ISSN 1022-7954, Russian Journal of Genetics, 2007, Vol. 43, No. 2, pp. 97–105. © Pleiades Publishing, Inc., 2007.
Original Russian Text © J.A. Makarova, D.A. Kramerov, 2007, published in Genetika, 2007, Vol. 43, No. 2, pp. 149–158.

THEORETICAL PAPERS AND REVIEWS

Small Nucleolar RNA Genes

J. A. Makarova and D. A. Kramerov

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia; fax: (495)135-14-05; e-mail: kramerov@eimb.ru

Received May 3, 2006

Abstract—Small nucleolar RNAs (snoRNAs) are one of the most numerous and well-studied groups of non-protein-coding RNAs. In complex with proteins, snoRNAs perform the two most common nucleotide modifications in rRNA: 2'-OH-methylation of ribose and pseudouridylation. Although the modification mechanisms and snoRNP structures are highly conserved, the snoRNA genes are surprisingly diverse in organization. In addition to genes transcribed independently, there are genes that are in introns of other genes, form clusters transcribed from a common promoter, or clusters in introns. Interestingly, one type of gene organization usually prevails in different taxa. Vertebrate snoRNAs mostly originate from introns of protein-coding genes; a small group of snoRNAs are encoded by introns of genes for noncoding RNAs.

DOI: 10.1134/S1022795407020019

INTRODUCTION

A substantial part of transcribed genome sequences do not code for proteins. Interest in such sequences greatly increased over the past several years, as new classes of noncoding RNAs were revealed. These RNAs proved to perform a surprisingly wide range of functions in the cell, utilizing mechanisms unknown before. In particular, RNAs are involved in the regulation of transcription and translation, suppression of mobile genetic elements, and other regulatory pathways: all these processes were attributed exclusively to proteins earlier. According to different estimates, noncoding RNAs identified to date are only the tip of the iceberg (e.g., see [1, 2]).

Small nucleolar RNAs (snoRNAs) are a particular class of noncoding RNAs. Several snoRNAs are necessary for pre-rRNA cleavage, while most snoRNAs serve to determine the rRNA, small nuclear RNA, and mRNA nucleotides that are subject to one of the two modifications, methylation of the ribose moiety at the 2'-OH group or conversion of uridine to pseudouridine [3]. These modifications are most common: more than one hundred nucleotides are modified only in rRNA. According to the type of modification and the conserved elements present in the nucleotide sequence, snoRNAs are divided into two families: C/D (box C/D snoRNAs) and H/ACA (box H/ACA snoRNAs). Some C/D snoRNAs are involved in pre-rRNA cleavage, while others guide 2'-O-methylation. The nucleotide sequences of all C/D snoRNAs harbor conserved elements C (UGAUGA) and D (CUGA) and their copies C' and D', which are often degenerate. In addition, C/D snoRNA has a so-called antisense element of 10–20 nt, which is fully complementary to the corresponding region of the target RNA and interacts with it. 2'-O-methylation affects the nucleotide that is contained in the resulting RNA/RNA helix and is 4 nt away from element D and/or D' [4]. Methylation is catalyzed by proteins of C/D snoRNP.

Almost all snoRNAs of the H/ACA family determine the RNA nucleotide subject to pseudouridylation, and only a few members of the family are involved in pre-rRNA cleavage. All H/ACA snoRNAs contain conserved sequences H (ANANNA) and ACA (ACA) and, like C/D snoRNAs, one or, in rare cases, two antisense elements. As in the case with C/D snoRNAs, the uridyl residue subject to modification is determined via complementary interactions of the antisense element with the nucleotide sequence of the target RNA; pseudouridylation is catalyzed by one of the proteins associated with H/ACA snoRNA [5].

In addition, the snoRNA group includes the RNA component of RNase MRP and, sometimes, the similar RNA component of RNase P. RNase MRP cleaves pre-rRNA at the A3 site of the first internal transcribed spacer to yield the mature 5.8 S rRNA [6]; yet its functions are not restricted to this reaction. RNase P is involved in removing the 5'-terminal leader sequence from pre-tRNA [6].

The structure and functions of snoRNA are considered in more detail elsewhere [7]. Here we focus on the variants of snoRNA gene organization in several taxonomic groups.

1. GENES OF snoRNAs

The snoRNA genes are extremely diverse in organization: some are individual transcription units, while some others occur in introns of protein-coding genes and are processed during splicing. In addition, snoRNA genes can form clusters, which are transcribed to yield a polycistronic transcript. Such a cluster can be located

