<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Emerging functions of alternative splicing coupled with nonsense-mediated decay</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Hamid, Fursham M.; Makeyev, Eugene V.</td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>2014</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10220/38868">http://hdl.handle.net/10220/38868</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>© The Authors Journal compilation © 2014 Biochemical Society</td>
</tr>
</tbody>
</table>
Emerging functions of alternative splicing coupled with nonsense-mediated decay

Fursham M. Hamid* and Eugene V. Makeyev†1
†School of Biological Sciences, Nanyang Technological University, Singapore
†MRC Centre for Developmental Neurobiology, King’s College London, London SE1 1UL, U.K.

Abstract
Higher eukaryotes rely on AS (alternative splicing) of pre-mRNAs (mRNA precursors) to generate more than one protein product from a single gene and to regulate mRNA stability and translational activity. An important example of the latter function involves an interplay between AS and NMD (nonsense-mediated decay), a cytoplasmic quality control mechanism eliminating mRNAs containing PTCs (premature translation termination codons). Although originally identified as an error surveillance process, AS-NMD additionally provides an efficient strategy for deterministic regulation of gene expression outputs. In this review, we discuss recently published examples of AS-NMD and delineate functional contexts where recurrent use of this mechanism orchestrates expression of important genes.

Introduction
Most pre-mRNAs (mRNA precursors) in the higher eukaryotes contain spliceosomal introns and up to 95% of these transcripts may undergo AS (alternative splicing) to generate more than one mature mRNA product [1,2]. AS increases eukaryotic protein diversity and may contribute to the evolutionary elaboration trend particularly evident in the metazoan clade [3,4]. Moreover, AS appears to be widely used in the higher eukaryotes to regulate mRNA stability, translational efficiency and intracellular localization [1,5,6].

One example of the latter regulation strategy relies on coupling between AS and NMD (nonsense-mediated decay), a quality control mechanism destabilizing mRNAs containing PTCs (premature translation termination codons) [7–9]. NMD requires a pioneering round of mRNA translation in the cytoplasm and depends on an intricate interplay between evolutionarily conserved Upf1-3 and Smg proteins as well as mRNP (messenger ribonucleoproteins) and components of the cellular translation machinery [7–9] (Figure 1A). A key step in this process is recognition of PTCs that can be realized through several distinct mechanisms depending on the species and mRNA identity.

In mammals, PTCs are commonly recognized through interaction of the NMD machinery with EJCs (exon–exon junction complexes) that are deposited on to newly spliced mRNAs in the nucleus [7–9] (Figure 1A). Following mRNA export from the nucleus to the cytoplasm, EJCs bound within or immediately downstream of an ORF are dislodged by translating ribosomes along with the associated NMD factors, whereas EJCs associated with exon–exon junctions occurring >50–55 nt downstream of a stop codon are typically left undisturbed. Consequently, stop codons positioned >50–55 nt of the last exon–exon junction tend to be recognized as PTCs and trigger mRNA decay [7–9] (Figure 1A).

The relationship between AS and NMD is two-fold. On the one hand, NMD functions as a surveillance mechanism that eliminates aberrant AS products appearing as a result of splicing errors. On the other hand, mounting evidence suggests that coupling of regulated AS events with NMD is used as an efficient strategy for deterministic control of gene expression (Figure 1B). Here we focus on the latter AS-NMD modality and delineate functional themes recurring in recently published studies.

Co-ordinating expression of functionally linked components of RNA metabolism
One of the most compelling lines of evidence suggesting that AS–NMD functions as a bona fide gene regulation mechanism came from identification of auto-regulation feedbacks maintaining cellular homoeostasis of RBPs (RNA-binding proteins) [10–14]. Interestingly, some of these studies pointed out that AS–NMD can additionally mediate cross-regulation between related RBPs including paralogous polypyrimidine tract-binding proteins and members of the SR (serine/arginine-rich) protein family [11,14].

