Non-protein coding RNAs: making their own path losing translation

Chinmoy Mishra*, Lipismita Samal

Indian Veterinary Research Institute, Izatnagar- 243 122, India

Abstract

Ribonucleic acid (RNA) can serve genetic, catalytic, structural, and regulatory roles. Non-coding RNA (ncRNA) genes make transcripts that function directly as RNA, rather than encoding proteins. Transfer RNA and ribosomal RNA are well known examples of non-coding RNA. Around 98% of all transcriptional output in humans is non-coding RNA. RNA-mediated gene regulation is widespread in higher eukaryotes and complex genetic phenomena like RNA interference, co-suppression, transgene silencing, imprinting, methylation, and possibly position-effect variegation and transvection, all involve intersecting pathways based on or connected to RNA signaling. Due to its high abundance, non-coding RNAs have several important biological roles in living organisms.

Keywords: RNA; Non-coding; Transcription; RNA interference; Imprinting

Introduction

Over the past decade there has been an explosion of large-scale genome sequencing, which leads to a dramatic change in our understanding of genome organization. Genome which was earlier considered as a linear arrangement of protein coding regions with a small set of non-coding regions is now being believed to contain chiefly two distinct classes of RNAs: messenger RNAs (mRNAs), which are translated into proteins, and the nonprotein-coding RNAs (ncRNAs), which function at the RNA level. Earlier there were only a few ncRNAs (e.g. transfer RNAs (tRNAs), ribosomal RNAs (rRNAs)and spliceosomal RNAs) were identified. However, advancement of technology reveals a much more complex genomic picture through understanding the function of small ncRNAs. Around 98% of all transcriptional output in humans is non-coding RNA (Carninci et al., 2008) and only 2% are coding part. Given the importance and ubiquity of ncRNAs and RNA-based mechanisms in all extant life forms, it is surprising that we still know relatively little about the most of RNA classes.

*Corresponding author: Chinmoy Mishra
Address: Indian Veterinary Research Institute, Izatnagar- 243 122, India
E-mail address: drchinmoymishra@gmail.com

ISSN: 2090-6277/2090-6269, www.advancedvetresearch.com
croRNAs (miRNAs) and piwi-interacting RNAs (piRNAs), have important roles as regulators for a number of biological processes. Some lncRNAs (>200 nt) have only recently emerged as a major class of eukaryotic transcripts (Bertone et al., 2004, Birney et al., 2007, Kapranov et al., 2007). Although lncRNAs have frequently been disregarded as artifacts of chromatin remodeling or transcriptional ‘noise’ (Brosius, 2005), there is substantial evidence to suggest that they mirror protein coding genes. While the majority of the identified lncRNAs are polyadenylated, an increasing number of non-polyadenylated transcripts have also been identified in recent years. In addition, many lncRNAs encompass repetitive elements, such as telomeric RNAs, satellite RNAs, long terminal repeat retrotransposon elements (LINEs) and short interspersed nuclear elements (SINEs).

**Transfer RNA**

Transfer RNA (tRNA) molecules function as carriers for the amino acids and transfer them to the machinery site of protein synthesis in the cell (Ribosome). Transfer RNA constitutes 15% of all RNA in the cell. There is, at least, a specific tRNA for each amino acid, which means that there are, at least, 20 different tRNA molecules in every cell. Each tRNA is specialized to carry one of the 20 different amino acids required for the process of protein synthesis. Transfer RNA molecule has a clover-leaf structure presenting three loops and having 2 free ends, the 3’ and 5’ ends (Fig. 1). The CCA end acts as an amino acid acceptor site.

**Ribosomal RNA**

Cell growth is intimately linked to ribosome biogenesis. The ribosome is a cytoplasmic nucleoprotein structure which acts as the machinery for the synthesis of proteins. Ribosome is essentially made up of ribosomal RNA and protein. Ribosomal RNA constitutes 80% of total RNA present in the cell. In prokaryotic cell three (i.e. 5S, 16S and 23S) types of rRNA are found, whereas in eukaryotic cells four (i.e. 5S, 5.8S, 18S and 28S) types of RNA are found.

