Novel examples of NMD escape through alternative intronic polyadenylation

Maria Vlasenok[1], Antonina Kuznetsova[1], Dmitry A. Skvortsov[1,2], and Dmitri D. Pervouchine[1,*]

[1]Center for Molecular and Cellular Biology, Moscow 143026, Russia

[2]Moscow State University, Moscow 119991, Russia

[*]e-mail: pervouchine@gmail.com; tel/fax +7 (916) 220 69 99

August 27, 2025

Abstract

The nonsense-mediated mRNA decay (NMD) surveillance system detects and selectively degrades transcripts with premature stop codons (PTC). A stop codon is recognized as premature if it is followed by an exon-exon junction in >50 nucleotides downstream. Pruning of the 3'-untranslated region containing such junctions by alternative polyadenylation (APA) could provide a mechanism of NMD escape. Here, we systematically examine exons that carry a PTC (poison exons) for the presence of polyadenylation sites (PAS) in

the downstream intron. Using the GTEx consortium data, we found that poison exons are followed by an active PAS more frequently than cassette exons, and identified tissue-specific switches between NMD-target and NMD-escape isoforms in a number of human genes. Among them are the tropomyosin 2 (*TPM2*) gene switching between NMD-target and NMD-escape isoforms in skeletal muscle and heart, and the vaccinia-related kinase 3 (*VRK3*) with a similar kidney-specific switch. Blocking the polyadenylation signal and the cleavage site in *VRK3* by antisense oligonucleotides in the A549 cell line led to a switch from NMD-escape to NMD-target isoform that was accompanied by a drop in the expression level. This study suggests that NMD escape via alternative polyadenylation could be a widespread and currently overlooked post-transcriptional regulatory mechanism.

Keywords: NMD escape, splicing, polyadenylation, *TPM2*, *VRK3*.

Introduction

Alternative splicing (AS) and alternative polyadenylation (APA) are two major steps of post-transcriptional pre-mRNA processing that substantially alter the aminoacid sequence of the gene product [1, 2, 3]. While AS operates by selecting different exon combinations, APA involves the use of different polyadenylation sites (PAS) resulting not only in alterations of the 3'-untranslated regions (3'UTRs) but also in the generation of proteins with different C-termini. The coordination between AS and APA is important in many biological processes including development [4, 5], tumorigenesis [6], and functioning of the nervous system [7].

Interestingly, both AS and APA carry out important roles beyond proteome diversification. APA significantly impacts mRNA stability, localization and translation through modulating the

3'UTR sequences, which contain cis-regulatory elements of microRNAs and RNA-binding proteins [8, 9]. AS is involved in the so-called unproductive splicing, a process in which a premature termination codon (PTC) is purposely introduced into the mRNA to trigger its degradation by the nonsense-mediated decay (NMD) pathway [10]. This way of regulation is particularly widespread among RNA-binding proteins [11]. Curiously, population genomics studies conclude that AS exerts greater impact on NMD-induced changes in gene expression than on expanding protein diversity [12].

One of the current NMD models postulates that exon-exon junction complexes (EJCs) that are deposited on the mRNA during splicing are displaced after the first round of translation, while EJCs that remain bound outside of the reading frame serve as a signal that mRNA contains a PTC [13]. Yet, PTC-containing transcripts can evade degradation by NMD in a process called NMD escape [14, 15]. Several stress response pathways inhibit NMD by affecting the translation efficiency [16]. Inefficient translation termination at the PTC (readthrough), or reinitiation of translation downstream of the PTC, evicts downstream EJCs also leading to NMD escape, for instance in *ATRX* and *ATP7A* genes [17, 18]. Some RNA-binding proteins such as PABPC1, PTBP1 and hnRNPL are capable of coordinating with NMD factors to protect specific transcripts from degradation [19].

Interestingly, AS and APA can act synergistically to trigger NMD escape. If an exon containing a PTC (poison exon) is skipped, or a sole intron downstream of a PTC is inefficiently spliced, then the mRNA will no longer undergo NMD [14]. The same will happen if a PTC is followed by a PAS located within the downstream intron because after cleavage and polyadenylation there will be no EJCs downstream of the PTC, and it will be perceived as a normal stop

codon. Examples of NMD escape through APA include translation of the truncated form of the *Tau* mRNA in Alzheimer's disease (AD) patients [20] and the *HFE* gene, which has a premature PAS within the last post-PTC intron [21]. Another example is a newfound cryptic poison exon in the *CSF3R* gene, in which the activation of a downstream intronic PAS allows the transcript to evade NMD [22].

NMD serves not only as an mRNA quality control system but also as a regulatory mechanism to maintain the abundances of many physiological transcripts [23]. This motivated us to investigate whether NMD escape through alternative polyadenylation could represent a more widespread regulatory mechanism than it is believed currently. To address this question, we juxtaposed the existing catalog of poison exons with the published PAS datasets and GTEx transcriptomic data to detect tissue-specific switches to the NMD isoform and simultaneous activation of an intronic PAS, a pattern that is expected for the proposed mechanism. As a proof of principle, we conducted an experiment in human A549 cells, in which the inhibition of intronic polyadenylation signal and the cleavage site by antisense oligonucleotides (AON) led to the suppression of the APA isoform and upregulation of the NMD isoform in the *VRK3* gene.

Methods

Genome assembly, transcript annotation, and PAS

The GRCh38 assembly of the human genome and comprehensive CHESS transcript annotation v3.1.3 were downloaded from Genome Reference Consortium [24] and CHESS websites [25], respectively. The lists of PAS were obtained from the PolyASite 2.0 database [26] and from

a study of short reads containing non-templated adenines (polyA reads) [27]. In the latter, the genomic coordinates of PAS from GTEx were lifted over to the GRCh38 assembly; the TCGA dataset was processed using the GRCh38 assembly [27]. The two PAS sets were merged, and polyA read support values from GTEx and TCGA datasets were added.

RNA-seq data

The RNA-seq data from 9,423 samples of the GTEx project v8 were obtained from the dbGaP portal and aligned to the human genome assembly GRCh38 (hg38) using STAR aligner v2.7.8a in paired-end mode [28]. The results of NMD inactivation experiments in human 293T cells, iPS-bronchioalveolar organoid, HeLa and A549 cells followed by RNA-seq were obtained from the Gene Expression Omnibus under the accession numbers GSE126520, GSE217099, GSE125086 and GSE270310, respectively. The data were downloaded in FASTQ format and aligned to the GRCh38 (hg38) assembly using STAR aligner v2.7.8a in paired-end mode [28]. Sequencing results of transcripts bound to EJCs in HEK293 cells (RIPiT protocol [29]) together with corresponding control RNA-seq datasets were downloaded from GEO under accession number GSE115788 and processed in the same way.

Poison and cassette exons

A cassette exon was defined as an exon that is included in a group of annotated protein-coding or NMD transcripts, and skipped in another such group. Exons that share an upstream or downstream splice junctions with another non-terminal exon were not considered. Thus, each cassette exon was characterized by the uniquely defined upstream, downstream and skipping splice

junctions (I_1, I_2, and E, respectively, in Figure 1A). Poison exons were defined as cassette exons that contain a stop codon of an annotated NMD isoform 50 or more nts upstream of exon's 3'-end (the 50-nt rule). Out of 12,528 cassette exons, 1388 contained a stop codon, and of those 813 obeyed the 50-nt rule.

Read coverage and gene expression levels

The read coverage (BPM) values were calculated genomewide for each GTEx sample using bamCoverage utility [30] considering only uniquely mapped reads. The read coverage values were averaged across samples within each detailed tissue (SMTSD) using wiggletools utility [31]. These values were used for visualization using the Gviz package [32]. The mean read coverage was computed for 150-nt windows at the end and at the beginning of the upstream and the downstream constitutive exons flanking each cassette exon, respectively (referred to as e_1 and e_2). If the upstream or downstream exon was shorter than the window size, the mean was computed for its genomic coordinates. The $\delta = \frac{e_2}{e_1}$ metric was used only when $e_1 > 0.2$. Gene expression values (counts) were computed for each GTEX sample based on the primary GENCODE v43 annotation downloaded from the EMBL-EBI portal with featureCounts utility [33]. The counts were normalized to TPM, and the resulting expression levels were averaged between samples within each detailed tissue type (SMTSD). In what follows, the term 'tissue' refers to SMTSD.