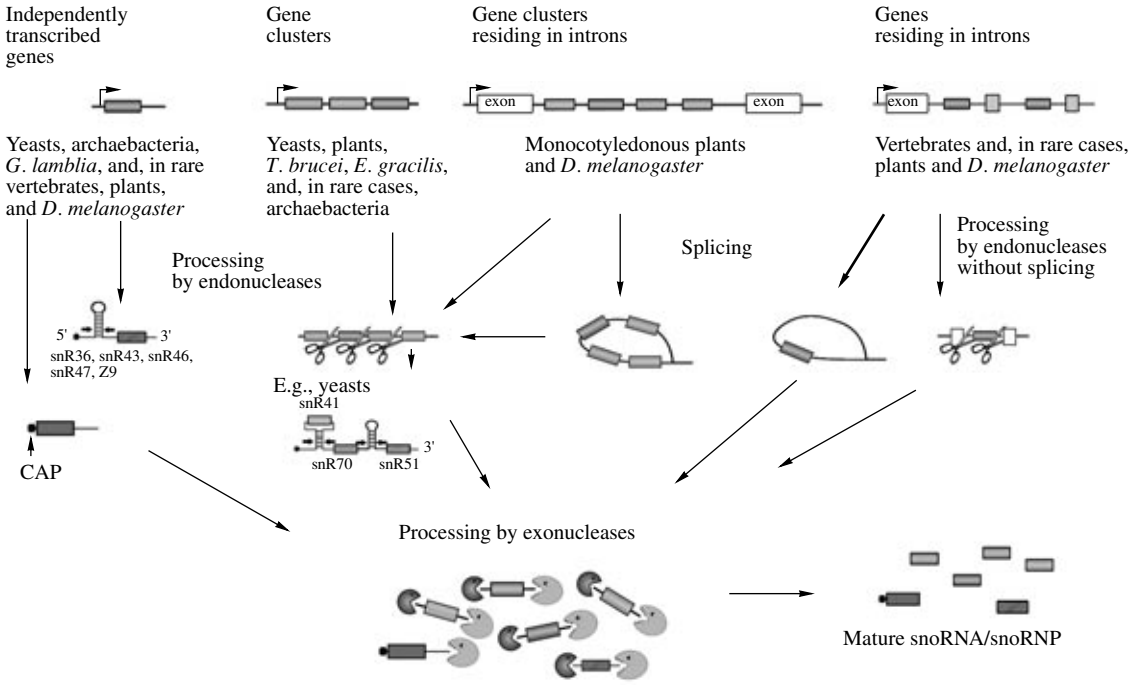

Fig. 1. Variants of gene organization and snoRNA processing in different taxonomic groups. Small arrows indicate the cleavage sites for endoribonuclease Rnt1p (see Section 4). Modified from [8, 9].

in an intron of another gene. Even more intricate variants are known. Interestingly, only one type of gene organization prevails in a particular taxon. For instance, almost all vertebrate snoRNAs are encoded by single-copy genes located in introns; most snoRNA genes have independent locations in yeasts and cluster in plants and protozoa (Fig. 1).

1.1. Vertebrate snoRNA Genes

Almost all vertebrate snoRNAs are unusually encoded: their genes are in introns of protein coding genes, each intron harboring only one snoRNA gene (Fig. 1). Only the genes for U3, U8, and U13 and for the RNA components of RNases P and MRP are transcribed from their own promoters by RNA polymerases II and III, respectively [10]. Two other genes transcribed from their own promoters by RNA polymerase II code for unusual C/D snoRNAs with additional copies of elements C and D [11]. Most snoRNA genes occur in single copies. A few exceptions include the multicopy vital U3 RNA gene [10].

1.2. Yeast snoRNA Genes

As in animals, almost all snoRNA genes occur in single copies in yeasts. However, unlike in animals, the majority of yeast snoRNA genes occur as independent transcription units; only seven genes located in introns and five polycistronic clusters (two dicistronic, two tricistronic, and one heptacistronic) are known [12]. An interesting case is presented by the U86 RNA, which is encoded within an open reading frame [13]. The U86 RNA and mRNA are produced from this reading frame via alternative processes. The vertebrate U86 RNA is encoded in an intron.

1.3. Plant snoRNA Genes

In plants, most genes for C/D and H/ACA snoRNAs are clustered (Fig. 1). Some clusters are homogeneous, consisting of two to four copies of one snoRNA gene; some others are heterogenous, containing genes for several different snoRNAs. One or more genes can occur in several copies in a heterogeneous cluster. Clusters mostly harbor two to five genes and are scattered through all chromosomes [14]; some clusters occur in several (two to five) copies [15]. In dicotyledonous plants, almost all snoRNA gene clusters are in intergenic spacers and are transcribed from their own promoters to yield polycistronic RNAs. For instance, 133 of 175 known *Arabidopsis thaliana* snoRNA genes are organized in 49 clusters (including a few clusters each occurring in two or three copies), of which only four are in introns of protein-coding genes. A cluster organization of snoRNA genes is similarly predominant in monocotyledonous plants, but the distribution pattern is different: more than half of the clusters are within introns of protein-coding genes [15]. This unique type of gene organization is absent from vertebrates and yeasts (Fig. 1). Plants of both classes have a relatively

few independent snoRNA genes, which reside in introns, as is the case with most vertebrate snoRNA genes [8].

Several snoRNA genes (about ten [4]) are transcribed from their own promoters. For instance the U3 RNA and RNase MRP genes occur as independent transcription units in all eukaryotes. However, the U3 RNA gene is transcribed by RNA polymerase III in plants, rather than by RNA polymerase II as in animals and yeasts. The promoter is similar to the promoters of plant small nuclear RNA genes in having the so-called upstream sequence element (USE) and the TATA box. Note, however, that promoters for RNA polymerase II and III are structurally similar to a certain extent in plants [16].

Another unusual variant of snoRNA genes was found in plants. Both dicotyledonous (*Arabidopsis thaliana*) and monocotyledonous (*Oryza sativa*) plants have several dicistronic genes consisting of tRNA and snoRNA genes (the tRNA gene is upstream of the snoRNA gene). These genes are transcribed from the tRNA gene promoter [17].

Unlike in vertebrates and yeasts, about half of plant snoRNAs are each encoded by several genes. For instance, approximately 65% of all snoRNA genes occur in more than one copy in *A. thaliana* [18]. This can be explained in part by numerous duplications of genome fragments: genes coding for one snoRNA often belong to duplicated clusters. In addition, a snoRNA gene can undergo tandem duplication within one cluster. Similarity between such duplicated genes greatly varies and, probably, is inversely proportional to the age of duplication. In many cases, copies differ from each other only by a few substitutions and/or inserts/deletions. The changes are usually beyond the conserved sequences (elements C, D, C', and D' and the antisense element) and, in rare cases, occur within such sequences. Most substitutions are insignificant and can hardly cause a loss of function by the snoRNA gene. Yet such minor changes result in new rRNA modification sites in rare cases. For instance, the differences between two copies of the *A. thaliana* snoR20 gene include a single-nucleotide insert between box D and the antisense element. The insert results in a new site of rRNA modification; i.e., the two variants of snoR20 guide modification of two different (adjacent) nucleotides in rRNA [14]. Dramatic differences between two copies of a snoRNA gene are also possible; e.g., one copy can actually represent a pseudogene fragment [14].

