

## NIH Public Access

**Author Manuscript** 

*Bioessays*. Author manuscript; available in PMC 2014 January 01.

### Published in final edited form as:

Bioessays. 2013 January ; 35(1): 46-54. doi:10.1002/bies.201200117.

## Processing of snoRNAs as a new source of regulatory noncoding RNAs snoRNA fragments form a new class of functional RNAs

#### Marina Falaleeva and Stefan Stamm

Department of Molecular and Cellular Biochemistry; B283 Biomedical Biological Sciences Research Building; 741 South Limestone; University of Kentucky; College of Medicine, Lexington, KY 40536-0509, USA; Phone: 859-323-0896

Stefan Stamm: stefan@stamms-lab.net

## Abstract

Recent experimental evidence suggests that most of the genome is transcribed into non-coding RNAs. The initially made transcripts undergo further processing generating shorter, metabolically stable RNAs with diverse functions.

Small nucleolar RNAs (snoRNAs) are non-coding RNAs acting in modification of rRNAs, tRNAs and snRNAs that were considered stable. We review evidence that snoRNAs undergo further processing. High-throughput sequencing and RNase protection experiments showed widespread expression of snoRNA fragments, called sdRNAs for snoRNA derived RNAs. Some sdRNAs resemble miRNAs, associate with argonaute proteins and influence translation. Other sdRNAs are longer, form complexes with hnRNPs and influence gene expression. C/D box snoRNA fragmentation patterns are conserved across multiple cell types, suggesting a processing event, rather than degradation. The loss of expression from genetic loci that generate canonical snoRNAs and processed snoRNAs results in diseases, such as the Prader-Willi Syndrome, indicating possible physiological roles for processed snoRNAs.

We propose that processed snoRNAs acquire new roles in gene expression and represent a new class of regulatory RNAs distinct from canonical snoRNAs.

## INTRODUCTION

# Identification of an increasing number of non-coding RNAs (ncRNAs), derived from longer RNA precursors

Previous work on gene expression was centered on protein coding mRNAs. The use of new RNA detection techniques, such as deep sequencing and microarray analyses showed that most of the human genome is transcribed as RNA although generally at lower levels than the 1.2% of the genome that encode proteins [1,2]. Thus, non-protein coding RNA represents the most abundant form of gene expression. Probably all initial transcripts are shortened through more than 20 different RNases [3].

The processing of RNA precursors results in an increasing number of non-coding RNA families, summarized in Tables 1 and 2. All these RNAs derive from a longer precursor through the action of various RNases (Figure 1). These RNases are diverse and include exoand endonucleases acting on single RNA strands, as well as RNases cleaving double strands. The mature non-coding RNA escapes enzymatic degradation through a variety of mechanisms. The formation of double stranded RNA structures at the ends protects them from exonucleases that cleave single stranded RNAs. Base modifications, such as 2'-O-methylation and pseudouridinylation decrease RNase sensitivity; modification of the ends by capping, adenylation or uridinylation protects against exonucleases. Most frequently forming ribonuclear protein complexes composed of RNA and proteins protect the RNA from further nuclease action.

The fragmentation of longer RNA precursors is not random, and selectivity for the cleavage sites is achieved by three major avenues: guiding RNAs, such as snoRNAs mark the site of cleavage, which is best understood for ribosomal RNAs. The second principle for marking the processing site is secondary structures, best understood for miRNAs, where double stranded regions in the RNA determine the site of cleavage by the Drosha/DGCR8 complex. Finally, proteins can mark cleavage sites by either protecting RNAs from further exonuclease activity, as discussed in the processing of snoRNAs below or by marking endonuclease sites, as exemplified by DGCR8 [4].

Through this interplay of nuclease action and RNA protection mechanism, metabolically stable shorter RNAs are formed. These RNAs generally show base modifications and are associated with proteins. Their function is subject to intense investigations. Some abundant non-coding RNAs have clear defined functions (Table 1), whereas roles for a large number of non-coding transcripts are not clear, but new functional classes are emerging (Table 2). It is currently investigated whether all non-coding RNAs have a function, represent 'transcriptional noise' or possibly represent a reservoir for future evolutionary selection [5].

Recently, fragments of medium size non-coding RNAs, snoRNAs and tRNAs have been detected [6,7]. This indicates that even RNAs that have previously been considered metabolically stable can be further processed to generate new RNA forms.

We review here the evidence that RNA fragments derived from snoRNAs influence gene expression, adding a new example of smaller, functional RNAs made from larger precursors.

#### Canonical snoRNAs: A classic example for functional ribonuclear complexes (RNPs)

Small nucleolar RNAs (snoRNAs) are 60–300 nt long non-coding RNAs that accumulate in the nucleolus. Based on characteristic sequence elements, snoRNAs are classified as C/D box and H/ACA box snoRNAs. In humans, these RNAs are usually derived from intronic regions. After the splicing reaction, introns are excised as lariats that are opened and subsequently degraded. snoRNAs escape this degradation by forming a protein complex. Most snoRNAs are located in a conserved area of vertebrate introns, which is 70–90 nt upstream of the 3' splice site [8] (Figure 2A). In addition, several snoRNAs in humans (for example U3, U8 and U13 snoRNAs) are transcribed as separate units from polymerase II promoters.

snoRNA particles (snoRNPs) contain proteins with enzymatic activities. C/D box snoRNAs contain fibrillarin that promotes the 2'-O-methylation of the RNA targets. H/ACA box snoRNPs contain dyskerin that catalyzes the conversion of uridine to pseudouridine. The overall function of snoRNAs in this process is twofold: the RNA forms a scaffold for the assembly of the RNP complex and in addition guides the enzymatic activity to distinct nucleotides of the target RNAs recognized by hybridization to specific parts of the snoRNAs. For C/D box snoRNAs, the assembly of proteins is aided by C (RUGAUGA, R=purine) and D (CUGA) boxes usually present in duplicates (C' and D' boxes) that generate a distinct secondary structure. An antisense box hybridizes to the target RNA. H/ACA box snoRNAs contain two stems that form a pseudouridinilation pocket and two single stranded regions that contain the conserved H (ANANNA) and ACA elements. Due to the bound proteins with enzymatic activity, snoRNAs modifies other non-coding RNAs, such as ribosomal RNA [9], U5 and U6 small nuclear RNAs (snRNAs) [10,11] and transfer RNA in archea [12]. C/D box snoRNAs perform 2'-O-methylation and H/ACA box snoRNAs

perform pseudouridinilation in the nucleolus. The function, biogenesis, snoRNP composition and targets of these canonical snoRNA, have been covered in excellent reviews [13–19] and are available in the LBME snoRNA database [20].