Splicing metrics

To quantify the inclusion level of poison and cassette exons in NMD inactivation experiments, we computed the split read counts using the IPSA pipeline with the default settings [34]. For the GTEx data, we used split read counts computed earlier [35]. The counts of split reads were pooled within each tissue type to characterize each exon with $\psi_1 = \frac{I_1}{I_1+E}$, $\psi_2 = \frac{I_2}{I_2+E}$, and $\varphi = \frac{I_1}{I_1+I_2}$ metrics, where I_1 and I_2 are the number of split reads supporting splice junctions with the upstream exon and the downstream exons, respectively, and E is the numbers of split reads supporting exon skip (Figure 1A). The values of ψ_i with the denominator below 10 were discarded as unreliable. The counts of continuous short reads overlapping the splice sites (r_1, r_2) were computed via the same pipeline and analyzed using the same metrics. To assess splicing changes, we used $\Delta\varphi = \varphi_{exp} - \varphi_{ctrl}$ and $\Delta\psi_2 = \varphi_{2,exp} - \varphi_{2,ctrl}$, where 'exp' and 'ctrl' refer to EJC vs. RNA-seq experiments for RIPiT, and CHX treatment vs. untreated cells for the NMD inactivation experiments.

Statistical analysis

The data were analyzed using R statistics software version 3.6.3. One-sided non-parametric tests were performed using normal approximation with continuity correction. For proportion comparisons, a hypergeometric test was used with Poisson or normal approximations when appropriate. P-values are denoted by P throughout the paper. In all figures, the significance levels of 5%, 1% and 0.1% are denoted by *, ** and ***, respectively; whiskers denote 95% confidence intervals.

Antisense oligonucleotides

AONs with 2'O methyl modified bases on a phosphorothioate backbone (2'OMe) were designed based on recommendations from [36]. Synthesis of AONs was carried out by Syntol JSC (Moscow, Russia). The sequences of used AONs are listed in Table S1.

Cell cultures, transfections and treatments

Human A549 lung adenocarcinoma cells were maintained in Dulbecco's modified Eagle's medium/Nutrient Mixture F-12 with 10% v/v fetal bovine serum (FBS), 1% GlutaMAX (Thermo Fisher Scientific), 0.05 mg/ml streptomycin and 50 units/ml penicillin (all products from Thermo Fisher Scientific) at 37°C in 5% CO_2. Cells were plated at a density of 120,000 cells per well in a 12-well plate.

AONs treatment was performed using Lipofectamine RNAiMAX (Invitrogen) on 50–70% confluent cells. Cells were harvested after 24h of treatment. To inactivate NMD, cycloheximide (CHX) was added to the cells 3h before harvest giving a final concentration of 300 mg/ml in the growth medium. Experiments were made in at least five independent biological replicates.

RNA purification and cDNA synthesis

Total RNA was isolated by a guanidinium thiocyanate-phenol-chloroform method using ExtractRNA Reagent (Evrogene) in accordance with the manufacturer's protocol. One microgram of total RNA was first subjected to RNase-free DNase I digestion (Thermo Fisher Scientific) at 37°C for 30 min to remove contaminating genomic DNA. Next, 500 ng of total RNA were used for complementary DNA (cDNA) synthesis using Magnus First Strand cDNA Synthesis

Kit (Evrogene) for reverse transcription-quantitative PCR (RT-qPCR) to a final volume of 20 µl. cDNA was diluted 1:5 with nuclease-free water for quantitative PCR (qPCR) and endpoint PCR.

Endpoint PCR and qPCR

qPCR reactions were performed in triplicates in a final volume of 12 µl in 96-well plates with 420 nM gene-specific primers and 2 µl of cDNA using 5XqPCRmix-HS SYBR reaction mix (Evrogen). Primers for PCR and qPCR are listed in Table S2; their annealing positions are shown in Figure S8A. A sample without reverse transcriptase enzyme was included as a control to verify the absence of genomic DNA contamination. Amplification of the targets was carried out on CFX96 Real-Time System (Bio-Rad), with the following parameters: 95°C for 5 min, followed by 39 cycles at 95°C for 20 s, 58°C for 20 s and 72°C for 20 s, ending at 72°C for 5 min. For each pair of primers in qPCR analysis, the primer efficiencies were estimated using a calibration curve. The expression of isoforms was calculated by the efficiency method with primers efficiency more than 90% in all cases.

Endpoint PCR was carried out under the following conditions: denaturing at 95°C for 3 min, 35 cycles of denaturing at 95°C for 30 s, annealing at 54°C for 30 s, and extension at 72°C for 30 s, followed by extension at 72°C for 5 min. The resulting products were analyzed on a 2% agarose gel stained with ethidium bromide and visualized using ChemiDoc XRS+ (Bio-Rad).

Results

Poison exons are often alternative terminal exons

When a PTC is followed by a PAS in the downstream intron, it will be interpreted as a regular stop codon by the NMD machinery. To assess the extent to which these situations occur, we assembled a catalog of poison exons, defined as cassette exons that carry a PTC with respect to the coding frame of the major transcript isoform (a broader definition also includes generation of a PTC by a frameshift, but it is not considered here) and obey the so-called 50-nt rule, which postulates the existence of at least one EJC >50 nts downstream of the PTC [37]. Thus, genes with poison exons (Figure 1A) have a protein-coding isoform, in which the poison exon is skipped (skip), and a NMD isoform, which includes the poison exon (NMD target). The occurrence of an intronic PAS downstream introduces the third possibility, a truncated transcript, where the extended poison exon becomes an alternative terminal exon, and PTC becomes a normal stop codon (NMD escape).

The list of poison exons and the reference list of protein-coding cassette exons without a PTC were compiled from the CHESS annotation database [38]. Interestingly, 64 out 814 (7.9%) unique poison exons in the initial list contained an annotated transcript end in the downstream intron, while only 325 out of 12641 (2.6%) non-poison cassette exons did so ($P \approx 10^{-14}$). Next, we performed a similar analysis for intronic PAS from the PolyASite 2.0 database [26] and PAS identified from short reads with non-templated adenines (polyA reads) that were predicted for GTEx and TCGA datasets earlier [27]. For PAS with substantial read support (≥50 polyA reads), we found that 191 out of 814 poison exons (23.5%) were associated with a PAS in the

downstream intron (Figure 1B), while only 2487 out of 12,641 cassette exons (19.7%) did so ($P = 0.005$). To reduce the level of noise in polyA read data, we imposed stricter thresholds on the PAS support level and found an even greater prevalence of PAS after poison exons, especially for PAS preceded by the canonical polyadenylation signal (Figure 1C left, Figure S1A,B). A similar trend was seen when strengthening the threshold for PAS from the PolyASite database (Figure 1C, middle). Overall, the odds ratio computed for different data sources unambiguously suggested that transcription termination happens more frequently after poison exons than after non-poison cassette exons (Figure 1C, right).

If a transcript terminates within or immediately after the poison exon, the counts of split reads supporting its splice junctions at the 5'- and the 3'-end must be distributed differently (I_1 and I_2 in Figure 1A). To characterize this difference, we used the metric $\varphi = \frac{I_1}{I_1+I_2}$. The value $\varphi \approx 0$ indicates that $I_2 \gg I_1$, $\varphi \approx 1$ indicates that $I_1 \gg I_2$, while in most cases we expect $I_1 \approx I_2$ and $\varphi \approx 0.5$. In applying this metric to human tissue transcriptomes from the GTEx project (Figure 1D, left), we expectedly found $\varphi \approx 0.5$ for the majority of cassette exons, while for some of them the support from either end was missing ($\varphi = 0$ or 1). A pattern remarkably different from this and from that of poison exons without a PAS downstream (Figure 1D, middle) was observed for poison exons followed by a PAS, where the distribution had a pronounced peak at $\varphi \approx 1$ indicating a lack of I_2 reads, especially for poison exons followed by a high-confidence, annotated transcript ends (Figure 1D, right). The prevalence of I_1 over I_2 also grew with increasing PAS support (Figure S1C). Repeating the same comparison for continuous reads overlapping splice sites revealed more frequent retention at the 3'-end for poison exons followed by a PAS, further supporting their transformation into alternative terminal exons (Figure S2).

