing splice

events they considered were legitimate, compared to 20% that appeared aberrant.[21] Thus, most of the conserved PTC-producing splice events they observed were not likely due to missplicing.

Pan and coworkers used an experimental approach to understand whether the prevalence of PTC^+ transcripts is a result of functional gene regulation or splicing noise.[23] As previously described, they developed an alternative splicing microarray platform to detect the relative abundance of PTC^+ versus PTC^- isoforms for over a thousand cassette-exon type alternative splicing events in mouse. Their study showed that in 10 diverse untreated mouse tissues where NMD is active, most PTC^+ isoforms represent less than 50% of the steady state pool of transcripts from a given gene. The low abundance is consistent either with a reduction in the levels of PTC^+ isoforms due to the action of NMD, or with infrequent occurrence of the alternative splice events that produces the PTC^+ isoforms.

To address these two possibilities in one cell type, Pan et al measured the changes in relative abundance of PTC^+ isoforms upon NMD inhibition in HeLa cells, using a microarray profiling 3055 human cassette exons. A small percentage of PTC^+ isoforms were upregulated after NMD inhibition, suggesting that their unproductive splicing could affect gene expression. Nonetheless, because the majority of PTC+ isoforms are present at low abundance even when NMD is inhibited, Pan et al inferred that most PTC^+ isoforms may not contribute important functional roles. One cannot exclude the possibility that subtle changes in the abundance of some PTC^+ isoforms have functional consequences, perhaps in different tissues. Nevertheless, this study suggests that the majority of PTC^+ isoforms may simply be due to infrequent splicing events and represent potential cellular noise cleared by the NMD machinery.[43]

If many PTC-containing transcripts are a result of splicing noise, their prevalence in a wide variety of genes could reflect a functional role that allows for the evolution of new gene functions via alternative splicing.[40,44] The existence of NMD could have led to an increase in alternative splicing, because any splicing errors that introduced PTCs would be removed by NMD, reducing the harmful effects of missplicing. As a result, the pressure to recognize splice sites perfectly would be lowered. Functional alternative splice forms could arise through splicing errors and then become established by sequence changes that strengthen their splice sites or add regulatory elements.

Are PTC^+ mRNA Isoforms Important for the Regulation of Gene Expression?

There are many examples of specific transcripts that are regulated by the coupling of alternative splicing and NMD (Table 1). This process provides an additional level of regulatory circuitry to help the cell achieve the proper level of expression for a given protein. The cell could change the level of productive mRNA after transcription by shunting some fraction of the already-transcribed pre-mRNA into an unproductive splice form and then to the decay pathway (Fig. 1). In the simplest case, some constant fraction of a gene's pre-mRNA is spliced into an unproductive, NMD-targeted form. In other cases, the proportion of transcripts targeted for degradation is regulated by an external input. Finally, autoregulatory loops can arise in which a protein affects the splicing pattern of its own pre-mRNA. The process of gene regulation through the coupled action of alternative splicing and NMD has been termed "Regulated Unproductive Splicing and Translation," or RUST.

Constitutive Unproductive Splicing

The simplest type of coupled alternative splicing and NMD is one in which the ratio of productive to unproductive splice forms is not significantly variable. In this case, the combined effect of alternative splicing with NMD reduces message abundance by a more or less constant factor. An apparent example of this is the *Calpain-10* gene, which encodes a ubiquitously expressed protease and is alternatively spliced to produce eight mRNA isoforms.[20,45] An analysis of these isoforms using SWISS-PROT and genomic sequences showed that four contained PTCs. An expression study by Horikawa et al showed that the four PTC^+ isoforms were "less abundant" in vivo than the other four.[45] Further experiments showed that the PTC^+ isoforms increased in abundance relative to the PTC^- isoforms when cells were treated with cycloheximide, which blocks translation and

thereby inactivates the NMD pathway.[46] This result confirmed that all eight mRNA isoforms are produced but that the four PTC$^+$ isoforms are degraded by NMD. Other experimentally confirmed examples in the literature reflect apparent constitutive unproductive splicing (Table 1a). Such cases may not be regulation, but simply cellular noise, with the unproductive splicing providing little or no selective advantage or function. Of course, in each of these cases, there may be as yet unknown regulatory inputs that impact the splicing process and alter the isoform proportions.

Regulated Unproductive Splicing

There are many examples of regulated alternative splicing, particularly in tissue-specific alternative splicing events, e.g., references 47-49. The role of regulated alternative splicing is emerging as an important layer of gene regulation, much like gene regulation at the transcriptional and translational levels.[39] Twenty-three examples of regulated alternative splicing leading to NMD are shown in Table 1. In addition to changing the relative abundance of functional isoforms, changes in the splicing environment may increase or decrease the production of translated isoforms relative to PTC$^+$ isoforms that are degraded by NMD (Fig. 1).

The 5' and 3' splice sites recognized by the spliceosome have a range of "strengths" or binding affinities for the core spliceosome components. Selection of splice sites is also under the control of a host of regulatory splicing factors which bind to specific sequence signals in the pre-mRNA. These sequences may be exonic or intronic and may be associated with enhancement or suppression of splicing at nearby (and sometimes at distant) splice sites. *Cis*-regulatory sequences, such as exonic splicing enhancers (ESEs), are frequently found in clusters, suggesting a combinatorial regulation of splicing by complexes of splicing factors.[50,51] (Refer to the chapter by L. Chasin in this volume.)