Recent publications extend this theme and suggest that AS–NMD might co-ordinate expression of functionally linked components of cellular RNA metabolism on a substantially larger scale. A compelling example is provided by the regulation of spliceosomal protein levels. The spliceosome is an elaborate RNP (ribonucleoprotein) complex catalysing excision of most eukaryotic introns [15,16]. Two types of the spliceosome have been described, the canonical spliceosome
Figure 1 | Mammalian NMD and its relationship with AS

(A) Outline of NMD in mammalian cells. (B) AS-NMD strategies used to control gene expression.

responsible for splicing of >99% introns and the non-canonical one excising a minor proportion of introns [15–17]. The two spliceosomes are assembled from partially overlapping sets of snRNAs (small nuclear RNA) (U1, U2, U4, U4 and U6 in the canonical spliceosome and U11, U12, U4atac, U5 and U6atac in the non-canonical spliceosome) and multiple protein subunits.

The Frilander and Bindereif laboratories have recently explained how relative abundance of these components could be co-ordinated to achieve the required stoichiometries [18–20]. One of these studies describes a cross-regulation circuitry co-ordinating expression of U1C and U1-70K proteins critical for the U1 snRNP (small nuclear ribonucleoprotein) assembly and its function in 5′ splice site recognition [18] (Figure 2A). U1-70K pre-mRNA can be alternatively spliced to generate a functional mRNA containing the exon 7–8 junction or a non-productive form additionally containing a PTC-containing cassette exon (exon 7a). The choice between skipping and inclusion of the NMD-inducing (‘poison’) exon 7a is regulated by the levels of U1 snRNP that includes U1C protein. When expressed at sufficiently high levels, U1C-containing U1 snRNPs associate with cryptic 5′ splice sites and promote utilization of the upstream 3′ splice site of exon 7a potentially through a variant of the exon definition mechanism. However, decreased expression of U1C (and as a result, of U1C-containing U1 snRNP) encourages skipping of the U1-70K poison exon thus up-regulating U1-70K abundance. Notably, the authors show that U1-70K positively regulates U1C by either stabilizing it at the protein level or promoting its translation. Combined with the known role of U1-70K in stimulating U1C incorporation into the U1 snRNP [21], this provides a potent AS–NMD-based mechanism for regulating U1 snRNP structure and function.

Similar mechanisms link the expression of non-canonical spliceosome proteins U11-48K and U11/U12-65K with the abundance of functional U11/U12 di-snRNP complex [19,20] (Figure 2B). U11 interaction with its cognate sequence element called USSE (U11 snRNP-binding splicing enhancer) allows utilization of upstream canonical spliceosome-dependent 3′ splice sites thus promoting inclusion of the frame-shifting exon 4i in the case of U11-48K or altering the identity of the 3′-terminal exon in the case of U11/U12-65K. Since both of these events reduce mRNA stability, through NMD in the case of U11-48K and a yet-to-be understood mechanism in the case of U11/U12-65K, they provide a homoeostatic feedback regulating overall functionality of the U11/U12 di-snRNP. Two additional lines of evidence hint at the highly regulated modular nature of this circuitry. First, at least one additional U11/U12 di-snRNP component, U11-35K (that is homologous to its canonical U1-70K counterpart), regulates U11/U12-65K
expression in a reciprocal manner, possibly through the USSE-dependent AS switch [19]. Secondly, the activation effect of U11/U12 di-snRNP bound to USSE can be modulated by hnRNPH1/H2 (hnRNP is heterogeneous nuclear ribonucleoprotein) proteins and/or U1 snRNP bound to the adjacent sequence elements [20].

Another relevant example has been reported by Saltzman et al. [22] who examined the network of genes regulated by SmB/B’, a protein subunit of the Sm complex associated with U1, U2, U4, U5 as well as U11, U12 and U4atac snRNAs (Figure 2C). Similar to the auto-regulation mechanisms reported earlier for individual RBPs [10,11], SmB/B’ homeoeostatically controls its own expression by stimulating the inclusion of a PTC-containing exon (exon A). However, an important nuance is that the SmB/B’ regulation appears to depend on changes in snRNA expression levels caused by reduced functionality of the entire Sm complex. Further RNA-seq analyses allowed the authors to uncover hundreds of alternative exons that are regulated by this mechanism. Notably, the affected alternative exons were enriched in genes encoding proteins involved in RNA processing and RNA binding, and many of these exons were predicted to regulate gene expression through NMD [22].