Some snoRNA genes have orthologs in vertebrates, yeasts, and plants. Similarity is usually restricted to the conserved sequences and the antisense element; i.e., the term orthologs is rather conventional in such cases. The situation is even more intricate in some cases where a genome has not only an ortholog, but also another gene with an extra antisense element differing from that of the ortholog. For instance, the U33 RNA occurs both in vertebrates and in plants. However, plants have another RNA, snoR34, which contains the antisense element of the U33 RNA together with a different one [14]. In some cases, two gene copies share one antisense element and differ in another within one species. For instance, snoR16.1 and snoR16.2 both guide U48 methylation in the 25S rRNA, but the former additionally guides U2445 methylation and the latter, U36 methylation in the same rRNA [14].

Consider another interesting example. Chromosome 2 of *A. thaliana* harbors a cluster of four snoRNA genes, including the genes for U55 and U16 RNAs (Fig. 2). The U55 and U16 RNAs each contain one antisense element. Chromosome 1 has a cluster of two genes; one of these codes for snoR15, which contains the antisense elements of both U55 and U16 RNAs. It is clear that the clusters are related, because they have one gene (snoR14) in common. The authors who have identified the clusters believe that the U55 and U16 RNAs result from duplication of the snoR15 gene located on chromosome 1 [8, 14]. An alternative explanation suggests that chromosome 1 cluster initially contained both U55 and U16 RNA genes. The U55 RNA gene acquired the antisense element characteristic of the U16 RNA: this required only three substitutions. The U16 RNA gene was then lost [14].

Thus, the plant snoRNA genes provide a wide area for speculations about possible ways of gene evolution.

1.4. Protozoan snoRNA Genes

As in plants, the snoRNA genes of *Trypanosoma brucei* (class Mastigophora, subclass Zoomastigina, order Kinetoplastidae) form clusters, which contain snoRNA genes of both families and are usually repeated several times in the genome [19] (Fig. 1). All clusters are in intergenic spacers; their full-length transcripts are synthesized by RNA polymerase II. A similar organization of the snoRNA genes is characteristic of *Euglena gracilis*, which belongs to the same class but to a different subclass, Phytomastigina [20]. However, another organization was observed for the order Diplomonadida, which is closer to *Trypanosoma* and belongs to the same subclass Zoomastigina. In *Giardia lamblia*, a representative of this order, the genes for 20 known snoRNAs are independent, occur in single copies, reside in spacers between protein-coding genes, and are transcribed from their own promoters by RNA polymerase II [21]. This observation is of interest because all earlier data suggested that the closer the taxa, the greater their similarity in the predominant variant of snoRNA gene organization.

1.5. Drosophila snoRNA Genes

It becomes evident now that the organization of snoRNA genes in animals is more diverse than believed earlier. In addition to snoRNA genes residing in introns (one gene in one intron) [22] as characteristic of animals and a few independent genes, snoRNA gene clusters were recently found in *Drosophila melanogaster* [23] (Fig. 1). A cluster includes up to nine genes. Usu-

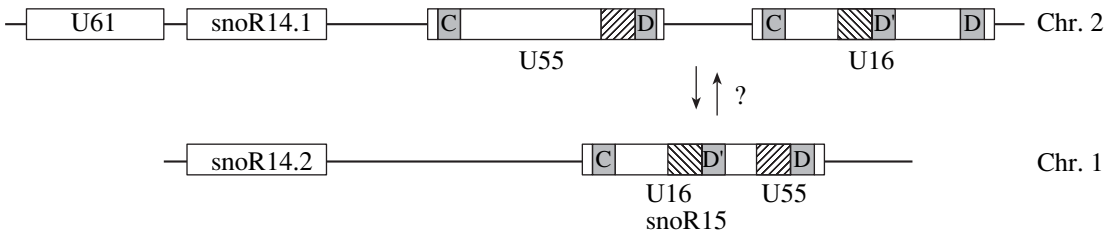

Fig. 2. Structure of two snoRNA gene clusters of *A. thaliana.* The antisense elements are crosshatched. Modified from [8, 14].

ally, these are two or three genes occurring in several copies. Such an organization was earlier known only for plants, mostly monocotyledonous. Another unusual variant of snoRNA genes was observed in *Drosophila*: at least five genes reside in exons of protein-coding genes; i.e., snoRNA and mRNA are produced via alternative processes. The only example of such an organization of snoRNA genes in eukaryotes was previously known only for yeasts (see Section 1.2). In addition, this type of organization is utilized by archaebacteria. Several other snoRNA genes reside in exons subject to alternative splicing: snoRNA is incorporated in an exon in some cases and in an intron in some others. One snoRNA was found at the exon–intron boundary [24].

1.6. Archaebacterial snoRNA Genes

While completely absent from bacteria, snoRNAs were found in archaebacteria, far as they are from eukaryotes. Archaebacterial snoRNA genes occur in single copies and are mostly independent; a few examples are known where several genes are close together and are probably transcribed to yield a common precursor. All snoRNA genes reside in spacers between protein-coding genes and seem to have their own promoters. An unusual location was observed for 10–20% of snoRNA genes: their 5'-terminal region overlaps the 3' end of an open reading frame. Thus, the sequence coding for several C-terminal amino acid residues simultaneously codes for the 5' region of snoRNA. Finally, a case is known where a snoRNA gene is in a tRNA intron [25].

1.7. Diverse Organization of snoRNA Genes

Large taxa, such as vertebrates and plants, usually differ in the organization of genes coding for the same snoRNAs. However, one organization variant is mostly used for the same snoRNA within a taxon. For instance, the U23 snoRNA is encoded by intron 12 of the nucleolin gene in vertebrates of all classes. Yet it is not infrequent that snoRNA changes the host gene. The U15 snoRNA is encoded by introns of the ribosomal protein S3 gene in human and the ribosomal protein S1 gene in *Xenopus laevis.* Moreover, snoRNA can have different host genes even in one species, as is the case with the *X. laevis* U14 snoRNA [10]. Such findings suggest that snoRNA originate from mobile elements. However, snoRNA genes lack terminal repeats implied by such a scenario [10].

