ISSN 1022-7954, Russian Journal of Genetics, 2018, Vol. 54, No. 8, pp. 879–887. © Pleiades Publishing, Inc., 2018.
Original Russian Text © E.N. Filatova, O.V. Utkin, 2018, published in Genetika, 2018, Vol. 54, No. 8, pp. 869–878.

REVIEWS
AND THEORETICAL ARTICLES

The Role of Noncoding mRNA Isoforms in the Regulation of Gene Expression

E. N. Filatova[a, *] and O. V. Utkin[a, b]

[a] *Blokhina Scientific Research Institute of Epidemiology and Microbiology, Nizhny Novgorod, 603950 Russia*
[b] *Nizhny Novgorod State Medical Academy, Nizhny Novgorod, 603950 Russia*
**e-mail: filatova@nniiem.ru*

Received August 31, 2017; in final form, February 1, 2018

Abstract—A majority of mammalian and human protein-coding genes undergo alternative splicing with the formation of mRNA isoforms. It was established that 30–40% of the formed mRNA isoforms fell into a special category of bifunctional molecules, "coding–noncoding RNAs." One possible explanation for the presence of such a large number of unproductive mRNAs is that these molecules are involved in basic processes of gene expression regulation. In this review, the concept of regulated unproductive splicing and translation is considered, which implies a close relationship between the processes of alternative splicing, formation of noncoding mRNA isoforms, and their subsequent degradation, which determines the proportion of productive mRNA transcripts of a gene and the level of its expression in the cell. Modern concepts of noncoding mRNA isoforms of protein-coding genes and their role in the regulation of gene expression under certain physiological and pathophysiological conditions are presented.

Keywords: noncoding mRNA isoforms, alternative splicing, nonsense-mediated decay, unproductive splicing, coding–noncoding RNAs, cncRNA, NMD, RUST

DOI: 10.1134/S1022795418080057

INTRODUCTION

Wide application of advanced technologies such as RNA sequencing and ribosomal profiling for transcriptome analysis led to the discovery of a special class of bifunctional RNAs [1, 2]. These RNAs function as protein-coding molecules or as noncoding RNAs. In the literature, such molecules are called "coding–noncoding RNAs" (cncRNA). The functions of cncRNAs are not completely known, but there is no doubt that they are the inherent part of the most complex mechanisms for regulating gene expression [3].

Among the cncRNAs, some noncoding mRNA isoforms (ncimRNA) of protein-coding genes quantitatively predominate. The ncimRNAs are mostly formed as a result of alternative splicing owing to the insertion of a premature termination codon or open reading frame disruption [4]. The critical role of ncimRNAs in the regulation of cell differentiation, intracellular signaling, and RNA splicing was established [5], as well as their role in the development of some severe pathophysiological conditions [3, 6, 7]. Thus, the identification of new ncimRNAs together with determination of their functional potential is one of the promising tasks of biology and medicine.

CODING–NONCODING RNAs

Among the whole variety of cncRNAs, several major groups are distinguished [8]. The first group includes long noncoding RNAs (lncRNAs), which contain information on micropeptides. In most cases, long noncoding RNAs do not encode the proteins, but are the most important regulators of gene expression during transcription, splicing of pre-mRNA, and translation. Long noncoding RNAs are involved in the regulation of the cell cycle, proliferation, apoptosis, differentiation, and maintenance of the cell pluripotency (for more details, see [9]). The structure of some long noncoding RNA contains short (less than 100 codons) open reading frames, from which functional micropeptides are actively translated. These micropeptides have no homology with known proteins, are highly conserved, are characterized by random structure, and are enriched in the protein interacting sites [1, 10, 11]. One example of peptides encoded by long noncoding RNAs is myoregulin, which is specifically expressed in mouse skeletal muscles, directly interacts with the endoplasmic reticulum (EPR) Ca^{2+} ATPase, and participates in the regulation of muscle fiber contraction [12].

The second group includes protein-coding mRNAs that perform regulatory functions regardless of the protein they encode. Acting as noncoding RNAs, these RNAs participate in the regulation of

transcription, translation, and folding, as well as in determination of the intracellular protein localization. It should be noted that untranslated regions (UTRs) of all protein-coding mRNAs contain *cis*-regulatory elements that determine mRNA stability and its translation efficiency [13–15]. In the view of the foregoing, all mRNA encoding proteins can be considered bifunctional to a certain extent. The UTRs of coding mRNAs can also contain *trans*-regulatory elements that affect expression of other genes [9]. For example, the 3' UTR region of MDFIC (MyoD family inhibitor domain-containing protein, also known as human I-mfa domain-containing protein, HIC) mRNA contains a fragment capable of competing with 7SK RNA in the complex inhibiting the activity of CDK9 elongation factor (Cyclin-dependent kinase 9, also known as positive transcription elongation factor P-TEFb). Insertion of HIC mRNA into the complex results in the loss of its inhibitory properties. For HIV type 1, CDK9-dependent transcription of viral genes was demonstrated, which could be stimulated by the MDFIC-encoding mRNA [16].

The third largest group of cncRNAs comprises noncoding mRNA isoforms of translated genes. A certain proportion of noncoding isoforms may arise owing to genetic mutations leading to the loss of the gene coding potential. The mutations leading to the formation of ncimRNAs of protein-coding genes can be represented by single nucleotide substitutions, causing the appearance of the premature termination codon or alternative splicing site, and small insertions or deletions that disturb the reading frame, as well as considerable deletions that remove the protein-coding sequence [17]. These gene changes occur in healthy people and often do not have phenotypic consequences [18]. Many mutations that cause the appearance of ncimRNAs of translated genes lead the appearance of loss-of-function genetic variants and are associated with such severe diseases as cystic fibrosis, Duchenne muscular dystrophy, Hurler's syndrome, nephrogenic diabetes insipidus, and certain types of tumors [19]. In other cases, the formation of ncimRNAs of protein-coding genes, on the contrary, has a protective effect, limiting the expression of defective or potentially harmful proteins [19]. For example, homozygous nonsense mutation in the proprotein convertase subtilisin/kexin type 9 (PCSK9) gene leads to transcription of ncimRNA, which results in a decrease in the levels of plasma low-density lipoproteins and reduces the risk of developing cardiovascular diseases [20].

