CHAPTER 2

RNA INTERACTIONS

Manja Marz*[,1,2] and Peter F. Stadler[1,3-5]

[1]Bioinformatics Group, Department of Computer Science and Interdisciplinary Center for Bioinformatics, University of Leipzig, Leipzig, Germany; [2]RNA-Bioinformatics Group, Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marburg, Germany; [3]Institute for Theoretical Chemistry, University of Vienna, Wien, Austria; [4]Fraunhofer Institute for Cell Therapy und Immunology, Leipzig, Germany; [5]Santa Fe Institute, Santa Fe, New Mexico, USA

**Corresponding Author: Manja Marz—Email: manja@staff.uni-marburg.de*

Abstract: Noncoding RNAs form an indispensible component of the cellular information processing networks, a role that crucially depends on the specificity of their interactions among each other as well as with DNA and protein. Patterns of intramolecular and intermolecular base pairs govern most RNA interactions. Specific base pairs dominate the structure formation of nucleic acids. Only little details distinguish intramolecular secondary structures from those cofolding molecules. RNA-protein interactions, on the other hand, are strongly dependent on the RNA structure as well since the sequence content of helical regions is largely unreadable, so that sequence specificity is mostly restricted to unpaired loop regions. Conservation of both sequence and structure thus this can give indications of the functioning of the diversity of ncRNAs.

INTRODUCTION

The discovery of a universe of noncoding RNAs (ncRNAs) of unexpected diversity during the last decade has profoundly changed our understanding of the role RNA in particular in eukaryotic cells. Merely a decade ago, RNA was mostly thought of as a (rather boring) intermediate in protein expression (mRNAs). A few classes of ancient RNAs, some with catalytic functions (RNase P RNA, ribosomal RNAs and self-splicing introns), were seen as remnants of an ancient RNA World, an evolutionary dead end in protein-dominated modern cells. In striking contrast, we now see ncRNAs as an integral part of the regulatory machinery in modern cells. In fact, there is mounting evidence that

RNA Infrastructure and Networks, edited by Lesley J. Collins.

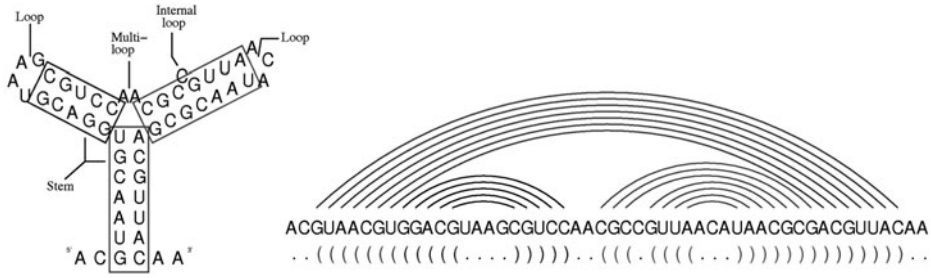

Figure 1. Example of an RNA secondary structure in three different representations: secondary structure graph (left), in bracket notation (right, bottom) and as arc plot (right, top). A pseudoknot is most easily defined by crossing arcs in its arc plot.

RNAs form a complex network of regulatory interactions among each other, with the genome and with a large variety of proteins.[1,2]

The spatio-temporal structure of the cellular "RNA infrastructure"[2] is shaped by the diversity of interactions of RNA molecules. Formation of base pairs between complementary nucleic acids is the key for both the structure formation of individual RNAs and for interactions between RNA and/or DNA molecules. The patterns of base pairing constitute the secondary structures (Fig. 1). They characterize functional classes of ncRNAs and are often well-conserved over large evolutionary time-scales (reviewed in see refs.3,4). Specific base pairing also contributes to the tertiary structure of the RNA.[5] The importance of specific base pairing is not limited to the structure of a single RNA, however. It has a crucial impact on the tertiary (3D) structure of the RNA and thus on way in which RNA is recognized by protein partners. Secondary structures are also formed by interacting RNAs, thus determining both strength and exact position of hybridization between two or more partners.