Aside from the spliceosomal core, splicing regulator Rbfox2 has been shown recently to fine-tune homeostatic AS–NMD circuits controlling expression levels of ∼70 distinct RBPs [23]. Similarly, NMD-mediated feedback loops are known to co-ordinate relative abundance of several NMD factors [24,25]. Although molecular details of the latter mechanism are not understood completely, the mRNAs encoding NMD components might be naturally prone to NMD due to their unusual 3’UTR structure and possibly the presence of upstream ORFs in the 5’UTR [25].

**Orchestrating gene expression during development**

The importance of NMD for normal development has been established in earlier genetic studies that uncovered major defects in embryos lacking key NMD factors [26]. Further underscoring its developmental significance, NMD is known to co-operate with AS to control expression of important lineage-specific genes [5,6]. Moreover, the NMD pathway itself is extensively regulated in a cell- and tissue-specific manner [27–31].

An exciting recent advance in this field has been identification of 86 functionally related genes regulated by AS–NMD during mammalian granulocyte development (granulopoiesis) [32] (Figure 3A). The primary transcripts of these genes are spliced completely at an early (promyelocyte) stage, resulting in the accumulation of translation-competent mRNAs and the corresponding protein products. However, at a later (granulocyte) stage, the specific introns are retained within these transcripts thus giving rise to PTC-containing mRNA forms that are cleared by NMD following their export to the cytoplasm. The progressive increase in the retention status of the regulated introns was attributed to their elevated GC-content and diminishing expression of core spliceosomal factors during granulopoiesis. Importantly, the AS–NMD targets are enriched in transcripts encoding proteins modulating the nuclear shape and constitutive expression of an intronless mRNA of at least one of these factors, Lmnb1 (lamin B1), which interfered with normal granulopoiesis [32].

AS–NMD has additionally been shown to contribute to global changes in gene expression during terminal erythropoiesis [33]. Using RNA-seq analysis the authors identified multiple AS events during erythroid lineage development and linked a fraction of these changes with NMD. Interestingly, expression of a large group of genes encoding RBPs and DNA-binding proteins as well as chromatin modifiers appeared to be under AS–NMD control, with the relative abundance of PTC-containing mRNA species increasing at later stages of erythropoiesis. Since these mRNAs were further up-regulated upon treating late erythroblasts with the translational inhibitor cycloheximide.
the authors argued that their developmental dynamics was mediated by corresponding AS changes rather than reduced NMD efficiency. In any case, further studies will be required to elucidate molecular mechanisms underlying this AS–NMD circuitry as well as its contribution to erythroblast differentiation.

Colak et al. [34] identified a novel function of AS–NMD in regulation of Robo3, a modulator of midline crossing by commissural axons (Figure 3B). Two alternative splice forms of Robo3 had been reported earlier to sequentially promote axon attraction to and crossing of the midline by inhibiting Slit-mediated repulsion (Robo3.1 form) and then blocking midline recrossing by stimulating Slit repulsion (Robo3.2 form) [35]. Of the two, Robo3.1 is fully spliced and efficiently translated in pre-crossing axons but undergoes translational repression after midline crossing. Robo3.2 mRNA retains a PTC-containing intron between exons 26 and 27 but does not undergo NMD before crossing since it is kept translationally silent in this compartment. Strikingly, Robo3.2 translation is activated in the post-crossing compartment which ultimately triggers NMD after delivering a localized burst of Robo3.2 protein production. Arguing for the importance of this mechanism for appropriate nervous system development, commissural neurons lacking NMD factor Upf2 or expressing a dominant negative mutant of NMD factor Upf1 develop abnormal axonal trajectories due to Robo3.2 overexpression in the post-crossing segment [34].