Most often ncimRNAs of translated genes arise from alternative splicing. The ratio of coding and noncoding isoforms fluctuates and differs depending on the type of cells and their physiological state, which has apparent functional consequences [8, 21].

ALTERNATIVE SPLICING AND NONSENSE-MEDIATED mRNA DECAY

Alternative splicing is the main source of protein diversity in eukaryotic cells, since it leads to the formation of a pool of heterogeneous mRNAs of a single gene, without altering structural organization of the latter [21, 22]. In the mammalian genome, mRNA isoforms resulting from alternative splicing predominate [23, 24]. In humans, more than 90% of genes contain several exons and undergo alternative splicing. However, the functions of most of the isoforms are not fully known [25, 26].

To date, most known alternative splicing regulators are RNA-binding proteins, namely, representatives of the SR protein family (proteins containing domains rich in serine and arginine residues, SR proteins) and heterogeneous nuclear ribonucleoproteins. SR proteins activate splicing through specific binding of pre-mRNA and facilitating the spliceosome assembly [27, 28]. Heterogeneous nuclear ribonucleoproteins, on the contrary, suppress splicing, directly interfering with the interaction of the spliceosome components with the splice sites [29]. The combination of RNA-binding proteins and their quantitative ratio and activity determine the pre-mRNA region recognized by the spliceosome and regulate the incorporation of nucleotide sequences into the mature transcript [21, 27]. In other words, the balance of regulatory proteins determines the set of mRNA isoforms formed depending on the type of cells, tissue, and the stage of development of the organism [30–33].

Alternative splicing can give rise to the events as follows: exon skipping, where one or more exons and flanking introns are sliced out from the mature mRNA sequence, which is the most common form of alternative splicing in higher eukaryotes; inclusion of mutually exclusive exons, which is a kind of exon skipping, where only one of the two exons is included in the mature mRNA; the use of alternative 3' and 5' splice sites, which leads to the increase or decrease in the exon size in the transcript; intron retention, which is the retention of an intron or its part in mature mRNA [22, 34]. Different alternative splicing events have different structural and functional consequences. Alternative splicing in the 3' or 5' UTR region affects mRNA stability and its intracellular localization [35, 36]. Changes in the mRNA coding region result in the reading frame shift and expression of proteins with similar or antagonistic functions, as well as in the formation of ncimRNA [37]. A separate group is represented by processed transcripts, in the structure of which, as a result of alternative splicing, a premature termination codon (PTC) appears, which also results in the formation of ncimRNA.

In mammals, the mRNA termination codon is regarded as premature if it is located more than 50 nt upstream of the 3'-most exon–exon junction (as a variant, the position of last intron) [38, 39]. Alterna-

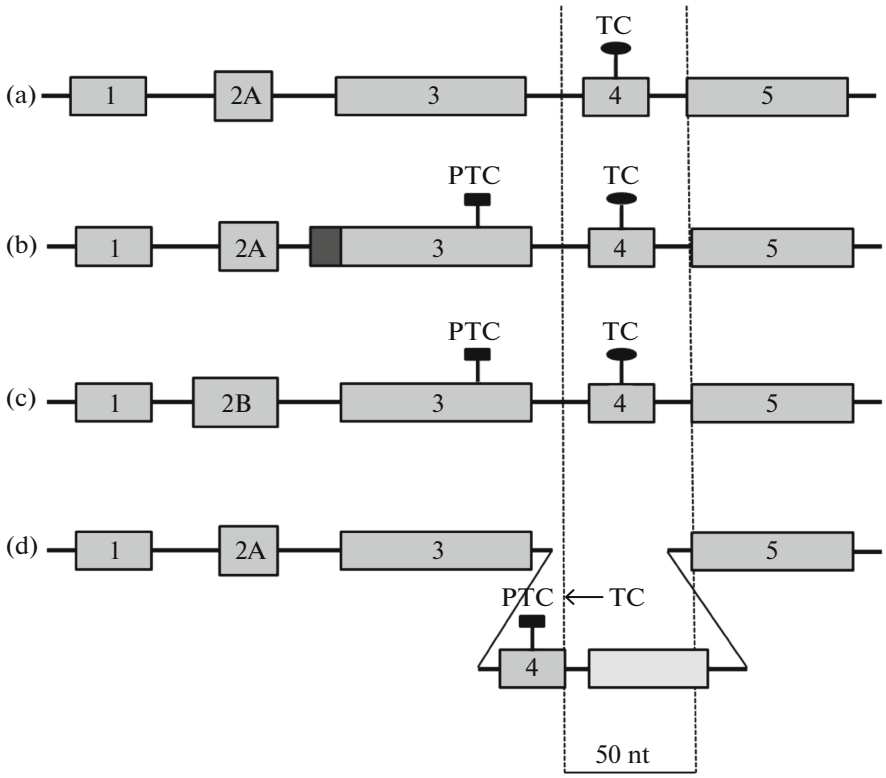

Fig. 1. Formation of premature termination codon in alternative splicing. (a) mRNA isoform that does not contain premature termination codon (PTC). The isoform contains five consecutive exons (designated by figures); termination codon (TC) of the translated region is located within 50 nt upstream of the 3'-most exon–exon junction. (b) The use of an alternative splice site results in a change in the length of exon 3 (indicated by a dark rectangle) and the reading frame shift. This results in the appearance of PTC in the mRNA structure 50 nt upstream of the 3'-most exon–exon junction. (c) Triggering of alternative exon 2B results in a reading frame shift and the appearance of PTC. (d) Intron preservation (indicated by light rectangle) in the 3'-untranslated region leads to movement of the TC away from the 3'-most exon–exon junction over a distance of more than 50 nt and the transformation of TC into PTC.

tive splicing leads to the appearance of PTC in several ways (Fig. 1): by using an alternative splice site, inclusion or splicing out of an alternative exon (this shifts the reading frame in the downstream exons) upon the intron preservation in the 3' UTR (in this case, the classical termination codon becomes the PTC) [40]. The PTC-containing mRNA isoforms are considered as potential targets for elimination by nonsense-mediated mRNA decay (NMD) [41], a universal defense mechanism directed against the synthesis of defective proteins. According to the 50 nt rule, about 30–45% of all mRNA isoforms expressed in human cells and tissues should undergo NMD [41, 42].