Taken together, these observations indicate that poison exons are frequently followed by a PAS in the downstream intron and tend to be extended at the 3'-end being converted to alternative terminal exons.

Tissue-specific switch between NMD-target and NMD-escape isoforms

To identify genes that evade NMD through intronic polyadenylation in a tissue-specific manner, we analyzed the panel of RNA-seq experiments provided by the GTEx project, focusing on switching between transcript isoforms with skipped poison exons (skip), included poison exons (NMD target), and truncated isoforms with poison exon being a part of an alternative terminal exon (NMD escape).

A widely used metric to quantify exon inclusion, percent-spliced-in (PSI, ψ), combines split read counts from both sides of the exon ($\psi = \frac{I_1+I_2}{I_1+I_2+2E}$). Therefore, it is not optimal for distinguishing between the NMD-target and NMD-escape isoforms. Following the logic of intron-centric estimation of AS [34], we considered two separate metrics, $\psi_1 = \frac{I_1}{I_1+E}$ and $\psi_2 = \frac{I_2}{I_2+E}$ (Figure 2A). Since the NMD-target and NMD-escape isoforms share the splice junction upstream of the poison exon, ψ_1 measures their combined inclusion ratio, while ψ_2 represents the inclusion of the poison exon alone. However, since NMD-sensitive isoforms are degraded in tissues as evidenced by low ψ_2 values (Figure S3A), another metric is needed to discriminate between NMD-target and NMD-escape isoforms.

Towards this goal, we introduced the metric $\delta = \frac{e_2}{e_1}$, which is a proxy for measuring the polyadenylation rate as the read coverage drop between the upstream and the downstream constitutive exons (Figure 2A). The value $\delta \approx 0$ indicates a prevalence of intronic polyadenylation,

while $\delta \approx 1$ indicates its lack. Note that δ can take arbitrary non-negative values, while ψ_1 and φ are restricted to the interval from 0 to 1.

In the proposed mechanism, APA converts NMD-target isoforms to NMD-escape isoforms. However, since the former are degraded, the observable signature of NMD escape must be a switch from the skip isoform to the NMD-escape isoform, both of which are stable. Such switches are manifested by anticorrelated tissue-specific changes in ψ_1 and δ, where ψ_1 increases, indicating splicing at the exon's 5'-end, while δ decreases, indicating termination at the intronic PAS. This logic is illustrated in Figure 2B.

In order to define the δ values that would robustly indicate an APA event, we surveyed the distribution of δ values across tissues for cassette exons followed by an annotated transcript end (Figure S3C). To find poison exons that are a part of the truncated transcript in at least one tissue, we selected exons that are characterized by $\delta < 0.6$, the value that separates the two modes in the δ distribution, which yielded 52 exons (Figure S3C). On the other hand, 76 out of 191 poison exons showed tissue-specific alternative splicing, as reflected by a large ψ_1 range (Figure S3B). The overlap between these sets comprised 23 exons (Figure 2C). Imposing the additional constraint that ψ_1 and δ be negatively correlated yielded 14 exons that represent potential NMD-escape events. Their characteristics including the associated diseases (according to MalaCards [39]) are listed in Table 1; the coordinates and tissue-specificity are shown in Figure S4.

The selected examples include intronic polyadenylation of the poison exon 9 of the Tropomyosin 2 (*TPM2*) gene (Figure 3A). This famous gene encodes β-tropomyosin, which stabilizes actin filaments and regulates muscle contraction. It is implicated in cardiac diseases, my-

opathies, and atherosclerosis [40, 41]. According to our findings, transcript isoforms of *TPM2* that skip the poison exon are dominant in cervix and prostate, while NMD-escape isoform with an alternative terminal exon dominates in muscle, in agreement with the high level of expression of *TPM2* in this tissue (Figure 3A, S5). Remarkably, in heart, we observe upregulation of the NMD-target isoform, which potentially explains how cardiac expression of *TPM2* is suppressed [42].

Another interesting example is the gene encoding vaccinia-related kinase 3, *VRK3*. The kinase activity of *VRK3* is involved in the regulation of the cell cycle [43]. *VRK3* has an additional role in preventing neuronal apoptosis caused by prolonged oxidative stress, where CDK5 activates neuroprotective signaling via *VRK3* phosphorylation [44]. Our analysis shows that about half of the *VRK3* transcripts undergo NMD escape in cervix and esophagus, which is evident from a lower read coverage in the next constitutive exon, the read coverage drop at PAS and low I_2 values (Figure 3B). On the other hand, in adrenal glands and kidneys, the exon skipping isoform prevails ($\psi_1 = 0.22$ and $\psi_1 = 0.25$, respectively), and the read coverage doesn't drop in the downstream exon. This model potentially explains increased *VRK3* production in kidneys [42].

Interestingly, *VRK3* has different intronic polyadenylation rates in brain tissues but doesn't have variability in ψ_1 values. We discovered that another PAS upstream of the *VRK3* poison exon is activated in brain tissues (Figure S6). Thus, the *VRK3* gene may have a distinct mechanism of NMD escape in the brain, consistent with its separate function in neural tissues [44].

Response of NMD escape events to NMD inactivation

In order to confirm the proposed mechanism experimentally, we chose a strategy of inhibiting the production of the truncated isoform by antisense oligonucleotides (AON) that block the polyadenylation signal and the cleavage site. In response, we expect a switch to the production of the NMD-target isoform and, consequently, a lower gene expression level. However, the isoform with a poison exon can be observed at detectable levels only when the NMD system is inactivated. Therefore, we reanalyzed previously conducted RNA-seq experiments, with and without NMD inactivation by the translation inhibitor cycloheximide (CHX).

We are interested in candidates that express both NMD-escape and NMD-target isoforms containing poison exons (Figure 2D). The latter were selected based on $\Delta\psi_2$ (the change of ψ_2 under CHX treatment) with larger $\Delta\psi_2$ suggesting higher expression of the NMD-target isoform. Additionally, a decrease in the φ metric ($\Delta\varphi$) under the CHX treatment was required to capture an increase in the relative abundance of NMD-target isoforms among NMD-target and NMD-escape isoforms combined. A high φ value or low δ value in the control experiment without CHX were used as an indicator that NMD-escape isoforms are sufficiently abundant when the NMD system is active (polyA$_{\text{ctrl}}$ in Figure 2D).

To select reasonable thresholds, we analyzed the set of cassette exons containing a stop codon followed by a PAS, which included both poison exons and non-poison exons with stop codons not obeying the 50-nt rule. The distribution of φ had a pronounced peak at $\varphi = 1$ that was accompanied by a notable increase at $\varphi = 0.5$ upon CHX treatment for poison exons (Figures S7A). The $\Delta\psi_2$ distribution was skewed towards positive values indicating that the inclusion rate of NMD-target isoform also increases (Figure S7B). These observations confirm

that the relative abundance of the NMD-target isoform with respect to both NMD-escape and skip isoforms increases under NMD inhibition.

However, CHX treatment blocks the elongation phase of eukaryotic translation, thereby affecting other pathways besides NMD. An alternative approach to detect NMD-sensitive transcripts is to pull down transcripts bound to EJCs using the RIPiT protocol, which isolates post-splicing, pre-translational mRNA-EJC particles [29]. To increase the power of detecting NMD-target transcripts, we used the same approach to analyze the dataset obtained with RIPiT in HEK293 cells [29]. Although the RIPiT dataset yielded fewer genes expressing both NMD-escape and NMD-target isoforms, the results showed substantial overlap between experiment types (Figure S7C). For example, 32 of 47 poison exons with higher inclusion in the EJC-bound fraction also displayed increased inclusion upon CHX treatment ($P < 10^{-9}$). The same trend was observed for the other two characteristics. While the resulting lists intersected only by two poison exons, the overlap was statistically significant ($P = 0.027$). These findings support the robustness of our observations across different experimental approaches.

