Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 363 (2005) 51–60
GENE
SECTION
FUNCTIONAL GENOMICS
www.elsevier.com/locate/gene

Noncoding RNA of U87 host gene is associated with ribosomes and is relatively resistant to nonsense-mediated decay

Julia A. Makarova, Dmitri A. Kramerov *

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russia

Received 15 June 2005; received in revised form 17 August 2005; accepted 18 August 2005
Available online 14 October 2005
Received by S.M. Mirkin

Abstract

Non-coding RNAs are involved in many cellular processes. In particular, most of C/D box small nucleolar RNAs (snoRNAs) function as guide RNAs in site-specific 2′-*O*-methylation of rRNAs. While most snoRNA genes reside in introns of protein-coding genes, here we demonstrated an unusual snoRNA gene occupying an intron of a previously unknown non-protein-coding gene U87HG. We characterized this host gene in human, mouse, rat, and dog. It is a member of 5′TOP gene family, which includes many translation apparatus genes. U87HG RNA carried multiple stop-codons and was associated with ribosomes, suggesting that it may be a target for nonsense-mediated mRNA decay (NMD), a process that eliminates transcripts bearing nonsense mutations. Surprisingly, we found that U87HG RNA was hardly susceptible to NMD. Possible mechanisms (translation reinitiation, ribosomal leaky scanning, and low efficiency of translation) of this phenomenon are discussed. Unlike transcripts of four other known non-protein-coding host genes, U87HG RNA shows a relatively high degree of conservation suggesting a selective pressure and a possible functional activity of U87HG apart from producing U87 snoRNA.

Keywords: ncRNA; snoRNA; 5′TOP gene; NMD

1. Introduction

Recently, a growing number of non-coding RNAs (ncRNAs) have been identified in eukaryotic cells. Functions of many of them have been established and turned out to be surprisingly diverse. However, functions of a large number of ncRNAs remain enigmatic and probably even more ncRNAs are waiting to be discovered. The progress in this field is retarded because many ncRNAs are undetectable by gene-finding programs and because unprecedented functions of these RNAs are hard to uncover.

Nucleoli contain a subset of small noncoding RNAs involved in different steps of rRNA processing. Based on their structure these small nucleolar RNAs (snoRNAs) fall into two classes: C/D box snoRNAs and H/ACA box snoRNAs. Vast majority of the C/D box snoRNAs function as guide RNAs in site-specific 2′-*O*-methylation of rRNAs whereas most of the H/ACA box snoRNAs direct site-specific pseudouridylation of rRNAs (Bachellerie et al., 2002).

In vertebrates, almost all of the snoRNA genes reside within introns of protein-coding genes. Genes hosting snoRNAs usually encode nucleolar proteins or proteins involved in translation (e.g., elongation factors and ribosome proteins) (Maxwell and Fournier, 1995). With a few exceptions, mRNAs of these genes share a common structural feature: they begin with a C residue followed by a stretch of 4–14 pyrimidines (Pelczar and Filipowicz, 1998). Another common feature of these genes is their coordinated regulation at the translational level. Because of the presence of the 5′ terminal olygopyrimidine tract (5′TOP) they are referred to as 5′TOP genes (Amaldi and Pierandrei-Amaldi, 1997).

Strikingly, a few snoRNA host genes do not code for any protein. Although their RNAs are spliced and polyadenylated, they contain multiple stop-codons in all three reading

Abbreviations: ncRNA, noncoding RNA; snoRNA, small nucleolar RNA; UTR, untranslated region; SINE, short interspersed element; NMD, nonsense-mediated decay.

* Corresponding author. 31 Vavilov st., Moscow, 119991, Russia. Tel.: +7 95 1359997; fax: +7 95 1351405.
E-mail address: kramerov@eimb.ru (D.A. Kramerov).

0378-1119/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2005.08.010

frames. Only five such genes are known so far: UHG, U17HG, U19HG, gas5, and U50HG (Tycowski et al., 1996; Pelczar and Filipowicz, 1998; Bortolin and Kiss, 1998; Smith and Steitz, 1998; Tanaka et al., 2000). The latter gene was not well studied. The only conserved portion of these genes is the snoRNA sequences located in their introns. Transcripts of human and mouse orthologs differ in length and are not conserved in sequence. All these unusual features allowed Smith and Steitz (1998) to hypothesize that the only function of these genes is snoRNAs production.

Except for U19HG RNA, which is mostly located in nucleoplasm, the other three studied RNAs (UHG, U17HG, and gas5) are located in cytoplasm. UHG and gas5 RNAs are associated with ribosomes and are thought to degrade by nonsense-mediated mRNA decay system (NMD). NMD is an mRNA surveillance mechanism, which degrades mRNAs with nonsense mutations. This degradation requires translation of the nonsense-contained message (Maquat, 2004).

Previously we have described a novel C/D box snoRNA U87 (Gogolevskaya et al., 2002).[1] Here we showed that U87 gene is located within an intron of a novel non-protein-coding gene, named U87HG (host gene). U87HG is a housekeeping gene of the 5′TOP family. We characterized human, mouse and rat U87HG RNAs and showed their association with ribosomes. Unlike transcripts of other non-protein-coding host genes, U87HG RNA is more conserved in sequence. Pattern and degree of conservation of the U87HG RNA resemble those of UTRs of protein-coding genes. Although U87HG RNA is translated, it is very ineffectively degraded by NMD. We suggest that U87HG serves not only to produce U87 RNA but carries other functions, in particular, it may be involved in translation or its regulation.

2. Materials and methods

2.1. Drugs treatment, cell fractionation, and RNA isolation

Cells were incubated with cycloheximide (20 μg/ml), puromycin (200 μg/ml) or actinomycin D (5 μg/ml) for time indicated. HeLa cells were fractionated into nuclear and cytoplasmic fractions as described (Lemieux and Beaud, 1982). Nuclear RNA was isolated by extraction with hot (65 °C) phenol saturated with sodium acetate, pH 5.0. Cytoplasmic RNA was isolated by phenol (saturated with Tris–HCl, pH 8.5)/chloroform extraction. Total RNA was isolated by the guanidinium thiocyanate–phenol–chloroform extraction.

2.2. Analysis of cell extracts on sucrose gradients

HeLa or SC1 cells (2×10^7 cells) were treated with 0.05 μg/ml cycloheximide for 30 min to enhance translation of 5′TOP mRNAs (Pierandrei-Amaldi et al., 1991; Walden et al., 1981). Cycloheximide (100 μg/ml) was added to the cells for 5 min prior harvesting. No drugs were added to the control cells. Cells were homogenized in 2 ml of lysis buffer (50 mM Tris–HCl, pH 8.0; 0.1 M NaCl; 5 mM $MgCl_2$; 0.5% NP40; 2000 U RNAsin) with 15 strokes of a tight-fitting Dounse homogenizer. At this step EDTA (10 mM) or cycloheximide (100 μg/ml) was added to the lysates except for the control lysate. After 2 min incubation on ice the lysates were centrifuged for 15 min at 4 °C at 12 000 g. Heparin (0.1 mg/ml) was added to each supernatant. The supernatants (1 ml) were loaded onto 10 to 50% linear sucrose gradients prepared in 50 mM Tris–HCl, pH 8.0 – 0.1 M NaCl –5 mM $MgCl_2$ –0.1 mg/ml heparin. For EDTA and cycloheximide treated lysates, the same sucrose gradients were prepared with EDTA (10 mM) or cycloheximide (100 μg/ml), respectively. The gradients were centrifuged at 4 °C for 15 h at 23 000 rpm in an SW40 rotor (Beckman). Optical density (D_{260}) of fractions was measured and RNA was isolated from individual fractions by phenol–chloroform extraction.