In this chapter we focus on how ncRNAs interact with their partners in a large number of different molecular and functional contexts. Base-pairing patterns play a central role because of their large impact on the structure of binding sites for proteins, or more directly, because the interaction consists of intermolecular base pairs. We organize our contribution by the composition of the players: RNA with DNA, RNA with RNA and RNA with proteins. Our presentation does not strive to be exhaustive and attempts to give a broad overview rather than an in-depth discussion of specific examples. Wherever possible, we thus cite recent reviews rather than the original literature.

ncRNA-DNA INTERACTIONS

RNA:DNA hybrids play a crucial role in transcription termination in bacteria: the formation of the terminator hairpin in the nascent RNA transcripts shortens the RNA:DNA duplex associated with the polymerase complex and facilitates dissociation.[6] Another important case are the RNA primers of the Okazaki fragment on the lagging strand during DNA replication.[7] Despite their use in biotechnology,[8,9] very little is known about the potential of triple-helices as form of direct RNA:DNA interaction.

Comprehensive surveys such as the ENCODE and FANTOM projects[10,11] demonstrated that the genomes of higher eukaryotes are pervasively transcribed. In mammals, many of these large intergenic noncoding RNAs (lincRNAs) associate with chromatin-modifying complexes and affect gene expression.[12-15] Apparently, there are diverse roles for lincRNAs in processes from embryonic stem cell pluripotency to cell proliferation, based on the observation that these transcripts are differentially regulated by key transcription factors such as p53, NFkappaB, or Nanog. Similar mechanisms are at work also in the yeast *Saccharomyces cerevisiae*[16] and in plants.[17,18] A significant fraction of long noncoding RNAs are subject to at least moderate stabilizing selection (where mutants of well adapted functional molecules are associated with a detrimental effect on fitness and hence are selected against) on the exon,[19] corroborating the functionality of these transcripts. Even transcripts with little or no sequence conservation may be functional as shown by the deep conservation of the gene structure itself.[20]

At present, multiple models for the mode of action of these long ncRNA transcripts have been proposed, reviewed e.g., in[21] and there is at least circumstantial evidence that different transcripts may be governed by different mechanisms (Fig. 2). Epigenetic regulator proteins directly or indirectly mediate chromatin regulation. By the first mechanism they are recruited from transcribed ncRNA or released from the acting site during transcription. By the latter mechanism regulative epigenetic markers may interact directly or via antisense RNA mediator with ncRNAs in order to bind to the target DNA. Therewith expression states might be switched and epigenetic memory is established.

Y RNAs were originally discovered as an RNA component of Ro RNPs, binding to Ro60 which at first is not related to DNA interaction. This RNP is active in RNA quality control[22] requiring RNA-RNA interactions between Y RNA and a variety of target RNAs. However, more recently Y RNAs were shown to be essential for chromosomal DNA replication.[23] Here, a cluster of Y RNAs is associated with the protein RPA and the sliding clamp PCNA and other proteins.[24-26] Y RNAs are essential for DNA replication. Not surprisingly, these ncRNAs are over expressed in many human cancer cells. So far, Y RNAs have been described in vertebrates and nematodes.[27-29]

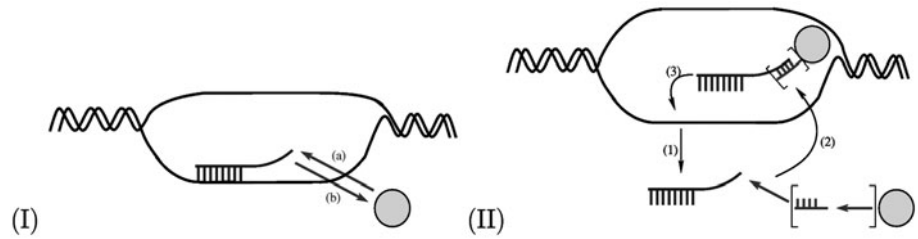

Figure 2. Putative mechanisms of chromatin-associated RNAs (modified from ref. 21: (I) Direct RNA mediated chromatin regulation during transcription. Epigenetic regulator proteins (circles) are (a) recruited directly from transcribed ncRNA or released from acting site during transcription (b). (II) Indirect association to DNA regulatory elements after transcription. After transcription of long ncRNAs from DNA (1), regulative epigenetic regulators (circles) may interact directly (without brackets) or via antisense RNA mediator (brackets) with ncRNAs (2) in order to bind to DNA (3). With this mechanism the recruitment of epigenetic regulatory proteins, switches off gene expression states and maintenance of epigenetic memory is performed.