UNPRODUCTIVE SPLICING

To explain the need for synthesis, splicing, and subsequent degradation of such a large number of ncimRNAs containing PTC, the hypothesis of unproductive splicing (UPS, regulated unproductive splicing and translation, RUST) was proposed that described coupling of alternative splicing and NMD into a single mechanism of gene expression regulation [41]. The backbone of the hypothesis is that RUST directs the already transcribed pre-mRNA along an unproductive pathway, i.e., the formation of noncoding isoform that undergoes degradation, which makes it possible to regulate the level of protein expression at the posttranscriptional level. The UPS process is evolutionarily conserved and is described in yeast, mammals, and plants [43–45]. In the literature, constitutive, regulated, and autoregulatory UPS is distinguished [40] (Fig. 2).

Constitutive UPS provides a stable ratio of coding and noncoding mRNA isoforms in cells and tissues (Fig. 2a). A typical example is the variety of mRNA isoforms of the CAPN10 gene (Caplain-10) encoding the ubiquitous protease. CAPN10 is involved in the cytoskeleton reconstruction and the intracellular signal transduction. It plays an important role in insulin-dependent glucose uptake. As a result of alternative splicing, eight isoforms of CAPN10 mRNA are continuously produced in cells, four of which contain PTC and undergo NMD [38, 42, 46]. Single nucleotide substitution in the third intron of the gene is described, which leads to the decrease in the total level

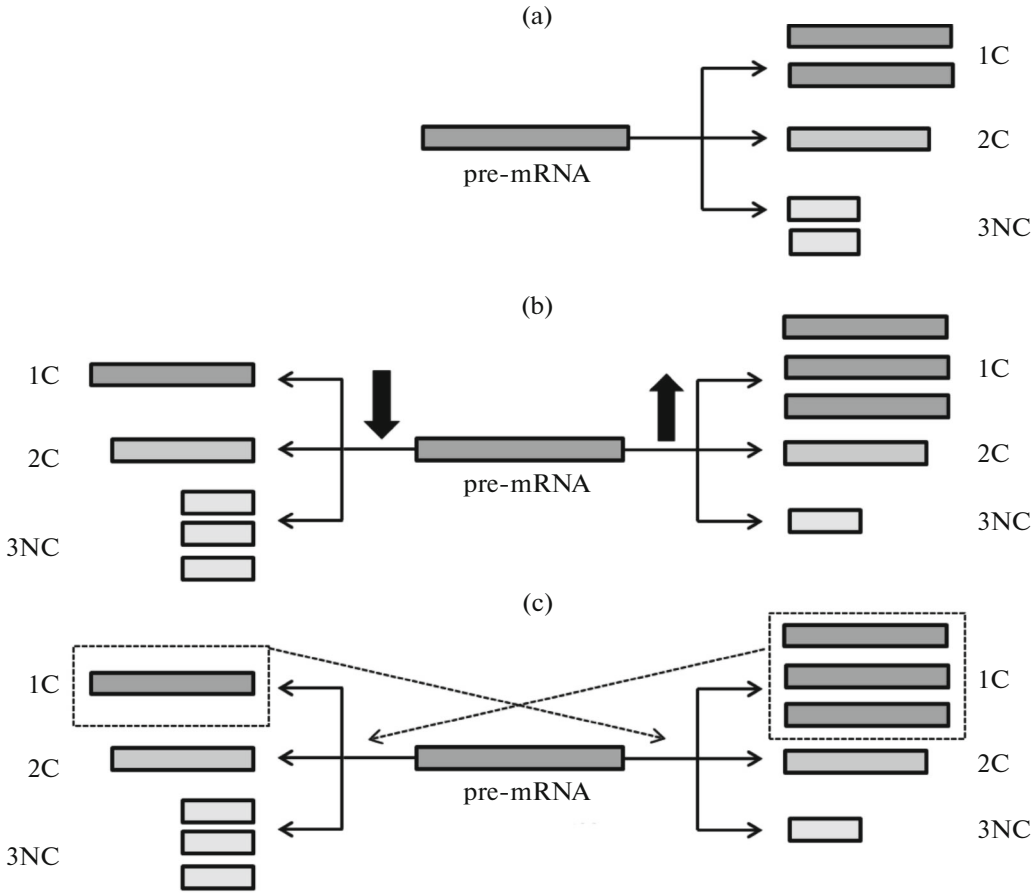

Fig. 2. Types of unproductive splicing. (a) Constitutive unproductive splicing (UPS). Constitutive UPS results in that, in the cell, three types of mRNA isoforms are constantly formed: protein-coding isoform 1 (1C), protein-coding isoform 2 (2C), and noncoding isoform 3 (3NC). (b) Regulated UPS. Under the influence of regulatory factors (indicated by large dark arrows), gene expression is suppressed owing to the increase in the proportion of 3NC isoform and the decrease in the proportion of 1C isoform (left side), or vice versa, gene expression is increased owing to the increase in the proportion of 1C isoform along with the reduction of the proportion of 3NC isoform (right side). (c) Autoregulated UPS. An increase or suppression of gene expression takes place in a way similar to scheme (b). In this case, the role of a regulatory factor is played by the protein product of a gene, which affects alternative splicing of its own pre-mRNA through a negative feedback mechanism (indicated by dashed arrows).

of CAPN10 mRNA expression and the development of insulin resistance in skeletal muscle cells [47]. In individuals homozygous for this mutation, the risk of developing type 2 diabetes mellitus is increased [46]. The researchers believe that the described polymorphism can lead to a shift in the ratio of CAPN10 mRNA isoforms toward the predominance of noncoding transcripts degraded by NMD [38].

It was demonstrated that, in humans, expression of the genes encoding the universal transporter ABCC4 (ATP binding cassette subfamily C member 4, also known as multidrug resistance-associated protein 4, MRP4) [48], purine metabolism enzyme, human hypoxanthine phosphoribosyl transferase 1 (HPRT1), DNA polymerase beta (POLB) [49], the T-cell receptor beta (TRB) [50], and other mostly ubiquitous proteins [40, 51] was regulated by constitutive UPS. The factors regulating splicing of pre-mRNA can themselves become targets of constitutive UPS. For instance, for the *U2AF1* gene encoding a small subunit of the spliceosome component U2AF, three mRNA isoforms were identified. Two productive mRNA isoforms encode proteins that, as the spliceosome component, stimulate its binding to pre-mRNA and provide the activity of the whole complex. The third alternative noncoding mRNA isoform is expressed at a high level, but contains PTC and undergoes NMD [52]. Thus, constitutive UPS can be attributed to the housekeeping processes that provide cellular homeostasis.