2.3. RACE analysis of human and rat U87HG RNA

Human or rat poly(A)$^+$ RNA were utilized for cDNA synthesis using M-MLV RT (“Gibco-BRL”) and $XbaT_{15}$ primer GTCGACTCTAGATTTTTTTTTTTTTT. cDNA was purified by electrophoresis in 4% NuSieve agarose gel.

Primers for U87HG cDNA were designed to get overlapping 5′ RACE and 3′ RACE products. In 3′ RACE, human cDNA was amplified by semi-nested PCR with primer $XbaT_{15}$ and primer #1 TGGCATGCAGCATTTTGA for the first round, and $XbaT_{15}$ and primer #2 T*GGATCC*AGGAAAATTGAAGACGTG (*Bam*HI site is underlined) for the second round. Primers used for PCR with rat cDNA were $XbaT_{15}$ and #3 (*Xho*I site is underlined): T*CTCGAG*TGCAACAAAACCTGTCT.

For 5′ RACE 3′ end of human or rat cDNA was tailed with poly(dG) using terminal deoxynucleotidyl transferase (“USB-Amersham”). Human cDNA was then amplified by semi-nested PCR with the following primers: $EcoRC_{10}$ CG*GAATTC*GTCCCCCCCCCC (*Eco*RI site is underlined) and primer # 4 GAAAACACTTTATTGTTCAG for the first round and primer $EcoRC_{10}$ and primer #5 T*CTCGAG*CTCCATAATTTATTGTGA (*Xho*I site is underlined) for the second round. Primers used for PCR with rat cDNA were $EcoRC_{10}$ and #6 T*CTCGAG*ACCCAAAACAACGAGGATT (*Xho*I site is underlined). Human and rat PCR products were cloned into pBS (SK)+vector. Obtained libraries were screened by hybridization with probes for human and rat U87HG (see Section 2.4).

2.4. Primer-extension and Northern analyses

Primer-extension was carried out as described previously (Gogolevskaya et al., 2002) with the following primers: #7 ACCACACCGGCATGACTAAC; #8 GCCCAAGGACACGCCGGGTTC, and #9 GCCGGAGGACACGCCGGGTTC complementary to exon 1 of human, mouse and rat U87HG gene, respectively.

Total, nuclear and cytoplasmic RNA were separated on a 1% or 1.5% agarose–formaldehyde gel, blotted onto a Hybond-N

[1] Huttenhofer et al. (2001) have cloned incomplete cDNA derived from this RNA and named it MBII-276 snoRNA.

nylon membrane, and hybridized with labeled probe under standard conditions. Hybridization signals were detected using X-ray film or Cyclone Storage Phosphor System.

To prepare a probe for human U87HG RNA, total genomic DNA or cloned human U87HG 3′RACE product was used in PCR with primers #2 and #4. The amplified PCR product was purified on a 4% NuSieve agarose gel and 0.5% of isolated DNA was reamplified in the reaction with the same primers and 25 μCi of α-[^{32}P] dATP. Probes for mouse and rat U87HG were prepared in the same way using previously cloned U87 gene locus (Gogolevskaya et al., 2002). The following primers were used: primers #3 and 5′ GGCAAATGTAGTCCTGGC specific for the second exon of mouse U87HG; primers 5′ ATGTA-GGTAGCTGCAGT and #6 for rat third U87HG exon. A probe for the U87 snoRNA was synthesized as described previously (Gogolevskaya et al., 2002).

2.5. Sequence analysis of U87HG and accession numbers

Multiple alignment of U87HG RNA sequences was performed with Clustal V (Higgins et al., 1992). Pairwise alignments of U87HG loci and U87HG RNA were generated using OWEN (Ogurtsov et al., 2002). Possible transcription factors binding sites were identified in U87HG promoter region using the Transcription Element Search System (http://www.cbil.upenn.edu/tess.) CpG island in U87HG promoter was predicted by CpG Island Searcher (Takai and Jones, 2002). RNA secondary structure predictions were made with the *mfold* program version 3.1. (Zuker, 2003).

The nucleotide sequences of human and rat U87HG RNA have been submitted to GenBank under accession nos. *AY264285* and *AY264286*, respectively. The nucleotide sequences of mouse and rat U87HG have been submitted to GenBank under accession nos. *AF396686* and *AF396685*, respectively.

3. Results

3.1. U87 snoRNA is encoded within an intron of a non-protein-coding gene

Previously we have identified a novel C/D box snoRNA called U87 that directs 2′-*O*-ribose methylation of a G_{3468} in rat 28S rRNA. We have also cloned and sequenced U87 gene-containing *Hin*d III/*Eco*RI fragments of rat and mouse DNA (3.0 and 3.5 kb, respectively) (Gogolevskaya et al., 2002). As snoRNAs are usually encoded within introns of protein-coding genes, we attempted to find such a gene within cloned mouse and rat U87 locus. Analysis of the cloned sequences revealed no exons of known protein-coding genes. Gene prediction programs (GenScan, Fgenesh++, and Twinscan) also failed to find a gene within the cloned fragments as well as in larger human, mouse, and rat U87 sequences deposited in GenBank. However, in EST division of GenBank we found multiple human, mouse and rat transcripts that matched U87 locus. In all cases U87 snoRNA gene was located in an intron of a hypothetical gene. To better characterize the structure of this gene, hereafter named U87HG (host gene), we performed 3′ and 5′ RACE analyses of U87HG RNA. PCR primers were designed so that the 3′ and 5′ RACE products overlapped. Mouse U87HG RNA sequence was deduced from alignment with rat U87HG RNA and from multiple mouse ESTs (UniGene cluster *Mm.28311*). 5′ termini of human, mouse, and rat cDNAs were additionally characterized by primer-extension analysis (see Section 3.2). Progress in dog genome sequencing project allowed us to deduce dog U87HG RNA sequence as well. Dog U87HG locus has been already sequenced, but no ESTs are still available. However, human U87HG RNA is readily aligned with dog U87HG locus. Human, mouse, rat and presumptive dog U87HG RNA sequences are shown in Fig. 1. U87HG gene of these four species consists of four exons, and U87 snoRNA is encoded within its second intron. The length of human, mouse, rat and dog U87HG RNA is 472, 472, 659, and 446 nt, respectively. Rat U87HG RNA is longer than the other orthologs due to the insertion of the B2 SINE in its fourth exon (Fig. 1). U87HG RNA was found to be polyadenylated. Human U87HG RNA possesses two polyadenylation signals in positions 394 and 444. The latter is utilized predominantly: it was used in 8 out of 9 sequenced clones. Rat U87HG RNA also has two polyadenylation signals (positions 566 and 640) but the first signal is utilized more frequently (in 4 out of 5 clones). There are two polyadenylation signals in mouse U87HG RNA and both are used in ESTs. Dog U87HG RNA has only one polyadenylation signal.

Surprisingly, U87HG RNA sequence is poorly conserved in comparison with mRNAs of protein-coding genes (Fig. 1). To better characterize the degree of its conservation, we performed pairwise alignments of human, mouse, rat and dog U87HG RNAs and calculated sequence similarity (Table 1). The mean 3′ UTRs sequence identity as well as the mean sequence identity in synonymous sites for these species are shown for comparison (Table 1). One can see that the degrees of U87HG RNA sequence similarity exceed that for selectively neutral sequences (although these values are close for the rat–mouse and human–dog pairs) and are very close to similarity between UTRs.