A change in the abundances of splicing factors can shift the balance of splicing patterns towards the production of NMD-targeted isoforms, thereby reducing the abundance of productive transcripts and hence the rate of protein production. In this way, splicing factors can act as regulatory inputs to alter gene expression in a manner analogous to transcription factors. An example of this intriguing mode of gene regulation is *MID1*, which encodes a microtubule-associated protein involved in triggering the degradation of phosphatase 2A.[52] *MID1* is ubiquitously transcribed, but it is spliced in a tissue- and development-specific manner. Winter and coworkers observed numerous alternatively spliced transcripts which included novel alternative exons, in addition to nine previously known constitutive exons. Most of the transcripts with novel exons contained in-frame stop codons and subsequent alternative poly(A+) tails; the alternative polyadenylation meant that the stop codons were not premature, allowing for translation of a C-terminally truncated protein. A second class of alternative transcripts contained stop codons closely followed by an in-frame start codon, suggesting the possibility of translation reinitiation and production of N-terminally truncated protein. A third class of alternative transcripts contained premature stop codons that were associated neither with an alternate poly(A+) signal nor with an alternate start site. These transcripts were predicted to be subject to NMD according to the 50-nucleotide rule. Consistent with this prediction, Winter et al found that the abundance of human *MID1* transcripts including exon 1c (an alternative exon introducing a PTC) increased in the presence of the NMD inhibitor cycloheximide.[52] Finally, Winter et al used RT-PCR to observe that different *MID1* isoforms are produced in different tissues and at different developmental stages in both human and mouse. For instance, the PTC-introducing exon 1a was observed in five distinct transcripts in human fetal brain cells, two transcripts in fetal liver cells and none in fetal fibroblasts. These results strongly suggest that alternative splicing and NMD are being employed to regulate the overall abundance of productive *MID1* transcripts.

The gene encoding spermidine/spermine N^1-acetyltransferase (SSAT), an enzyme that regulates the intracellular levels of the polyamines spermidine and spermine, provides an interesting example of RUST.[53] SSAT acetylates spermidine and spermine, which are then excreted out of the cell.[53] Polyamines were known to regulate gene expression of *SSAT* at the level of transcription and stabilization of the mRNA[53] and Hyvönen et al present evidence that polyamines regulate

gene expression of *SSAT* by promoting the exclusion of an exon containing three in-frame PTCs. Upregulation of the PTC$^+$ isoform occurs when polyamine levels are low—a condition where the enzyme is not needed. Upon depletion of polyamines in mouse embryonic stem cells, Hyvönen et al observed an increase in the relative abundance of the PTC$^+$ isoform of *SSAT*, termed *SSAT-X*. Conversely, after treating cells with DENSpm, a polyamine analog, they observed a decrease in the amount of *SSAT-X* mRNA relative to the normal, PTC$^-$ *SSAT* mRNA.[53] To demonstrate that *SSAT-X* mRNAs are degraded by NMD, Hyvönen et al inhibited NMD in fetal fibroblasts by treatment with cycloheximide or an siRNA targeted to *UPF1*. Under both conditions, they observed an increase in the relative amount of *SSAT-X* mRNA, thus providing evidence that *SSAT-X* is a target of NMD. Furthermore, they observed a decrease in the relative amount of *SSAT-X* in cells treated with cycloheximide and DENSpm compared to cells treated with only cycloheximide, indicating that the addition of polyamines does not enhance NMD activity, but does affect alternative splicing.[53] Polyamines also affected the splicing of a PTC$^+$ isoform of an unrelated gene, *Clk1,* but did not affect the splicing of three other genes, indicating that polyamines may specifically regulate other transcripts as well. Although the mechanism which enables polyamines to regulate changes in alternative splicing of *SSAT* is not known, the results of the above experiments show a clear and novel example of RUST.

Defects in the regulation of unproductive splicing can lead to disease. Myotonic dystrophy (DM), an autosomal dominant disease, is the most common form of adult-onset muscular dystrophy. DM has been shown to be caused by either of two repeat expansions whose presence in an mRNA affect the function of several splicing factors[54] such as CUG-BP1 and thus induce splicing changes in several genes.[55,56] (Refer to the chapter by T. Cooper in this volume.) The mechanism by which these repeat expansions affect the function of CUG-BP1 is not clear; however, the misregulation of CUG-BP1 has downstream effects that contribute to DM. Patients develop myotonia from lack of muscle-specific chloride channel 1 (ClC-1). The misexpression of CUG-BP1 in DM tissue results in the mis-splicing of the *ClC-1* pre-mRNA.[57]

The normal developmental splicing pattern for *ClC-1* has a PTC$^+$ splice form in embryos but a productive splice form in adult cells. In DM tissue, *ClC-1* splicing reverts to its embryonic, PTC-containing splicing pattern—which is greatly reduced in abundance, likely as a consequence of NMD.[58] Tissues from DM patients have increased steady-state levels of CUG-BP1 protein and the overexpression of CUG-BP1 in mouse skeletal muscle and heart tissues results in the embryonic splicing pattern of *ClC-1*.[59,60] In addition, expression of the CUG-BP1 protein is decreased in mouse skeletal muscle and heart tissues shortly after birth,[60,61] providing evidence that the CUG-BP1 protein influences the splicing of the PTC$^+$ isoform. Thus, it appears that normal *ClC-1* expression is governed by RUST and that the DM disease is caused by undermining the proper function of splicing factors with consequent disruption of RUST.