#### Orphan snoRNAs suggest non canonical snoRNA functions

The conserved overall structure of snoRNAs allows the identification of their target RNAs. However, numerous snoRNAs without target RNAs have been identified [21–24] and were termed "orphan snoRNAs". Despite the absence of an RNA that binds to their antisense sequences, these snoRNAs have physiological roles. For example, the SNORD114-1 snoRNA promotes cell cycle regulation through G0/G1 to S phase transition and is deregulated in cancer cells [17].

Another C/D box snoRNA, SNORD 115 (HBII-52), showed sequence complementarity to the serotonin receptor 2C pre-mRNA and influenced alternative splicing of this pre-mRNA [25]. Unexpectedly, the snoRNAs SNORD32A (U32A), SNORD33 (U33) and SNORD35A (U35A) accumulate in the cytosol under cellular stress conditions. As the latter are canonical snoRNAs with known ribosomal targets, it is possible snoRNAs could function outside the nucleolus (Figure 2H) [26].

These findings indicated that snoRNAs could have functions other than the nucleolar modification of ribosomal RNA (rRNA) and snRNAs. Furthermore, if snoRNAs travel at least temporarily into the cytosol, they could be subjected to non-nuclear RNases.

## snoRNAs GIVE RISE TO SHORTER FRAGMENTS (sdRNAs) CONSISTING OF miRNAs AND PROCESSED snoRNAs (psnoRNAs)

#### Identification of snoRNA fragments suggest a new class of non-coding RNAs

Deep sequencing experiments that identified shorter snoRNA fragments provided insights into how snoRNAs could perform new functions. These fragments were found in all species tested, ranging from mammalian species to *Giardia lamblia* [27] and *Epstein-Barr* virus [28]. H/ACA box snoRNAs give rise to RNAs of 17–19 nt in length. C/D box snoRNAs generate fragments longer than 27 nt. This length distribution argues that most of C/D box snoRNA fragments are different from miRNAs that have a medium length of 21–22 nt [6,29–31]. A recent thorough bioinformatic analysis supported by experimental validation suggested that about half of all C/D box snoRNAs generate shorter fragments. These fragments mostly contain C and D boxes and are conserved across multiple cell lines, indicating they are not the result of random degradation [31,32]. Well-characterized fragments are shown in Figure 3 and are listed in Supplemental Table 1.

#### Technical improvements allow detection of an increasing number of RNA fragments

Similar to other small non-coding RNAs, snoRNA fragments were discovered only recently and are part of ongoing research. In evaluating the literature, the techniques employed need to be taken into account. snoRNA fragments were detected by high-throughput sequencing and RNase protection analysis, which are more sensitive than previously employed conventional cloning techniques or oligonucleotide-based detection by Northern blot. Initially, snoRNA fragments were discovered from cDNA libraries constructed to identify miRNAs. To enrich for the expected miRNAs, the cDNA was generated from RNA selected to be less than 30–35 nts long [33] or from RNAs associated with argonaute proteins [34], which resulted in a bias towards miRNAs.

RNase protection assays offer an unbiased way to detect snoRNA fragments. In these experiments, a uniformly labeled antisense probe is hybridized to the snoRNA. RNAse

protection is more sensitive, as the probes have a higher specific activity compared to endlabeled oligonucleotide probes. For example, a 100 nt long probe has a 25 fold higher specific activity than a single end-labeled oligonucleotide. A RNA protection probe covers the whole area of a precursor RNA, which allows detecting fragments whose lengths indicate the RNA fragments. Since snoRNA fragments are often less than half the size of the full-length snoRNAs, it is important to use high-specific activity probes for detection [35] as the signal strength is directly proportional to the RNA length. End modifications of RNA can interfere with linker addition necessary for cDNA synthesis. The synthetic antisense RNA used for protection is not modified, which allows cloning of RNA fragments independent of modifications. Using a new cloning technique the sequences of snoRNA fragments generated from the C/D box snoRNAs SNORD 116 (HBII-85) and SNORD115 (HBII-52) were determined [35]. In agreement with earlier deep-sequencing studies [36], it was found that these snoRNAs generate fragments larger than miRNAs.

Together, the application of new experimental techniques showed the existence of metabolically stable snoRNA fragments. snoRNA fragments were termed psnoRNAs for processed snoRNAs [37] or sdRNAs for snoRNA derived RNAs [6]. We suggest to use the term psnoRNAs for snoRNA derived RNAs larger than 22 nt and use sdRNAs to describe both psnoRNAs and snoRNA derived miRNAs (Figure 2E).

#### psnoRNAs associate with different proteins than canonical snoRNAs

Canonical C/D and H/ACA box snoRNA associate with a defined set of proteins (fibrillarin, NOP56/58, 15.5 kD for C/D box proteins; dyskerin, NHP2, GAR1 and NOP10 for H/ACA box snoRNAs) that form the mature snoRNP [19,38]. The mechanism of snoRNP formation was studied for U3, a C/D box snoRNA acting in pre-rRNA processing that is not intron encoded, but driven by its own promoter [39]. There are at least two protein complexes associating with the U3 snoRNA during maturation. The complexes contained the core C/D box snoRNP proteins 15.5 kD, Nop 56/58 and fibrillarin. Additional proteins were present that function in snoRNP assembly (TIP48, TIP49, Nopp140), RNA processing (TGS1, La, Lsm4, Rrp46) and RNA localization (CRM1, PHAX) [40]. These findings showed that non-canonical proteins associate with snoRNAs during their maturation.