This regulation is evocative of the previously published NMD mechanism limiting stability of Arc and potentially other PTC-containing mRNAs in the dendritic compartment of mammalian neurons [36]. However, unlike Robo3, the NMD-promoting features of the Arc mRNA appear to be generated through constitutive splicing. Another important lesson from the study by Colak et al. [34] is that key components of the NMD machinery (Upf1, Upf2 and Smg1) localize to the growth cones of both central and peripheral axons [34] (Figure 3B). Together with the studies reporting net down-regulation of NMD factors (including Upf1 and Smg1) in developing nervous system [29,30], this indicates that NMD might be progressively repurposed to function primarily as a gene regulation rather than an RNA surveillance mechanism in this developmental context.

**Mediating physiological responses**

A previously discussed advantage of AS–NMD as a regulation strategy is that, at least theoretically, it can be used to rapidly change gene expression outputs, e.g. through post-translational modifications or changes in intracellular localization of AS or NMD factors. Given that both AS and NMD are extensively regulated by environmental cues [2,27,28,37], this would provide an ideal mechanism for adjusting gene expression output in response to changing environment. Recent studies begin providing some experimental support for this interesting prediction.
In one such study, Darnell and co-workers [38] used a combination of exon microarrays and HTS-CLIP (high-throughput sequencing of RNA isolated by cross-linking immunoprecipitation) to uncover >200 mRNAs encoding proteins enriched in synaptic functions that significantly changed their steady-state levels in mouse brain lacking the RBPs Nova1 and Nova2 (Figure 3C). At least a subset of these transcripts was regulated through Nova-stimulated or Nova-repressed alternative exons promoting NMD. Importantly, this regulation was partially recapitulated in the wild-type mouse brain in response to pilocarpine-induced seizures, possibly as a result of altered subcellular localization of the Nova protein. These data, along with the fact that Nova-haploinsufficient mice spontaneously developed epileptic symptoms, led the authors to propose that Nova-controlled AS–NMD events might mediate homeostatic plasticity of the brain in response to seizures.

Similarly, recent global transcriptome analyses in Arabidopsis thaliana uncovered a large number of AS transcripts up-regulated following the genetic ablation of the NMD factors UPF1 and/or UPF3 [39,40] (Figure 3D). Interestingly, some of these transcripts corresponded to known stress response genes [40] and relative expression of a subset of NMD-promoting AS variants increased in response to salt-induced stress [39] hinting at the importance of AS–NMD regulation in this physiological context [41]. Although additional work will be required to better characterize AS–NMD circuitries involved in stress response in both animals and plants, these studies appear to open a new chapter in our understanding of how cells and organisms respond to external cues.

**Future outlook**

In conclusion, the current literature corroborates the increasingly prevalent view that, in addition to its relevance to RNA surveillance, AS–NMD provides an important means for regulating gene expression programmes in multiple biological contexts. What transpires is that this molecular strategy is particularly suitable for co-ordinating expression outputs of multiple functionally linked genes, a recurring feature in complex biological processes, be it cellular RNA homeostasis, development or a stress response.

As with many other areas of biology, progress in the AS–NMD field has been dramatically accelerated by recent advances in the next-generation sequencing that allowed interrogating transcriptome compositions and RBP specificities at an individual nucleotide level [42,43]. Further use of these approaches and their improved derivatives, including single-cell resolved analyses [44], should result in the discovery of new AS–NMD circuitries and better description of the already known ones. Knockout and RNAi studies have been exceptionally informative for elucidating gene circuitries regulated by specific AS regulators and NMD components (e.g. [12,13,38–40,45–48]) and this work will almost certainly be extended in the future using recently introduced genome manipulation tools [49].

All in all, we predict that further investigation of the interface between regulated pre-mRNA processing and cellular RNA quality controlled mechanisms will continue providing valuable molecular insights into normal development and physiology as well as biological processes associated with disease for years to come.

**Acknowledgements**

We thank Snezhka Oliferenko for discussion and apologize to our colleagues whose work we could not mention owing to space limitations.

**Funding**

This work was supported by the National Medical Research Council [grant number NMRC/CBRG/0028/2013 (to E.V.M.)] and Nanyang President Graduate Scholarship (to F.M.H.).

**References**