RNA-RNA INTERACTIONS

General Properties

RNA-RNA interactions provide one of the fundamental mechanisms of cellular regulation. Single-stranded nucleic acids readily form complex interaction structures (cofolds) stabilized by complementary base pairing, thereby achieving a high sequence specificity. This recognition principle is utilized for wide variety of biological functions, including the decoding of the genetic code. Each codon is recognized by a complementary anti-coding presented by the corresponding tRNA. In this case, the interaction covers only three nucleotides, which is stabilized by the surrounding ribosomal machinery.[30] The direct binding of small RNAs to each other also plays a crucial role for the catalytic activity of the spliceosome, where snRNAs and the pre-mRNA cofold.[31]

At present, we lack high-throughput methods to assess RNA-RNA interactions both in vitro and in vivo. The direct proof of RNA-RNA binding in vivo is at least very difficult so that most if not all experimental reports on two interacting RNA-RNA molecules come from in vitro experiments. To this end chemical probing[32] is used to determine the secondary structure of both the isolated binding partners and their interaction structures. In practice, chemical probing, which determines by means of context-specific degradation whether a nucleotide is paired or unpaired only provides constraints which in some cases imply an unique RNA secondary structure model, while in other cases ambiguities remain.

In contrast to the experimental difficulties, RNA secondary structures and RNA-RNA interactions can be predicted efficiently by computational methods. The folding problem is posed as the following combinatorial matching problem with certain constraints. A matching in a graph consists of edges (here representing the base pairs) that have no vertices (here the nucleotides) in common. The task of computing a matching with the maximal number of edges is a classical problem in combinatorial optimization that can be solved efficiently. The nucleotides of the two sequences are represented as vertices of a graph, whose edges encode the logically allowed Watson-Crick and GU base pairs. We then search for a matching in this graph (i.e., a subset of edges such that every vertex/nucleotides take part in only one edge) that satisfies further structural constraints and maximizes an energy function. For RNA secondary structure prediction, for instance, one requires that base-pairs do not cross each other. The energy function accounts for base-pair stacking and loop entropies, i.e., it favours parallel base pairs and discounts long unpaired regions.[33] Although the RNA-RNA interaction problem (RIP) and the closely related RNA folding problem with arbitrary pseudoknots (see Fig. 1) are NP-complete (i.e., not be solved efficiently as the required computation scales exponentially with the length of the RNA sequence) in their most general forms,[34] efficient polynomial-time dynamic programming algorithms (a solution is recursively composed of solutions of smaller sub-programs) can be derived by restricting the space of allowed configurations.

The simplest approaches concatenate the two interacting sequences and subsequently employ a slightly modified standard secondary structure folding algorithm,[35] possibly allowing some pseudoknots.[36] The resulting model, however, still does not generate all relevant interaction structures. Alternatively, internal base-pairs in the interaction partners are neglected.[37] RNAup[38] and intaRNA[39] restrict interactions to a single contact interval; this model class has proved particularly useful for bacterial sRNA/mRNA interactions. To-date only a handful of interaction structures are known that are even more complex, some of which we will encounter later in this section.

MicroRNAs and RNA Interference

MicroRNAs were the first small regulatory RNAs found in animals. During the past few years, a wide variety of additional classes of small RNAs associated with the RNA interference (RNAi) pathways[40] were discovered throughout Eukarya. They share many functional properties and utilize at least in part the same processing machinery (see refs. 41, 42 for recent reviews). Their size of about 20-30nt and their final destination define these RNAs as a reasonably homogeneous group. Without going into detail about their biochemical processing, they guide large protein complexes to their targets, thus comprising the "RNA sensor" allowing sequence specific binding of the proteins. Both miRNAs and siRNAs form subclasses of this large class of small ncRNAs involved predominantly in gene silencing. MiRNAs stand out from the other small RNAs in several ways: they are processed from an extremely stable hairpin-structured precursor and they are typically highly conserved over long evolutionary timescales,[43] indicating crucial regulatory functions.