Another question is why different snoRNA genes are organized in different ways in one organism. The answer is obvious only for snoRNA genes expressed independently. For instance, all vertebrate independent genes code for snoRNAs that are involved in rRNA cleavage and, consequently, are essential. There are some questions even in this case. Two C/D snoRNAs recently found in vertebrates are encoded by independent genes but guide 2'-O-methylation of small nuclear RNAs [11]; in yeasts, the majority of snoRNA genes are independent. Another problem is a biological significance, if any, of the fact that many orthologous snoRNAs are differently encoded in different large taxa. The simplest assumption is that the variant of snoRNA gene organization is selected at random. However, this assumption disagrees with the observation that most host genes of vertebrate snoRNAs are somehow involved in translation (see Section 2). Finally, it is unclear why the organization of snoRNA genes is so diverse: nothing of the kind is known for other gene groups.

It is commonly accepted that the polycistronic organization of genes allows a coordinated regulation of their expression. Plant snoRNA genes are clustered, and vertebrate genes reside in introns of the genes that mostly code for translation proteins. Thus, such genes each code for more than one product and can be considered polycistronic. It seems that the two taxa differently solve the problem of coordinated expression of functionally related genes. Note that successful processing of almost all snoRNA genes requires splicing in vertebrates. In plants, snoRNA processing is splicing-independent, allowing a diverse organization of snoRNA genes. Plant snoRNAs are produced even when splicing is suppressed (e.g., in heat shock [26]). A greater tolerance of snoRNA production in plants compared with vertebrates is possibly related to the fact that plants are to a greater extent exposed to extreme factors.

2. HOST GENES OF VERTEBRATE snoRNAs

Most genes whose introns give origin to snoRNAs code for translation proteins (initiation and elongation factors, ribosomal proteins, etc.) or proteins involved in

the function of the nucleolus. Such a location is hardly occasional and probably serves to coordinate the expression of transcripts synthesized by RNA polymerases I and II: the intensity of rRNA production and, eventually, generation of ribosomes is thereby associated with the level of mRNAs for translation proteins [10].

Almost all host genes of vertebrate snoRNAs belong to the so-called 5'-TOP (5'-terminal oligopyrimidine) family of housekeeping genes [27]. The family includes the genes for ribosomal proteins and translation factors together with some other genes. The first nucleotide of their mRNAs is cytosine, which is followed by a short (4–13 nt) pyrimidine tract. This structure, in combination with specific sequences of the 5'-untranslated region of mRNA, allows coordinated expression of 5'-TOP family genes in response to external stimuli (e.g., growth factors). As a result, the intensity of translation is increased [28].

Interestingly, some snoRNA host genes (UHG, U17HG, U19H, gas5, U50HG, and U87HG) of the 5'-TOP family have exons with multiple stop codons scattered throughout their lengths and, consequently, are incapable of coding for a protein [27, 29–33]. Upon transcription, the introns give origin to snoRNAs, while the noncoding host gene transcript is polyadenylated and is either retained in the nucleus (U19H RNA) or exported into the cytoplasm. Three transcripts (UHG, gas5, and U87HG) are associated with ribosomes, and the UHG and gas5 "mRNAs" undergo rapid degradation, which depends on the activity of translation machinery. The U87HG RNA is also associated with ribosomes, but is far more stable [34]. Note that the transcripts of noncoding host genes display a very low conservation, comparable with that of neutral sequences [31]. This prompted the idea that the only role of such genes is generation of snoRNA, while their exons have no functional significance [27, 29, 30, 32]. The only exception is the U87HG gene transcript: the extent of its conservation is higher than that of the other noncoding host genes and corresponds to that of untranslated mRNA regions (sequence similarity among mammals is meant in all cases). This suggests an independent function for the U87HG gene transcript [34], especially as a similar extent of conservation is characteristic of several noncoding RNAs with well-known functions [35, 36]. As for the other noncoding host genes, the following should be noted. Recent studies demonstrate that noncoding RNAs can be functional even in the absence of conservation, and the number of such RNAs is continuously increasing [35]. Such RNAs were not identified until recently mostly because nonconserved sequences were not searched for genes. A rapid growth in the number of such RNAs can be expected for the nearest future. Hence, the accepted view that exons of noncoding host genes lack functional significance is at least tentative.

There are several other cases where snoRNA genes reside in introns of noncoding genes, but none of these belongs to the 5'-TOP family. For instance, the rat RBII-36 snoRNA, which is expressed exclusively in the brain and lacks regions complementary to rRNA or small nuclear RNA, is encoded by an intron of the gene for the noncoding Bsr RNA, which is expressed only in the nervous system [37]. The genes for three other C/D snoRNAs found recently reside in introns of the Irm noncoding gene, which is subject to imprinting. Interestingly, homologs of these snoRNAs were not found in human cells [38]. Another example is provided by the *SNURF–SNURPN* imprinted gene, which is expressed from the paternal allele in all tissues [39]. The gene is approximately 450 kb in size and contains more than 150 exons. Its dicistronic mRNA codes consecutively for the SNURF protein (SNRPN upstream reading frame; exons 1–3) and the SmN spliceosomal protein (exons 4–10). All other exons are noncoding, and their introns harbor snoRNA genes at one gene per intron: introns 21–52 contain 27 copies of the HBII-85 snoRNA gene, introns 63–142 contain 47 copies of the HBII-52 snoRNA gene, and the remaining introns contain the genes coding for five other snoRNAs and occurring in fewer copies. These snoRNAs are devoid of sequences complementary to rRNA or small nuclear RNA and are expressed predominantly or exclusively in the brain. The 3'-terminal noncoding region of the giant transcript serves as an antisense RNA for the *UBE3A* protein-coding gene, which is transcribed in the opposite direction [40]. Deletion of the *SNURF–SNURPN* gene together with all noncoding exons and snoRNA genes is responsible for the majority of cases of the Prader–Willi syndrome (PWS) [41], a severe hereditary disease characterized by mental retardation, obesity, muscular hypotonia, and several other signs. Recent data demonstrate that deletion of snoRNA sequences, in particular, HBII-85, plays an important, if not critical, role in the pathogenesis of this disease (e.g., see [42, 43]).