Autoregulatory Unproductive Splicing

There is abundant evidence that RUST is used for autoregulation, most prevalently of proteins that are part of the splicing or translation machinery. In some fascinating cases, proteins that are not generally involved in mRNA processing bind specifically to their own transcripts to affect their splicing and elicit NMD. One example of this is found in yeast. Yeast genes are generally unspliced, but in the few intron-containing genes, intron inclusion can introduce an in-frame stop codon and target the transcript for NMD. The yeast ribosomal protein RPL30 binds to its own pre-mRNA to prevent intron removal. This retained intron contains a PTC that triggers NMD.[62] The mRNAs of other yeast ribosomal protein genes, including *RPL28* (*CYH2*) and *RPS17B* (*RP51B*), sometimes retain their introns and become natural NMD targets, leaving open the possibility that their splicing is also regulated to elicit NMD.[63]

Some ribosomal proteins in *C. elegans* are similarly autoregulated. A screen for natural targets of NMD identified the genes for ribosomal proteins L3, L7a, L10a and L12. Each of these genes can be alternatively spliced to generate either a productive isoform or an unproductive isoform that contains a PTC and is therefore degraded by NMD. The ratio of productive to unproductive

alternative splicing of *rpl-12* is affected by levels of RPL-12 protein, indicating that unproductive splicing of *rpl-12* is under feedback control.[64] More recently, NMD-target isoforms of the human ribosomal protein genes *rpL3* and *rpL12* were identified and the unproductive splicing of *rpL3* was shown to be autoregulated by rpL3 protein in a negative feedback loop.[99] RUST thus seems to play a similar role in the regulation of ribosomal proteins in species from yeast to human.

A striking number of splicing factors and elements of the splicing machinery are autoregulated through RUST (Fig. 5 and Tables 1c and 1d). One such example is the polypyrimidine tract binding protein (PTB), a protein whose function is to inhibit splicing by competing with U2AF for the polypyrimidine tract and perhaps through other mechanisms as well (reviewed in refs. 50,65).

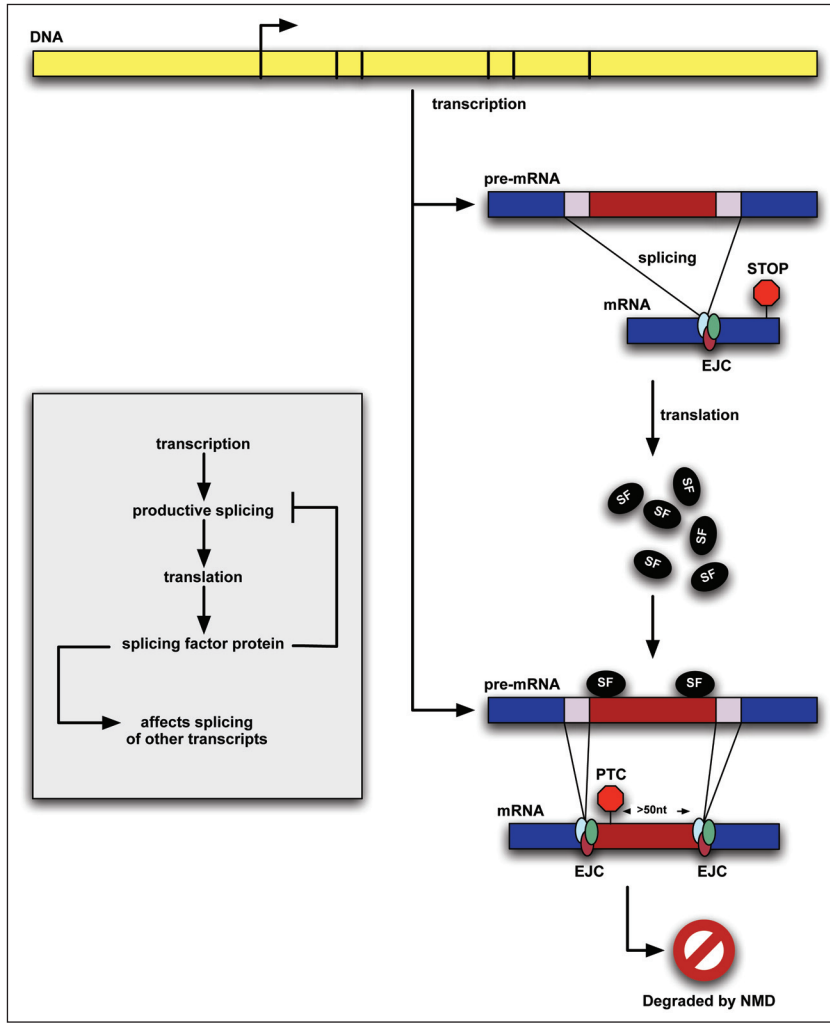

Figure 5. Autoregulatory unproductive splicing. Some splicing factors, such as PTB and SC35, regulate the splicing of their own transcripts so as to alter the proportion of unproductive isoforms.[10,67] This creates a negative feedback loop, stabilizing the concentration of the splicing factor over time. Autoregulated splicing factors are generally not specific to their own transcripts, however; they impinge on the splicing of many other pre-mRNAs as well.

PTB is alternatively spliced to produce two major productive isoforms (one of which lacks exon 9),[66,67] one minor productive isoform lacking exons 3-9,[67,68] and two unproductive isoforms lacking exon 11. Removing exon 11 causes a frameshift leading to a downstream PTC. PTB protein has been found to promote the removal of exon 11 from its own transcripts.[67] Consequently, when PTB levels are high, PTB production is slowed by targeting *PTB* transcripts for NMD and when PTB levels are low, production is accelerated by reducing the proportion of transcripts that are degraded.[67,69,70]

A similar autoregulatory process has been reported for members of a family of splicing factors known as SR proteins. Overexpression of the SR protein SC35 upregulates the splicing of its own NMD-targeted isoform to reduce protein production.[10] Intriguingly, similar unproductive splicing is found in all human SR genes and some hnRNPs,and the alternative splicing events that create the PTC-containing isoforms are conserved in mouse orthologs.[107, 108] Remarkably, some of the most conserved regions of the human and mouse genomes are associated with this unproductive splicing.