By combining the lists obtained for CHX and RIPiT protocols with the results of tissue-specific analysis, we selected five genes expressing both the NMD-escape and NMD-target isoforms in cell lines and demonstrating a tissue-specific switch between them (*VRK3*, *TPM2*, *SYNGAP1*, *BDP1*, and *NMGN1*). Due to technical difficulties in assessing the polyadenylated isoform in other genes, only the *VRK3* gene was chosen as an example to validate the mechanism of NMD escape through intronic polyadenylation.

Polyadenylation inhibition promotes NMD isoform expression

To inhibit intronic polyadenylation, we designed AON sequences that disrupt the polyadenylation signal and the cleavage site in *VRK3* by complementary base pairing (Figure 4A). The abundances of all three isoforms were assessed with, without, and with a mock AON treatments, and with and without NMD inhibition by CHX in human adenocarcinoma A549 cells.

RT-PCR analysis revealed that the transfection of 25 nM of the AON blocking the cleavage and polyadenylation site alone, or in combination with the second AON blocking the AAUAAA signal efficiently induced the inclusion of the poison exon and suppressed the production of NMD-escape isoform, while the transfection of the control AON blocking an off-target site didn't lead to a noticeable difference from the untreated cells. Moreover, upon NMD inhibition with CHX, the observed effects were greater (Figure 4B).

Quantitative RT-PCR (qRT-PCR) with isoform-specific primers confirmed an increase and a reduction in relative abundances of NMD-target and NMD-escape isoforms, respectively, in response to AON treatment (Figure 4C). As in endpoint PCR, the effects were more significant in cells treated with CHX. Notably, the treatment with AON masking the PAS was more effective than with AON targeting the polyadenylation signal. The measurement of *VRK3* mRNA levels showed a drop in the expression of the dominant NMD-escape isoform and, consequently, of the gene expression level upon treatment with AON mix when cells were not treated with CHX (Figure 4D and Figure S8B). This is an expected result since mRNA degradation occurs more effectively in the conditions when the NMD system is not inactivated.

In sum, these results confirm that the fraction of transcripts of *VRK3* gene degraded via NMD pathway can be increased by inhibition of the premature polyadenylation at a proximal

downstream PAS.

Discussion

Current estimates suggest that at least one-third of alternatively spliced transcripts contain a PTC and may be targeted for degradation by the NMD pathway [45]. However, whether all of these transcript isoforms or which fraction of them are actually degraded remains an open question. A study of human genotyping datasets found that 52.3% of highly common nonsense variants escape NMD, mostly via mRNA isoform switching [46]. Another study showed that the variance in NMD efficiency can only partly be explained by the 50-nt rule, and >30% of NMD variability is still unexplained [47], suggesting that NMD escape could be much more widespread than it is currently appreciated.

In this work, we focused on the mechanism, in which APA removes the part of the 3'-UTR that signals for the presence of a PTC, thus converting it into a normal stop codon. Only a few genes were previously suggested to undergo NMD escape via this mechanism. By using a number of metrics including φ, which captures the imbalance between 5'- and 3'- splice junctions, and δ, which reflects the drop in read coverage due to APA, we demonstrated a tendency for poison exons to be followed by an active intronic PAS indicating that NMD escape through APA could be a factor responsible for the unexplained variability of NMD efficiency.

Using stringent filtering criteria, we examined tissue-specific splicing and APA and identified 14 genes with tissue-specific NMD escape. However, they may represent only the tip of the iceberg as none of the genes with a known such mechanism have been detected. The poison exon in the HFE gene [21] was not annotated, and we didn't observe upregulation of

splicing under CHX treatment when specifically looking at its 3'-UTR. Two other examples, in the CSF3R and Tau genes, were specific to CD34+ immune cells and Alziher disease patients, respectively. The thresholds used in this work could be excessively strict, e.g., the requirement of high polyA read support selects strong candidates but excludes PAS expressed in only a few tissues; $\delta < 0.6$ aims at minimizing non-uniform read coverage artifacts but is too restrictive in the case of 3'-bias. Furthermore, the diversity of NMD isoforms is not limited to poison exons only, and some APA events could actually switch a cryptic poison exon [35]. Consequently, we likely underestimated the true rate of tissue-specific APA-mediated NMD escape.

Our analysis predicted that approximately a half of VRK3 transcripts undergo NMD escape in the cervix and esophagus, while the skip isoform dominates in other tissues such as the adrenal gland and kidney. As we show, VRK3 also harbors an upstream PAS, which is preferentially activated in the brain and could represent an alternative mode of NMD escape in neural tissues (Figure S6). This observation is consistent with VRK3 supporting neuronal survival under oxidative stress via CDK5-mediated phosphorylation [44]. We were unable to directly test the contribution of the brain-specific upstream PAS since our experiments were performed in an adenocarcinoma cell line, where only the downstream PAS is active.

The most remarkable example of tissue-specific NMD-escape is the poison exon in the β-tropomyosin gene (*TPM2*), which is a paralog of α-tropomyosin (*TPM1*). Relative expression levels of tropomyosins in cardiac tissues are thoroughly regulated [48], with increase in β-tropomyosin fraction leading to severe cardiac abnormalities and thrombus formation [49]. We observe that the *TPM2* poison exon 9 becomes the terminal exon in skeletal muscles. This exon is skipped in the majority of non-brain tissues, except for the salivary gland, blood and

muscles. Inclusion of this exon in heart muscles and salivary gland is accompanied by substantial drop of its expression level (Figure S5), while activation of PAS within the exon in skeletal muscles coincides with its increase. This clearly suggests that NMD escape through alternative polyadenylation in skeletal muscles could be the mechanism of fine-tuning its expression levels in various muscle tissue types.

Interestingly, both *TPM1* and *TPM2* have striated muscle-specific APA isoforms and are upregulated in skeletal muscle tissues [50]. In the heart, TPM1 maintains the same isoform and expression profile as in skeletal muscle, whereas in *TPM2* the inclusion of poison exon 9 coincides with a marked reduction in expression. This pattern suggests a two-layer regulatory model. First, a mechanism common to striated muscle tissues promotes APA isoforms. Second, an additional heart-specific layer selectively downregulates *TPM2*, likely through inhibition of polyadenylation at the PAS within exon 9. The latter would prevent the formation of the NMD-escape isoform and thus limit *TPM2* expression in cardiac tissue.

Several aspects of APA-mediated NMD escape remained beyond the scope of this study. In particular, we surveyed poison exons that become terminal terminal exons, but the opposite transformation is also possible. Indeed, alternative terminal exons may contain unannotated splice sites which convert them into poison exons. We also noted that poison exons with a downstream PAS are more likely to also harbor an upstream PAS, hence it may be informative to explore intronic pruning upstream of poison exons. Extending the analysis framework presented here to capture such events will be critical for a more comprehensive understanding of APA-mediated NMD escape and its contribution to gene expression regulation.

Competing interests

The authors declare no competing interests.

Funding

This work was supported by Russian Science Foundation grant 22-14-00330.

Authors' contributions

DP and MV designed the study; MV performed data analysis; AK and DS performed the experiments; MV and DP wrote the first draft of the manuscript. All authors edited the final version of the manuscript.

Acknowledgments

The authors cordially thank Prof. Vera Rybko for introducing us to this research area.

References

[1] Ule, J. and Blencowe, B. J. (2019) Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution. *Mol Cell,* **76**, 329–345.

[2] Tian, B. and Manley, J. L. (2017) Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol,* **18**, 18–30.

[3] Zhang, X., Liu, F., and Zhou, Y. (2024) Coupling of alternative splicing and alternative polyadenylation. *Acta Biochim Biophys Sin (Shanghai),* **57**, 22–32.

[4] Agarwal, V., Lopez-Darwin, S., Kelley, D. R., and Shendure, J. (2021) The landscape of alternative polyadenylation in single cells of the developing mouse embryo. *Nat Commun,* **12**, 5101.

[5] Sun, Y., Pang, Y., Wu, X., Zhu, R., Wang, L., Tian, M., He, X., Liu, D., and Yang, X. (2024) Landscape of alternative splicing and polyadenylation during growth and development of muscles in pigs. *Commun Biol,* **7**, 1607.