The protein composition of psnoRNAs from the SNORD115 (MBII-52) clusters was analyzed by pull down experiments from soluble nuclear extracts using oligonucleotide capturing. The RNA associates with proteins involved in different pathways such as pre-mRNA splicing (hnRNPs A1, A2/B1, A3, D0, Matrin-3, ELAV-like protein I) transcriptional regulation (pur alpha, pur beta, ATP-dependent RNA helicase A and TDP-43), and chromatin organization (Centromere protein V, nucleolin) [37,41]. These experiments suggest that most of the SNORD115 RNAs form new RNPs.

However, these experiments cannot rule out that a part of SNORD115 RNAs associates with the canonical C/D box snoRNA proteins (fibrillarin, NOP56/58, 15.5 kD) in a fraction that is not captured in the oligonucleotide pull-down. Because of its role in the Prader-Willi Syndrome, the RNAs formed from the SNORD115 locus are under intense investigations. It was found that antisera against fibrillarin immunoprecipitate SNORD115 RNA from total cell lysates [42], which suggests the formation of a canonical snoRNP complex. However, when SNORD115-RNA containing protein complexes are isolated according to their density in glycerol gradients, these complexes show a broader distribution than fibrillarin. Importantly, the fibrillarin and SNORD115 profiles do not overlap, as expected for a uniform particle [41]. In fact, it was shown that fibrillarin was absent from a large number of SNOR115 containing fractions, strongly arguing the existence of a non-canonical RNA-protein complex [41]. It is possible that fibrillarin associated SNORD115 RNA resembles

'storage' particles that under certain conditions can be processed and reorganized into different RNA:protein complexes.

Together, these reports indicate that RNAs derived from the SNORD115 cluster associate with a variety of RNA-binding proteins. It is possible the RNA forms both a canonical, fibrillarin containing RNP and new hnRNP containing complexes. In summary, the evidence shows that psnoRNAs can form ribonuclear protein complexes different from canonical snoRNAs.

#### How are psnoRNAs generated?

It is still largely unknown how snoRNAs generate shorter RNAs. Best understood is the formation of miRNAs from snoRNAs. Most mammalian miRNAs are generated from primiRNAs that undergo nuclear cleavage by Drosha/DGCR8, resulting in pre-miRNAs that are exported into the cytosol and cleaved by DICER into mature miRNAs that are loaded onto argonautes [4]. A dependence of snoRNA-generated miRNAs on DICER was experimentally confirmed for a micro RNA generated by the H/ACA box snoRNA ACA45 [34]. Other H/ACA box snoRNAs seem to follow this canonical miRNA pathway, as deep sequencing results from wild-type and DICER and DGCR8 knockout embryonic stem cells that showed different length distribution for H/ACA box snoRNA fragments [6].

The situation is different for C/D box snoRNAs as deep sequencing showed no difference in length distribution between wild-type and DICER/DGCR8 knockout cells [6]. Work in the protozoan model organism *Gardia lamblia* suggests a DICER involvement in generating miRNAs from snoRNAs. Twelve out of 25 snoRNAs found in *Gardia lamblia* generated smaller (about 30 nt long) RNAs, six of which are Ago-associated [44–47]. *Gardia lamblia* does not have homologues of DROSHA/DGCR8 and exportin 5, but has functional homologs for Dicer (GlDcr) and Argonate (GlAgo) [48]. One of the snoRNAs generating smaller fragments in *Gardia lamblia* is GlsR2, which is similar to the yeast and vertebrate U14 snoRNAs [49]. *In vitro* data and the *in vivo* GlDcr knockdown showed that the GLsR2 snoRNA is processed by unknown factors to an intermediate RNA (pre-miR5) that forms the substrate for *Gardia lamblia* DICER to form the miR5 micro RNA [45]. miR5 is loaded on *Gardia lamblia* argonaute and functions as micro RNA. This indicates miRNA generation from C/D box snoRNAs could share some, but not all features of classical miRNA formation. Since *Gardia lamblia* is an early eukaryote [50], snoRNAs could be an evolutionally conserved source of miRNAs.

Global cross-linking of cellular RNAs to DGCR8 followed by deep sequencing surprisingly identified snoRNAs bound to DGCR8, indicating that some psnoRNAs could be generated through a novel DGCR8 dependent pathway [43]. Importantly, these snoRNAs are dependent on DGCR8, but not drosha, which could indicate that DGCR8 acts with other RNAses in psnoRNA generation. The knock-down of DGCR8 increased mature snoRNAs but did not affect pre-snoRNAs abundance. This suggests that DGCR8 acts on some mature snoRNAs rather than snoRNAs precursors.

For other C/D box snoRNAs, it remains to determined whether shorter RNA fragments are generated from mature snoRNAs or from their precursors. Although there is not much known about snoRNA degradation, these RNAs likely undergo a turnover in mammals. SNORD115 resides in an actively expressed gene in non-dividing neurons, which make a degradation pathway necessary to keep a steady state concentration. psnoRNAs could represent metabolically stable forms of this degradation, which is reminiscent to miRNAs generated from precursor RNAs through RNase action (Figure 2H).

The alternative for psnoRNA generation is that the hosting intron can generate either a canonical snoRNA or a psnoRNA. The knock down of the core C/D box snoRNA protein NOP58 in yeast resulted in the accumulation of snoRNA precursor [51]. The U3 maturation showed that these precursors associate with a variety of RNA binding proteins [40], which could help in the formation and subsequent stabilization of psnoRNAs. In this model, canonical snoRNAs and psnoRNAs share a common precursor RNA (Figure 2C).

The data suggest that H/ACA and C/D box snoRNAs undergo various different processing pathways to generate miRNAs and psnoRNAs. It is possible that psnoRNAs can be formed by multiple pathways.

## FUNCTION OF psnoRNAs

Proof of principle studies in yeast and human cell lines showed that altered canonical snoRNAs could change gene expression by modifying pre-mRNAs. For example, the antisense box of a C/D box snoRNA can be modified to target a region in the pre-mRNA adjacent to the branch point, which causes 2'-O-methylation of the branch point adenosine and reduces recognition of the downstream 3' splice site [52,53].