The CDC-like kinases (Clks), which regulate SR proteins, seem to be affected by RUST as well.[20] RUST appears to regulate the Clk1 protein through an indirect feedback mechanism. Clk1 has been shown to indirectly modify splicing of its own transcript, most likely through phosphorylation of SR proteins.[71] Thus, as a variation of the autoregulatory circuit described above, increased Clk1 activity may result in changes in the activity of one or more SR proteins. These SR proteins in turn affect the splicing of *Clk1* pre-mRNA to favor a PTC^+ transcript that is predicted to undergo NMD. This PTC^+ transcript is stabilized by cycloheximide, providing evidence that it is indeed normally degraded by NMD.[72] RUST regulation of SR proteins, Clks and PTB may have downstream effects on many pre-mRNAs. Thus, RUST can regulate factors that control alternative splicing of many other gene products.

Finally, splicing factors that are autoregulated by RUST may also be subject to RUST that is triggered by heterologous factors rather than autoregulation. This is seen in the alternative splicing of *PTB*, which can be affected by the splicing regulators raver1 and CELF4, forming a network of regulatory factors contributing to RUST.[67]

Conservation of RUST

The coordinated use of alternative splicing and NMD is seen not only in mammals but in organisms as distant as yeast[62] and plants.[73] The mechanism of PTC recognition differs between mammals and other species, where it does not seem to depend on the location of the stop codon relative to exon junctions.[74] There have been significant advances recently in elucidating the recognition mechanism in flies and yeast,[75,76] but the rules are not clear enough to allow for computational identification of NMD targets. Nonetheless, NMD affects gene expression in a variety of different organisms.[24,77]

In several of the examples discussed above, analysis of orthologous and paralogous sequences suggests that splicing to generate PTC^+ alternative isoforms and thus RUST regulation, is shared across species and across protein families. In the case of *PTB*, the sequence and upstream regulatory elements of alternatively included PTC-containing exon 11 are very similar between the *Fugu rubripes* ortholog and the analyzed human gene, as well as in the human *nPTB* paralog.[67] Mouse and monkey orthologs of the human multidrug resistance associated transporter *ABCC4* share highly conserved PTC-containing exons that are orthologous to the alternatively included exons of human *ABCC4*, another apparent RUST target.[78] Particularly strong evidence of conservation of RUST is found in the Clks. Alternative splicing to exclude exon 4, introducing a frameshift and PTC, is conserved among three human paralogs (*Clk1, Clk2* and *Clk3*), the three orthologs of these genes in mouse and even the sole copy of the gene in the sea squirt *Ciona intestinalis*.[20] One SR protein in *Ciona* also has unproductive splice patterns matching those seen in human and mouse.[107]

The action of NMD on a gene can be retained even when the specific alternative splicing events that elicit NMD are not conserved. As discussed above, MID1 is a human RUST target. Interestingly,

while PTC$^+$ isoforms of MID1 were found in human, mouse and fugu, the responsible stop codons were introduced by alternative exons that showed no homology between these species.[52] Thus, in this case, it appears that the RUST mode of regulation was maintained while the specific sequence elements triggering it were not. This suggests that RUST has often and readily evolved to regulate specific classes of genes in organisms that already have both alternative splicing and NMD.

Why RUST?

A substantial portion of alternatively-spliced mRNAs seem to be targets of NMD. We have discussed possible explanations for the prevalence of unproductive splicing: do these splice forms represent biological noise or are they produced to regulate protein expression? The relatively low abundance and lack of conservation of many PTC$^+$ isoforms suggests that many of these isoforms are nonfunctional or cellular noise, but the growing body of examples nevertheless suggests that RUST plays a significant regulatory role in the cell.

Many truncated proteins encoded by alternative transcripts would be nonfunctional even if their transcripts were not removed by NMD. Is the combination of alternative splicing and NMD inherently different from alternative splicing that produces nonfunctional protein? Or does alternative splicing alone provide the important regulatory step, with NMD acting only as a convenient but inessential cleanup mechanism? Some proven cases of RUST illustrate that the coordinated action of both pathways is required for regulation. As described above, expression of the SR protein SC35 is autoregulated by RUST; its alternative splicing occurs in the 3' UTR to create an exon junction downstream of the original stop codon without changing the open reading frame.[10] The alternative splicing seems to have no role other than causing the original termination codon to be recognized as premature. Without NMD, the alternative mRNA would still encode the full-length protein, so the alternative splice event alone could not be used to regulate protein levels. It seems, then, that some genes have evolved to take advantage of the combination of alternative splicing and NMD in a role different from those filled by either process alone.

RUST seems, at first, to be a wasteful process. A gene is transcribed and spliced, only to be degraded before it can produce a protein. Yet we know that there are functional cases of RUST. The cost to the cell of transcribing apparently-extraneous RNA is clearly not prohibitive. In humans, roughly 85-95% of transcribed sequence is spliced out as introns and discarded.[79] Evidently, transcription of intron sequence is not a significant selective disadvantage and intron splicing may even provide some general selective advantage. Similarly, the cost of transcribing a pre-mRNA only to splice it into an unproductive form must be balanced by the advantages of an additional layer of regulation of gene expression or the flexibility to evolve new gene isoforms without harmful effects.