[6] Xiang, Y., Ye, Y., Lou, Y., Yang, Y., Cai, C., Zhang, Z., Mills, T., Chen, N.-Y., Kim, Y., Muge Ozguc, F., Diao, L., Karmouty-Quintana, H., Xia, Y., Kellems, R. E., Chen, Z., Blackburn, M. R., Yoo, S.-H., Shyu, A.-B., Mills, G. B., and Han, L. (2018) Comprehensive Characterization of Alternative Polyadenylation in Human Cancer. *J Natl Cancer Inst,* **110**, 379–389.

[7] Lee, S., Aubee, J. I., and Lai, E. C. (2023) Regulation of alternative splicing and polyadenylation in neurons. *Life Sci Alliance,* **6**, e202302000.

[8] Turner, R. E., Pattison, A. D., and Beilharz, T. H. (2018) Alternative polyadenylation in the regulation and dysregulation of gene expression. *Semin Cell Dev Biol,* **75**, 61–69.

[9] Mittleman, B. E., Pott, S., Warland, S., Zeng, T., Mu, Z., Kaur, M., Gilad, Y., and Li, Y. (2020) Alternative polyadenylation mediates genetic regulation of gene expression. *Elife,* **9**, e57492.

[10] Lareau, L. F., Inada, M., Green, R. E., Wengrod, J. C., and Brenner, S. E. (2007) Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature,* **446**, 926–9.

[11] Müller-McNicoll, M., Rossbach, O., Hui, J., and Medenbach, J. (2019) Auto-regulatory feedback by RNA-binding proteins. *J Mol Cell Biol,* **11**, 930–939.

[12] Fair, B., Buen Abad Najar, C. F., Zhao, J., Lozano, S., Reilly, A., Mossian, G., Staley, J. P., Wang, J., and Li, Y. I. (2024) Global impact of unproductive splicing on human gene expression. *Nat Genet,* **56**, 1851–1861.

[13] Hug, N., Longman, D., and Cáceres, J. F. (2016) Mechanism and regulation of the nonsense-mediated decay pathway. *Nucleic Acids Res,* **44**, 1483–95.

[14] Dyle, M. C., Kolakada, D., Cortazar, M. A., and Jagannathan, S. (2020) How to get away with nonsense: Mechanisms and consequences of escape from nonsense-mediated RNA decay. *Wiley Interdiscip Rev RNA,* **11**, e1560.

[15] Sherlock, M. E., Baquero Galvis, L., Vicens, Q., Kieft, J. S., and Jagannathan, S. (2023) Principles, mechanisms, and biological implications of translation termination-reinitiation. *RNA,* **29**, 865–884.

[16] Goetz, A. E. and Wilkinson, M. (2017) Stress and the nonsense-mediated RNA decay pathway. *Cell Mol Life Sci,* **74**, 3509–3531.

[17] Howard, M. T., Malik, N., Anderson, C. B., Voskuil, J. L. A., Atkins, J. F., and Gibbons, R. J. (2004) Attenuation of an amino-terminal premature stop codon mutation in the ATRX gene by an alternative mode of translational initiation. *J Med Genet,* **41**, 951–6.

[18] Paulsen, M., Lund, C., Akram, Z., Winther, J. R., Horn, N., and Møller, L. B. (2006) Evidence that translation reinitiation leads to a partially functional Menkes protein containing two copper-binding sites. *Am J Hum Genet,* **79**, 214–29.

[19] Kishor, A., Ge, Z., and Hogg, J. R. (2019) hnRNP L-dependent protection of normal mRNAs from NMD subverts quality control in B cell lymphoma. *EMBO J,* **38**, e99128.

[20] Ngian, Z.-K., Tan, Y.-Y., Choo, C.-T., Lin, W.-Q., Leow, C.-Y., Mah, S.-J., Lai, M. K.-P., Chen, C. L.-H., and Ong, C.-T. (2022) Truncated Tau caused by intron retention is enriched in Alzheimer's disease cortex and exhibits altered biochemical properties. *Proc Natl Acad Sci U S A,* **119**, e2204179119.

[21] Martins, R., Proença, D., Silva, B., Barbosa, C., Silva, A. L., Faustino, P., and Romão, L. (2012) Alternative polyadenylation and nonsense-mediated decay coordinately regulate the human HFE mRNA levels. *PLoS One,* **7**, e35461.

[22] Druhan, L. J., Lance, A., Hamilton, A., Steuerwald, N. M., Tjaden, E., and Avalos, B. R. (2020) Altered splicing and intronic polyadenylation of CSF3R via a cryptic exon in acute myeloid leukemia. *Leuk Res,* **92**, 106349.

[23] Zavileyskiy, L. G. and Pervouchine, D. D. (2024) Post-transcriptional Regulation of Gene Expression via Unproductive Splicing. *Acta Naturae,* **16**, 4–13.

[24] Schneider, V. A., Graves-Lindsay, T., Howe, K., Bouk, N., Chen, H.-C., Kitts, P. A., Murphy, T. D., Pruitt, K. D., Thibaud-Nissen, F., Albracht, D., Fulton, R. S., Kremitzki, M., Magrini, V., Markovic, C., McGrath, S., Steinberg, K. M., Auger, K., Chow, W., Collins, J., Harden, G., Hubbard, T., Pelan, S., Simpson, J. T., Threadgold, G., Torrance, J., Wood, J. M., Clarke, L., Koren, S., Boitano, M., Peluso, P., Li, H., Chin, C.-S., Phillippy, A. M., Durbin, R., Wilson, R. K., Flicek, P., Eichler, E. E., and Church, D. M. (2017) Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. *Genome Res,* **27**, 849–864.

[25] Pertea, M., Shumate, A., Pertea, G., Varabyou, A., Breitwieser, F. P., Chang, Y.-C., Madugundu, A. K., Pandey, A., and Salzberg, S. L. (2018) CHESS: a new human gene catalog curated from thousands of large-scale RNA sequencing experiments reveals extensive transcriptional noise. *Genome Biol,* **19**, 208.

[26] Herrmann, C. J., Schmidt, R., Kanitz, A., Artimo, P., Gruber, A. J., and Zavolan, M. (2020) PolyASite 2.0: a consolidated atlas of polyadenylation sites from 3' end sequencing. *Nucleic Acids Res,* **48**, D174–D179.

[27] Vlasenok, M., Margasyuk, S., and Pervouchine, D. D. (2023) Transcriptome sequencing suggests that pre-mRNA splicing counteracts widespread intronic cleavage and polyadenylation. *NAR Genom Bioinform,* **5**, lqad051.

[28] Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics,* **29**, 15–21.

[29] Kovalak, C., Donovan, S., Bicknell, A. A., Metkar, M., and Moore, M. J. (2021) Deep sequencing of pre-translational mRNPs reveals hidden flux through evolutionarily conserved alternative splicing nonsense-mediated decay pathways. *Genome Biol,* **22**, 132.

[30] Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A., and Manke, T. (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res,* **42**, W187–91.

[31] Zerbino, D. R., Johnson, N., Juettemann, T., Wilder, S. P., and Flicek, P. (2014) WiggleTools: parallel processing of large collections of genome-wide datasets for visualization and statistical analysis. *Bioinformatics,* **30**, 1008–9.

[32] Hahne, F. and Ivanek, R. (2016) Visualizing Genomic Data Using Gviz and Bioconductor. *Methods Mol Biol,* **1418**, 335–51.

[33] Liao, Y., Smyth, G. K., and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics,* **30**, 923–30.

[34] Pervouchine, D. D., Knowles, D. G., and Guigó, R. (2013) Intron-centric estimation of alternative splicing from RNA-seq data. *Bioinformatics,* **29**, 273–4.

[35] Margasyuk, S., Kuznetsova, A., Zavileyskiy, L., Vlasenok, M., Skvortsov, D., and Pervouchine, D. D. (2024) Human introns contain conserved tissue-specific cryptic poison exons. *NAR Genom Bioinform,* **6**, lqae163.

[36] Naveed, A., Cooper, J. A., Li, R., Hubbard, A., Chen, J., Liu, T., Wilton, S. D., Fletcher, S., and Fox, A. H. (2021) NEAT1 polyA-modulating antisense oligonucleotides reveal

opposing functions for both long non-coding RNA isoforms in neuroblastoma. *Cell Mol Life Sci,* **78**, 2213–2230.