**Micro RNA**

Micro RNAs are endogenously encoded approximately 20-24 nucleotide (nt) long (Fig. 2) single-stranded RNAs (Winter et al., 2009), expressed in a highly tissue or developmental-stage-specific fashion that specifically control translation of target mRNAs by binding to sites of antisense complementarily in 3′ untranslated regions (Grishok et al., 2001). Evidence also indicates that miRNAs regulate gene expression at posttranscriptional levels in various organisms (Reinhart et al., 2002). One of the first miRNAs discovered is let-7. The let-7 family of miRNAs is highly conserved and functions as a master temporal regulator of development and differentiation, both in early embryos and complex adult tissues (Roush and Slack, 2008).
Small interfering RNA

Small interfering RNAs (siRNAs) are 21-25 nucleotide (nt) long double stranded RNA and bear 2 nt, 3’ overhanging ends. Each siRNA has a 5’ phosphate and 3’ hydroxyl termini. RNase III family enzymes such as Dicer cleave double stranded RNA into 21 nt fragments (Ghildiyal and Zamore, 2009). Small interfering RNAs act as functional intermediate in RNA interference (RNAi) (Brodersen and Voillnet, 2006) to induce target mRNA cleavage by RNA-induced silencing complex (RISC) (Fig. 3). RNA-induced silencing complex then uses unwound antisense siRNA strand as a guide to find target mRNA and induces its endonucleolytic cleavage (Carthew and Sontheimer, 2009). Although exogenous siRNAs were discovered a decade ago, endogenous siRNAs have only recently been identified in fruit-flies and mammals (Okamura and Lai, 2008, Okamura et al., 2008).

PIWI-interacting RNAs

PIWI-interacting RNAs (piRNAs) are the Dicer-independent 26-30 nt small RNAs. They are mostly restricted to the border of germ line and somatic cells and associated with PIWI-clade Argonaute proteins regulating transposon activity and chromatin state (Malone and Hannon, 2009). Mutations in Piwi lead to defects in oogenesis and a depletion of germ line stem cells (Cox et al., 1998). In human millions of unique piRNA sequences are identified.

MicroRNA-offset RNAs

MicroRNA-offset RNAs (moRNAs) are small RNAs of 20 nt long. They are derived from the regions next to pre-miRNAs without any known function until the date (Shi et al., 2009, Langenberger et al., 2009).

Ribonuclease P and Ribonuclease mitochondrial RNA processing

Ribonuclease P (RNases P) and mitochondrial RNA processing (MRP) are ribonucleoproteins act as endoribonucleases in tRNA and rRNA processing, respectively. Ribonuclease P helps generation of the mature 5’ ends of tRNAs (Fig. 4) (Walker and Engelke, 2006). Ribonuclease MRP is eukaryote-specific, which cleaves leading to the maturation of the 5’end of 5.8S rRNA, generates RNA primers for mitochondrial DNA replication, and is involved in the degradation of certain mRNAs. The high similarity of P and MRP RNA secondary structures and the protein contents suggest that P and MRP RNAs are paralogs. RNase P RNA is found almost ubiquitously. However, MRP RNA is only found in plants, including green algae, and red algae (Piccinelli et al., 2005).

U7 small nuclear RNA

The U7 small nuclear RNA (snRNA) is the smallest polymerase II transcripts that play the key role...
in 3’ end processing (Fig. 5) of histone gene transcript (Spycher et al., 1994), which is not polyadenylated, ending instead in a conserved stem-loop sequence. A detailed analysis shows that each of its three major domains, the histone binding region, the Sm-binding sequence, and the 3’ stem-loop structure, exhibits substantial variation in both sequence and structural details across the species (Mowry and Steitz, 1987).

Promoter-associated RNAs

Promoter-associated RNAs (PARs) are broadly defined as short transcripts (250-500 nucleotides) within a few hundred bases of protein-coding or non-coding RNA transcription start sites (Fejes-Toth et al., 2009). Promoter-associated RNAs were first identified in yeast and until the date, they are described in all major eukaryotes (Taft et al., 2009b).

Telomere small RNAs

Extensive studies (mostly in yeast) demonstrated that small RNA-dependent processes are central for the formation of constitutive heterochromatin (Moazed et al., 2006, Grewal and Elgin, 2007). A methylation mechanism for heterochromatin assembly in mammalian cells by telomere small RNAs (tel-sRNAs) were suggested, but have yet to be firmly established. It remains controversial whether they are Dicer-dependent small RNAs in association with repeat sequences and heterochromatin (Cao et al., 2009).

Centrosome-associated RNAs

Centrosome-associated RNAs (crasiRNAs) are group of 34-42 nt small RNAs, derived from centrosomes that show evidence of guiding local chromatin modifications (Carone et al., 2009).