The identification of psnoRNAs expanded these new functions and raised the question whether psnoRNAs simply represent degradation forms of psnoRNAs or are functional, which is suggested by their abundance and their association with proteins. Some snoRNA expressing units were shown to change expression levels rapidly in response to physiological stimuli. For example, the C/D box snoRNAs SNORD116 and SNORD115 expression levels changes in the brain in learning paradigms [54], which are not known to affect ribosomal modifications. This suggests that SNORD116 and SNORD115 loci could play roles outside canonical snoRNA functions, which could be fulfilled by psnoRNAs.

#### snoRNAs can be the source for miRNAs regulating translation

The best mechanistically understood role of snoRNA fractions is their function as miRNAs (reviewed in [55]), (Figure 2H). Numerous miRNAs were generated from snoRNAs. The functionality of several snoRNA derived miRNAs was confirmed in reporter gene assays [34,45]. Interestingly, some canonical miRNAs are located in the nucleolus and not in the cytosol as most other miRNAs. This could indicate a nucleolar function of some snoRNA derived miRNAs [56].

Several of the snoRNA derived miRNAs originate from orphan snoRNAs, such as SNORD83A, GlsR17, GlsR16 (Supplemental Table 1). This could indicate that some orphan snoRNAs function as a substrate for miRNA production [33]. In this model, the mature, nucleolar orphan snoRNA would act as a storage form for a miRNA precursor. Furthermore, several snoRNAs shown to act on ribosomal RNA generate miRNAs, which could indicate that these snoRNAs act on multiple aspects of translational regulation (examples in Supplemental Table 1).

#### Regulation of overall gene expression

Canonical C/D box snoRNAs interact with their target RNAs through their 10–20 nt long antisense-box elements. In contrast, psnoRNAs can interact with other sequence elements with their targets, as they form a different RNP. For example, the C/D box snoRNA SNORD88C (HBII-180C) predicted to target 28S rRNA [21] is processed into shorter fragments [31] and exhibits complementarity to fibroblast growth receptor-3 (FGFR-3) pre-mRNA in an area between the D' and C' box, called the M-box for mRNA complementarity. The M-box sequence is active, as it can change GFP reporter gene expression in a sequence dependent way [57]. These studies indicate a snoRNA mediated

#### psnoRNAs regulate alternative splicing

SNORD115 (HBII-52) is a C/D box snoRNAs that is processed into shorter fragments [25]. The full-length snoRNA shows a perfect 18 nt long complementarity between its antisense box and the alternative exon of the serotonin receptor 2C pre-mRNA. Using the antisense box sequence as a guide, six more alternative exons were identified that are regulated by SNORD115 [37], suggesting that this snoRNA can regulate multiple transcripts. Since the SNORD115 clusters generate both full-length snoRNAs and psnoRNAs, it is difficult to determine which is the functional RNA form. It is possible that the shorter psnoRNAs represents degradation products. However, since psnoRNAs are metabolically stable and form complexes with hnRNPs involved in alternative splicing regulation, it is likely that the shorter RNAs are functional in splice site selection.

SNORD88C (HBII-180C) is another snoRNA that is processed and that exhibits complementarity to a pre-mRNA [31]. In addition to changing reporter gene expression, HBII-180C expression increases exon inclusion in the FGFR-3 gene, likely by an interaction between its M-box and an intronic FGFR-3 element [32]. The reports show that psnoRNAs can change splicing through interactions of antisense and M-box elements with pre-mRNA regulatory elements (Figure 2G).

#### psnoRNAs In Disease

Given the multiple functional roles snoRNAs and their processed fragments play, it is not surprising they are involved in human diseases. HBII-52 and HBII-85 are two orphan snoRNAs expressed from an imprinted region on chromosome 15, the Prader-Willi critical region. Loss of gene expression from this region leads to Prader-Willi syndrome that is characterized by hyperphagia, hormonal imbalances, short stature and mild mental retardation. Recent genetic studies identified three patients where only the units expressing HBII-85 and HBII-52 are deleted. One of these patients has only a micro deletion of HBII-85 [58,59], suggesting the loss of these psnoRNA-expressing clusters is a decisive cause for the syndrome that could be enhanced by the loss of other proteins in the region. Prader-Willi subjects show symptoms of the autistic spectrum. Interestingly, overexpression of MBII-52 generates mice that show similar symptoms in behavioral tests [60].

As described above, several target genes for HBII-52 have been identified, but HBII-85 targets remain to be identified. Direct cloning of protected fragments [35,37] and deep sequencing of RNPs [36] indicated that HBII-52 and HBII-85 are processed into shorter fragments. However, conventional cloning suggested that HBII-52 generated only on canonical snoRNA and a minor shorter fragment of 49 nt containing the antisense box [42]. The reasons for these differences remain to be determined. They could include different techniques or biological differences in the material used.

There are a growing number of reports that a change in snoRNA expression occurs during cancer development [17,61–65]. Given the widespread processing of snoRNAs, it is likely that some of these snoRNAs act through their processed forms.

### CONCLUSION

The accumulating evidence clearly shows that most snoRNAs are processed into smaller RNAs, collectively termed sdRNAs (snoRNA derived RNAs). Similar to other non-coding RNAs, it is difficult to say whether sdRNAs represent RNAs on their way to degradation or have a cellular function. It is possible that sdRNAs represent both: fragments of longer

RNAs generated by a degradation mechanism that can be selected to function in RNA metabolism, which is indicated by their stability.

Some snoRNAs generate miRNAs, whereas others give rise to longer RNAs, psnoRNAs. The snoRNA derived miRNAs fulfill traditional cytosolic functions. Several canonical snoRNAs, which were experimentally validated to edit rRNA, generated miRNAs that influenced translation. This could indicate that snoRNAs influence protein synthesis on both rRNA processing and translation. The conversion of snoRNAs to miRNAs through RNAses could be a new integration mechanism that could become a general principle.

The processed snoRNAs have new functions in pre-mRNA processing and in controlling of gene expression. Since psnoRNAs do not form canonical snoRNPs, they can likely interact with their targets in any part of their sequence. Bioinformatic predictions for target genes should take this into account.