Human, mouse, rat, and dog U87HG RNAs contain multiple stop codons in all three reading frames (Fig. 2). The longest peptide that could be produced from U87HG RNA is 30, 54, 44 and 30 amino acids for human, mouse, rat and dog, respectively. Efficiency of the translation initiation is known to be affected by the sequences flanking an AUG codon—the so-called “context” (Kozak, 1986). The most conserved nucleotides are the R at −3 and the G at +4. A strong consensus sequence contains both of these important nucleotides, an adequate sequence contains only one of them, and a weak sequence loses both. In nearly all mRNAs, the AUG start codon occurs in a strong or at least an adequate context. Rare mRNAs in which the start codon is located in a weak context have the special property of initiation translation at the first, second, and, sometimes, downstream AUG codons due to leaky scanning mechanism (Meijer and Thomas, 2002). The longest open reading frame (ORF) of mouse U87HG RNA starts at the first AUG (position 42) located in a strong translation context

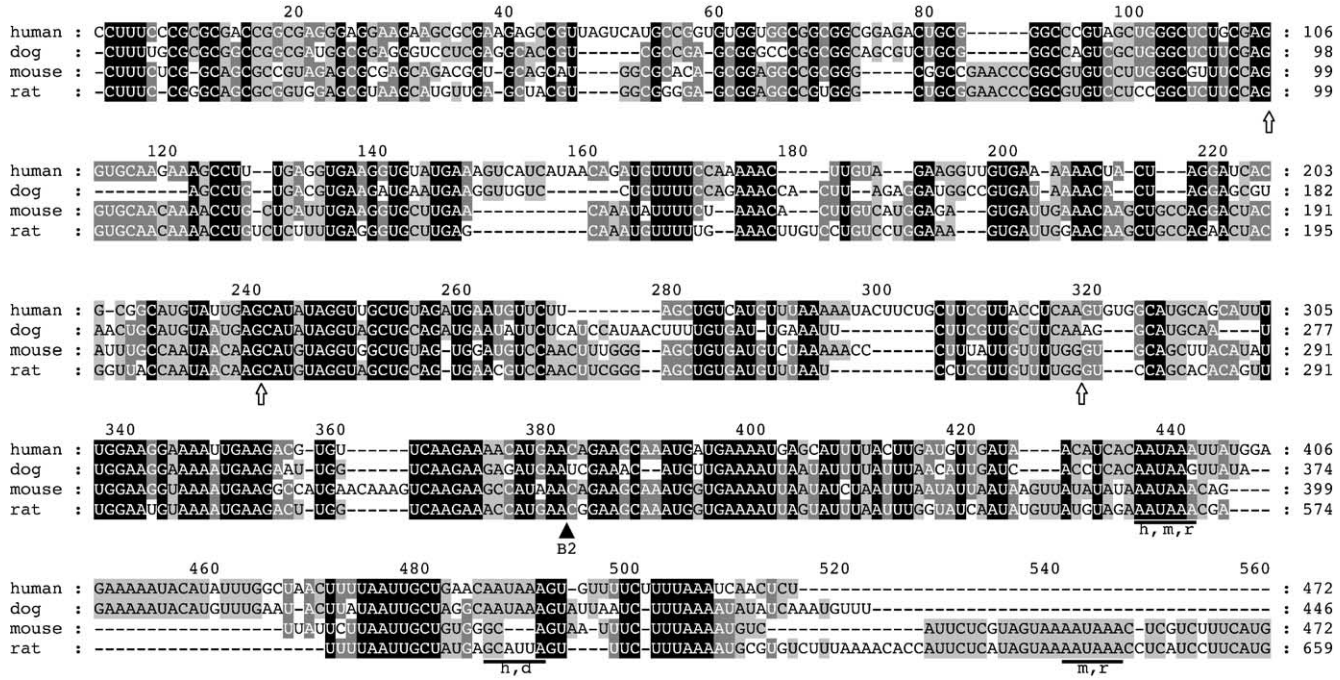

Fig. 1. Alignment of human, dog, mouse, and rat U87HG RNAs. Exon borders are marked by arrows. Polyadenylation signals in human (h), dog (d), mouse (m), and rat (r) RNAs are underlined. Location of B2 SINE, which was removed from the rat RNA sequence to save space, is indicated by a filled triangle.

GC*A*GC**AUG***G*. However, the longest ORF of rat RNA starts at the seventh AUG within the SINE B2, whereas the longest authentic ORF starts at the first AUG (position 30) predicted to be only an adequate translation start site (context UA-*A*GC**AUG***U)*. In human U87HG RNA, the longest ORF starts at the second AUG (position 136) surrounded by a weak translation context (GG*U*GU**AUG***A*). The longest ORF of dog U87HG RNA starts at the first initiation codon (position 20) located in a strong translation context CG*G*CG**AUG***G*. Thus, the longest ORF (sometimes not being the first one) starts in different translation contexts, and has different length in three of four species. All short peptides encoded by U87HG RNA are not conserved. None of these peptides produced any hits in protein databases searches. The same results were obtained for ORFs started with alternative initiation codons (GUG, UUG, CUG and ACG) (Spirin, 1999). Therefore, U87HG RNA hardly codes for any functionally important peptides and appears to be a non-protein-coding RNA.

Table 1
Sequence similarity (%) between U87HG RNAs from four mammalian species[a]

	Mouse	Rat	Dog
Human	**68**/*68*/44	**72**/*67*/43	**75**/*76*/73
Mouse	–	**83**/*86*/81	**71**/*64*/53
Rat	–	–	**69**/*62*/nd

[a] Human, mouse, rat, and dog U87HG RNAs were pairwise aligned and sequence similarity (bold figures) were calculated excluding indels. The mean sequence similarity (italic) for 3′ UTR of mRNAs (Xie et al., 2005) are shown for comparison. The third figure in each cell is the similarity between selectively neutral sequences (it was calculated as the average similarity in synonymous sites of mRNAs from the same four species (Kirkness et al., 2003; Gibbs et al., 2004).

3.2. U87HG is a housekeeping gene of 5′TOP family

The presence of the U87HG RNA in different tissues was studied by primer-extension analysis. Total rat RNA from liver, kidney, brain, and muscle was analyzed. The signal intensity was similar in all tissues (data not shown). The broad distribution of the U87HG RNA in human and mouse tissues was supported by multiple human and mouse ESTs from more than 15 tissues (UniGene clusters *Hs.372680* and *Mm.28311*). Thus, U87HG appears to be a housekeeping gene (Lercher et al., 2002). This is not surprising, because it codes for at least one

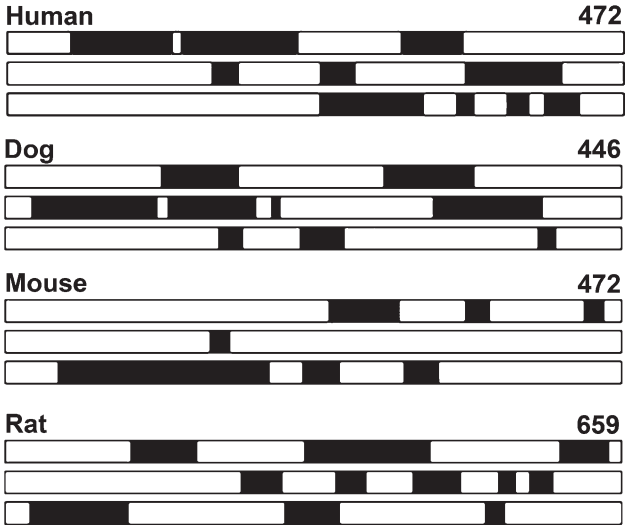

Fig. 2. U87HG has no long ORF. ORFs of U87HG RNAs from four mammalian species are shown as black boxes. The length (nt) of each RNA is indicated.

essential product, namely, the U87 snoRNA, which is required for ribosome biogenesis.