Although microRNAs clearly interact by basepairing with their targets, our understanding of the rules that govern the recognition of microRNA target sites is limited. Surprisingly, even exact complementarity is insufficient in some cases, while rather poor interactions are still functional in some cases. For animal microRNAs, the "seed rule", for instance, postulates that exact complementarity between nucleotides 2-7 with the mRNAs is necessary. There are several well-documented exceptions, however[44] (Fig. 3). In case of *C.elegans* insertions in the seed region are known as well as for plants mismatches beside several GU pairs. On the other hand, no mRNA cleavage was shown for perfect helix formations (mir-834/CIP4.1).

MicroRNAs may act in two distinct ways to downregulate protein expression either by translational repression or by *slicing*, i.e., endonuclease cleavage of the target mRNA. The latter pathway is the principle mode of action of siRNAs and possibly other sub-classes of short RNAs. MicroRNA target complementarity (Fig. 3) does not necessarily predict the regulatory output of the interaction. In particular complementarity does not imply slicing, while central mismatches prevent slicing, reviewed in reference 44. This observation is

Figure 3. Structures of miRNA/target interactions adapted from reference 44. Let-7/lin-41, mir-398a/CSD1 and mir-298/Nanog violate the seed hypothesis by having mismatches or GU pairs in the (shaded) seed region. The mir-2a/Tom and lin-4/lin-14 interactions are functional but show little complementarity outside the seed region. Although mir-834/CIP4.1 forms a perfect helix, it does not lead to mRNA cleavage, mir-398a/CSD1, on the other hand, can act according to both mechanisms. The vertical line indicates the putative mRNA cleavage position.

consistent with structural models that suggest that the mRNA faces the Ago RNase H active site about 10nt from the beginning of the miRNA–mRNA duplex.[45] Imperfectly paired target sites in animal transcripts often occur multiple times in the same mRNA. Cooperative action of multiple sites has also been observed.

Small Bacterial RNAs and Translational Control

Small noncoding RNAs (sRNAs) form an important class of regulators of gene expression. In bacteria, their role in the cell was not fully appreciated until the discovery of hundreds of sRNAs in *E. coli* and many other species.[46,47] While the chaperone protein Hfq plays an important role in facilitating sRNA/mRNA interactions, sRNAs are also abundant in species such as *Helicobacter pylori* that lack Hfq[48] [(Fig. 4)(see Chapter 14, pages 224-227 for more detail on Hfq related regulation)].

The exact location and structure of the sRNA/target duplex has a decisive impact on the function of the sRNA. While most sRNAs act to downregulate protein expression exceptions are well known. Typically, the sRNAs regulate translation by influencing the accessibility of the Shine-Dalgarno (SD) sequence. If the SD is contained in a stable secondary structure, translation is inhibited. Most sRNAs achieve this by binding directly to the SD, thus covering the SD by a sRNA/mRNA interaction.[49] Several sRNAs, however, (among them DsrA, GlmZ, RNAIII, RprA, RyhB and Qrr) activate translation by binding the 5′ mRNA region in such a way that the mRNA refolds to liberate a sequestered ribosome binding site.[50] The pairing of GadY sRNA to the 3′-end alters processing and increases mRNA stability.

The binding of a small RNA to an mRNA can have long-range effects through a refolding of the mRNA. The binding affinity of *HuR* protein to human mRNAs, for example, can be tuned by artificial "opener" and "closer" RNAs that interact far away from the *HuR* binding motif.[51,52] So far, it remains open to what extent such induced RNA refolding plays a role in eukaryotes. Changes in RNA secondary structure can also be induced by temperature changes or metabolite binding. A wide variety of procaryotic riboswitches, reviewed in,[53] utilizes this mechanism to control either translation (by hiding or exposing the SD) or transcription (by forming a premature terminator signal upstream of the coding sequence). Complex sensors consisting of elaborate RNA structures also play a role in recently-evolved eukaryotic mRNAs.[54,55] A recent study of the primary transcriptome of *Helicobacter pylori*[48] demonstrated an unexpectedly high number of anti-sense transcripts and alternative transcription start sites, indicating that there may be an additional layer of regulatory complexity based on these novel transcripts.