It is likely there is no uniform pathway for the formation of psnoRNAs and snoRNA derived miRNAs. Both RNA forms could derive from mature snoRNAs or their precursors through the action of a competing set of RNAses.

Similar to canonical snoRNAs, psnoRNAs likely serve as a scaffold to form protein complexes and also as a targeting device. Given the multiple functions described for psnoRNAs, it is likely they work with different mechanisms in various aspects of gene expression: gene transcription, pre-mRNA processing and possibly chromatin modification. Therefore, psnoRNA could play an important regulatory role in gene expression.

#### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was supported by R01GM083187.

## REFERENCES

- Clark MB, Amaral PP, Schlesinger FJ, Dinger ME, Taft RJ, et al. The reality of pervasive transcription. PLoS Biol. 2011; 9:e1000625. discussion e1102. [PubMed: 21765801]
- Pennisi E. Genomics. ENCODE project writes eulogy for junk DNA. Science. 2012; 337:1159. 61. [PubMed: 22955811]
- 3. Tomecki R, Dziembowski A. Novel endoribonucleases as central players in various pathways of eukaryotic RNA metabolism. Rna. 2010; 16:1692–1724. [PubMed: 20675404]
- Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet. 2010; 11:597–610. [PubMed: 20661255]
- Kowalczyk MS, Higgs DR, Gingeras TR. Molecular biology: RNA discrimination. Nature. 2012; 482:310–311. [PubMed: 22337043]
- Taft RJ, Glazov EA, Lassmann T, Hayashizaki Y, Carninci P, et al. Small RNAs derived from snoRNAs. Rna. 2009; 15:1233–1240. [PubMed: 19474147]
- Lee YS, Shibata Y, Malhotra A, Dutta A. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). Genes Dev. 2009; 23:2639–2649. [PubMed: 19933153]
- Hirose T, Steitz JA. Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98:12914–1299. [PubMed: 11606788]

- Kiss-Laszlo Z, Henry Y, Bachellerie JP, Caizergues-Ferrer M, Kiss T. Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. Cell. 1996; 85:1077–1088. [PubMed: 8674114]
- Jady BE, Kiss T. A small nucleolar guide RNA functions both in 2'-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. Embo J. 2001; 20:541–551. [PubMed: 11157760]
- Ganot P, Jady BE, Bortolin ML, Darzacq X, Kiss T. Nucleolar factors direct the 2'-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. Molecular and cellular biology. 1999; 19:6906–6917. [PubMed: 10490628]
- Clouet, d'Orval B.; Bortolin, ML.; Gaspin, C.; Bachellerie, JP. Box C/D RNA guides for the ribose methylation of archaeal tRNAs. The tRNATrp intron guides the formation of two ribosemethylated nucleosides in the mature tRNATrp. Nucleic acids research. 2001; 29:4518–4529. [PubMed: 11713301]
- Kiss T, Fayet-Lebaron E, Jady BE. Box H/ACA small ribonucleoproteins. Mol Cell. 2010; 37:597– 606. [PubMed: 20227365]
- Smith CM, Steitz JA. Sno storm in the nucleolus: new roles for myriad small RNPs. Cell. 1997; 89:669–672. [PubMed: 9182752]
- Brown JW, Marshall DF, Echeverria M. Intronic noncoding RNAs and splicing. Trends Plant Sci. 2008; 13:335–342. [PubMed: 18555733]
- Terns M, Terns R. Noncoding RNAs of the H/ACA family. Cold Spring Harb Symp Quant Biol. 2006; 71:395–405. [PubMed: 17381322]
- 17. Valleron W, Laprevotte E, Gautier EF, Quelen C, Demur C, et al. Specific small nucleolar RNA expression profiles in acute leukemia. Leukemia. 2012
- Matera AG, Terns RM, Terns MP. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. Nat Rev Mol Cell Biol. 2007; 8:209–220. [PubMed: 17318225]
- Reichow SL, Hamma T, Ferre-D'Amare AR, Varani G. The structure and function of small nucleolar ribonucleoproteins. Nucleic acids research. 2007; 35:1452–1464. [PubMed: 17284456]
- 20. Lestrade L, Weber MJ. snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. Nucleic acids research. 2006; 34:D158–D162. [PubMed: 16381836]
- Huttenhofer A, Kiefmann M, Meier-Ewert S, O'Brien J, Lehrach H, et al. RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. Embo J. 2001; 20:2943–2953. [PubMed: 11387227]
- 22. Jady BE, Kiss T. Characterisation of the U83 and U84 small nucleolar RNAs: two novel 2'-Oribose methylation guide RNAs that lack complementarities to ribosomal RNAs. Nucleic acids research. 2000; 28:1348–1354. [PubMed: 10684929]
- Cavaille J, Buiting K, Kiefmann M, Lalande M, Brannan CI, et al. Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. Proceedings of the National Academy of Sciences of the United States of America. 2000; 97:14311–14316. [PubMed: 11106375]
- 24. Vitali P, Royo H, Seitz H, Bachellerie JP, Huttenhofer A, et al. Identification of 13 novel human modification guide RNAs. Nucleic acids research. 2003; 31:6543–6551. [PubMed: 14602913]
- Kishore S, Stamm S. The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. Science. 2006; 311:230–232. [PubMed: 16357227]
- Michel CI, Holley CL, Scruggs BS, Sidhu R, Brookheart RT, et al. Small nucleolar RNAs U32a, U33, and U35a are critical mediators of metabolic stress. Cell Metab. 2011; 14:33–44. [PubMed: 21723502]
- Saraiya AA, Wang CC. snoRNA, a novel precursor of microRNA in Giardia lamblia. PLoS Pathog. 2008; 4:e1000224. [PubMed: 19043559]
- Hutzinger R, Feederle R, Mrazek J, Schiefermeier N, Balwierz PJ, et al. Expression and processing of a small nucleolar RNA from the Epstein-Barr virus genome. PLoS Pathog. 2009; 5:e1000547. [PubMed: 19680535]
- 29. Kawaji H, Nakamura M, Takahashi Y, Sandelin A, Katayama S, et al. Hidden layers of human small RNAs. BMC Genomics. 2008; 9:157. [PubMed: 18402656]
- Scott MS, Avolio F, Ono M, Lamond AI, Barton GJ. Human miRNA precursors with box H/ACA snoRNA features. PLoS Comput Biol. 2009; 5:e1000507. [PubMed: 19763159]