Another snoRNA encoded by this gene, HBII-52/MBII-52 (human/mouse) contains an extended (18-nt) region fully complementary to the mRNA for the serotonin 5-HT2C receptor, which is expressed exclusively in the brain [44]. Interestingly, the genes for MBII-52 and the 5-HT2C receptor are both expressed in the brain, but their coexpression is restricted to certain sections. The potential target of 2'-O-methylation potentially guided by the HBII-52/MBII-52 snoRNA is an adenine residue, which is usually subject to editing and is converted to inosine (I). Recent studies with mice provided convincing evidence that MBII-52 does guide methylation of this adenine, which is accompanied by a substantial decrease in the efficiency of its deamination (A-I) [45]. In other words, 2'-O-methylation decreases the efficiency of editing. During translation, inosine is decoded as guanosine. Possibly, this mechanism is responsible for the fine regulation of the interaction between the serotonin receptor with the G-protein. Interestingly, mice knocked-out in the 5-HT2C receptor gene have defects similar to PWS signs.

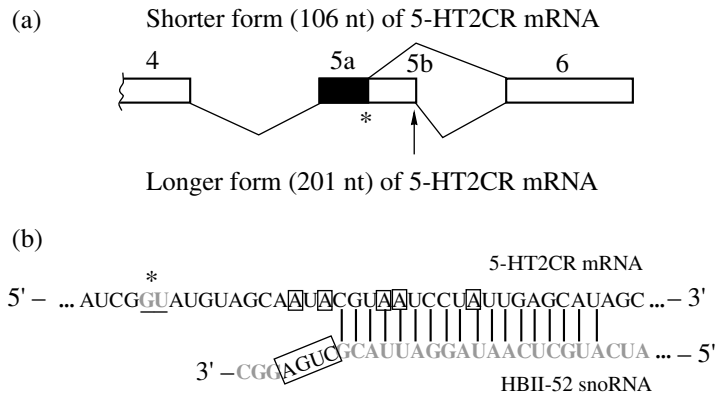

Fig. 3. Complementary interactions are possible between the HBII-52 snoRNA and the serotonin 5-HT2C receptor mRNA. (a) Fragment of the human 5-HT2C receptor mRNA. Exons (bars) are indicated. The proximal and distal splicing sites are shown with an asterisk and an arrow, respectively. (b) Complementary interactions between the HBII-52 snoRNA and the 5-HT2C receptor mRNA. The five nucleotides subject to editing (deamination of A to produce I) are framed. The proximal splicing site is indicated with an asterisk and underlined. In snoRNA, box D is framed. Modified from [46].

Recent studies yielded the following data. It is known that alternative splicing results in two isoforms of the 5-HT2C receptor mRNA, which differ in the size of exon 5. The functional receptor is synthesized only from the longer isoform. The downstream region of exon 5 (5b, Fig. 3) contains a sequence (silencer) that prevents 5b incorporation in mRNA and thereby determines production of the shorter isoform of the 5-HT2C receptor mRNA. Five nucleotides of region 5b are subject to editing. Editing ensures incorporation of region 5b in mRNA, and the full-length isoform is produced. However, editing reduces the sensitivity of the 5-HT2C receptor to serotonin by one or two orders of magnitude. The HBII-52/MBII-52 snoRNA, whose antisense element is complementary to fragment 5b, ensures incorporation of region 5b in mRNA and, consequently, editing-independent production of the long isoform [46]. The protein synthesized from the resulting unedited isoform is highly sensitive to serotonin. The cell pool of 5-HT2C receptor mRNAs normally includes both edited and unedited forms, which is apparently necessary for regulating the serotonin sensitivity of nerve cells. Thus, not only do PWS patients lack the HBII-52 snoRNA, but they also have a dramatically reduced fraction of the 5-HT2C receptor mRNA with the unedited sequence of region 5b [46]. The latter circumstance certainly contributes to the pathogenesis of PWS. Recent studies offer new lines for developing a therapeutic strategy in PWS.

3. TRANSCRIPTION OF snoRNA GENES

Individual snoRNA genes are transcribed by RNA polymerases II (e.g., U3, U8, and U13 RNAs) and III (RNA component of RNase MRP, plant U3 RNA). The 5' ends of the transcripts carry a 2,2,7-trimethylguanosine cap or γ-monomethyl phosphate, respectively. In contrast, the 5' end is nonmodified in snoRNAs encoded by introns or synthesized within a polycistronic transcript. Such snoRNAs are excised from a precursor; then, their ends are formed by exonucleases.

The 3' end is not modified in all known snoRNAs [10].

4. MATURATION AND TRANSPORT OF snoRNAs

Several pathways are known for snoRNA maturation. This is explained, first and foremost, by the diverse organization of snoRNA genes, which are transcribed independently to yield mono- and polycistronic transcripts or reside in introns. Whichever the case, free 5' and 3' ends are generated in a snoRNA precursor when it is processed by specific nucleases or cleaved at the branching point. The free ends are attacked by 5'–3' and 3'–5' exonucleases, which eventually form the ends of mature snoRNA. Complete exonucleolytic degradation of snoRNAs is prevented by their binding with core proteins.

Endonucleases are responsible for initial processing in the case of independently transcribed genes. In yeasts, most snoRNA are transcribed from such independent genes and included in mono- and polycistronic transcripts. Most of these transcripts contain recognition sites for endoribonuclease Rnt1p [47] (Fig. 1). Rnt1p is active towards double-stranded RNA and, according to this feature, is assigned to a family of proteins collectively known as RNases III [48]. RNases of this family occur in all living organisms, playing a role in the processing of rRNA, tRNA, mRNA, and microRNA and initiating mRNA degradation [49]. Apart from snoRNA, Rnt1p processes rRNA and some small nuclear RNAs (U5 and U2) [50, 51]. During snoRNA processing, Rnt1p cleaves both polycistronic [51] and monocistronic [52] transcripts (Fig. 1). The snoRNA products have immature free ends, which are truncated by exonucleases: the 5' region is attacked by 5'–3' exonucleases Rat1p and Xrn1p [53], while the 3' region is processed by the exosome and, in particular,

its component Rrp6p [54]. Exosomes are complexes of many (at least ten) cell 3'–5' exonucleases and occur both in the nucleus and in the cytoplasm [55] to form the 3' ends of many cell RNAs and to degrade cytoplasmic mRNA and the between-cistron sequences of pre-rRNA [56].