Small Nucleolar RNAs and Chemical Modifications

Chemical modifications of RNAs are an evolutionarily old phenomenon. In particular, tRNAs are heavily edited. In total, over 100 structurally distinguishable modified nucleosides have been encountered so far in different types of RNAs from many diverse organisms of all domains of life.[56,57] These modifications increase the chemical diversity of RNA and are beneficial or even critical for proper folding and function of the RNA molecule. Most types of modifications are carried out by specialized, site-specific enzymes. A large number of pseudouridine and 2′-O-methylated nucleotides however, are synthesized by generic RNP complexes. They recognize their target sites in rRNAs, snRNAs and tRNAs through their RNA components, the H/ACA and C/D snoRNAs,

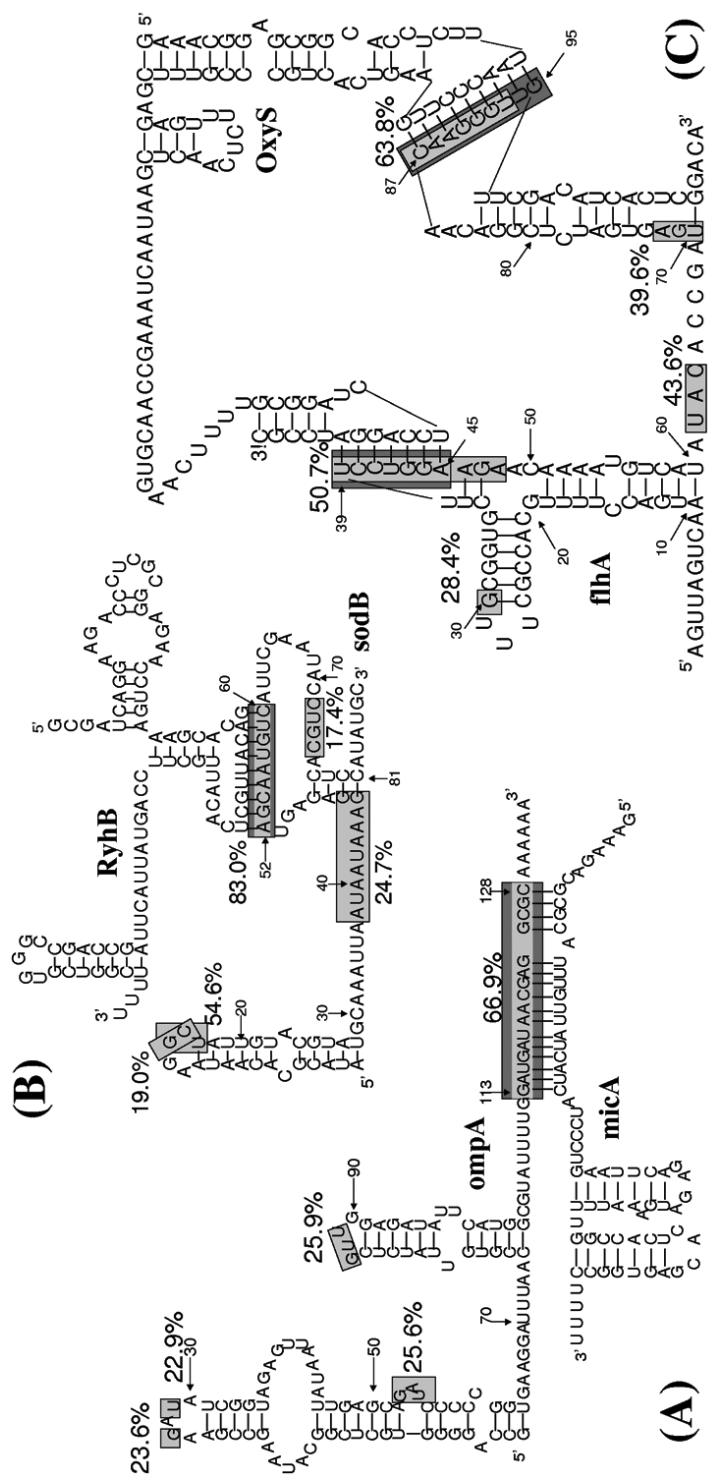

Figure 4. Three examples of interaction between bacterial sRNA and mRNAs. The primary interaction region(s), corresponding to the published structures are highlighted with double boxes, hybridization probabilities computed by rip2[135] are annotated by shaded boxes, identifying additional hybridization regions that may further stabilize the interaction. A) The small RNA micA accumulates in *E. coli* upon entry into stationary phase and down-regulates expression of ompA (outer membrane protein A) by binding across the Shine-Dalgarno sequence of this mRNA. B) RyhB represses translation of the superoxide dismutase SodB in response to low iron concentration in several bacteria. C) OxyS is induced in response to oxidative stress in *E. coli* and represses the translation of fhlA, a transcriptional activator for formate metabolism.