Regulated UPS determines the ratio of coding and noncoding mRNA isoforms depending on the cellular or tissue assignment, as well as the stage of development and the physiological state of the organism (Fig. 2b). The balance of functional and PTC-containing isoforms depends on the environmental conditions that affect the alternative splicing regulatory factors. Thus, shifting the balance toward increasing the production of

ncimRNA of the target gene will lead to a decrease in the expression of productive transcripts and, as a result, to a decrease in the amount of synthesized target protein [40]. An example of regulated UPS is the polymorphism of mRNA isoforms of the *MID1* gene (midline 1) in mammalian and human cells. The MID1 protein plays an important role in cellular metabolism, associating as a heterodimer with cytoplasmic microtubules and participating in the degradation of protein phosphatase 2. MID1 is characterized by the presence of mRNA pool formed as a result of alternative splicing. The expression of both coding and noncoding isoforms of the *MID1* mRNA is tissue-specific and depends on the stage of development of the organism. In fetal brain and liver cells, five and two noncoding isoforms were detected, respectively, and in fetal fibroblasts, no noncoding isoforms were detected. The authors consider expression of PSC-containing MID1 isoforms and their subsequent degradation by NMD as the main mechanism for regulation of tissue-specific gene expression [53].

Regulated UPS provides variable expression levels of the fumarylacetoacetate hydrolase (FAH) minor isoforms in heart, liver, kidney, spleen, and other human and mouse tissues [54]. In embryonic and adult rat neural tissues, an alternative form of niscarin (transmembrane glycoprotein regulating presenelin activity) is specifically expressed in neurons, while in astrocytes and glial cells the protein-coding mRNA isoform is recognized as unproductive and is subjected to NMD. It is suggested that the disturbance of the described regulated UPS in the cells of the nervous system is one of the possible mechanisms for the development of Alzheimer's disease [55]. Thus, along with transcription factors, regulated UPS is considered as one of the main intracellular mechanisms for quantitative regulation of gene expression.

Autoregulatory UPS is usually observed in genes encoding splicing factors and spliceosome components. A distinguishing feature of autoregulated UPS is the ability of gene protein products to specifically bind to their pre-mRNAs and to influence the process of their alternative splicing, leading to the formation of noncoding isoforms (Fig. 2c). A good illustration is the expression regulation of the polypyrimidine tract binding protein 1 (PTBP1), a ubiquitous protein involved in the regulation of splicing, polyadenylation, translation, mRNA stability, etc. It was demonstrated that, in HeLa cells, the PTBP1 overexpression promoted alternative splicing of its own mRNA through a negative feedback mechanism. Alternative splicing results in skipping of the transcript exon 11, which leads to the appearance of the PTC and degradation of the noncoding isoform by NMD. This pathway consumes more than 20% of the total pool of PTBP1 mRNA in the cell [54]. The authors note that the described mechanism can also be used to enhance the PTBP1 expression level by decreasing the proportion of degraded transcripts [56, 57]. The regulation of the expression of other key splicing elements by autoregulatory RUST was shown: SRSF2 proteins (also known as SC35) [58] and TRA2B [59], pair of RNA-binding proteins TIA-1 and TIAL1 [60], trans-regulatory factor HNRNPD (also knows as AUF1) [61], etc. [40].

It should be noted that autoregulatory UPS can be realized indirectly. For instance, representatives of the family of CDC-like kinases (CLK) participate in the regulation of alternative splicing by activating a group of SR proteins by means of their phosphorylation. Activation of SR proteins, in turn, leads to the enhancement of alternative splicing of CLK pre-mRNA and the formation of predominantly PTC-containing noncoding isoforms [38, 62]. Owing to the fact that SR proteins and other alternative splicing regulators have a wide range of targets, autoregulatory UPS has an effect on alternative splicing and expression of a great number of genes.

THE ROLE OF NONCODING mRNA ISOFORMS IN DEVELOPMENT OF SOME DISEASES

The change in the set of mRNA isoforms of protein-coding genes can be both the cause of the development of the disease and its consequence [63]. Although most pathological alterations in the gene structure or changes in alternative splicing are associated with the formation of an alternative protein isoform [64], the formation of predominantly ncimRNA in the cell generally has more severe consequences. For example, cystic fibrosis is a serious hereditary disease caused by over 2000 mutations in the gene of the transmembrane conductance regulator (CFTR) protein. In 70% of cases, mutations affect folding and stability of the protein product, reducing its activity. In 10% of patients, mutations that lead to the appearance of PTC in the transcript structure and its subsequent degradation are revealed. In the latter case, standard protocols of treatment, implying the use of modulators of cystic fibrosis transmembrane conductance regulator, have no therapeutic effect [65, 66].

Another example is Duchenne muscular dystrophy, a severe disease inherited in X-linked pattern. The disease is caused by the mutation in the dystrophin gene at the Xp21 locus, the protein product of which is an important structural component of muscle tissue. Mutation leads to the appearance of PTC and the protein degradation. In homozygous carriers, these defects are accompanied by a complete loss of motor activity and death at the age of 20–30 years. Becker muscular dystrophy is a benign form of Duchenne dystrophy, characterized by slower disease progression and longer life expectancy of patients. The disease develops upon different mutations of the dystrophin gene, often leading to the occurrence of considerable deletions in the final transcript, but without the formation of PTC [42].

At present, a variant of therapy of such severe conditions by preventing the decay of PTC-containing transcripts of mutated genes by suppressing NMD has been suggested. It is supposed that preservation of noncoding mRNA isoforms in the cell leads to the production of a partially or fully functional alternative protein isoform [19]. The effectiveness of the described approach with the use of antibiotics of aminoglycoside series decreasing the level of NMD was demonstrated in model systems and in small groups of patients with cystic fibrosis [65, 67] and muscular dystrophy [68, 69].