Some yeast snoRNAs lack sequences recognized by Rnt1p. The free 5' and 3' ends are formed in the precursors of such snoRNAs by proteins that introduce a break during mRNA cleavage/polyadenylation [9].

Intron-encoded snoRNAs are processed via two strategies. Maturation of most snoRNAs requires splicing and subsequent hydrolysis of the 2'–5' phosphodiester bond in the branching point. The resulting product (snoRNA with the flanking sequences) is attacked by exonucleases, which produce mature snoRNA [57] (Fig. 1). The processing of a few intron-encoded snoRNAs in various organisms, including vertebrates [58], does not depend on splicing: exonucleases sites result from endonucleolytic cleavage of nonspliced pre-mRNA (Fig. 1). In yeasts, this process involves Rnt1p, which cleaves a short double-stranded stem formed by snoRNA-flanking intron sequences, while exonucleolytic degradation is performed by Rat1, Xrn1 [53], and the exosome [54], as is the case with independent snoRNA genes. The processes yielding snoRNA and mRNA are alternative. The question arises as to how the cell chooses between them. Almost all few intronic snoRNAs are processed following this scenario in yeasts. It was found that five out of six genes for these snoRNAs are flanked by sequences that cannot form a double-stranded region providing a substrate for Rnt1p. These sequences are still capable of complementary interaction, but their interaction yields a noncanonical substrate, which can be recognized by Rnt1p only in the presence of Nop1p, a core protein of snoRNP. Thus, snoRNA processing occurs only in the presence of snoRNP proteins; otherwise, splicing yields the host gene mRNA [59]. Other experimental data support the involvement of snoRNP core proteins actively participate in processing [60].

Maturation of intron snoRNAs takes place in the vicinity of their genes [61], while polycistronic transcripts are processed in so-called Cajal bodies (CBs) and in the nucleolus [62]. CBs are often associated with the nucleolus, are 0.3–0.5 μm in diameter, and consist of helical fibrillar chains formed mostly by p80 coilin [63]. The available data suggest that CBs is the scene of maturation and assembly of small nuclear RNPs and transcription complexes, which then function in other regions of the nucleus [63]. Assembly and maturation of some snoRNPs continue in CBs, as is the case with U3 [64] and several plant snoRNAs [60]. In addition, CBs retain defective snoRNPs [65]. After assembly, snoRNPs are transported into the nucleolus, although some maturation steps can proceed there as well [60].

To conclude, C/D and H/ACA snoRNAs occur in all eukaryotes, are associated with the same set of homologous proteins, and can fold into similar secondary structures. Thus, the snoRNP structure and the mechanisms of snoRNP-guided modification of cell RNAs are highly conserved. The extent of their conservation is in striking contrast with diversity of the organization of snoRNA genes. Further studies will elucidate this interesting phenomenon.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 05-04-49553).

REFERENCES

1. Claverie, J.M., Fewer Genes, More Noncoding RNA, *Science*, 2005, vol. 309, pp. 1529–1530.
2. Vaughn, M.W. and Martienssen, R., It's a Small RNA World, After All, *Science*, 2005, vol. 309, pp. 1525–1526.
3. Kiss, T., Small Nucleolar RNA-Guided Post-Transcriptional Modification of Cellular RNAs, *EMBO J.*, 2001, vol. 20, pp. 3617–3622.
4. Tollervey, D., Small Nucleolar RNAs Guide Ribosomal RNA Methylation, *Science*, 1996, vol. 273, pp. 1056–1057.
5. Ganot, P., Bortolin, M.L., and Kiss, T., Site-Specific Pseudouridine Formation in Preribosomal RNA Is Guided by Small Nucleolar RNAs, *Cell*, 1997, vol. 89, pp. 799–809.
6. Xiao, S., Scott, F., Fierke, C.A., et al., Eukaryotic Ribonuclease P: A Plurality of Ribonucleoprotein Enzymes, *Annu. Rev. Biochem.*, 2002, vol. 71, pp. 165–189.
7. Makarova, Yu.A. and Kramerov, D.A., Small Nucleolar RNAs, *Mol. Biol.* (Moscow) 2007, vol. 41 (in press).
8. Brown, J.W., Echeverria, M., and Qu, L.H., Plant snoRNAs: Functional Evolution and New Modes of Gene Expression, *Trends Plant. Sci*, 2003, vol. 8, pp. 42–49.
9. Fatica, A., Morlando, M., and Bozzoni, I., Yeast snoRNA Accumulation Relies on a Cleavage-Dependent/Polyadenylation-Independent 3'-Processing Apparatus, *EMBO J.*, 2000, vol. 19, pp. 6218–6229.
10. Maxwell, E.S. and Fournier, M.J., The Small Nucleolar RNAs, *Annu. Rev. Biochem.*, 1995, vol. 64, pp. 897–934.
11. Tycowski, K.T., Aab, A., and Steitz, J.A., Guide RNAs with 5' Caps and Novel Box C/D snoRNA-Like Domains for Modification of snRNAs in Metazoa, *Curr. Biol.*, 2004, vol. 14, pp. 1985–1995.
12. Lowe, T.M. and Eddy, S.R., A Computational Screen for Methylation Guide snoRNAs in Yeast, *Science*, 1999, vol. 283, pp. 1168–1171.
13. Filippini, D., Renzi, F., Bozzoni, I., et al., U86, a Novel snoRNA with an Unprecedented Gene Organization in Yeast, *Biochem. Biophys. Res. Commun.*, 2001, vol. 288, pp. 16–21.
14. Brown, J.W., Clark, G.P., Leader, D.J., et al., Multiple snoRNA Gene Clusters from *Arabidopsis*, *RNA*, 2001, vol. 7, pp. 1817–1832.
15. Chen, C.L., Liang, D., Zhou, H., et al., The High Diversity of snoRNAs in Plants: Identification and Compara-

tive Study of 120 snoRNA Genes from *Oryza sativa*, *Nucleic Acids Res.*, 2003, vol. 31, pp. 2601–2613.