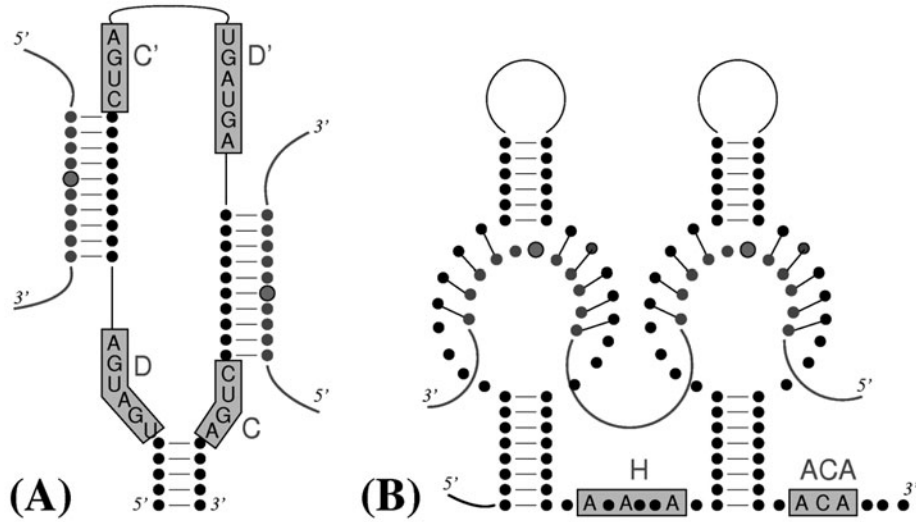

Figure 5. The two classes of snoRNA have different structures and different interactions with their targets. Box C/D snoRNAs (A) guide 2′ O-methylation to the fifth nucleotide from the D and D′ boxes. Some C/D snoRNAs lack a discernible D′ box and modify only a single target. In contrast, the target site of box H/ACA snoRNAs (B) is determined by the structure. The target uracil (shaded) is positioned by two specific interactions of the flanking target RNA sequence with the complementary sequence of the recognition loop of the snoRNA.[136]

respectively.[58] These two classes of enzymatically active RNPs are ubiquitous in eukaryotic and archaeal organisms but are not found in bacteria. H/ACA and C/D snoRNAs are evolutionarily and structurally unrelated (Fig. 5).

A small class of hybrid snoRNAs contain an H/ACA and a box C/D motif. An example is U85, guiding both the pseudouridylation of base U46 and the 2′O-ribose methylation of base C45 of the U5 snRNA.[59] Several snoRNAs, in particular those of hybrid structure but also several otherwise canonical ones, contain a CAB box (AGAG, typically located in a hairpin loop). The signal directs the snoRNAs to the Cajal body, where they guide modifications of the RNA polymerase II-transcribed snRNAs (U1, U2, U4 and U5). In Archaea, the snoRNAs also target tRNA precursors.[60]

An increasing number of orphan snoRNAs, i.e., snoRNAs lacking targets in rRNAs or snRNAs, has been described in different eukaryotes. In particular, a subgroup of snoRNAs expressed in the mammalian brain does not appear to be involved in modifications of rRNAs or snRNAs. Instead some of these snoRNAs make use of the same type of interaction to target specific mRNAs and appear to interfere with A-to-I editing, which in turn can alter the functional properties of proteins, silence constitutive activity and modulate RNA translation, localization and stability.[61,62]

RNAs in RNA End-Processing

The overwhelming majority of protein-coding pol-II transcripts ends in a poly(A) tail that is generated by endonucleolytic cleavage followed by polyadenylation.[63] This is not an absolute rule, however. The mRNAs of the replication-dependent histones lack