Primer-extension analysis was also used to characterize 5′ terminus of the U87HG RNA. Total RNA from human HeLa cells, mouse and rat liver was studied. Human U87HG RNA was found to start at two adjacent cytidine residues. Surprisingly, besides the cytidine start site that was found in cDNA sequences, rat and mouse U87HG RNAs appear to have heterogeneous 5′ ends (Fig. 3).

Promoter of the U87HG gene does not contain neither TATA box nor initiator sequence (5′)YYA^{+1}NT/AYYY(3′). It is located within a CpG island and contains Sp1-binding sites (Fig. 3). These features are typical of housekeeping gene promoters (Smale and Kadonaga, 2003). Housekeeping genes often have multiple transcription start sites thus yielding a subset of mRNAs with different 5′ termini. U87HG gene seems to be a rare example of a gene possessing only two adjacent transcription start sites in one mammalian species and multiple scattered start sites in two other species.

The vast majority of mRNAs start with a purine nucleotide. Human, mouse and rat U87HG cDNAs start with a C residue located within a pyrimidine tract followed by GC-rich sequences. This unusual 5′ end structure is a characteristic feature of 5′TOP family of housekeeping genes largely encoding ribosomal proteins and some translation factors (Amaldi and Pierandrei-Amaldi, 1997). Many genes harboring snoRNAs in their introns also belong to the 5′TOP family (Pelczar and Filipowicz, 1998). Involvement of the vast majority of known snoRNAs in rRNA modification led to speculation that being the members of the 5′TOP family, the snoRNA host genes allow the cell to coordinate the regulation of genes involved in translation (Maxwell and Fournier, 1995).

3.3. U87HG RNA is located in the cytoplasm and is stable

We investigated the cellular localization of the U87HG RNA. Human HeLa cells were fractionated into nuclear and cytoplasmic fractions. U87HG RNA was detected mainly in the cytoplasm (Fig. 4).

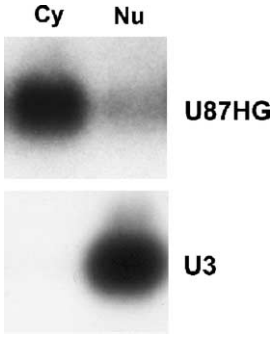

Fig. 4. Cellular localization of U87HG RNA. HeLa cells were fractionated into cytoplasmic (Cy) and nuclear (Nu) fractions. Total cell RNA was analyzed by Northern hybridization with probes for U87HG RNA and U3 snoRNA (control).

Non-protein-coding host genes are believed to serve only to produce snoRNAs and their transcripts are shown to rapidly degrade in cytoplasm (Tycowski et al., 1996). To test whether U87HG RNA is rapidly degraded, we treated HeLa cells with transcription inhibitor actinomycin D (Fig. 5). UHG RNA, another non-protein-coding RNA, was used as a control (Tycowski et al., 1996). Introns of UHG gene encode snoRNAs similar to U87HG. Short-lived histone H1F1 mRNA served as another control. In contrast to rapidly degraded H1F1 mRNA, U87HG RNA appears to be very stable (with half-life more than 22 h) (Fig. 5). To our surprise, UHG RNA, which was previously reported to be unstable in the presence of actinomycin D (Tycowski et al., 1996), turned out to be quite stable. The same results were obtained for human MCF-7, mouse SC1 and rat Rat1 cells (data not shown).

3.4. U87HG RNA is associated with ribosomes

As the U87HG RNA possesses characteristic features of mRNAs except long ORFs, we looked to see whether it is translated. Cytoplasmic extracts were prepared from human

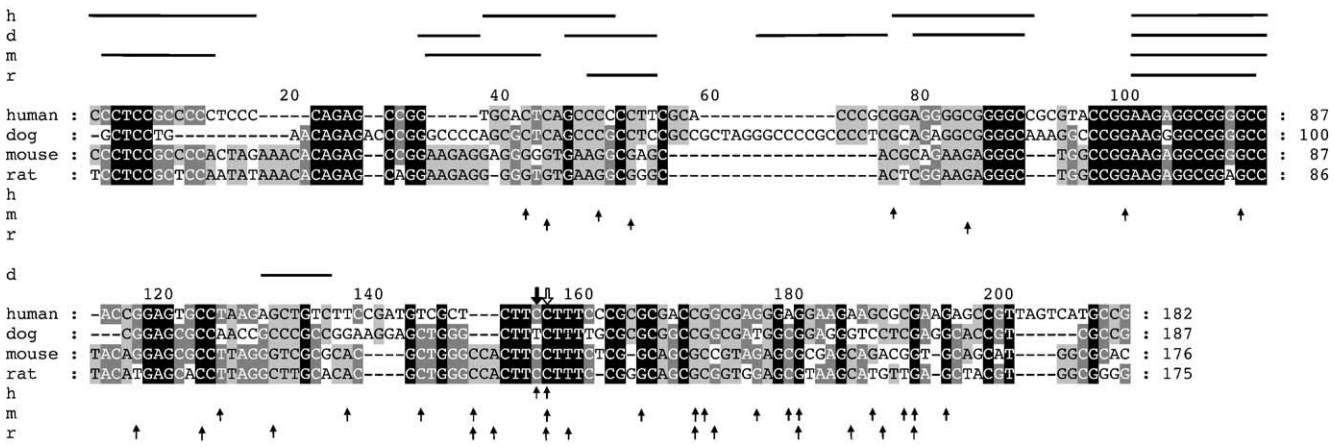

Fig. 3. Alignment of U87HG promoter region. Sp1 binding sites predicted for human (h), dog (d), mouse (m), and rat (r) U87HGs are marked by lines above the alignment. The position corresponding to the 5′ end of U87HG RNA established by cDNA cloning is indicated by filled (human) and empty (rat) arrows (above the alignment). Transcription start sites in human (h), mouse (m), and rat (r) U87HGs determined by primer extension are marked by thin arrows (bellow the alignment).

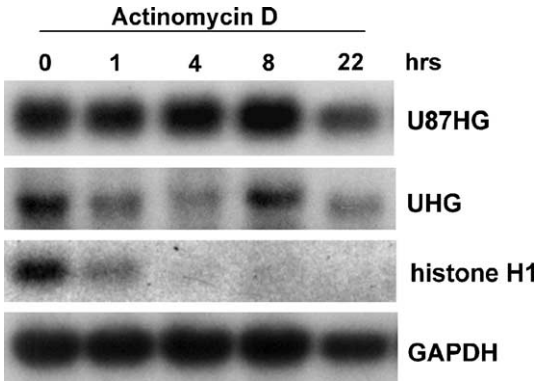

Fig. 5. Analysis of U87HG RNA and UHG RNA stability. HeLa cells were treated with transcription inhibitor actinomycin D (5 μg/ml) for the time indicated. Equal amount (10 μg) of total cellular RNA was analyzed by Northern hybridization with the indicated probes. Short-lived (histone H1F1) and long-lived (GAPDH) mRNAs (Sabath et al., 1990) were used as controls.