16. Kiss, T., Marshallsay, C., and Filipowicz, W., Alteration of the RNA Polymerase Specificity of U3 snRNA Genes during Evolution and in Vitro, *Cell*, 1991, vol. 65, pp. 517–526.
17. Kruszka, K., Barneche, F., Guyot, R., et al., Plant Dicistronic tRNA–snoRNA Genes: A New Mode of Expression of the Small Nucleolar RNAs Processed by RNase Z, *EMBO J.*, 2003, vol. 22, pp. 621–632.
18. Vision, T.J., Brown, D.G., and Tanksley, S.D., The Origins of Genomic Duplications in *Arabidopsis*, *Science*, 2000, vol. 290, pp. 2114–2117.
19. Uliel, S., Liang, X.H., Unger, R., et al., Small Nucleolar RNAs That Guide Modification in Trypanosomatids: Repertoire, Targets, Genome Organization, and Unique Functions, *Int. J. Parasitol*, 2004, vol. 34, pp. 445–454.
20. Russell, A.G., Schnare, M.N., and Gray, M.W., Pseudouridine-Guide RNAs and Other Cbf5p-Associated RNAs in *Euglena gracilis*, *RNA*, 2004, vol. 10, pp. 1034–1046.
21. Yang, C.Y., Zhou, H., Luo, J., et al., Identification of 20 snoRNA-Like RNAs from the Primitive Eukaryote, *Giardia lamblia*, *Biochem. Biophys. Res. Commun.*, 2005, vol. 328, pp. 1224–1231.
22. Tycowski, K.T. and Steitz, J.A., Non-Coding snoRNA Host Genes in *Drosophila*: Expression Strategies for Modification Guide snoRNAs, *Eur. J. Cell Biol.*, 2001, vol. 80, pp. 119–125.
23. Huang, Z.P., Zhou, H., Liang, D., et al., Different Expression Strategy: Multiple Intronic Gene Clusters of Box H/ACA snoRNA in *Drosophila melanogaster*, *J. Mol. Biol.*, 2004, vol. 341, pp. 669–683.
24. Accardo, M.C., Giordano, E., Riccardo, S., et al., A Computational Search for Box C/D snoRNA Genes in the *Drosophila melanogaster* Genome, *Bioinformatics*, 2004, vol. 20, pp. 3293–3301.
25. Dennis, P.P., Omer, A., and Lowe, T., A Guided Tour: Small RNA Function in Archaea, *Mol. Microbiol.*, 2001, vol. 40, pp. 509–519.
26. Bond, U., Heat Shock but Not Other Stress Inducers Leads to the Disruption of a Sub-Set of snRNPs and Inhibition of *in vitro* Splicing in HeLa Cells, *EMBO J.*, 1988, vol. 7, pp. 3509–3518.
27. Pelczar, P. and Filipowicz, W., 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.*, 1998, vol. 18, pp. 4509–4518.
28. Amaldi, F. and Pierandrei-Amaldi, P., TOP Genes: a Translationally Controlled Class of Genes Including those Coding for Ribosomal Proteins, *Prog. Mol. Subcell. Biol.,* 1997, vol. 18, pp. 1–17.
29. Tycowski, K.T., Shu, M.D., and Steitz, J.A., A Mammalian Gene with Introns Instead of Exons Generating Stable RNA Products, *Nature*, 1996, vol. 379, pp. 464–466.
30. Bortolin, M.L. and Kiss, T., Human U19 Intron-Encoded snoRNA Is Processed from a Long Primary Transcript that Possesses Little Potential for Protein Coding, *RNA*, 1998, vol. 4, pp. 445–454.
31. Smith, C.M. and Steitz, J.A., 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.*, 1998, vol. 18, pp. 6897–6909.
32. Tanaka, R., Satoh, H., Moriyama, M., et al., 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*, 2000, vol. 5, pp. 277–287.
33. Makarova J.A. and Kramerov D.A. Mammalian 87 Small Nucleolar RNA and Its Host Gene, *Mol. Biol.* (Moscow), 2005, vol. 39, pp. 655–663.
34. Makarova, J.A. and Kramerov, D.A., Noncoding RNA of U87 Host Gene Is Associated with Ribosomes and Is Relatively Resistant to Nonsense-Mediated Decay, *Gene*, 2005, vol. 363, pp. 51–60.
35. Mattick, J.S., The Functional Genomics of Noncoding RNA, *Science*, 2005, vol. 309, pp. 1527–1528.
36. Tam, W., Identification and Characterization of Human BIC, a Gene on Chromosome 21 that Encodes a Noncoding RNA, *Gene,* 2001, vol. 274, pp. 157–167.
37. Cavaille, J., Vitali, P., Basyuk, E., et al., A Novel Brain-Specific Box C/D Small Nucleolar RNA Processed from Tandemly Repeated Introns of a Noncoding RNA Gene in Rats, *J. Biol. Chem.*, 2001, vol. 276, pp. 26 374–26 383.
38. Xiao, Y., Zhou, H., and Qu, L.H., Characterization of Three Novel Imprinted snoRNAs from Mouse *Irm* Gene, *Biochem. Biophys. Res. Commun.,* 2006, vol. 340, pp. 1217–1223.
39. Gray, T.A., Saitoh, S., and Nicholls, R.D., An Imprinted, Mammalian Bicistronic Transcript Encodes Two Independent Proteins, *Proc. Natl. Acad. Sci. USA*, 1999, vol. 96, pp. 5616–5621.
40. Runte, M., Huttenhofer, A., Gross, S., et al., The IC–SNURF–SNRPN Transcript Serves as a Host for Multiple Small Nucleolar RNA Species and as an Antisense RNA for UBE3A, *Hum. Mol. Genet.*, 2001, vol. 10, pp. 2687–2700.
41. Ohta, T., Gray, T.A., Rogan, P.K., et al., Imprinting-Mutation Mechanisms in Prader-Willi Syndrome, *Am. J. Hum. Genet.*, 1999, vol. 64, pp. 397–413.
42. Ding, F., Prints, Y., Dhar, M.S., et al., Lack of Pwcr1/MBII-85 snoRNA Is Critical for Neonatal Lethality in Prader-Willi Syndrome Mouse Models, *Mamm. Genome*, 2005, vol. 16, pp. 424–431.
43. Schule, B., Albalwi, M., Northrop, E., et al., Molecular Breakpoint Cloning and Gene Expression Studies of a Novel Translocation t(4;15)(q27;q11.2) Associated with Prader–Willi Syndrome, *BMC Med. Genet.*, 2005, vol. 6, p. 18.
44. Cavaille, J., Buiting, K., Kiefmann, M., et al., Identification of Brain-Specific and Imprinted Small Nucleolar RNA Genes Exhibiting an Unusual Genomic Organization, *Proc. Natl. Acad. Sci. USA*, 2000, vol. 97, pp. 14 311–14 316.
45. Vitali, P., Basyuk, E., Le Meur E., et al., ADAR2-Mediated Editing of RNA Substrates in the Nucleolus Is Inhibited by C/D Small Nucleolar RNAs, *J. Cell Biol.*, 2005, vol. 169, pp. 745–753.
46. Kishore, S. and Stamm, S., The snoRNA HBII-52 Regulates Alternative Splicing of the Serotonin Receptor 2C, *Science*, 2006, vol. 311, pp. 230–232.