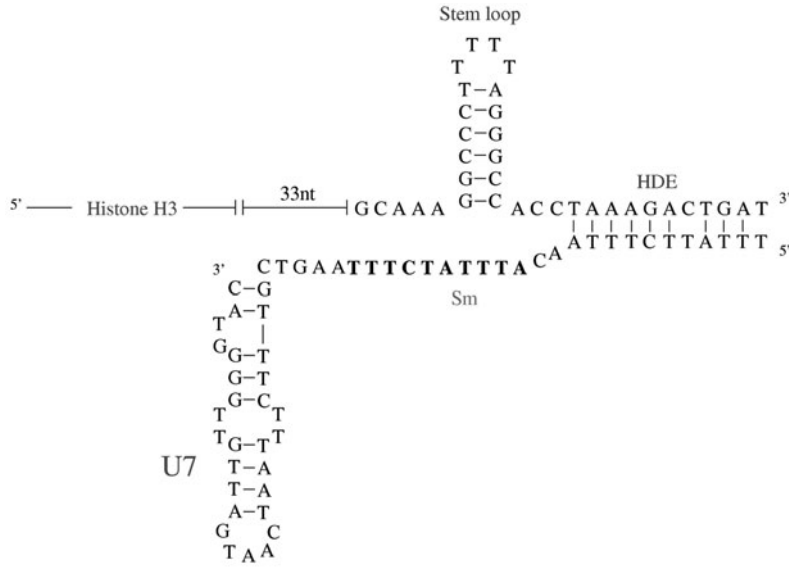

Figure 6. U7-directed processing of histone 3′ ends. U7 snRNA interacts upstream of the Sm protein binding site with histone downstream element (HDE), about 30nt downstream of histone-pre-mRNA.

poly(A) and instead feature a highly conserved stem–loop structure in their 3′ UTRs. It binds the stem–loop binding protein that ensures RNA stability and enhances translational efficiency. The 3′ end is determined by base-pairing of the histone-down-stream element (HDE) with the U7 snRNA (Fig. 6) which directs endonucleolytic cleavage.[64] In contrast, the 3′ ends of the pol-II transcribed snRNAs are produced by the *Integrator*, a specialized protein complex.[65] The 3′ end of the telomerase RNA is processed by the spliceosome in *Schizosaccharomyces pombe*.[66] Finally, RNase P RNA processes the 3′ ends of the two long ncRNAs MALAT-1 and MENε, utilizing a tRNA-like element to attract the tRNA processing machinery. Besides tRNAs, RNase P RNA in addition cleaves the yeast HRA1 ncRNAs, bacterial riboswitches and possibly also some box C/D snoRNAs [(see ref. 67 for a review)(see Chapter 1, Figure 2, page 5 for a summary of RNase P activity, and Chapter 14, pages 222-223 for more detail on RNase P in prokaryotes)].

The rRNA operon is transcribed by pol-I. The maturation of the ribosomal RNAs involves a complex cascade of processing steps. In Eukarya and possibly Archaea, this involves also the use of the atypical box C/D U3 snoRNA[68] that acts as an RNA-chaperone mediating the correct structure conformations of the pre-rRNA for endonuclease cleavage.[69] Unlike other small nucleolar RNAs, U3 snoRNA acts not by a direct 2′-O-methylation, instead the 5′ domain (consisting of A and A′ box) and the single stranded hinge region (connecting the the 5′ and 3′ domain of U3 snoRNA) bind to the prerRNA. The 3′ domain of U3 snoRNA hypermethylates itself, arranges RNA stability and has essential conserved protein binding sites.

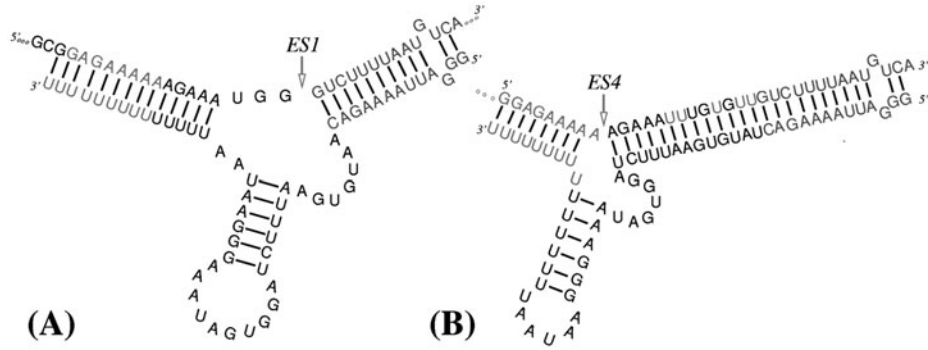

Figure 7. U insertion editing of the CYb mRNA of *Trypanosoma*.[137] A single guide RNA (bottom strand) encodes multiple editing sites. The gRNA and its mRNA target (upper strand) form a 3-stem structure that is refolded as the processing moves from one editing site to the next.