HeLa cells and analyzed on sucrose gradients. Since the U87HG RNA contains only short ORFs, it is expected to be associated predominantly with monosomes. To better distinguish between ribosome-associated and free U87HG RNA, we used 10% to 50% sucrose gradient and extended centrifugation time (see Section 2.2). While major part of the U87HG RNA was found close to the top of the gradient, a fraction of it was associated with monosomes. This fraction was released after the EDTA treatment of the extracts resulting in ribosomes dissociation (Fig. 6). Phosphorimager analysis showed that 24% of U87HG RNA was associated with ribosomes in HeLa cells. Thus, U87HG RNA is not translated very efficiently. The same observation has been made for other 5′TOP mRNAs: on average, depending on intensity of cell proliferation only 25–65% of them are engaged with ribosomes compared with 90% of the other housekeeping mRNAs (Amaldi and Pierandrei-Amaldi, 1997). As has been shown previously, treatment of cells with low concentration (0.05 μg/ml) of translation elongation inhibitor cycloheximide results in increase of 5′TOP mRNA fraction associated with ribosomes

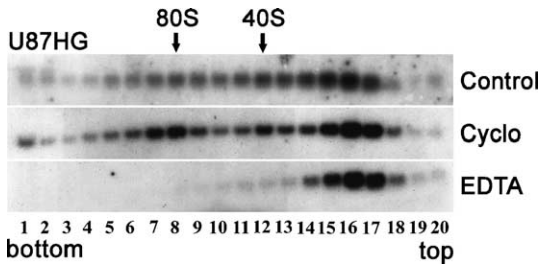

Fig. 6. Distribution of U87HG RNA on sucrose gradient after centrifugation of HeLa cytoplasmic extracts. The cytoplasmic fraction was isolated from 10^7 HeLa cells untreated (control) or treated (cyclo) with low concentration (0.05 μg/ml) of cycloheximide for 30 min. The same amount of cytoplasmic fraction from control cells was treated with 10 mM EDTA. The extracts were centrifuged through a 10% to 50% sucrose gradient. RNA from individual fractions was analyzed by Northern hybridization with U87HG RNA-specific probe. Fraction #1 is RNA from pellet on centrifuge tube bottom. RNA from fractions 1 to 9 was considered as ribosome-associated. Positions of 80S ribosomes and 40S ribosomal subunits are indicated above.

(Pierandrei-Amaldi et al., 1991; Walden et al., 1981). We incubated HeLa cells in 0.05 μg/ml cycloheximide for 0.5 h before harvesting. Indeed, after this treatment a fraction of U87HG RNA cosedimenting with ribosomes increased up to 35% (Fig. 6). In mouse SC1 cells, a fraction of U87HG RNA engaged with ribosomes increased from 17% in control cells to 30% in cycloheximide treated cells (data not shown). This effect of low cycloheximide concentration is typical of templates that do not initiate very frequently (Walden et al., 1981). Thus, U87HG RNA like other 5′TOP RNAs appears to be rather a 'weak' template for translation.

3.5. Effect of translation inhibitors on U87HG RNA concentration

A special mechanism, called NMD (*n*onsense-*m*ediated mRNA *d*ecay) that recognizes mRNA with premature stop-codons and degrades them has been described (Maquat, 2004). This degradation requires ongoing translation. As U87HG RNA contains multiple stop-codons and is associated with ribosomes, it is predicted to be a target for NMD. Amount of an mRNA degraded by NMD increases when translation is inhibited (Maquat, 2004 and references therein). To test whether U87HG RNA is a target for NMD, we treated HeLa cells with translation elongation inhibitors puromycin and cycloheximide. Besides U87HG RNA we analyzed UHG RNA that has been shown to get accumulated after cycloheximide treatment, to be associated with ribosomes and, therefore, is thought to be degraded by NMD (Tycowski et al., 1996). To our surprise, in contrast with UHG RNA, U87HG RNA was barely accumulated in HeLa cells after translation was halted (Fig. 7). Similar data were obtained for human MCF-7, mouse SC1 and rat Rat1 cells (data not shown). These results suggest that U87HG RNA is less sensitive to NMD than UHG RNA.

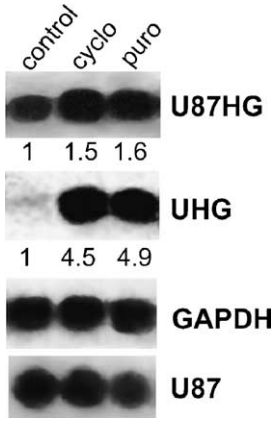

Fig. 7. Effect of translation inhibition on the level of U87HG and UHG RNAs. HeLa cells were treated with 20 μg/ml of cycloheximide (cyclo) or 200 μg/ml puromycin (puro) for 4h before collection. Total cellular RNA (10 μg) was resolved on 1.5% agarose formaldehyde gel, transferred onto nylon membrane and hybridized with probes specific for U87HG and UHG RNAs, GAPDH mRNA (loading control), and U87 snoRNA. Blots were analyzed using a Phosphorimager. Numbers show increase of hybridization signals relative to control.

Susceptibility of those two RNAs to nonsense-mediated decay, although very different, seems to disagree with stability of U87HG and UHG RNAs demonstrated in the experiment with actinomycin D. However, treatment of cells with actinomycin D was previously shown to block NMD (Kessler and Chasin, 1996). Taken together these two sets of data may indicate that NMD is the key mechanism of degradation of those RNAs.

3.6. Different association of U87HG RNA and UHG RNA with ribosomes contributes to the difference of their response to cycloheximide treatment

To address the question whether different response of U87HG and UHG RNAs to cycloheximide treatment is related to the difference in their association with ribosomes, we determined the fraction of UHG RNA cosedimenting with ribosomes and compared it with the corresponding fraction of U87HG RNA. We were also interested to discover whether U87HG and UHG RNAs accumulated following inhibition of translation are actually associated with ribosomes. To this end, HeLa cells were treated with cycloheximide for 4 h and cytoplasmic extracts were fractionated on the sucrose gradient. Here, we used the high concentration (20 μg/ml) of cycloheximide that efficiently inhibits translation elongation and abolish NMD (Smith and Steitz, 1998). We found that in cells treated with cycloheximide, the amount of both U87HG and UHG RNAs increased only in fractions cosedimenting with ribosomes, and for UHG RNA this effect was more pronounced (Fig. 8).

After incubation with cycloheximide the portion of U87HG RNA associated with ribosomes increased threefold (from 22% up to 61%), although total amount of U87HG RNA did not change significantly. This increase was mostly due to redistribution between free and ribosome-associated U87HG RNA (Fig. 8). Thus, free U87HG RNA is not irreversibly sequestered from translation.

41% of UHG RNA was found in association with ribosomes in control cells. After cycloheximide treatment, the fraction of UHG RNA loaded on ribosomes increased to 76%, and all UHG RNA accumulated for 4 h was observed in ribosome fractions (Fig. 8). This is consistent with the current view that cycloheximide abolishes both translation and coupled degradation of NMD targets (Maquat, 2004).

U87HG RNA and UHG RNA are not translated very efficiently when compared with most mRNAs. However, in control cells the amount of U87HG RNA associated with ribosomes is two times less than that of UHG RNA indicating that U87HG RNA is translated less efficiently. That may contribute to a less pronounced accumulation of U87HG RNA after inhibition of translation, as compared with UHG RNA.