- Ono M, Scott MS, Yamada K, Avolio F, Barton GJ, et al. Identification of human miRNA precursors that resemble box C/D snoRNAs. Nucleic acids research. 2011; 39:3879–3891. [PubMed: 21247878]
- Scott MS, Ono M, Yamada K, Endo A, Barton GJ, et al. Human box C/D snoRNA processing conservation across multiple cell types. Nucleic acids research. 2012; 40:3676–3688. [PubMed: 22199253]
- Brameier M, Herwig A, Reinhardt R, Walter L, Gruber J. Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. Nucleic acids research. 2011:675–686. [PubMed: 20846955]
- 34. Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, et al. A human snoRNA with microRNA-like functions. Mol Cell. 2008; 32:519–528. [PubMed: 19026782]
- 35. Shen M, Eyras E, Wu J, Khanna A, Josiah S, et al. Direct cloning of double-stranded RNAs from RNase protection analysis reveals processing patterns of C/D box snoRNAs and provides evidence for widespread antisense transcript expression. Nucleic acids research. 2011; 39:9720–9730. [PubMed: 21880592]
- Rederstorff M, Bernhart SH, Tanzer A, Zywicki M, Perfler K, et al. RNPomics: defining the ncRNA transcriptome by cDNA library generation from ribonucleo-protein particles. Nucleic acids research. 2010; 38:e113. [PubMed: 20150415]
- Kishore S, Khanna A, Zhang Z, Hui J, Balwierz PJ, et al. The snoRNA MBII-52 (SNORD 115) is processed into smaller RNAs and regulates alternative splicing. Hum Mol Genet. 2010; 19:1153– 1164. [PubMed: 20053671]
- Matera AG, Terns RM, Terns MP. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. Nat Rev Mol Cell Biol. 2007; 8:209–220. [PubMed: 17318225]
- Terns MP, Terns RM. Small nucleolar RNAs: versatile trans-acting molecules of ancient evolutionary origin. Gene Expr. 2002; 10:17–39. [PubMed: 11868985]
- Watkins NJ, Lemm I, Ingelfinger D, Schneider C, Hossbach M, et al. Assembly and maturation of the U3 snoRNP in the nucleoplasm in a large dynamic multiprotein complex. Mol Cell. 2004; 16:789–798. [PubMed: 15574333]
- 41. Soeno Y, Taya Y, Stasyk T, Huber LA, Aoba T, et al. Identification of novel ribonucleo-protein complexes from the brain-specific snoRNA MBII-52. Rna. 2010; 16:1293–3000. [PubMed: 20484469]
- Bortolin-Cavaille ML, Cavaille J. The SNORD115 (H/MBII-52) and SNORD116 (H/MBII-85) gene clusters at the imprinted Prader-Willi locus generate canonical box C/D snoRNAs. Nucleic acids research. 2012; 40:6800–6807. [PubMed: 22495932]
- Macias S, Plass M, Stajuda A, Michlewski G, Eyras E, et al. DGCR8 HITS-CLIP reveals novel functions for the Microprocessor. Nat Struct Mol Biol. 2012; 19:760–766. [PubMed: 22796965]
- Saraiya AA, Wang CC. snoRNA, a novel precursor of microRNA in Giardia lamblia. PLoS Pathog. 2008; 4:e1000224. [PubMed: 19043559]
- 45. Li W, Saraiya AA, Wang CC. Gene regulation in Giardia lambia involves a putative microRNA derived from a small nucleolar RNA. PLoS Negl Trop Dis. 2011; 5:e1338. [PubMed: 22028939]
- Li W, Saraiya AA, Wang CC. The profile of snoRNA-derived microRNAs that regulate expression of variant surface proteins in Giardia lamblia. Cell Microbiol. 2012; 14:1455–1473. [PubMed: 22568619]
- Saraiya AA, Li W, Wang CC. A microRNA derived from an apparent canonical biogenesis pathway regulates variant surface protein gene expression in Giardia lamblia. Rna. 2011; 17:2152– 2164. [PubMed: 22033329]
- Macrae IJ, Zhou K, Li F, Repic A, Brooks AN, et al. Structural basis for double-stranded RNA processing by Dicer. Science. 2006; 311:195–198. [PubMed: 16410517]
- Yang CY, Zhou H, Luo J, Qu LH. Identification of 20 snoRNA-like RNAs from the primitive eukaryote, Giardia lamblia. Biochemical and biophysical research communications. 2005; 328:1224–1231. [PubMed: 15708007]
- Laguri C, Duband-Goulet I, Friedrich N, Axt M, Belin P, et al. Human mismatch repair protein MSH6 contains a PWWP domain that targets double stranded DNA. Biochemistry. 2008; 47:6199–6207. [PubMed: 18484749]