Disturbance of the ratio between productive and nonproductive mRNA isoforms of a wide range of regulatory genes can cause the development of different pathologies. The genetic causes of myotonic dystrophy, the most common form of adult muscular dystrophy, are well studied. Myotonic dystrophy type 1 is associated with the increased number of tandem CTG repeats in the 3' UTR of the *DMPK* gene (myotonic dystrophy protein kinase), while myotonic dystrophy type 2 is associated with the increased number of CCTG repeats in intron 1 of the *CBPB* gene (CCHC-type zinc finger nucleic acid binding protein; another name is zinc finger protein 9). In the disease, repeats in the mRNA structure form hairpins, as a result of which molecules form aggregates at the periphery of nuclear speckles, where splicing factors are associated and disrupt their work [3]. Similar disorders in the cells of adult tissues are accompanied by the realization of alternative splicing along the embryonic pathway, i.e., formation of a spectrum of mRNA isoforms characteristic of the developing embryo and normally absent in adults.

For example, in patients with myotonic dystrophy type 1, impaired alternative splicing of the steroid receptor RNA activator 1 (*SRA1*) gene is observed. The *SRA1* gene has two transcripts, protein-coding and noncoding, formed as a result of preservation of the first intron and disruption of the open reading frame during alternative splicing. SRA1 ncimRNA enhances the activity of the myogenic regulatory factor MYOD1 (myogenic differentiation 1); its proportion increases upon myogenic differentiation of human primary satellite cells. In patients with type 1 myotonic dystrophy, the ratio of productive and unproductive isoforms of SRA1 is preserved at the level of embryonic tissues [6]. In patients with myotonic dystrophy type 1 and type 2, decreased expression of the CLCN1 protein (chloride voltage-gated channel 1), the main protein of the skeletal muscle membrane chloride channel, was observed. The mouse model demonstrated that the increased number of CTG repeats triggered alternative splicing of the CLCN1 pre-mRNA following the neonatal pattern with the formation of a large number of noncoding isoforms [70]. It should be noted that the similar mechanism of increasing the number of RNA-binding triplets that trigger aberrant alternative splicing is associated with other diseases, FXTAS syndrome [71] and Friedreich ataxia [72].

In addition to the above-described participation of the noncoding SRA1 mRNA isoform in muscle cell development, this molecule plays the role of coactivator of a great number of nuclear factors, including chromatin-modulating proteins of polycomb group, Trithorax group of transcriptional regulators, stem cell pluripotency factor NANOG [73], and others. The RNA form of SRA1 is a potential regulator of proliferation and differentiation of various types of cells, while the protein form can inhibit its activity [74]. It is not surprising that disturbance of the ratio between coding and noncoding SRA1 mRNA isoforms can stimulate the development of malignant tumors, determine the type of cancer, and influence its metastasis [3]. For example, enhancement of the SRA1 ncimRNA expression is observed in breast cancer cell lines. Moreover, invasive cancers are characterized by a higher proportion of unproductive isoform compared to noninvasive malignancies [75, 76].

The appearance of neoplasia may be often associated not only with the expression of allele-specific isoforms of oncogenesis regulating genes [7] but also with disturbances in the mechanisms of NMD that prevent accumulation of defective mRNAs [8, 77]. The NMD pathway triggers degradation of the PTC-containing ncimRNAs of the *BRCA1* gene, associated with breast cancer [78], and the PTC-containing mRNA isoforms of the *TP53* gene (tumor protein p53) in bladder cancer cells [79]. It was demonstrated that pancreatic carcinoma was often associated with mutations of the *UPF1* gene encoding the central component of NMD. Impaired function of the *UPF1* gene leads to a decrease in the level of NMD in tumor cells. The process is accompanied by the preservation and accumulation of ncimRNAs of different genes [80]. To date, more than 750 mRNA isoforms of the genes regulating cell division and differentiation that undergo NMD in normal cells but are overexpressed in tumor cells of various origin in conditions of decreased activity of NMD have been identified [5]. It is suggested that preservation of ncimRNAs results in expression from their sequences of alternative proteins possessing oncogenic properties. In this case, activation of ncimRNA degradation by NMD can be considered a component of anticancer therapeutics [19].

Thus, to date, the role of ncimRNAs of protein-coding genes in the development of different diseases has been demonstrated. In most cases, the pathogenetic effect of noncoding isoforms is determined by genetic mutations affecting alternative splicing of both single and multiple genes. The result of the changes is an alteration of the ratio between productive and nonproductive mRNA isoforms. In this context, NMD, which is a natural cellular mechanism for regulating the ncimRNA content, can be considered as a promising therapeutic target in certain diseases.

CONCLUSIONS

Noncoding mRNA isoforms of protein-coding genes are bifunctional molecules. On one hand, as in the case of RNA isoform of the *SRA1* gene, nonproductive sequences can interact with a wide range of regulatory ribonucleoproteins and influence their activity. On the other hand, those ncimRNAs whose regulatory potential has not yet been confirmed indirectly determine expression of the protein product of their own gene, participating in the regulated UPS.

The concept of regulated UPS implies a close relationship between the processes of alternative splicing and NMD, which carry out posttranscriptional regulation of the expression of most mammalian and human genes. Within the framework of the concept, the process of alternative splicing determines the future fate of the transcript, forming either a productive mRNA isoform or a noncoding isoform, which subsequently undergoes NMD. Thus, the UPS can purposefully reduce the production of targeted protein owing to the increase of the proportion of ncimRNA. Reduction of the production of noncoding isoforms, on the contrary, causes an increase in gene expression at the protein level. The use of the described mechanism of expression regulation plays a key role in the case where the requirement of a cell in a protein changes already after the beginning of the coding gene transcription [40].