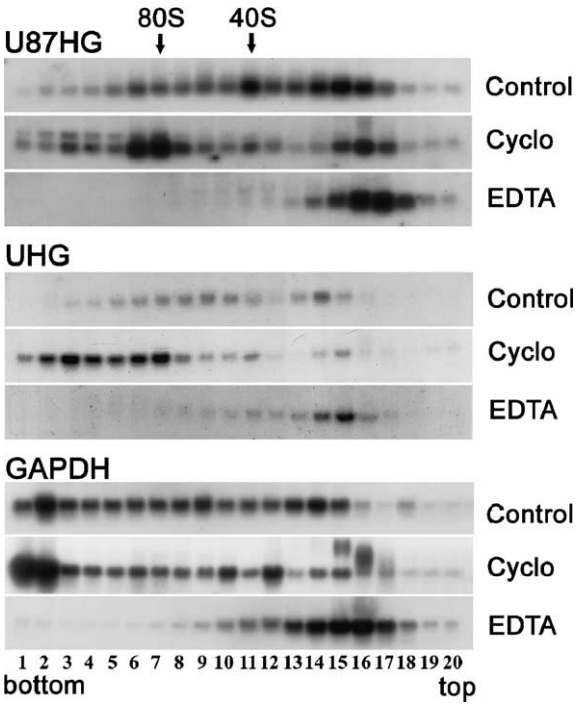

Fig. 8. Different portions of U87HG RNA and UHG RNA are associated with ribosomes. Cytoplasmic fraction of Hela cells untreated (control) or treated (cyclo) with high concentration (20 μg/ml) of cycloheximide for 4 h was isolated. The same amount of cytoplasmic fraction from control cells was treated with 10 mM EDTA. The extracts were centrifuged and RNA isolated from individual fractions were analyzed by Northern hybridization with U87HG, UHG, and GAPDH probes. The latter was used as a control, to demonstrate distribution of a typical mRNA. For details, see Fig. 5.

4. Discussion

A novel non-protein-coding gene, called U87HG, was described for humans, mouse, rat and dog. U87HG is a housekeeping gene of 5′ TOP family and encodes for two noncoding RNAs. The first one, U87 C/D box snoRNA (Gogolevskaya et al., 2002), is processed from its second intron. Another one, U87HG RNA, is encoded by the exons of U87HG. Although U87HG RNA does not contain long ORFs, it is polyadenylated, spliced and its fraction is associated with ribosomes. As U87HG RNA contains multiple stop-codons and its substantial portion is associated with ribosomes, it should be a subject of the nonsense-mediated decay (NMD), an RNA surveillance mechanism that degrades mRNAs with premature stop codons.

During the splicing of an mRNA, an exon–exon junction complex (EJC) consisting of at least five proteins is deposited 20 to 24 nt upstream of each exon–exon junction. Some of the EJC components remain associated with the mRNA in the cytoplasm. During the first round of translation, the EJCs are displaced from the transcript by ribosomes. However, if translation is prematurely terminated, the EJC proteins remain associated with the mRNA and act as markers for the NMD degradation factors, some of which (e.g., Upf1) are thought to be associated with post-terminated 40S subunit. Only stop-codons located more than 50–55 nt upstream of an exon–exon junction are generally recognized as premature. Translation inhibitors are known to abolish NMD. Therefore, mRNAs subjected to NMD accumulate when translation is inhibited (reviewed in Maquat, 2004).

To our surprise, when translation in HeLa cells was halted, U87HG RNA was barely accumulated, in contrast with UHG RNA. This implies either a more efficient translation of UHG RNA or the ability of U87HG RNA to escape NMD.

We found that the fraction of U87HG RNA loaded on ribosomes is two times less than that of UHG RNA. This implies that U87HG RNA is translated less efficiently than UHG RNA. Therefore, different rate of translation seems to contribute to the difference in susceptibility to NMD.

As was mentioned above, mRNAs containing premature stop-codons are degraded by NMD after the first round of translation. Therefore, all U87HG RNA molecules associated with ribosomes (20% of total) should undergo the first round of translation and then degrade. Translation is known to be a rapid process: typical rate of the eukaryotic elongation is 0.1 to 0.5 s per codon (Spirin, 1999). It means that all U87HG RNA molecules (as well as UHG RNA) present in cell at a given moment should undergo translation for much less period of time than 4 h. Therefore, after inhibition of translation during four hours, one could expect a significant increase of the U87HG RNA level. However, in contrast with UHG RNA, the level of U87HG RNA did not change significantly. This allowed us to suggest that low susceptibility of U87HG RNA to NMD is due not only to its relatively low level of translation. Below we hypothesize that different susceptibility of U87HG RNA and UHG RNA to NMD is partially determined by the differences in their ORF patterns.

Inefficient degradation of U87HG RNA by NMD may be due to reinitiation of translation on downstream ORFs. It was shown, that translation reinitiation abrogates NMD (Zhang and Maquat, 1997). Recently, it has become clear that translation reinitiation is not a rare event: about 40% of cellular mRNAs possess one or more upstream ORFs (uORFs) (reviewed in Peri and Pandey, 2001). uORFs generally encode short conservative peptides up to 100 amino acids. In order to translate the main ORF, ribosome has to reinitiate translation, sometimes more than once. Efficiency of reinitiation is usually quite low and depends on different factors, many of which have not been well studied yet (Morris and Geballe, 2000). The lengths of the uORF is one of the most important factors: the longer the uORF is, the less efficiently the main ORF is translated. For example, lengthening uORF from 13 to 33 codons causes a 3-fold decrease in reinitiation rate (Kozak, 2001), possibly because of dissociation of translation initiation factors. The distance between ORFs is also important: 50–150 nt usually produce maximum reinitiation rate (Kozak, 1987), probably because in this case ribosome has enough time to reacquire initiation complex with Met-tRNA. However, short spacers (e.g., 11 nt) (Kozak, 1987) can also provide sufficient reinitiation rate as demonstrated for a number of mRNAs. Interestingly, shortening of the spacer from 7 to 6 and even 4 nt proved to substantially increase translation of the main ORF, suggesting that the efficiency of reinitiation depends on this sequence (Lincoln et al., 1998). Moreover, Wang and Rothnagel (2004) observed that the 12 nt spacer provided more efficient reinitiation than the longer 52 nt spacer. Translation context of the uORFs also contributes to the efficiency of reinitiation. Weak and, to a lesser degree, suboptimal contexts allow ribosomal leaky scanning. For example, in constructs containing up to five uORFs located in a weak context, 50–65% of 40S subunits are still able to initiate at the main, sixth, ORF (Wang and Rothnagel, 2004).

In U87HG RNA, the first ORF (80 nt long) is located in suboptimal context, while two subsequent ORFs (90 and 27 nt) spaced by 6 and 10 nt, respectively, are in a weak context. We speculate that similarly to many other mRNAs, after translation of the first ORF some ribosomes could reinitiate on downstream ORFs of U87HG RNA. It was shown, that after translation of 84 nt uORF, efficiency of reinitiation was 50% (Luukkonen et al., 1995). Translation efficiency for downstream ORFs of U87HG RNA can be further enhanced by leaky scanning, because the ORFs discussed are located in a weak or suboptimal context. After translation of the third ORF, ribosome is located 24 nt upstream of the last exon–exon junction. Once in such close vicinity to the last exon–exon junction, stop codons fail to trigger NMD (Maquat, 2004). Therefore, a portion of U87HG RNA becomes immune to NMD (Fig. 9). Yet a fraction of U87HG RNA is still subjected to NMD, since reinitiaton is not a very efficient process.