- Lafontaine DL, Tollervey D. Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. Rna. 1999; 5:455–467. [PubMed: 10094313]
- Semenov DV, Vratskih OV, Kuligina EV, Richter VA. Splicing by exon exclusion impaired by artificial box c/d RNA targeted to branch-point adenosine. Ann N Y Acad Sci. 2008; 1137:119– 124. [PubMed: 18837934]
- Zhao X, Yu YT. Targeted pre-mRNA modification for gene silencing and regulation. Nat Methods. 2008; 5:95–100. [PubMed: 18066073]
- Rogelj B, Hartmann CE, Yeo CH, Hunt SP, Giese KP. Contextual fear conditioning regulates the expression of brain-specific small nucleolar RNAs in hippocampus. The European journal of neuroscience. 2003; 18:3089–3096. [PubMed: 14656304]
- Scott MS, Ono M. From snoRNA to miRNA: Dual function regulatory non-coding RNAs. Biochimie. 2011; 93:1987–1992. [PubMed: 21664409]
- Politz JC, Hogan EM, Pederson T. MicroRNAs with a nucleolar location. Rna. 2009; 15:1705– 1715. [PubMed: 19628621]
- 57. Ono M, Yamada K, Avolio F, Scott MS, van Koningsbruggen S, et al. Analysis of human small nucleolar RNAs (snoRNA) and the development of snoRNA modulator of gene expression vectors. Molecular biology of the cell. 2010; 21:1569–1584. [PubMed: 20219969]
- Sahoo T, del Gaudio D, German JR, Shinawi M, Peters SU, et al. Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. Nat Genet. 2008; 40:719–721. [PubMed: 18500341]
- Duker AL, Ballif BC, Bawle EV, Person RE, Mahadevan S, et al. Paternally inherited microdeletion at 15q11.2 confirms a significant role for the SNORD116 C/D box snoRNA cluster in Prader-Willi syndrome. Eur J Hum Genet. 2010; 18:1196–1201. [PubMed: 20588305]
- Nakatani J, Tamada K, Hatanaka F, Ise S, Ohta H, et al. Abnormal behavior in a chromosomeengineered mouse model for human 15q11-13 duplication seen in autism. Cell. 2009; 137:1235– 1246. [PubMed: 19563756]
- 61. Mei YP, Liao JP, Shen J, Yu L, Liu BL, et al. Small nucleolar RNA 42 acts as an oncogene in lung tumorigenesis. Oncogene. 2012; 31:2794–2804. [PubMed: 21986946]
- Gee HE, Buffa FM, Camps C, Ramachandran A, Leek R, et al. The small-nucleolar RNAs commonly used for microRNA normalisation correlate with tumour pathology and prognosis. Br J Cancer. 2011; 104:1168–1177. [PubMed: 21407217]
- 63. Dong XY, Guo P, Boyd J, Sun X, Li Q, et al. Implication of snoRNA U50 in human breast cancer. J Genet Genomics. 2009; 36:447–454. [PubMed: 19683667]
- 64. Dong XY, Rodriguez C, Guo P, Sun X, Talbot JT, et al. SnoRNA U50 is a candidate tumorsuppressor gene at 6q14.3 with a mutation associated with clinically significant prostate cancer. Hum Mol Genet. 2008; 17:1031–1042. [PubMed: 18202102]
- Mannoor K, Liao J, Jiang F. Small nucleolar RNAs in cancer. Biochimica et biophysica acta. 2012; 1826:121–128. [PubMed: 22498252]
- 66. Watkins NJ, Bohnsack MT. The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. Wiley Interdiscip Rev RNA. 2012; 3:397–414. [PubMed: 22065625]
- 67. Phizicky EM, Hopper AK. tRNA biology charges to the front. Genes Dev. 2010; 24:1832–1860. [PubMed: 20810645]
- Kiss T. Biogenesis of small nuclear RNPs. Journal of cell science. 2004; 117:5949–5951. [PubMed: 15564372]
- 69. Huang Q, Jacobson MR, Pederson T. 3' processing of human pre-U2 small nuclear RNA: a basepairing interaction between the 3' extension of the precursor and an internal region. Molecular and cellular biology. 1997; 17:7178–7185. [PubMed: 9372950]
- 70. Okamura K. Diversity of animal small RNA pathways and their biological utility. Wiley Interdiscip Rev RNA. 2012; 3:351–368. [PubMed: 22086843]
- Meister G, Tuschl T. Mechanisms of gene silencing by double-stranded RNA. Nature. 2004; 431:343–349. [PubMed: 15372041]
- 72. Reddy R, Henning D, Busch H. Substitutions, insertions, and deletions in two highly conserved U3 RNA species. The Journal of biological chemistry. 1980; 255:7029–7033. [PubMed: 6771280]

Falaleeva and Stamm

- Moore MJ. From birth to death: the complex lives of eukaryotic mRNAs. Science. 2005; 309:1514–1518. [PubMed: 16141059]
- 74. Pillai RS, Chuma S. piRNAs and their involvement in male germline development in mice. Dev Growth Differ. 2012
- 75. Klattenhoff, C.; Theurkauf, W. Development. Vol. 135. Cambridge, England: 2008. Biogenesis and germline functions of piRNAs; p. 3-9.
- 76. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem. 2012; 81:145–166. [PubMed: 22663078]
- 77. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nat Rev Genet. 2009; 10:155–159. [PubMed: 19188922]
- 78. Yan BX, Ma JX. Promoter-associated RNAs and promoter-targeted RNAs. Cell Mol Life Sci. 2012; 69:2833–2842. [PubMed: 22415323]
- Project AET Project CSHLT. Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs. Nature. 2009; 457:1028–1032. [PubMed: 19169241]
- Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, et al. Divergent transcription from active promoters. Science. 2008; 322:1849–1851. [PubMed: 19056940]
- Sobala A, Hutvagner G. Transfer RNA-derived fragments: origins, processing, and functions. Wiley Interdiscip Rev RNA. 2011; 2:853–862. [PubMed: 21976287]
- Shi W, Hendrix D, Levine M, Haley B. A distinct class of small RNAs arises from pre-miRNAproximal regions in a simple chordate. Nat Struct Mol Biol. 2009; 16:183–189. [PubMed: 19151725]
- Taft RJ, Simons C, Nahkuri S, Oey H, Korbie DJ, et al. Nuclear-localized tiny RNAs are associated with transcription initiation and splice sites in metazoans. Nat Struct Mol Biol. 2010; 17:1030–1034. [PubMed: 20622877]
- Carone DM, Longo MS, Ferreri GC, Hall L, Harris M, et al. A new class of retroviral and satellite encoded small RNAs emanates from mammalian centromeres. Chromosoma. 2009; 118:113–125. [PubMed: 18839199]
- 85. Cao F, Li X, Hiew S, Brady H, Liu Y, et al. Dicer independent small RNAs associate with telomeric heterochromatin. Rna. 2009; 15:1274–1281. [PubMed: 19460867]
- Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science. 2007; 316:1484–1488. [PubMed: 17510325]
- Kapranov P, Ozsolak F, Kim SW, Foissac S, Lipson D, et al. New class of gene-termini-associated human RNAs suggests a novel RNA copying mechanism. Nature. 2010; 466:642–646. [PubMed: 20671709]
- 88. Jalali S, Jayaraj GG, Scaria V. Integrative transcriptome analysis suggest processing of a subset of long non-coding RNAs to small RNAs. Biol Direct. 2012; 7:25. [PubMed: 22871084]

Falaleeva and Stamm



#### Figure 1. Generation of non-coding RNAs from longer precursors

**rRNAs**: ribosomal RNAs are embedded in a larger precursor RNAs from which they are released through a series of endo and exo-nuclease activities [66].