How is a process like RUST beneficial to the cell? It provides an additional level of regulation. Transcriptional regulation is the most studied means of controlling gene expression, but in some cases additional control may be beneficial. Splicing regulation occurs after the decision to transcribe a region and RUST may provide a rapid way to change the levels of productive mRNA. In extreme cases such as the dystrophin gene, transcription can take many hours[80] and the requirements of the cell might change after transcription begins but before a critical splicing decision that determines whether or not to introduce a PTC. Even when temporal regulation is not necessary, an extra layer of regulation can help fine tune or amplify transcriptional and other regulation.

RUST could either increase or decrease protein expression from steady-state levels. The splicing factor PTB illustrates this point. At steady state, 20% of the transcribed pre-mRNAs of PTB are spliced to an unproductive form.[67] In general, we expect that a RUST-regulated gene is transcribed to produce more pre-mRNA than is needed at steady state and that in normal conditions there is a base level of downregulation by unproductive splicing. This fraction of "wasted" transcripts constitutes the headroom available to the regulatory system to increase levels of productive transcript.

In a system with prevalent alternative splicing, regulation by RUST may evolve easily. For any particular gene, there are many possible alternative splicing events that could introduce a PTC and elicit NMD (Fig. 1). If the sequence of the gene changes slightly to promote one of these splicing

events in certain splicing environments and the resulting downregulation of the gene by NMD is beneficial, then a basic sort of regulation has evolved. This has clearly occurred independently many times. Indeed, RUST seems to have evolved independently in every one of the SR genes.[107] Without NMD, alternative splicing can still regulate gene expression by producing nonfunctional proteins. The additional advantages of coupling splicing with NMD may be that it prevents accumulation of potentially harmful truncated proteins and that it reduces wasted translation, making unproductive splicing less costly.

Splicing factors such as PTB seem to be overrepresented among the known RUST targets. Is this a coincidence or acquisition bias, or is RUST used most often to regulate a small set of proteins that are already capable of binding pre-mRNAs? The latter may be the case for autoregulation by RUST. A protein that has an existing role in splicing may evolve autoregulation through splicing more easily than a non-RNA-binding protein. Indeed, this is a simple and elegant means of regulation for RNA binding proteins. There are only a handful of known cases in which a protein that is not a splicing factor is autoregulated by RUST and even these are predominantly ribosomal proteins that do bind RNA in other, nonsplicing contexts. However, autoregulation is by no means the only role of RUST and there is no reason for non-autoregulatory RUST to affect splicing factors preferentially. The examples listed in Table 1 indicate that RUST is involved in the regulation of a diverse set of proteins.

The potential for alternative splicing to regulate gene expression has been appreciated for many years. Bingham et al proposed that "on/off regulation at the level of splicing might be unexpectedly common," in a 1988 review featuring three cases of unproductive splicing in *Drosophila*.[81] An early paper about the splicing factor ASF discussed alternative splicing as a quantitative control of gene expression.[82] Nonsense-mediated decay adds an additional layer to the story[83]; many of the unproductive splice forms identified years ago are now known to be degraded rather than translated. Alternative splicing and NMD can be combined in an elegant way to regulate a wide variety of genes.

Acknowledgements

We would like to thank Maki Inada, Richard E. Green and Rajiv Bhatnagar for helpful discussions and comments. This work was supported by National Instituties of Health grants K22 HG00056 and R01 GM071655 and an IBM SUR grant. DAWS is supported by a predoctoral fellowship from the Howard Hughes Medical Institute. ANB is supported by the Chancellor's Fellowship for Graduate Study from the University of California, Berkeley. SEB is also supported by NIH/NIGMS R01 GM073109 and Sloan and Searle fellowships. This chapter represents an update of a version that was previously published.[109]

References

1. Boue S, Letunic I, Bork P. Alternative splicing and evolution. Bioessays 2003; 25:1031-1034.
2. Modrek B, Lee C. A genomic view of alternative splicing. Nat Genet 2002; 30:13-19.
3. Maniatis T, Tasic B. Alternative pre-mRNA splicing and proteome expansion in metazoans. Nature 2002; 418:236-243.
4. Garcia J, Gerber SH, Sugita S et al. A conformational switch in the piccolo C2A domain regulated by alternative splicing. Nat Struct Mol Biol 2004; 11:45-53.
5. Resch A, Xing Y, Modrek B et al. Assessing the impact of alternative splicing on domain interactions in the human proteome. J Proteome Res 2004; 3:76-83.
6. Xing Y, Xu Q, Lee C. Widespread production of novel soluble protein isoforms by alternative splicing removal of transmembrane anchoring domains. FEBS Lett 2003; 555:572-578.
7. Maquat LE. NMD in mammalian cells: A history. In: Maquat LE, ed. Nonsense-Mediated mRNA Decay. Georgetown: Landes Bioscience, 2005:45-58.
8. Nagy E, Maquat LE. A rule for termination-codon position within intron-containing genes: When nonsense affects RNA abundance. Trends Biochem Sci 1998; 23:198-199.
9. Lejeune F, Maquat LE. Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. Curr Opin Cell Biol 2005; 17:309-315.
10. Sureau A, Gattoni R, Dooghe Y et al. SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. EMBO J 2001; 20:1785-1796.