REFERENCES

1. Andrews, S.J. and Rothnagel, J.A., Emerging evidence for functional peptides encoded by short open reading frames, *Nat. Rev. Genet.*, 2014, vol. 15, no. 3, pp. 193—204. doi 10.1038/nrg3520
2. Mascarenhas, R., Pietrzak, M., Smith, R.M., et al., Allele-selective transcriptome recruitment to polysomes primed for translation: protein-coding and noncoding RNAs, and RNA isoforms, *PLoS One*, 2015, vol. 10, no. 9. 10. e0136798. doi 10.1371/journal.pone.0136798
3. Kumari, P. and Sampath, K., cncRNAs: bi-functional RNAs with protein coding and non-coding functions, *Semin. Cell Dev. Biol.*, 2015, vols. 47—48, pp. 40—51. doi 10.1016/j.semcdb.2015.10.024
4. Gonzalez-Porta, M., Frankish, A., Rung, J., et al., Transcriptome analysis of human tissues and cell lines reveals one dominant transcript per gene, *Genome Biol.*, 2013, vol. 14, no. 7, p. R70. doi 10.1186/gb-2013-14-7-r70
5. Wang, D., Zavadil, J., Martin, L., et al., Inhibition of nonsense-mediated RNA decay by the tumor microenvironment promotes tumorigenesis, *Mol. Cell. Biol.*, 2011, vol. 31, no. 17, pp. 3670—3680. doi 10.1128/MCB.05704-11
6. Hube, F., Velasco, G., Rollin, J., et al., Steroid receptor RNA activator protein binds to and counteracts SRA RNA-mediated activation of MyoD and muscle differentiation, *Nucleic Acids Res.*, 2011, vol. 39, no. 2, pp. 513—525. doi 10.1093/nar/gkq833
7. Mayba, O., Gilbert, H.N., Liu, J., et al., MBASED: allele-specific expression detection in cancer tissues and cell lines, *Genome Biol.*, 2014, vol. 15, no. 8, p. 405. doi 10.1186/s13059-014-0405-3
8. Nam, J.-W., Choi, S.-W., and You, B.-H., Incredible RNA: dual functions of coding and noncoding, *Mol. Cells*, 2016, vol. 39, no. 5, pp. 367—374. doi 10.14348/molcells.2016.0039
9. Karapetyan, A.R., Buiting, C., Kuiper, R.A., and Coolen, M.W., Regulatory roles for long ncRNA and mRNA, *Cancers*, 2013, vol. 5, no. 2, pp. 462—490. doi 10.3390/cancers5020462
10. Cheng, H., Chan, W.S., Li, Z., et al., Small open reading frames: current prediction techniques and future prospect, *Curr. Protein Pept. Sci.*, 2011, vol. 12, no. 6, pp. 503—507.
11. Mackowiak, S.D., Zauber, H., Bielow, C., et al., Extensive identification and analysis of conserved small ORFs in animals, *Genome Biol.*, 2015, vol. 16, p. 179. doi 10.1186/s13059-015-0742-x
12. Anderson, D.M., Anderson, K.M., Chang, C.-L., et al., A micropeptide encoded by a putative long noncoding RNA regulates muscle performance, *Cell*, 2015, vol. 160, no. 4, pp. 595—606. doi 10.1016/j.cell.2015.01.009
13. Jenny, A., Hachet, O., Zavorszky, P., et al., A translation-independent role of oskar RNA in early *Drosophila* oogenesis, *Dev. Camb. Engl.*, 2006, vol. 133, no. 15, pp. 2827—2833. doi 10.1242/dev.02456
14. Lim, S., Kumari, P., Gilligan, P., et al., Dorsal activity of maternal squint is mediated by a non-coding function of the RNA, *Dev. Camb. Engl.*, 2012, vol. 139, no. 16, pp. 2903—2915. doi 10.1242/dev.077081
15. Wan, Y., Qu, K., Ouyang, Z., et al., Genome-wide measurement of RNA folding energies, *Mol. Cell.*, 2012, vol. 48, no. 2, pp. 169—181. doi 10.1016/j.molcel.2012.08.008
16. Young, T.M., Tsai, M., Tian, B., et al., Cellular mRNA activates transcription elongation by displacing 7SK RNA, *PLoS One*, 2007, vol. 2, no. 10. e1010. doi 10.1371/journal.pone.0001010
17. Rivas, M.A., Pirinen, M., Conrad, D.F., et al., Impact of predicted protein-truncating genetic variants on the human transcriptome, *Science*, 2015, vol. 348, no. 6235, pp. 666—669. doi 10.1126/science.1261877
18. Lappalainen, T., Sammeth, M., Friedländer, M.R., et al., Transcriptome and genome sequencing uncovers functional variation in humans, *Nature*, 2013, vol. 501, no. 7468, pp. 506—511. doi 10.1038/nature12531
19. Holbrook, J.A., Neu-Yilik, G., Hentze, M.W., and Kulozik, A.E., Nonsense-mediated decay approaches the clinic, *Nat. Genet.*, 2004, vol. 36, no. 8, pp. 801—808. doi 10.1038/ng1403
20. Cohen, J.C., Boerwinkle, E., Mosley, T.H., and Hobbs, H.H., Sequence variations in PCSK9, low LDL, and protection against coronary heart disease, *N. Engl. J. Med.*, 2006, vol. 354, no. 12, pp. 1264—1272. doi 10.1056/NEJMoa054013
21. Nilsen, T.W. and Graveley, B.R., Expansion of the eukaryotic proteome by alternative splicing, *Nature*, 2010, vol. 463, no. 7280, pp. 457—463. doi 10.1038/nature08909
22. van den Hoogenhof, M.M.G., Pinto, Y.M., and Creemers, E.E., RNA splicing: regulation and dysregu-

lation in the heart, *Circ. Res.*, 2016, vol. 118, no. 3, pp. 454—468. doi 10.1161/CIRCRESAHA.115.307872