[37] Kurosaki, T., Popp, M. W., and Maquat, L. E. (2019) Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat Rev Mol Cell Biol,* **20**, 406–420.

[38] Varabyou, A., Sommer, M. J., Erdogdu, B., Shinder, I., Minkin, I., Chao, K.-H., Park, S., Heinz, J., Pockrandt, C., Shumate, A., Rincon, N., Puiu, D., Steinegger, M., Salzberg, S. L., and Pertea, M. (2023) CHESS 3: an improved, comprehensive catalog of human genes and transcripts based on large-scale expression data, phylogenetic analysis, and protein structure. *Genome Biol,* **24**, 249.

[39] Rappaport, N., Twik, M., Plaschkes, I., Nudel, R., Iny Stein, T., Levitt, J., Gershoni, M., Morrey, C. P., Safran, M., and Lancet, D. (2017) MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search. *Nucleic Acids Res,* **45**, D877–D887.

[40] Tajsharghi, H., Ohlsson, M., Palm, L., and Oldfors, A. (2012) Myopathies associated with β-tropomyosin mutations. *Neuromuscul Disord,* **22**, 923–33.

[41] Zhang, J., Zhang, C., Miao, L., Meng, Z., Gu, N., and Song, G. (2021) Abnormal TPM2 expression is involved in regulation of atherosclerosis progression via mediating RhoA signaling in vitro. *Arch Med Sci,* **20**, 1197–1208.

[42] Consortium, G. (2013) The Genotype-Tissue Expression (GTEx) project. *Nat Genet,* **45**, 580–5.

[43] Park, C.-H., Ryu, H. G., Kim, S.-H., Lee, D., Song, H., and Kim, K.-T. (2015) Presumed pseudokinase VRK3 functions as a BAF kinase. *Biochim Biophys Acta,* **1853**, 1738–48.

[44] Song, H., Kim, W., Choi, J.-H., Kim, S.-H., Lee, D., Park, C.-H., Kim, S., Kim, D.-Y., and Kim, K.-T. (2016) Stress-induced nuclear translocation of CDK5 suppresses neuronal death by downregulating ERK activation via VRK3 phosphorylation. *Sci Rep,* **6**, 28634.

[45] Lareau, L. F., Brooks, A. N., Soergel, D. A. W., Meng, Q., and Brenner, S. E. (2007) The coupling of alternative splicing and nonsense-mediated mRNA decay. *Adv Exp Med Biol,* **623**, 190–211.

[46] Jagannathan, S. and Bradley, R. K. (2016) Translational plasticity facilitates the accumulation of nonsense genetic variants in the human population. *Genome Res,* **26**, 1639–1650.

[47] Lindeboom, R. G. H., Vermeulen, M., Lehner, B., and Supek, F. (2019) The impact of nonsense-mediated mRNA decay on genetic disease, gene editing and cancer immunotherapy. *Nat Genet,* **51**, 1645–1651.

[48] Yin, Z., Ren, J., and Guo, W. (2015) Sarcomeric protein isoform transitions in cardiac muscle: a journey to heart failure. *Biochim Biophys Acta,* **1852**, 47–52.

[49] Muthuchamy, M., Grupp, I. L., Grupp, G., O'Toole, B. A., Kier, A. B., Boivin, G. P., Neumann, J., and Wieczorek, D. F. (1995) Molecular and physiological effects of overexpressing striated muscle beta-tropomyosin in the adult murine heart. *J Biol Chem,* **270**, 30593–603.

[50] Lin, J.-C. and Tarn, W.-Y. (2005) Exon selection in alpha-tropomyosin mRNA is regulated by the antagonistic action of RBM4 and PTB. *Mol Cell Biol,* **25**, 10111–21.

[51] Girotto, G., Abdulhadi, K., Buniello, A., Vozzi, D., Licastro, D., d'Eustacchio, A., Vuckovic, D., Alkowari, M. K., Steel, K. P., Badii, R., and Gasparini, P. (2013) Linkage study and exome sequencing identify a BDP1 mutation associated with hereditary hearing loss. *PLoS One,* **8**, e80323.

[52] Heliö, K., Mäyränpää, M. I., Saarinen, I., Ahonen, S., Junnila, H., Tommiska, J., Weckström, S., Holmström, M., Toivonen, M., Nikus, K., Hathaway, J., Siivonen, P., Muona, M., Sistonen, J., Salmenperä, P., Gentile, M., Paananen, J., Myllykangas, S., Alastalo, T.-P., Heliö, T., and Koskenvuo, J. (2021) GRINL1A Complex Transcription Unit Containing GCOM1, MYZAP, and POLR2M Genes Associates with Fully Penetrant Recessive Dilated Cardiomyopathy. *Front Genet,* **12**, 786705.

[53] Johnson, K. R., Cook, S. A., Bustin, M., and Davisson, M. T. (1992) Genetic mapping of the murine gene and 14 related sequences encoding chromosomal protein HMG-14. *Mamm Genome,* **3**, 625–32.

[54] Morimura, H., Saindelle-Ribeaudeau, F., Berson, E. L., and Dryja, T. P. (1999) Mutations in RGR, encoding a light-sensitive opsin homologue, in patients with retinitis pigmentosa. *Nat Genet,* **23**, 393–4.

Figures and Tables

Gene name	min(δ)	range(ψ_1)	cor(ψ_1, δ)	Associated diseases
AKR1E2	0.44	0.26	-0.58	Cataract [39]
BDP1	0.53	0.73	-0.44	Hereditary hearing loss [51]
CPED1	0.56	0.49	-0.73	N/A
GCOM1	0.55	0.49	-0.69	Cardiomyopathy [52]
HMGN1	0.18	0.42	-0.02	Down Syndrome [53]
MEGF11	0.56	0.50	-0.93	N/A
NPTN	0.03	0.92	-0.30	N/A
PARP16	0.14	0.31	-0.19	N/A
PWWP2A	0.22	0.39	-0.29	N/A
RGR	0.48	0.92	-0.20	Retinitis Pigmentosa and other [54]
SYNGAP1	0.54	0.76	-0.53	Intellectual developmental disorders [39]
TPM2	0.02	0.99	-0.89	Congenital myopathy, arthrogryposis [39]
TUBGCP5	0.06	1.00	-0.48	Autism spectrum disorders [39]
VRK3	0.39	0.31	-0.28	N/A

Table 1: Characteristics of tissue-specific NMD escape events. See also Figure S4.