11. Cali BM, Anderson P. mRNA surveillance mitigates genetic dominance in caenorhabditis elegans. Mol Gen Genet 1998; 260:176-184.
12. Maquat LE, Carmichael GG. Quality control of mRNA function. Cell 2001; 104:173-176.
13. He F, Jacobson A. Endogenous substrates of the yeast NMD pathway. In: Maquat LE, ed. Nonsense-Mediated mRNA Decay. Georgetown: Landes Bioscience, 2005:27-27-41.
14. Sharifi NA, Dietz HC. Physiologic substrates and functions for mammalian NMD. In: Maquat LE, ed. Nonsense-Mediated mRNA Decay. Georgetown: Landes Bioscience, 2005:97-97-109.
15. Mendell JT, Sharifi NA, Meyers JL et al. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat Genet 2004; 36:1073-1078.
16. Lewis BP, Green RE, Brenner SE. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc Natl Acad Sci USA 2003; 100:189-192.
17. Pruitt KD, Maglott DR. RefSeq and LocusLink: NCBI gene-centered resources. Nucleic Acids Res 2001; 29:137-140.
18. Boguski MS, Lowe TM, Tolstoshev CM. dbEST—database for "expressed sequence tags". Nat Genet 1993; 4:332-333.
19. Boeckmann B, Bairoch A, Apweiler R et al. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res 2003; 31:365-370.
20. Hillman RT, Green RE, Brenner SE. An unappreciated role for RNA surveillance. Genome Biol 2004; 5:R8.
21. Baek D, Green P. Nonsense-mediated decay, sequence conservation and relative isoform frequencies in evolutionarily conserved alternative splicing. Proc Natl Acad Sci USA 2005.
22. Wittmann H, Jack. hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay. Molecular and cellular biology 2006; 26:1272.
23. Pan Q, Saltzman AL, Kim YK et al. Quantitative microarray profiling provides evidence against widespread coupling of alternative splicing with nonsense-mediated mRNA decay to control gene expression. Genes development 2006; 20:153.
24. Rehwinkel J, Letunic I, Raes J et al. Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. RNA 2005; 11:1530.
25. Ishigaki Y, Li XJ, Serin G et al. Evidence for a pioneer round of mRNA translation: MRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. Cell 2001; 106:607-617.
26. Maquat LE. Nonsense-mediated mRNA decay: Splicing, translation and mRNP dynamics. Nat Rev Mol Cell Biol 2004; 5:89-99.
27. Cao D, Parker R. Computational modeling and experimental analysis of nonsense-mediated decay in yeast. Cell 2003; 113:533-545.
28. Chang AC, Sohlberg B, Trinkle-Mulcahy L et al. Alternative splicing regulates the production of ARD-1 endoribonuclease and NIPP-1, an inhibitor of protein phosphatase-1, as isoforms encoded by the same gene. Gene 1999; 240:45-55.
29. Veldhoen N, Metcalfe S, Milner J. A novel exon within the mdm2 gene modulates translation initiation in vitro and disrupts the p53-binding domain of mdm2 protein. Oncogene 1999; 18:7026-7033.
30. Zhang J, Maquat LE. Evidence that translation reinitiation abrogates nonsense-mediated mRNA decay in mammalian cells. EMBO J 1997; 16:826-833.
31. Chester A, Somasekaram A, Tzimina M et al. The apolipoprotein B mRNA editing complex performs a multifunctional cycle and suppresses nonsense-mediated decay. EMBO J 2003; 22:3971-3982.
32. Inácio A, Silva AL, Pinto J et al. Nonsense mutations in close proximity to the initiation codon fail to trigger full nonsense-mediated mRNA decay. J Biol Chem 2004; 279:32170-32180.
33. Stockklausner C, Breit S, Neu-Yilik G et al. The uORF-containing thrombopoietin mRNA escapes nonsense-mediated decay (NMD). Nucleic acids research 2006; 34:2355.
34. Danckwardt S, Neu-Yilik G, Thermann R et al. Abnormally spliced beta-globin mRNAs: A single point mutation generates transcripts sensitive and insensitive to nonsense-mediated mRNA decay. Blood 2002; 99:1811-1816.
35. Dreumont N, Maresca A, Boisclair-Lachance JF et al. A minor alternative transcript of the fumarylacetoacetate hydrolase gene produces a protein despite being likely subjected to nonsense-mediated mRNA decay. BMC Mol Biol 2005; 6:1.
36. McAdams HH, Arkin A. It's a noisy business! genetic regulation at the nanomolar scale. Trends Genet 1999; 15:65-69.
37. Jones RB, Wang F, Luo Y et al. The nonsense-mediated decay pathway and mutually exclusive expression of alternatively spliced FGFR2IIIb and -IIIc mRNAs. J Biol Chem 2001; 276:4158-4167.
38. Yeo GW, Van Nostrand E, Holste D et al. Identification and analysis of alternative splicing events conserved in human and mouse. Proc Natil Acad Sci USA 2005; 102:2850.
39. Blencowe BJ. Alternative splicing: New insights from global analyses. Cell 2006; 126:37.

