Available online at www.sciencedirect.com
ScienceDirect
Biochemical and Biophysical Research Communications 353 (2007) 487–492
BBRC
www.elsevier.com/locate/ybbrc

A novel function for alternative polyadenylation as a rescue pathway from NMD surveillance

Roi Gilat, Dorit Shweiki *

Bioinformatics Program, School of Computer Science, The Academic College of Tel Aviv-Yaffo, 4 Antokolsky St., Tel-Aviv 64044, Israel

Received 5 December 2006
Available online 15 December 2006

Abstract

Premature termination codon (PTC) containing transcripts are subjected to a rapid degradation via nonsense-mediated decay (NMD) surveillance mechanism. By and large degradation is desired in order to prevent the translation of truncated, most likely deleterious, protein. Nevertheless, several dissimilar NMD-rescue events, capable of turning NMD-candidates into NMD-immune, are described. Yet, the extent and nature of this phenomenon is unknown. We screened the human genome for NMD-candidates transcripts. Among which we sub-grouped "pseudo-NMD" genes, which all their annotated transcripts contain PTCs, and therefore allegedly are transcribed but never translated. Here we show that alternative polyadenylation can rescue prematurely terminated transcripts, by truncating the pre-mRNA so that the PTC is now "legally" positioned. ESTs-based analysis shows that NMD-rescued genes are indeed expressed in human tissues. Furthermore, predicted NMD-rescue variants' existence is computationally verified. Hence, we suggest a novel role for the exon-truncated class of alternative polyadenylation as an NMD-rescue regulatory mechanism.

Keywords: Nonsense-mediated mRNA decay; PTC-containing transcripts; Alternative polyadenylation; NMD-rescue mechanism

In mammals, prematurely translation-terminated mRNAs are subjected to mRNA degradation via nonsense-mediated mRNA decay (NMD) mechanism [1–3]. Premature termination codon (PTC) is identified as such if it positioned more than 50–55 nucleotides upstream of the terminal exon–exon junction in the mRNA [2,4]. Exon–exon junctions in spliced mRNAs are "marked" by exon junction protein complex (EJC), which is bound to the mRNA 20 to 24 nucleotides upstream to the junction [2,5]. A pioneer round of translation is responsible for the removal of the EJC from mRNAs which obey the stop codon position basic-rule, so that the mRNA molecule becomes immune to NMD. Un-removed EJCs, in prematurely translation-terminated mRNAs, are the trigger for Upf1-mediated mRNA degradation [1,6].

NMD was shown to be associated with several human genetic disorders either by preventing a translation of deleterious truncated protein, or alternatively by causing a loss-of-function like phenotype due to the reduction in mRNA levels [7,8]. Mutation-related, PTC-containing transcripts are not always eliminated by the NMD machinery. Mutations in an intronless gene, in a spliced gene in a proximity to the exon–exon junction in penultimate exon, or in the terminal exon, are all scenarios in which a truncated protein, functionally altered is expected to be synthesized.

Moreover, PTC-containing transcripts are not necessarily the outcome of genomic mutations. One-third of alternatively spliced transcripts are prematurely translation-terminated, therefore are target for NMD surveillance mechanism [9,10]. Upstream open reading frames (uORF) in the 5′ untranslated region (UTR) of mRNAs are known to regulate eukaryotic protein synthesis levels via the activation of NMD-mediated degradation [11,12]. Apolipoprotein B mRNA (apoB100) undergoes C to U RNA editing, to

* Corresponding author. Fax: +972 3 5211871.
E-mail address: dorits@mta.ac.il (D. Shweiki).

0006-291X/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2006.12.052

produce a PTC-containing mRNA which does not undergo NMD. Rather it is translated to produce a truncated functional apoB48 protein. It was shown that the apoB mRNA editing complex is responsible for the protection of the edited mRNA from NMD machinery, and for its export to the cytoplasm [13,14]. Rescue of "legitimate" PTC-containing mRNAs from NMD degradation is not limited to RNA editing regulation. The uORF-containing thrombopoietin mRNA is also known to undergo rescue from NMD degradation [15].