23. Barbosa-Morais, N.L., Irimia, M., Pan, Q., et al., The evolutionary landscape of alternative splicing in vertebrate species, *Science*, 2012, vol. 338, no. 6114, pp. 1587—1593. doi 10.1126/science.1230612
24. Merkin, J., Russell, C., Chen, P., and Burge, C.B., Evolutionary dynamics of gene and isoform regulation in mammalian tissues, *Science*, 2012, vol. 338, no. 6114, pp. 1593—1599. doi 10.1126/science.1228186
25. Pan, Q., Shai, O., Lee, L.J., et al., Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing, *Nat. Genet.*, 2008, vol. 40, no. 12, pp. 1413—1415. doi 10.1038/ng.259
26. Wang, E.T., Sandberg, R., Luo, S., et al., Alternative isoform regulation in human tissue transcriptomes, *Nature*, 2008, vol. 456, no. 7221, pp. 470—476. doi 10.1038/nature07509
27. Graveley, B.R., Sorting out the complexity of SR protein functions, *RNA*, 2000, vol. 6, no. 9, pp. 1197—1211.
28. Lin, S. and Fu, X.-D., SR proteins and related factors in alternative splicing, *Adv. Exp. Med. Biol.*, 2007, vol. 623, pp. 107—122.
29. Gerstberger, S., Hafner, M., and Tuschl, T., A census of human RNA-binding proteins, *Nat. Rev. Genet.*, 2014, vol. 15, no. 12, pp. 829—845.
30. Eperon, I.C., Makarova, O.V., Mayeda, A., et al., Selection of alternative 5' splice sites: role of U1 snRNP and models for the antagonistic effects of SF2/ASF and hnRNP A1, *Mol. Cell. Biol.*, 2000, vol. 20, no. 22, pp. 8303—8318.
31. Hall, M.P., Nagel, R.J., Fagg, W.S., et al., Quaking and PTB control overlapping splicing regulatory networks during muscle cell differentiation, *RNA*, 2013, vol. 19, no. 5, pp. 627—638. doi 10.1261/rna.038422.113
32. Hanamura, A., Caceres, J.F., Mayeda, A., et al., Regulated tissue-specific expression of antagonistic pre-mRNA splicing factors, *RNA*, 1998, vol. 4, no. 4, pp. 430—444.
33. Kalsotra, A., Xiao, X., Ward, A.J., et al., A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart, *Proc. Natl. Acad. Sci. U.S.A.*, 2008, vol. 105, no. 51, pp. 20333—20338. doi 10.1073/pnas.0809045105
34. Kim, E., Goren, A., and Ast, G., Alternative splicing: current perspectives, *BioEssays News Rev. Mol. Cell. Dev. Biol.*, 2008, vol. 30, no. 1, pp. 38—47. doi 10.1002/bies.20692
35. Garcia, J., Gerber, S.H., Sugita, S., et al., A conformational switch in the Piccolo C2A domain regulated by alternative splicing, *Nat. Struct. Mol. Biol.*, 2004, vol. 11, no. 1, pp. 45—53. doi 10.1038/nsmb707
36. 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, vol. 3, no. 1, pp. 76—83.
37. Clark, F. and Thanaraj, T.A., Categorization and characterization of transcript-confirmed constitutively and alternatively spliced introns and exons from human, *Hum. Mol. Genet.*, 2002, vol. 11, no. 4, pp. 451—464.
38. Hillman, R.T., Green, R.E., and Brenner, S.E., An unappreciated role for RNA surveillance, *Genome Biol.*, 2004, vol. 5, no. 2. R8. doi 10.1186/gb-2004-5-2-r8
39. Nagy, E. and Maquat, L.E., A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance, *Trends Biochem. Sci.*, 1998, vol. 23, no. 6, pp. 198—199.
40. Soergel, D.A.W., Lareau, L.F., and Brenner, S.E., *Regulation of Gene Expression by Coupling of Alternative Splicing and NMD*, Landes Bioscience, 2013. https://www.ncbi.nlm.nih.gov/book/NBK 6088/.
41. Lewis, B.P., Green, R.E., and Brenner, S.E., Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans, *Proc. Natl. Acad. Sci. U.S.A.*, 2003, vol. 100, no. 1, pp. 189—192.
42. Green, R.E., Lewis, B.P., Hillman, R.T., et al., Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes, *Bioinf. Oxf. Engl.*, 2003, vol. 19, suppl. 1, pp. i118—i121.
43. Conti, E. and Izaurralde, E., Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species, *Curr. Opin. Cell Biol.*, 2005, vol. 17, no. 3, pp. 316—325. doi 10.1016/j.ceb.2005.04.005
44. Staiger, D., Zecca, L., Wieczorek Kirk, D.A., et al., The circadian clock regulated RNA-binding protein AtGRP7 autoregulates its expression by influencing alternative splicing of its own pre-mRNA, *Plant J. Cell Mol. Biol.*, 2003, vol. 33, no. 2, pp. 361—371.
45. Vilardell, J., Chartrand, P., Singer, R.H., and Warner, J.R., The odyssey of a regulated transcript, *RNA*, 2000, vol. 6, no. 12, pp. 1773—1780.
46. Horikawa, Y., Oda, N., Cox, N.J., et al., Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus, *Nat. Genet.*, 2000, vol. 26, no. 2, pp. 163—175. doi 10.1038/79876
47. Baier, L.J., Permana, P.A., Yang, X., et al., A calpain-10 gene polymorphism is associated with reduced muscle mRNA levels and insulin resistance, *J. Clin. Invest.*, 2000, vol. 106, no. 7, pp. R69—R73.
48. Lamba, J.K., Adachi, M., Sun, D., et al., Nonsense mediated decay downregulates conserved alternatively spliced ABCC4 transcripts bearing nonsense codons, *Hum. Mol. Genet.*, 2003, vol. 12, no. 3, pp. 99—109.
49. Skandalis, A. and 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, vol. 32, no. 22, pp. 6557—6564. doi 10.1093/nar/gkh967
50. Chan, W.-K., Huang, L., Gudikote, J.P., et al., An alternative branch of the nonsense-mediated decay pathway, *EMBO J.*, 2007, vol. 26, no. 7, pp. 1820—1830. doi 10.1038/sj.emboj.7601628
51. Fatscher, T., Boehm, V., and Gehring, N.H., Mechanism, factors, and physiological role of nonsense-mediated mRNA decay, *Cell. Mol. Life Sci.*, 2015, vol. 72, no. 23, pp. 4523—4544. doi 10.1007/s00018-015-2017-9
52. Pacheco, T.R., Gomes, A.Q., Barbosa-Morais, N.L., et al., Diversity of vertebrate splicing factor U2AF35: identification of alternatively spliced U2AF1 mRNAs, *J. Biol. Chem.*, 2004, vol. 279, no. 26, pp. 27039—27049. doi 10.1074/jbc.M402136200
53. 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,