Small nucleolar RNAs (snoRNAs)

Small nucleolar ribonucleoprotein (snoRNP) particles are nucleolus-localized complexes consisting of snoRNAs and a few associated proteins. Most snoRNPs function in the modification and processing of pre-ribosomal RNA. There has been also report that several snoRNAs may as well target tRNAs and other snoRNAs (Zemann et al., 2006). Known snoRNAs can be grouped in two major classes referred to as ‘box C/D’ and ‘box H/ACA’ snoRNPs based on a conserved sequence elements and secondary structure (Fig. 6). C/D snoRNAs exhibit the “boxes” C and D with consensus sequence (A) UGAUGA and CUGA respectively. They frequently contain a second copy of these two boxes, usually designated C’ and D’, in the region between the C and D box. The H box of the H/ACA snoRNAs has the consensus sequence ANANNA. The ACA motif located at the 3’ end of the molecule is highly conserved. The two major classes direct distinct chemical modifications.

Most C/D snoRNAs determine specific target nucleotides for methylation of the 2’-hydroxyl posi-

Fig. 5. Processing of histone gene transcript by U7 Sn RNA.

Fig. 6. Small nucleolar RNA
tion of the sugar, while most of the H/ACA snoR-NAs specify uridines that are converted into pseudouridine by rotation of the base.

**Signal Recognition Particle RNA**

The signal recognition particle (SRP) is a RNA polymerase III transcript ribonucleoprotein complex that interacts with the ribosome during the synthesis and translocation of secretory proteins. It helps co-translational targeting of proteins that contain signal peptides to membranes, including the prokaryotic plasma membrane and the endoplasmic reticulum. The assembled SRP consists of two structurally and functionally distinct domains (Fig. 7). The small domain (Alu-domain of the SRP RNA and the proteins SRP9 and SRP14) modulates the elongation of the secretory protein and the larger domain (S-domain of the SRP RNA and several specific proteins) captures the signal peptide. Eukaryotic SRP RNA is also known as 7SL RNA.

**Telomerase RNA**

Maintenance of linear chromosomes requires a specific mechanism to compensate for the terminal sequence loss that occurs pursuant to conventional DNA replication in eukaryotes. Eukaryotes rely on a cellular reverse transcriptase, telomerase that synthesizes tandem tracts of a short guanine-rich sequence onto the 3’ overhangs of chromosomes (Schoeftner and Blasco, 2008). The main components of telomerase are its reverse transcriptase and telomerase RNA containing a short template sequence. The RNA component varies dramatically in sequence composition and size (Podlevsky et al., 2008). Although several telomerase RNAs have been identified, the known examples are restricted to four narrow phylogenetic groups: vertebrates, yeasts, ciliates, and plasmodia (Fig. 8).
Spliceosomal snRNAs

Removal of the intervening sequences (introns) from pre-mRNAs is a crucial and ubiquitous step in eukaryotic gene expression (Fig. 9). The process of intron splicing takes place in a massive ribonucleoprotein complex known as the spliceosome which consists of up to 300 proteins and five small non-coding RNAs (Nilsen, 2003).

Three distinct variants of spliceosomal machinery may be present in eukaryotic cells. The major group of spliceosome contains U1, U2, U4, U5 and U6 snRNAs and removes introns delimited by the canonical donor-acceptor pair GT-AT. The minor group of spliceosome is composed of U11, U12, U4atac, U5 and U6atac snRNAs (Patel and Steitz, 2003) and acts on AT-AC introns (Sheth et al., 2006). The snRNAs U11, U12, U4atac, and U6atac take on the roles of U1, U2, U4, and U6. Whereas, both U6 and U6atac are polymerase-III transcripts, all other spliceosomal snRNAs are transcribed by polymerase-II. The snRNAs themselves are not only part of the spliceosomes but are also involved in transcriptional regulation (Kwek et al., 2002).

Small nucleolar-derived RNAs

Small nucleolar-derived RNAs (sdRNAs) are small RNAs (some of which are Dicer-dependent), which are processed from snoRNAs. Some sdRNAs have been shown to function as miRNA-like regulators of translation (Taft et al., 2009a).

Transfer-messenger RNA

Translation of a messenger RNA (mRNA) lacking a stop codon often results in a ribosome stalled at the end of the mRNA with a peptidyl-tRNA in the P site and with an empty A site (Fig. 10). The transfer-messenger RNA (tmRNA) which is also known as 10Sa RNA or SsrA rescues these stalled ribosomes, promoting release factor binding and thus providing the template for a peptide-tag that then causes the rapid degradation of the incomplete translation product (Marzluff, 2005).