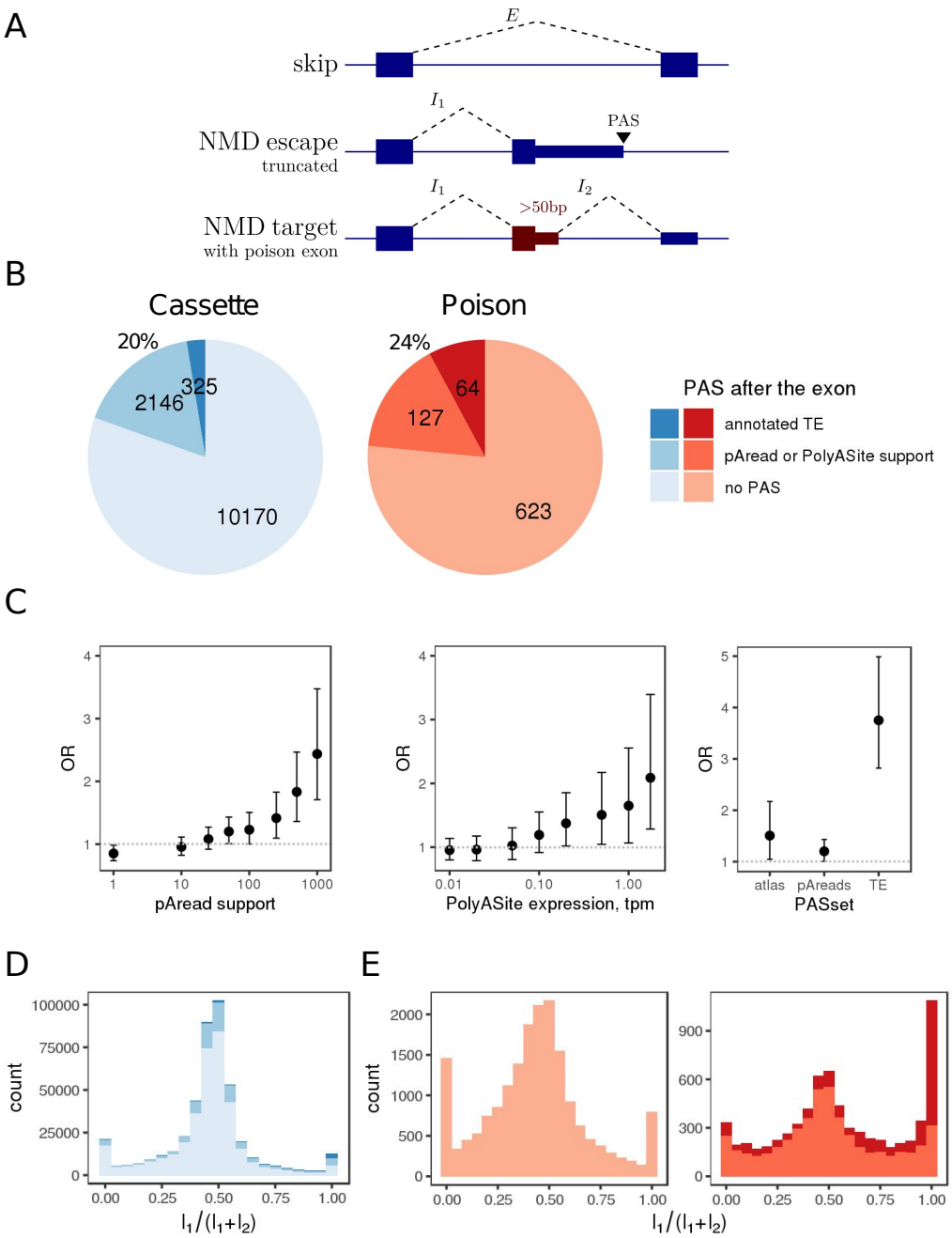

Figure 1: Poison exons are frequently followed by an intronic PAS. **(A)** A schema of transcript isoforms in a gene with a poison exon and a PAS in the downstream intron. The poison exon is shown in red. **(B)** The number of cassette and poison exons followed by a PAS. The color intensity represents the level of PAS support: an annotated transcript end, an unannotated PAS from the PolyAsite database or predicted from polyA reads (≥50 polyA reads or ≥0.5 tpm in polyASite), and no PAS support. **(C)** The odds ratio (*OR*) for the proportion of exons followed by a PAS in cassette vs. poison exons. Left: *OR* as a function of polyA read support. Middle: *OR* as a function of PAS expression in TPMs according to the PolyAsite database. Right: *OR* for PolyAsite database (atlas), polyA reads, and annotated transcript ends (TE). Whiskers represent 95% confidence intervals. **(D,E)** The distribution of $\varphi = \frac{I_1}{I_1+I_2}$ across GTEx tissues for cassette (D) and poison (E) exons. The colors are as in B.

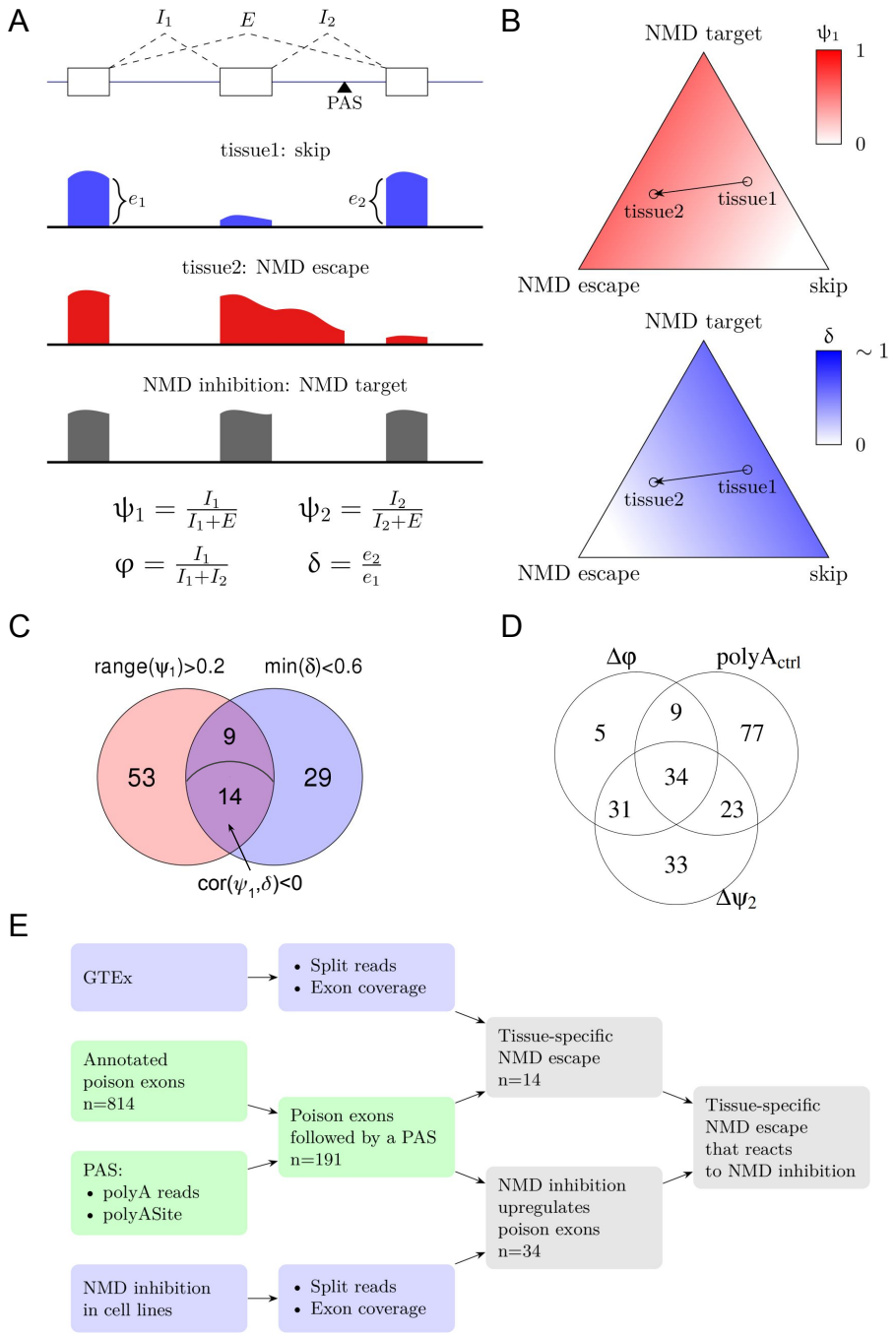

Figure 2: Identification of tissue-specific NMD escape events. **(A)** Transcript isoforms and metrics used for their assessment. e_1 and e_2 are the read coverage values per nucleotide. **(B)** A schema of tissue-specific changes in metrics ψ_1 and δ. Ternary diagrams represent the ψ_1 and δ distributions across isoforms. NMD escape corresponds to an increase in ψ_1 and a decrease in δ. **(C)** A Venn diagram for poison exons selected on the basis of ψ_1 and δ metrics. A part of the intersection with the negative correlation between ψ_1 and δ is indicated by an arrow. **(D)** A Venn diagram of the response to NMD inactivation for poison exons followed by a PAS. Δ denotes the changes of the respective metrics between NMD inactivation and control conditions (see Methods). The $polyA_{ctrl}$ denotes the category of high φ_{ctrl} or low δ_{ctrl} values in untreated cells. **(E)** A flowchart of the analysis. Green nodes use database information. Blue nodes use transcriptomic data. Gray nodes combine the two.