vol. 114, no. 6, pp. 541—552. doi 10.1007/s00439-004-1114-x

54. Dreumont, N., Maresca, A., Boisclair-Lachance, J.-F., 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, vol. 6, p. 1. doi 10.1186/1471-2199-6-1
55. 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, vol. 136, nos. 1—2, pp. 12—22. doi 10.1016/j.molbrainres.2004.12.022
56. Wollerton, M.C., Gooding, C., Wagner, E.J., et al., Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay, *Mol. Cell*, 2004, vol. 13, no. 1, pp. 91—100.
57. Spellman, R., Rideau, A., Matlin, A., et al., Regulation of alternative splicing by PTB and associated factors, *Biochem. Soc. Trans.*, 2005, vol. 33, no. 3, pp. 457—460. doi 10.1042/BST0330457
58. Sureau, A., Gattoni, R., Dooghe, Y., et al., SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs, *EMBO J.*, 2001, vol. 20, no. 7, pp. 1785—1796. doi 10.1093/emboj/20.7.1785
59. Stoilov, P., Daoud, R., Nayler, O., and Stamm, S., Human tra2-beta1 autoregulates its protein concentration by influencing alternative splicing of its pre-mRNA, *Hum. Mol. Genet.*, 2004, vol. 13, no. 5, pp. 509—524. doi 10.1093/hmg/ddh051
60. Le Guiner, C., Gesnel, M.-C., and Breathnach, R., TIA-1 or TIAR is required for DT40 cell viability, *J. Biol. Chem.*, 2003, vol. 278, no. 12, pp. 10465—10476. doi 10.1074/jbc.M212378200
61. Wilson, G.M., Sun, Y., Sellers, J., et al., Regulation of AUF1 expression via conserved alternatively spliced elements in the 3' untranslated region, *Mol. Cell. Biol.*, 1999, vol. 19, no. 6, pp. 4056—4064.
62. Menegay, H.J., Myers, M.P., Moeslein, F.M., and Landreth, G.E., Biochemical characterization and localization of the dual specificity kinase CLK1, *J. Cell Sci.*, 2000, vol. 113, part 18, pp. 3241—3253.
63. Lim, K.H., Ferraris, L., Filloux, M.E., et al., Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes, *Proc. Natl. Acad. Sci. U.S.A.*, 2011, vol. 108, no. 27, pp. 11093—11098. doi 10.1073/pnas.1101135108
64. Wang, G.-S. and Cooper, T.A., Splicing in disease: disruption of the splicing code and the decoding machinery, *Nat. Rev. Genet.*, 2007, vol. 8, no. 10, pp. 749—761. doi 10.1038/nrg2164
65. Bedwell, D.M., Kaenjak, A., Benos, D.J., et al., Suppression of a CFTR premature stop mutation in a bronchial epithelial cell line, *Nat. Med.*, 1997, vol. 3, no. 11, pp. 1280—1284.
66. Liang, F., Shang, H., Jordan, N.J., et al., High-throughput screening for readthrough modulators of CFTR PTC mutations, *SLAS Technol.*, 2017, vol. 22, no. 3, pp. 315—324. doi 10.1177/2472630317692561
67. Clancy, J.P., Bebok, Z., Ruiz, F., et al., Evidence that systemic gentamicin suppresses premature stop mutations in patients with cystic fibrosis, *Am. J. Respir. Crit. Care Med.*, 2001, vol. 163, no. 7, pp. 1683—1692. doi 10.1164/ajrccm.163.7.2004001
68. Barton-Davis, E.R., Cordier, L., Shoturma, D.I., et al., Aminoglycoside antibiotics restore dystrophin function to skeletal muscles of mdx mice, *J. Clin. Invest.*, 1999, vol. 104, no. 4, pp. 375—381.
69. Politano, L., Nigro, G., Nigro, V., et al., Gentamicin administration in Duchenne patients with premature stop codon: preliminary results, *Acta Myol. Myopathies Cardiomyopathies: Off. J. Mediterr. Soc. Myol.*, 2003, vol. 22, no. 1, pp. 15—21.
70. Mankodi, A., Takahashi, M.P., 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, vol. 10, no. 1, pp. 35—44.
71. Vittal, P., Pandya, S., Sharp, K., et al., Antisense FMR1 splice variant and loss of AGG interruptions are predictors of Fragile X-associated tremor/ataxia syndrome (FXTAS) (P3.002), *Neurology*, 2017, vol. 88, no. 16, suppl., p. 3.002.
72. Kumari, D. and Usdin, K., Is Friedreich ataxia an epigenetic disorder?, *Clin. Epigenet.*, 2012, vol. 4, no. 1, p. 2. doi 10.1186/1868-7083-4-2
73. Wongtrakoongate, P., Riddick, G., Fucharoen, S., et al., Association of the long non-coding RNA steroid receptor RNA activator (SRA) with TrxG and PRC2 complexes, *PLoS Genet.*, 2015, vol. 11, no. 10. e1005615. doi 10.1371/journal.pgen.1005615
74. Sampath, K. and Ephrussi, A., cncRNAs: RNAs with both coding and non-coding roles in development, *Development*, 2016, vol. 143, no. 8, pp. 1234—1241. doi 10.1242/dev.133298
75. Leygue, E., Dotzlaw, H., Watson, P.H., et al., Expression of the steroid receptor RNA activator in human breast tumors, *Cancer Res.*, 1999, vol. 59, no. 17, pp. 4190—4193.
76. Murphy, L.C., Simon, S.L., Parkes, A., et al., Altered expression of estrogen receptor coregulators during human breast tumorigenesis, *Cancer Res.*, 2000, vol. 60, no. 22, pp. 6266—6271.
77. Gardner, L.B., Nonsense mediated RNA decay regulation by cellular stress; implications for tumorigenesis, *Mol. Cancer Res.*, 2010, vol. 8, no. 3, pp. 295—308. doi 10.1158/1541-7786.MCR-09-0502
78. Perrin-Vidoz, L., Sinilnikova, O.M., Stoppa-Lyonnet, D., et al., The nonsense-mediated mRNA decay pathway triggers degradation of most BRCA1 mRNAs bearing premature termination codons, *Hum. Mol. Genet.*, 2002, vol. 11, no. 33, pp. 2805—2814.
79. Kawasaki, T., Tomita, Y., Watanabe, R., et al., mRNA and protein expression of p53 mutations in human bladder cancer cell lines, *Cancer Lett.*, 1994, vol. 82, no. 1, pp. 113—121.
80. Liu, C., Karam, R., Zhou, Y., et al., The UPF1 RNA surveillance gene is commonly mutated in pancreatic adenosquamous carcinoma, *Nat. Med.*, 2014, vol. 82, no. 1, pp. 596—598.

Translated by N. Maleeva