47. Chanfreau, G., Legrain, P., and Jacquier, A., Yeast RNase III as a Key Processing Enzyme in Small Nucleolar RNAs Metabolism, *J. Mol. Biol.*, 1998, vol. 284, pp. 975–988.
48. Elela, S.A., Igel, H., and Ares, M., Jr. RNase III Cleaves Eukaryotic Preribosomal RNA at a U3 snoRNP-Dependent Site, *Cell*, 1996, vol. 85, pp. 115–124.
49. Carmell, M.A. and Hannon, G.J., RNase III Enzymes and the Initiation of Gene Silencing, *Nat. Struct. Mol. Biol.*, 2004, vol. 11, pp. 214–218.
50. Abou, Elela, S. and Ares, M., Jr., Depletion of Yeast RNase III Blocks Correct U2 3' End Formation and Results in Polyadenylated But Functional U2 snRNA, *EMBO J.*, 1998, vol. 17, pp. 3738–3746.
51. Chanfreau, G., Rotondo, G., Legrain, P., et al., Processing of a Dicistronic Small Nucleolar RNA Precursor by the RNA Endonuclease Rnt1, *EMBO J.*, 1998, vol. 17, pp. 3726–3737.
52. Kufel, J., Allmang, C., Chanfreau, G., et al., Precursors to the U3 Small Nucleolar RNA Lack Small Nucleolar RNP Proteins but Are Stabilized by La Binding, *Mol. Cell. Biol.*, 2000, vol. 20, pp. 5415–5424.
53. Petfalski, E., Dandekar, T., Henry, Y., et al., Processing of the Precursors to Small Nucleolar RNAs and RRNAs Requires Common Components, *Mol. Cell. Biol.*, 1998, vol. 18, pp. 1181–1189.
54. Allmang, C., Kufel, J., Chanfreau, G., et al., Functions of the Exosome in rRNA, snoRNA and snRNA Synthesis, *EMBO J.*, 1999, vol. 18, pp. 5399–5410.
55. Van Hoof, A. and Parker, R., The Exosome: A Proteasome for RNA?, *Cell*, 1999, vol. 99, pp. 347–350.
56. Butler, J.S., The Yin and Yang of the Exosome, *Trends Cell. Biol.,* 2002, vol. 12, pp. 90–96.
57. Kiss, T. and Filipowicz, W., Exonucleolytic Processing of Small Nucleolar RNAs from Pre-MRNA Introns, *Genes Dev.*, 1995, vol. 9, pp. 1411–1424.
58. Caffarelli, E., Arese, M., Santoro, B., et al., In Vitro Study of Processing of the Intron-Encoded U16 Small Nucleolar RNA in *Xenopus laevis*, *Mol. Cell. Biol.,* 1994, vol. 14, pp. 2966–2974.
59. Giorgi, C., Fatica, A., Nagel, R., et al., Release of U18 snoRNA from Its Host Intron Requires Interaction of Nop1p with the Rnt1p Endonuclease, *EMBO J.*, 2001, vol. 20, pp. 6856–6865.
60. Filipowicz, W. and Pogacic, V., Biogenesis of Small Nucleolar Ribonucleoproteins, *Curr. Opin. Cell. Biol.*, 2002, vol. 14, pp. 319–327.
61. Samarsky, D.A., Fournier, M.J., Singer, R.H., et al., The snoRNA Box C/D Motif Directs Nucleolar Targeting and also Couples snoRNA Synthesis and Localization, *EMBO J.*, 1998, vol. 17, pp. 3747–3757.
62. Shaw, P.J., Beven, A.F., Leader, D.J., et al., Localization and Processing from a Polycistronic Precursor of Novel snoRNAs in Maize, *J. Cell Sci.*, 1998, vol. 111, pp. 2121–2128.
63. Gall, J.G., Cajal Bodies: The First 100 Years, *Annu. Rev. Cell. Dev. Biol.*, 2000, vol. 16, pp. 273–300.
64. Gerbi, S.A., Borovjagin, A.V., and Lange, T.S., The Nucleolus: A Site of Ribonucleoprotein Maturation, *Curr. Opin. Cell. Biol.,* 2003, vol. 15, pp. 318–325.
65. Narayanan, A., Speckmann, W., Terns, R., et al., Role of the Box C/D Motif in Localization of Small Nucleolar RNAs to Coiled Bodies and Nucleoli, *Mol. Biol. Cell*, 1999, vol. 10, pp. 2131–2147.