In UHG RNA translation contexts of the first and second ORFs are the same as in U87HG RNA. However, the first ORF of UHG RNA is longer: 130 nt. An uORF of comparable length (120 nt) was shown to almost completely block translation of a downstream ORF (Luukkonen et al., 1995). Thus, one may expect that reinitiation rate after the translation of the first ORF of UHG RNA will be extremely low. Even if reinitiation occurs, and the first and second ORFs of UHG RNA is translated, the 40S subunit will face with four EJCs scattered over a 160 nt spacer without ORFs (Fig. 9). Therefore, UHG RNA should be

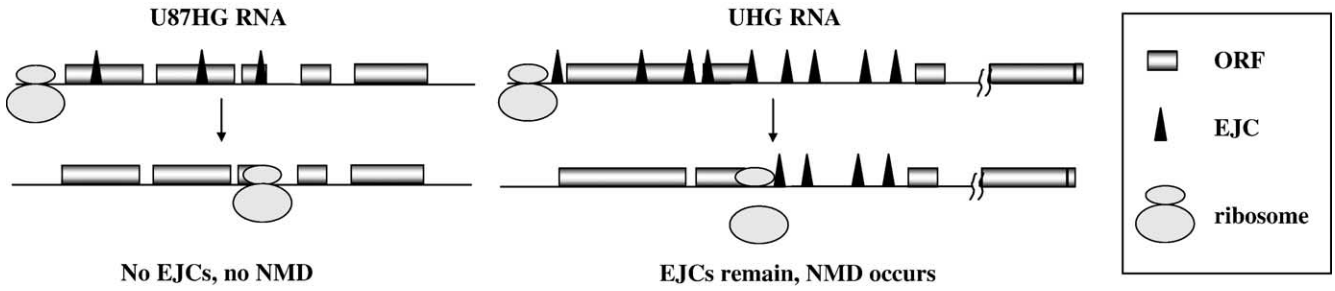

Fig. 9. Scheme illustrating different susceptibility of U87HG and UHG RNA to nonsense-mediated decay (NMD). In the course of translation ribosome can remove all exon junction complexes (EJCs) from U87HG RNA but not from UHG RNA resulting in less pronounced NMD of U87HG RNA. For details, see Section 4.

subjected to NMD much more efficiently than U87HG RNA and this suggestion is substantiated by the experimental data.

We believe that the difference in the susceptibility of U87HG and UHG RNAs to NMD is due to their different involvement in translation and to the ability of a fraction of U87HG RNA to escape NMD. Nevertheless, we cannot exclude that U87HG RNA form stable secondary structures which prevent elongation. However, we believe this is not the case for U87HG RNA, because search for conservative secondary structure elements in U87HG RNA did not produce any significant hits. We also cannot rule out the possibility that one of the peptides encoded by U87HG RNA slows translation down and prevents termination (which is known to be required to trigger NMD (Maquat, 2004)). For some mRNAs, ribosome stalling during either elongation or termination phase of the uORF translation is described (Morris and Geballe, 2000). Yet, it seems unlikely for U87HG RNA because peptides encoded by it are not conserved and because after inhibition of translation in mouse and rat cells it accumulates to the same degree as in human cells.

Currently, only five genes similar to U87HG are known. These are UHG, U17HG, U19HG, gas5 and U50HG (Bortolin and Kiss, 1998; Pelczar and Filipowicz, 1998; Smith and Steitz, 1998; Tanaka et al., 2000; Tycowski et al., 1996). Introns of these genes encode snoRNAs, and their spliced and polyadenylated transcripts appear to be non-protein-coding. Only two of them (UHG and gas 5) are completely independent transcription units, whereas a number of U17HG and U19HG transcripts are spliced with downstream protein-coding mRNAs of RCC1 and matrin 3, respectively. All non-coding snoRNA host genes studied are not conserved. Sequence similarity between exons of human and mouse orthologs is about 50% (gaps were excluded from all identity calculations), and, moreover, the orthologs differ in length 1.5–5 times. A similar degree of sequence divergence is characteristic of the alignable parts of human and mouse selectively neutral sequences (Gibbs et al., 2004). These observations allowed Smith and Steitz (1998), and Pelczar and Filipowicz (1998) to propose that processed noncoding host gene transcripts have no function. It should be noted, however, that host genes contain a few conserved short stretches, which in principal may be the functional sites of these RNAs.

Unlike the above mentioned noncoding host gene transcripts, which are hard to align due to multiple insertions and deletions (indels) and sequence divergence (data not shown), U87HG RNAs from the four species can be readily aligned (Fig. 1). Human and mouse U87HG RNA has the same length of 472 nt, the size of the dog U87HG RNA is similar (446 nt). Rat U87HG RNA is longer (659 nt) because of the insertion of B2 SINE. Thus, excluding rat, in other species the size of U87HG RNA is very similar. Notably, the larger size of the rat U87HG RNA is due to a single SINE insertion, rather than to multiple indels like in other non-coding host genes. Although the SINE invasions in exons are mostly harmful, there are a lot of examples of SINEs residing in 3′ UTRs of functional protein-coding genes (Pesole et al., 2002). Sometimes SINEs become even functional, providing the polyadenylation signal (Ryskov et al., 1983). Thus, SINE insertion in the exon sequence may be neutral or even useful.

The observed degree of similarity in pairwise alignments of human, mouse, rat, and dog U87HG RNA (Table 1) is higher than that of selectively neutral sequences of these species determined using Ks, which is the number of substitutions per a synonymous site (Gibbs et al., 2004; Kirkness et al., 2003). Thus U87HG exons seem to be under selective pressure.

We performed pairwise alignments for human, mouse, rat and dog U87HG genomic locus 100 kb in length (the size of U87HG gene varies from 2 to 3 kb). Human, rodent, and dog exons of U87HG were aligned with 65–90% sequence similarity whereas introns were aligned poorly if any (30–50% of similarity), except for intronic U87 snoRNA gene. This result also suggests the existence of selective constraints on U87HG RNA.

The degree of sequence conservation of U87HG RNA is similar to that of 3′ UTRs of protein-coding genes (Table 1). Like UTRs, U87HG RNAs of the four species discussed contain stretches with local similarity of 70–100% (Fig. 1). In UTRs these short fragments with high local similarity are usually the functional sites (Kuersten and Goodwin, 2003). For example, binding of the HuR protein to AUUUA sites stabilizes mRNA (Brennan and Steitz, 2001). It may also be true for U87HG RNA in which case some specific proteins could bind to the conservative regions. Interestingly, the insertion of B2 SINE into the rat U87HG RNA caused no changes in the pattern of conservative stretches of flanking sequences (Fig. 1). This may be due to the independent activity of certain protein-binding sites (Wilkie et al., 2003). Thus, the selective constraints that operate on U87HG RNA are quite similar to those that operate on UTRs.

Nowadays, more noncoding RNAs are being discovered. All of them are functional, although the degree of orthologs similarity is not always high. For example, human and mouse BIC RNAs which give rise to miR-155 microRNA (Lagos-Quintana et al., 2002) have overall 72% sequence similarity. U87HG and probably other noncoding host genes seem to belong to this group of moderately conserved noncoding RNA genes. As U87HG RNA contains many stretches with high local similarity interspersed among loosely conserved sequences, just like UTRs of protein-coding genes do, these short sequences may be functionally important. In addition to conserved regions, nonconserved species-specific sequences may also be essential for U87HG function. U87HG RNA may be required for translation considering that U87HG RNA is associated with ribosomes and that its gene belongs to the 5′TOP family most members of which are required for translation.