An intriguing possibility is that rescue from NMD degradation is not a sporadic event rather it is a well-coordinated regulatory mechanism/s embedded at the post-transcription trail. Thus, the extent to which NMD rescue events are physiologically expected or required should be assessed. One way of inferring on this matter is by evaluating the amount of "legitimate" PTCs encoded in the human genome. In this study, we aimed to screen the human genome for NMD-candidate transcripts, to characterize and sub-group them. Further, and based on the large amount of legitimate PTC-containing transcripts that we identified, we addressed the question of NMD-rescue mechanism.

Our lead candidate for NMD-rescue mechanism was alternative polyadenylation, a key regulatory process affecting transcripts' 3′-end formation. The 3′ UTRs of eukaryotic mRNAs contain valuable regulatory information, vastly affecting mRNA translation levels, mRNA stability, and mRNA cellular localization [16,17]. Thus, alternative 3′-end determination may significantly alter transcripts' cellular behaviour and characteristics, and consequently genes expression pattern.

The 3′-end formation involves the cleavage and polyadenylation of the pre-mRNA transcript downstream to a polyadenylation signal sequence. Two canonical signals (AAUAAA or AUUAAA) and additional nine alternative hexamers were described [18]. The recognition of polyadenylation signals was suggested to involve downstream elements as well as secondary structures formation in the pre-mRNA. Furthermore, several polyadenylation upstream enhancers were identified in viral and eukaryotic genes, functioning in a distance-dependent manner [19,20].

A significant fraction of pre-mRNAs was reported to be subjected to differential polyadenylation events, mostly in a tissue and a process specific manner [21,22]. Thus, and along with alternative polyadenylation main feature to alter transcripts' fate, the plausibility that this regulatory mechanism may play additional role in gene expression regulation was strengthen. In this study, we suggest an NMD-rescue function for alternative polyadenylation mechanism. Rescue is accomplished by truncating the pre-mRNA so that the PTC is now "legally" positioned. We identified a "pseudo-NMD" sub-group of genes, in which NMD-rescue is particularly expected to occur, due to the fact that all their annotated transcripts contain PTCs, and therefore are allegedly transcribed but never translated. By utilizing ESTs-based gene expression analysis we show that NMD-rescued genes are indeed expressed in human tissues. Furthermore, we verified the existence of the predicted NMD-rescue variants, utilizing poly(A)/poly(T) tailed ESTs.

Results

The majority of coding sequences (CDS) ends at the terminal exon

Human RefSeq transcripts were classified as non-NMD or NMD-candidates according to the annotated stop codon position (as elaborated in Methods). Out of the 18167 transcripts that were analyzed (constitute 14343 genes), in 16773 transcripts (92.3%; 13551 genes) the stop codon is positioned in the terminal exon, thus these transcripts are considered NMD-immune; this stand in line with a previous report by Nagy et al. [3]. In 834 transcripts (4.6%; 678 genes; 151 genes overlap with the previous group described) the stop codon is located within the range of 55 nucleotides from the last exon–exon junction; and 560 sequences (3.1%; 464 genes; 199 genes overlap with the non-NMD transcripts) were identified as PTC-containing transcripts. Namely, the annotated stop codon is positioned in internal exon or in distance greater than 55 nucleotide from the last exon–exon junction (Table 1).

"Pseudo-NMD" sub-group are candidates for NMD-rescue

From the NMD-candidates group we isolated a "pseudo-NMD" sub-group, which holds a built-in paradox, as all of their annotated transcripts are subjected to NMD degradation. The 242 "pseudo-NMD" genes (270 transcripts) are transcribed, but at least allegedly, never translated to produce a protein. Though it is possible that these genes undergo an un-annotated regulation (e.g., alternative splicing, RNA editing) to produce a non-NMD transcript, it is equally likely that at least some undergo a yet unreported NMD-rescue event. Thus, "pseudo-NMD" genes may serve as an excellent training-set for the identification of NMD-rescue mechanisms.

Alternative polyadenylation may rescue NMD-candidates transcripts

In order to test the plausibility of alternative polyadenylation mediated NMD-rescue, we analyzed NMD-candidates transcripts for the existence of alternative polyadenylation signals and for their NMD-rescue potential. Table 2 summarizes the number of predicted polyadenylation signals and of transcripts identified to contain predicted signals, in the NMD-candidate groups. In the general NMD-candidate population, out of a total of 560 transcripts, 188 transcripts contained 371 predicted polyadenylation signals. In the "pseudo-NMD" sub-group, out of a total of 270 transcripts, 90 transcripts contained 162 predicted polyadenylation signals. Namely, the ratio of