40. Lareau LF, Green RE, Bhatnagar RS et al. The evolving roles of alternative splicing. Curr Opin Struct Biol 2004; 14:273-282.
41. Modrek B, Lee CJ. Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation and/or loss. Nat Genet 2003; 34:177-180.
42. Kan Z, States D, Gish W. Selecting for functional alternative splices in ESTs. Genome Res 2002; 12:1837-1845.
43. Neu-Yilik G, Gehring NH, Hentze MW et al. Nonsense-mediated mRNA decay: From vacuum cleaner to swiss army knife. GenomeBiology.com 2004; 5:218.
44. Xing Y, Lee CJ. Negative selection pressure against premature protein truncation is reduced by alternative splicing and diploidy. Trends Genet 2004; 20:472-475.
45. Horikawa Y, Oda N, Cox NJ et al. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. Nat Genet 2000; 26:163-175.
46. Green RE, Lewis BP, Hillman RT et al. Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes. Bioinformatics 2003; 19(Suppl 1):118-121.
47. Ule J, Ule A, Spencer J et al. Nova regulates brain-specific splicing to shape the synapse. Nature genetics 2005; 37:844.
48. Yeo G, Holste D, Kreiman G et al. Variation in alternative splicing across human tissues. GenomeBiology.com 2004; 5:R74.
49. Sugnet CW, Srinivasan K, Clark TA et al. Unusual intron conservation near tissue-regulated exons found by splicing microarrays. PLoS Computational Biology 2006; 2:e4.
50. Black DL. Mechanisms of alternative premessenger RNA splicing. Annu Rev Biochem 2003; 72:291-336.
51. Wagner EJ, Baraniak AP, Sessions OM et al. Characterization of the intronic splicing silencers flanking FGFR2 exon IIIb. J Biol Chem 2005.
52. Winter J, Lehmann T, Krauss S et al. Regulation of the MID1 protein function is fine-tuned by a complex pattern of alternative splicing. Hum Genet 2004; 114:541-552.
53. Hyvönen MT, Uimari A, Keinänen TA et al. Polyamine-regulated unproductive splicing and translation of spermidine/spermine N1-acetyltransferase. RNA 2006; 12:1569.
54. Philips AV, Timchenko LT, Cooper TA. Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. Science 1998; 280:737-741.
55. Brook JD, McCurrach ME, Harley HG et al. Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family member. Cell 1992; 68:799-808.
56. Liquori CL, Ricker K, Moseley ML et al. Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. Science 2001; 293:864-867.
57. Charlet BN, Savkur RS, Singh G et al. Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. Mol Cell 2002; 10:45-53.
58. Mankodi A, Takahashi MP, Jiang H et al. Expanded CUG repeats trigger aberrant splicing of ClC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. Mol Cell 2002; 10:35-44.
59. Ho TH, Bundam D, Armstrong DL. Transgenic mice expressing CUG-BP1 reproduce splicing mis-regulation observed in myotonic dystrophy. Human molecular genetics 2005; 14:1539.
60. Ranum LPW, Cooper TA. RNA-mediated neuromuscular disorders. Annual review of neuroscience 2006; 29:259.
61. Ladd AN, Stenberg MG, Swanson MS et al. Dynamic balance between activation and repression regulates pre-mRNA alternative splicing during heart development. Developmental dynamics 2005; 233:783.
62. Vilardell J, Chartrand P, Singer RH et al. The odyssey of a regulated transcript. RNA 2000; 6:1773-1780.
63. He F, Peltz SW, Donahue JL et al. Stabilization and ribosome association of unspliced pre-mRNAs in a yeast upf1- mutant. Proc Natl Acad Sci USA 1993; 90:7034-7038.
64. Mitrovich QM, Anderson P. Unproductively spliced ribosomal protein mRNAs are natural targets of mRNA surveillance in C. elegans. Genes Dev 2000; 14:2173-2184.
65. Valcárcel J, Gebauer F. Post-transcriptional regulation: The dawn of PTB. Curr Biol 1997; 7:R705-8.
66. Ghetti A, Piñol-Roma S, Michael WM et al. hnRNP I, the polypyrimidine tract-binding protein: Distinct nuclear localization and association with hnRNAs. Nucleic Acids Res 1992; 20:3671-3678.
67. Wollerton MC, Gooding C, Wagner EJ et al. Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. Mol Cell 2004; 13:91-100.
68. Hamilton BJ, Genin A, Cron RQ et al. Delineation of a novel pathway that regulates CD154 (CD40 ligand) expression. Mol Cell Biol 2003; 23:510-525.