Acknowledgements

We thank Maria Rostovskaya, Drs. Dmitry Ivanov and Vladimir Prassolov for their help with cell cultures, to Dr. Alexey Kondrashov for helpful discussion, and to Dr. Andrei Kramerov for revising the English text. This work was supported by the Russian Foundation for Basic Research (project nos. 02-04-49772; 05-04-49553).

References

Amaldi, F., Pierandrei-Amaldi, P., 1997. TOP genes: a translationally controlled class of genes including those coding for ribosomal proteins. Prog. Mol. Subcell. Biol. 18, 1–17.

Bachellerie, J.P., Cavaille, J., Huttenhofer, A., 2002. The expanding snoRNA world. Biochimie 84, 775–790.

Bortolin, M.L., Kiss, T., 1998. Human U19 intron-encoded snoRNA is processed from a long primary transcript that possesses little potential for protein coding. RNA 4, 445–454.

Brennan, C.M., Steitz, J.A., 2001. HuR and mRNA stability. Cell. Mol. Life Sci. 58, 266–277.

Gibbs, R.A., et al., 2004. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature 428, 493–521.

Gogolevskaya, I.K., Makarova, J.A., Gause, L.N., Kulichkova, V.A., Konstantinova, I.M., Kramerov, D.A., 2002. U87 RNA, a novel C/D box small nucleolar RNA from mammalian cells. Gene 292, 199–204.

Higgins, D.G., Bleasby, A.J., Fuchs, R., 1992. CLUSTAL V: improved software for multiple sequence alignment. Comput. Appl. Biosci. 8, 189–191.

Huttenhofer, A., et al., 2001. RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. EMBO J. 20, 2943–2953.

Kessler, O., Chasin, L.A., 1996. Effects of nonsense mutations on nuclear and cytoplasmic adenine phosphoribosyltransferase RNA. Mol. Cell. Biol. 16, 4426–4435.

Kirkness, E.F., et al., 2003. The dog genome: survey sequencing and comparative analysis. Science 301, 1898–1903.

Kozak, M., 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44, 283–292.

Kozak, M., 1987. Effects of intercistronic length on the efficiency of reinitiation by eucaryotic ribosomes. Mol. Cell. Biol. 7, 3438–3445.

Kozak, M., 2001. Constraints on reinitiation of translation in mammals. Nucleic Acids Res. 29, 5226–5232.

Kuersten, S., Goodwin, E.B., 2003. The power of the 3′ UTR: translational control and development. Nat. Rev., Genet. 4, 626–637.

Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W., Tuschl, T., 2002. Identification of tissue-specific microRNAs from mouse. Curr. Biol. 12, 735–739.

Lemieux, R., Beaud, G., 1982. Expression of vaccinia virus early mRNA in Ehrlich ascites tumor cells: 2. Part of the polysomes at an early stage of virus infection are not bound to the cytoskeleton. Eur. J. Biochem. 129, 273–279.

Lercher, M.J., Urrutia, A.O., Hurst, L.D., 2002. Clustering of housekeeping genes provides a unified model of gene order in the human genome. Nat. Genet. 31, 180–183.

Lincoln, A.J., Monczak, Y., Williams, S.C., Johnson, P.F., 1998. Inhibition of CCAAT/enhancer-binding protein alpha and beta translation by upstream open reading frames. J. Biol. Chem. 273, 9552–9560.

Luukkonen, B.G., Tan, W., Schwartz, S., 1995. Efficiency of reinitiation of translation on human immunodeficiency virus type 1 mRNAs is determined by the length of the upstream open reading frame and by intercistronic distance. J. Virol. 69, 4086–4094.

Maquat, L.E., 2004. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. Nat. Rev., Mol. Cell Biol. 5, 89–99.

Maxwell, E.S., Fournier, M.J., 1995. The small nucleolar RNAs. Ann. Rev. Biochem. 64, 897–934.

Meijer, H.A., Thomas, A.A., 2002. Control of eukaryotic protein synthesis by upstream open reading frames in the 5′-untranslated region of an mRNA. Biochem. J. 367, 1–11.

Morris, D.R., Geballe, A.P., 2000. Upstream open reading frames as regulators of mRNA translation. Mol. Cell. Biol. 20, 8635–8642.

Ogurtsov, A.Y., Roytberg, M.A., Shabalina, S.A., Kondrashov, A.S., 2002. OWEN: aligning long collinear regions of genomes. Bioinformatics 18, 1703–1704.

Pelczar, P., Filipowicz, W., 1998. The host gene for intronic U17 small nucleolar RNAs in mammals has no protein-coding potential and is a member of the 5′-terminal oligopyrimidine gene family. Mol. Cell. Biol. 18, 4509–4518.

Peri, S., Pandey, A., 2001. A reassessment of the translation initiation codon in vertebrates. Trends Genet. 17, 685–687.

Pesole, G., et al., 2002. UTRdb and UTRsite: specialized databases of sequences and functional elements of 5′ and 3′ untranslated regions of eukaryotic mRNAs. Update 2002. Nucleic Acids Res. 30, 335–340.

Pierandrei-Amaldi, P., Campioni, N., Cardinali, B., 1991. Experimental changes in the amount of maternally stored ribosomes affect the translation efficiency of ribosomal protein mRNA in Xenopus embryo. Cell. Mol. Biol. 37, 227–238.

Ryskov, A.P., Ivanov, P.L., Kramerov, D.A., Georgiev, G.P., 1983. Mouse ubiquitous B2 repeat in polysomal and cytoplasmic poly(A)+RNAs: uniderectional orientation and 3′-end localization. Nucleic Acids Res. 11, 6541–6558.

Sabath, D.E., Broome, H.E., Prystowsky, M.B., 1990. Glyceraldehyde-3-phosphate dehydrogenase mRNA is a major interleukin 2-induced transcript in a cloned T-helper lymphocyte. Gene 91, 185–191.

Smale, S.T., Kadonaga, J.T., 2003. The RNA polymerase II core promoter. Ann. Rev. Biochem. 72, 449–479.

Smith, C.M., Steitz, J.A., 1998. Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5′-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. Mol. Cell. Biol. 18, 6897–6909.

Spirin, A.S., 1999. Initiation of translation. Ribosomes. Kluwer Academic Pub., p. 237.

Takai, D., Jones, P.A., 2002. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc. Natl. Acad. Sci. U. S. A. 99, 3740–3745.

Tanaka, R., et al., 2000. Intronic U50 small-nucleolar-RNA (snoRNA) host gene of no protein-coding potential is mapped at the chromosome breakpoint t(3;6)(q27;q15) of human B-cell lymphoma. Genes Cells 5, 277–287.

Tycowski, K.T., Shu, M.D., Steitz, J.A., 1996. A mammalian gene with introns instead of exons generating stable RNA products. Nature 379, 464–466.

Walden, W.E., Godefroy-Colburn, T., Thach, R.E., 1981. The role of mRNA competition in regulating translation: I. Demonstration of competition in vivo. J. Biol. Chem. 256, 11739–11746.

Wang, X.Q., Rothnagel, J.A., 2004. 5′-untranslated regions with multiple upstream AUG codons can support low-level translation via leaky scanning and reinitiation. Nucleic Acids Res. 32, 1382–1391.

Wilkie, G.S., Dickson, K.S., Gray, N.K., 2003. Regulation of mRNA translation by 5′- and 3′-UTR-binding factors. Trends Biochem. Sci. 28, 182–188.

Xie, X., et al., 2005. Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals. Nature 434, 338–345.

Zhang, J., Maquat, L.E., 1997. Evidence that translation reinitiation abrogates nonsense-mediated mRNA decay in mammalian cells. EMBO J. 16, 826–833.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31, 3406–3415.