**tRNAs**: transfer RNAs are generated from precursors that often contains introns, that are released through nuclease cleavage [67].

snRNAs: a primary pol II transcript is cleaved at the 3' end [68,69].

**snoRNAs**: in vertebrates, small nucleolar RNAs are located in introns from which they are released after lariat opening and exonuclease trimming [68].

**miRNAs**: microRNAs reside in double stranded structures, which are cleaved by Drosha/ DGCR8 and later DICER [4].

endosiRNAs: endogenous siRNAs are cleaved by DICER [70,71].

Stars indicate base modifications that have been well documented for rRNA [66], tRNA, snRNA [67] and snoRNA [72]. The dotted arrow indicates the generation of miRNAs from snoRNAs. It is currently not clear at what stage this further processing occurs.

Falaleeva and Stamm



#### Figure 2. Function and biogenesis of snoRNA derived RNAs (sdRNAs)

A: Most mammalian snoRNAs reside in introns from which they are released during the splicing reaction. After lariat opening, the intron is degraded by exonucleases (yellow). Proteins assembling on the snoRNA prevent its further degradation and mature snoRNAs accumulate in the nucleolus (**B**). It is currently unclear whether sdRNAs originate from a common precursor RNA (**C**) or are generated from snoRNAs through further processing (**D**). **E**: snoRNA give rise to psnoRNAs and miRNAs that together form snoRNA derived RNAs (sdRNAs). **F**: psnoRNAs influence gene expression and (**G**) alternative splicing. **H**: snoRNA derived miRNAs act in translational regulation. **I**: Canonical snoRNAs can accumulate in the cytosol under stress conditions.



#### Figure 3. Schematic alignment of psnoRNAs to a generic C/D box snoRNA

On the top is a hypothetical, generic C/D box snoRNA. The RNA is artificially set to 100 nt in length, which allows alignment of the functional motifs found in various snoRNAs of different lengths. The RNA elements are colored. C: C-box, D: D-box, D' and C': D' and C' boxes, AS: Antisense box. The yellow boxes at the end reflect the terminal stem structures. The psnoRNAs listed in Supplemental Table 1 are indicated schematically by showing the RNA elements they contain. These sequences were experimentally validated. Note that most include C and D boxes and lack the stems. The numbers indicate the occurrence of each form listed in Supplemental Table 1. More sequences have been identified in deep sequencing experiments [32].

#### Table 1

## Well-characterized classes of RNA in eukaryotes

Short name	Full name	Processing mechanism	Function	Reference
rRNA	ribosomal RNA	Series of exo- and endonuclease cleavages	mRNA translation	[66]
tRNA	transfer RNA	Rnase P and Z; pre-tRNA splicing	mRNA translation	[67]
snRNA	Small nuclear RNA	U1, U2, U4,U5: capping and 3'-end processing; U6: 3'-end processing	pre-mRNA splicing	[68,69]
snoRNA	Small nucleolar RNA	5'- and 3'-exonuclease trimming after pre-mRNA splicing and intron lariat opening	pre-rRNA processing, rRNA, tRNA, snRNA modifications	[68]
miRNAs	Micro RNA	Two-step cleavage by Drosha/DGCR8 and Dicer	Translation regulation	[4]
siRNAs	Small interfering RNA	Cleavage by Dicer	RNA cleavage	[70,71]
(a) Endo- siRNA	Endogenous trans- acting siRNA	Cleavage by Dicer	RNA cleavage	
(b) rasiRNA	Repeat associated siRNA	Cleavage by Dicer	DNA/histone modifications	
mRNA	messenger RNA	Splicing, capping, polyadenylation	translating the information encoded in DNA into proteins	[73]
piRNA	PIWI-interacting RNA	Aubergine, Ago3, and other exo- and endonucleases	epigenetic gene silencing in germ line cells	[74,75]
lincRNA	long intergenic non- coding RNA	diverse	Family of diverse RNA longer then 200 nt; regulation of transcription, pre-mRNA splicing, mRNA translation	[76,77]

Table 2

Emerging classes of non-coding RNAs

0	0				
Short name	Full name	Source	Processing enzyme	Function	Referencew
sdRNA	sno- derived RNAs	snoRNAs	unknown,	see Figure 2	[6,32,34]
PARs	Promoter-associated RNA			transcription regulation	[78]
a tiRNA	transcription initiation RNAs	RNA pol II transcription, backtracking on coding and non- coding genes	TFIIs transcription factor endonucleolitic cleavage	transcription regulation	[9]
b PASR	Promoter-associated small RNAs	Processing of mature mRNAs?	unknown	transcription regulation	[79]
c TSSa-RNAs	transcription start site associated RNAs	unknown	unknown	transcription regulation	[80]
tRFs	tRNA fragments	tRNA	Dicer, RNase Z or unknown	Endogenous targets have not demonstrated;	[81]
moRs	miRNA-offset RNAs	pre-miRNA	Unknown exonuclease and Drosha?	unknown	[82]
spliRNA	Splice-site RNAs	cleavage of nascent transcript?	unknown	3'-termini maps to pre-mRNA splice donor site	[83]
crasiRNA	Centromere- associated RNAs	transcription of centromeric repeats from retroviral promoters	processed from dsRNA by unknown pathway	Heterochromatin formation and centromer identity	[84]
tel-sRNA	Telomere small RNAs	long telomere RNA transcript?	Dicer independent	epigenetic regulation	[85]
TASR	gene-termini- associated human RNAs	uncharacterized pathway of copying 3'-genes terminies	unknown	unknown	[86][87]
÷	Processed long- noncoding RNAs	unknown	unknown	unknown	[88]