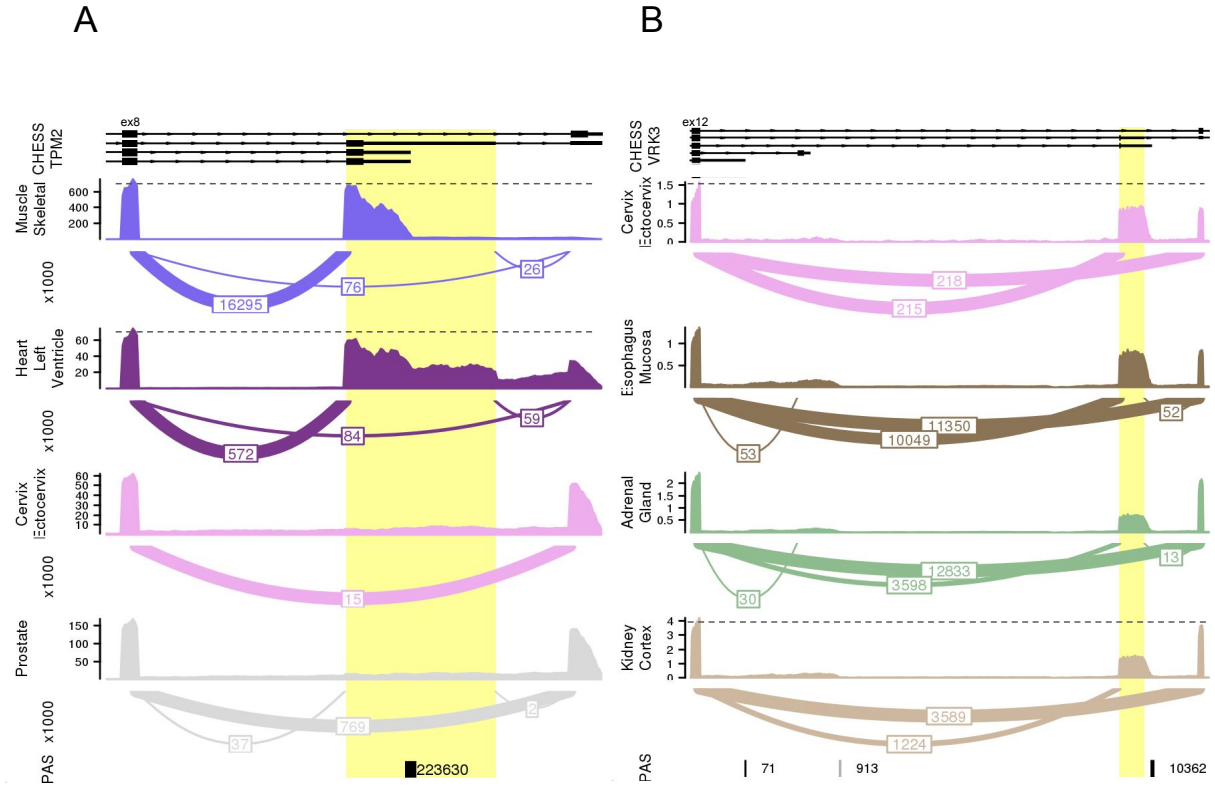

Figure 3: Examples of tissue-specific NMD escape. The read coverage profiles and sashimi plots for two examples of tissue-specific switches from skip to NMD-escape isoforms, *TPM2* (A) and *VRK3* (B). The poison exon followed by or containing a PAS is highlighted. The scale of the vertical axes is the read coverage density. The numbers on the sashimi plots are split read counts. Transcript models from CHESS are shown on the top. The PAS track shows the number of polyA reads supporting a PAS. The dotted lines illustrate the read coverage drop between the upstream and the downstream exons.

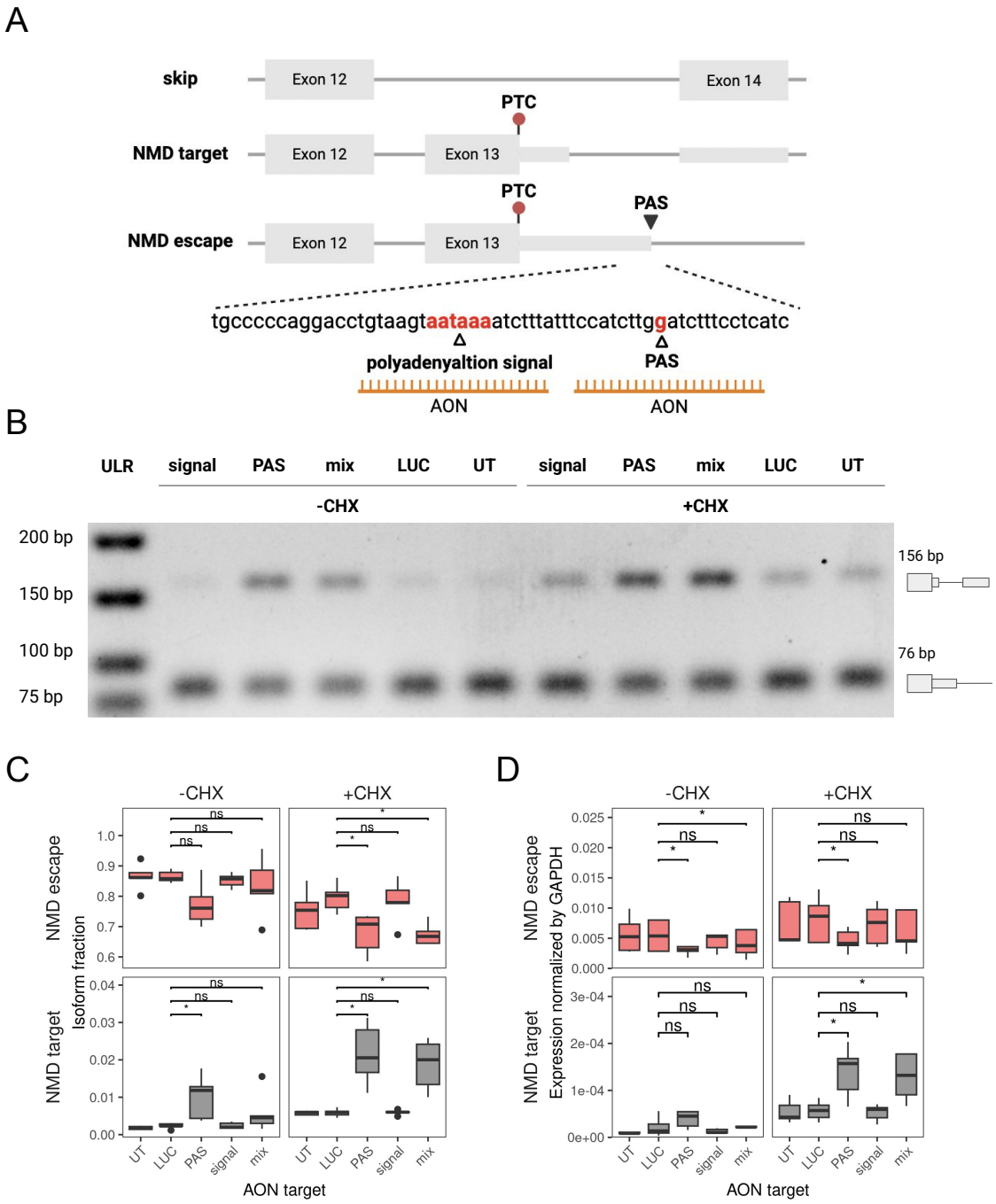

Figure 4: Validation of NMD escape in the *VRK3* gene. **(A)** The isoforms of the *VRK3* gene transcripts. Genomic sequences around the polyadenylation signal and PAS targeted by AONs are indicated. **(B)** The endpoint PCR: untreated cells (UT), cells treated with AON against luciferase (LUC), AON against polyadenylation signal (signal), AON against the cleavage site (PAS), and their mixture (mix), with (+CHX) and without (-CHX) cycloheximide treatment. **(C)** qPCR: the fraction of NMD-escape (red) and NMD-target (grey) isoforms among all three isoforms under AON treatment with and without CHX. **(D)** Expression levels of *VRK3* isoforms under AON treatment with and without NMD inhibition normalized by GAPDH expression; * denotes statistically significant differences at the 5% level; 'ns' denotes no statistically significant differences (one-sided Wilcoxon signed rank test for 5 biological replicates).