69. Rahman L, Bliskovski V, Reinhold W et al. Alternative splicing of brain-specific PTB defines a tissue-specific isoform pattern that predicts distinct functional roles. Genomics 2002; 80:245-249.
70. Spellman R, Rideau A, Matlin A et al. Regulation of alternative splicing by PTB and associated factors. Biochem Soc Trans 2005; 33:457-460.
71. Duncan PI, Stojdl DF, Marius RM et al. In vivo regulation of alternative pre-mRNA splicing by the Clk1 protein kinase. Mol Cell Biol 1997; 17:5996-6001.
72. Menegay HJ, Myers MP, Moeslein FM et al. Biochemical characterization and localization of the dual specificity kinase CLK1. J Cell Sci 2000; 113(Pt 18):3241-3253.
73. Staiger D, Zecca L, Wieczorek Kirk DA et al. The circadian clock regulated RNA-binding protein AtGRP7 autoregulates its expression by influencing alternative splicing of its own pre-mRNA. Plant J 2003; 33:361-371.
74. Conti E, Izaurralde E. Nonsense-mediated mRNA decay: Molecular insights and mechanistic variations across species. Curr Opin Cell Biol 2005; 17:316-325.
75. Amrani N, Ganesan R, Kervestin S et al. A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature 2004; 432:112-118.
76. Gatfield D, Unterholzner L, Ciccarelli FD et al. Nonsense-mediated mRNA decay in drosophila: At the intersection of the yeast and mammalian pathways. EMBO J 2003; 22:3960-3970.
77. Arciga-Reyes L, Wootton L, Kieffer M et al. UPF1 is required for nonsense-mediated mRNA decay (NMD) and RNAi in arabidopsis. The plant journal 2006; 47:480.
78. Lamba JK, Adachi M, Sun D et al. Nonsense mediated decay downregulates conserved alternatively spliced ABCC4 transcripts bearing nonsense codons. Hum Mol Genet 2003; 12:99-109.
79. Lander ES, Linton LM, Birren B et al. Initial sequencing and analysis of the human genome. Nature 2001; 409:860-921.
80. Tennyson CN, Klamut HJ, Worton RG. The human dystrophin gene requires 16 hours to be transcribed and is cotranscriptionally spliced. Nat Genet 1995; 9:184-190.
81. Bingham PM, Chou TB, Mims I et al. On/off regulation of gene expression at the level of splicing. Trends Genet 1988; 4:134-138.
82. Ge H, Zuo P, Manley JL. Primary structure of the human splicing factor ASF reveals similarities with drosophila regulators. Cell 1991; 66:373-382.
83. Hilleren P, Parker R. Mechanisms of mRNA surveillance in eukaryotes. Annu Rev Genet 1999; 33:229-260.
84. Skandalis A, Uribe E. A survey of splice variants of the human hypoxanthine phosphoribosyl transferase and DNA polymerase beta genes: Products of alternative or aberrant splicing? Nucleic Acids Res 2004; 32:6557-6564.
85. Carter MS, Li S, Wilkinson MF. A splicing-dependent regulatory mechanism that detects translation signals. EMBO J 1996; 15:5965-5975.
86. Wang J, Vock VM, Li S et al. A quality control pathway that down-regulates aberrant T-cell receptor (TCR) transcripts by a mechanism requiring UPF2 and translation. J Biol Chem 2002; 277:18489-18493.
87. Hovhannisyan RH, Carstens RP. A novel intronic cis element, ISE/ISS-3, regulates rat fibroblast growth factor receptor 2 splicing through activation of an upstream exon and repression of a downstream exon containing a noncanonical branch point sequence. Mol Cell Biol 2005; 25:250-263.
88. Mitrovich QM, Anderson P. mRNA surveillance of expressed pseudogenes in C. elegans. Current biology 2005; 15:963.
89. Blanchette M, Labourier E, Green RE et al. Genome-wide analysis reveals an unexpected function for the drosophila splicing factor U2AF(50) in the nuclear export of intronless mRNAs. Mol Cell 2004; 14:775-786.
90. Cazalla D, Newton K, Cáceres JF. A novel SR-related protein is required for the second step of pre-mRNA splicing. Mol Cell Biol 2005; 25:2969-2980.
91. Henscheid KL, Shin DS, Cary SC et al. The splicing factor U2AF65 is functionally conserved in the thermotolerant deep-sea worm Alvinella pompejana. Biochim Biophys Acta 2005.
92. Lallena MJ, Chalmers KJ, Llamazares S et al. Splicing regulation at the second catalytic step by sex-lethal involves 3' splice site recognition by SPF45. Cell 2002; 109:285-296.
93. Pacheco TR, Gomes AQ, Barbosa-Morais NL et al. Diversity of vertebrate splicing factor U2AF35: Identification of alternatively spliced U2AF1 mRNAS. J Biol Chem 2004; 279:27039-27049.
94. Confaloni A, Crestini A, Albani D et al. Rat nicastrin gene: CDNA isolation, mRNA variants and expression pattern analysis. Brain Res Mol Brain Res 2005; 136:12-22.
95. Engebrecht JA, Voelkel-Meiman K, Roeder GS. Meiosis-specific RNA splicing in yeast. Cell 1991; 66:1257-1268.
96. Tsyba L, Skrypkina I, Rynditch A et al. Alternative splicing of mammalian intersectin 1: Domain associations and tissue specificities. Genomics 2004; 84:106-113.

97. Petruzzella V, Panelli D, Torraco A et al. Mutations in the NDUFS4 gene of mitochondrial complex I alter stability of the splice variants. FEBS Lett 2005.
98. Screaton GR, Xu XN, Olsen AL et al. LARD: A new lymphoid-specific death domain containing receptor regulated by alternative pre-mRNA splicing. Proc Natl Acad Sci USA 1997; 94:4615-4619.
99. Cuccurese M, Russo G, Russo A et al. Alternative splicing and nonsense-mediated mRNA decay regulate mammalian ribosomal gene expression. Nucleic Acids Res 2005; 33:5965.
100. Le Guiner C, Gesnel MC, Breathnach R. TIA-1 or TIAR is required for DT40 cell viability. J Biol Chem 2003; 278:10465-10476.
101. Colwill K, Pawson T, Andrews B et al. The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. EMBO J 1996; 15:265-275.
102. Duncan PI, Stojdl DF, Marius RM et al. The Clk2 and Clk3 dual-specificity protein kinases regulate the intranuclear distribution of SR proteins and influence pre-mRNA splicing. Exp Cell Res 1998; 241:300-308.
103. Wilson GM, Sun Y, Sellers J et al. Regulation of AUF1 expression via conserved alternatively spliced elements in the 3' untranslated region. Mol Cell Biol 1999; 19:4056-4064.
104. Wilson GM, Brewer G. The search for trans-acting factors controlling messenger RNA decay. Prog Nucleic Acid Res Mol Biol 1999; 62:257-291.
105. Morrison M, Harris KS, Roth MB. Smg mutants affect the expression of alternatively spliced SR protein mRNAs in caenorhabditis elegans. Proc Natl Acad Sci USA 1997; 94:9782-9785.
106. Stoilov P, Daoud R, Nayler O et al. Human tra2-beta1 autoregulates its protein concentration by influencing alternative splicing of its pre-mRNA. Hum Mol Genet 2004; 13:509-524.
107. Lareau LF, Inada M, Green RE et al. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature 2007; 446:926-929.
108. Ni JZ, Grate L, Donahue JP et al. Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. Genes Dev 2007; 21:708-718.
109. Soergel DAW, Lareau LF, Brenner SE. Regulation of gene expression by the coupling of alternative splicing and nonsense-mediated mRNA decay. In: Maquat LE, ed. Nonsense-mediated mRNA decay. Goergetown: Landes Bioscience 2006; 175-196.