Transfer-messenger RNA that combines the functions of a tRNA and a mRNA has two distinct domains tRNA-like (TLD) and a mRNA-like (MLD) domain linked by RNA-helices and RNA-pseudoknots (Keiler et al., 1996). Transfer RNA-like domain binds to the protein SmpB and then aminoacylated by an alanyl-tRNA synthetase. A quaternary complex that contains elongation factor EF-Tu recognizes ribosome stalled at the 3’ end of a nonstop mRNA and like tRNA, enters the ribosome’s A site. The nascent polypeptide is transferred to the alanyl-tmRNA, which now switches to its mRNA-like mode of action by translocating to the P site of the ribosome where it places a TAG codon in the ribosome’s mRNA channel. This leads to the release of the defective mRNA and its subsequent selective degradation by RNase R. The ribosome continues translation with the tmRNA ORF as a surrogate template and terminates at a tmRNA-encoded stop codon, thereby releasing the nascent protein with the 11-amino acid degradation
tag, which contains epitopes for ubiquitin proteases for degradation.

Y RNAs

Y RNAs have so far been reported only in vertebrates, the nematode Caenorhabditis elegans and the prokaryote Deinococcus radiodurans. They are short RNA polymerase-III transcripts with characteristic secondary structure (Fig. 11) without any known function yet (Farris et al., 1999). However, Y RNAs are seemed to be component of Ro RNP particle and Y5 that are implicated in 5S rRNA quality control (Hogg and Collins, 2007).

6S RNA

The 6SRNA is one of the bacterial small RNAs that binds specifically to the bacterial RNA-Polymerase (RNAP) holoenzyme and selectively inhibits σ70 dependent transcription (Maeda et al., 2000). It seems to mimic the open σ70 dependent promoter complex (Cavanagh et al., 2008). This RNAP-6S complex is present when NTP concentrations are low and as soon as new NTPs become available to the RNAP-6S complex, the 6S RNA serves as a template for the transcription of the 14-20nt pRNAs (Wassarman, 2007). This results in the formation of an unstable 6S RNA-pRNA complex which releases the RNAP-holoenzyme (Wassarman and Saecker, 2006). 6S RNA has three domains closing stem central bubble and terminal loop (Fig. 12).

Vault RNAs

Vaults are large ribonucleoprotein particles that are ubiquitous in eukaryotic cells with characteristic barrel-like shape with poorly understood function in multi-drug resistance (van Zon et al., 2003). Vault RNAs are short (Fig. 13) polymerase-III transcripts having two internal polymerase-III promoter sequences, Box A and Box B, and a typical terminator sequence at the 3’ end (Mosig et al., 2007). The terminal stem is conserved among all known examples. They exhibit little sequence conservation beyond their Box A and Box B internal promoter elements.
7SK RNA

The 7SK snRNA is one of the polymerase-III transcripts, with a length of about 330nt (Wassarman and Steitz, 1991). It is highly abundant ncRNAs in vertebrate cells and function as a transcriptional regulator by inhibiting the transcription elongation factor P-TEFb, a critical regulator of RNA polymerase-II. 7SK RNA is highly variable in both sequence and structure (Fig. 14). Only two stem-loop structures, towards the 5’ and the 3’ end, are conserved. The 5’-terminal hairpin is responsible for HEXIM1 and PTEFb binding (Egloff et al., 2006), the 3’-stem interacts with P-TEFb.

Ornate large extremophilic RNA

The “ornate large extremophilic” (OLE) RNA with a length of approximately 610nt is highly conserved in both sequence and secondary structure non-coding RNA (Fig. 15). Ornate large extremophilic is found predominantly in extremophilic Gram-positive eubacteria. A functional link between OLE RNA and membrane protein has been suggested (Puerta-Fernandez et al., 2006).

Guide RNAs

Mitochondrial mRNAs of some protozoa need to undergo a post-transcriptional editing process before translation. Kinetoplastids of the trypanosomatid group possess two types of mitochondrial DNA molecules: Maxi-circles bear protein and ribosomal RNA genes. Mini-circles specify guide RNAs (gRNAs), with a typical length of about 50nt that mediate uridine insertion/deletion RNA editing. Following the hybridization of the 5’-anchor region of a gRNA to the 3’ end of its target mRNA, a U insertion and deletion is directed by sequential base pairing (Alatortsev et al., 2008).

Conclusion

Everyday there are new class of non-coding RNA are discovered. So it is very difficult to conduct both qualitative as well as quantitative study of non-coding RNA genes. However, completion of genome sequencing will help understand the genome in a more comprehensive way that may ease our effort to identify the non-coding RNA. New generation of sequencing technology will lead to understand the insight of non-coding RNA genes in a much better way. However, sharing of information should not be overlooked.

References

Alatortsev, V.S., Cruz-Reyes, J., Zhelonkina, A.G., Sollner-


in a broad range of eukaryotes. Nucleic Acids Research 33, 4485-4495.