Guide RNAs and RNA Editing

Many of the mRNAs of the kinetoplasts (mitochondria) of Trypanosomes and Leishmanias are edited by inserting and/or deletion of uridines, a process that depends on a specialized class on ncRNAs, the guide RNAs (gRNAs). In contrast to the ncRNA-guided editing system in Kinetoplastida, the C->U editing in plant organelles is directed by *cis*-acting elements and a plethora of PPR enzymes (reviewed in see ref. 70).

The gRNAs of kinetoplastids are typically 50-70nt long and contain three functional elements. A 5-21 nt region on their 5′ side acts as anchor specifically recognizing the target mRNA. The "guide region" in the middle of the molecule serves a template for editing. It is complementary (allowing GU pairs, however) to the mature, edited mRNA. The 3′-tail consists of posttranscriptionally added poly-U tail. The editing process takes place in the editosome, a complex structure comprising of more than 20 proteins arranged around the mRNA/gRNA pair.[71] The editosome typically performs several successive rounds of enzymatic reactions templated by a gRNA (Fig. 7). Most of the gRNAs are not encoded in the major maxicircle DNA, which also contains the mitochondrial rRNA and protein-coding genes. Instead, large numbers of gRNA genes are located on the minicircles.[72,73]

The U insertions and deletions can be extensive. The 1246 nts ND7 mRNA of *Trypanosoma brucei*, for instance, is processed by inserting 551 and deleting 86 uridines. It is no surprise, therefore, that the U insertions and deletions typically change the open reading frames of the mRNAs. Intriguingly, the premRNAs of the *Trypanosoma brucei* mitochondrion are edited in sometimes alternative ways to yield distinctive protein sequences. Multiple gRNAs that target the same primary transcript[72] expand the diversity of mitochondrial proteins since they can produce significantly different alternative ORFs. These editing variants are of functional importance as the case of AEP-1 (alternatively edited protein 1) shows, which is involved in the maintenance of kinetoplastid DNA.

CRISPRs

Immunity against viruses and plasmids in 40% of Eubacteria and 90% of Archaea is connected to clustered regularly interspaced short palindromic repeat sequences

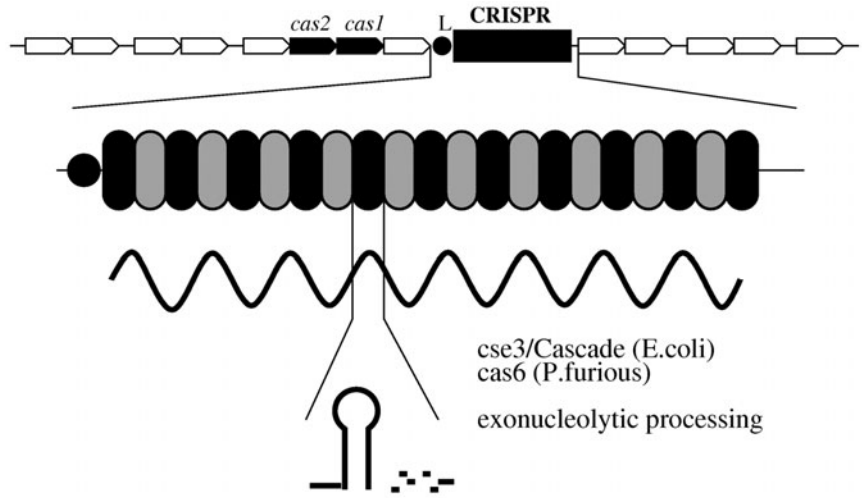

Figure 8. Processing of CRISPR content into crRNAs (redrawn from refs. 74,79). Cas genes (black constrained, white optionally) upstream and downstream of CRISPR cluster. The Leader sequence (L) is located directly upstream of CRISPR. Transcribed CRISPRs are cleaved within the repeat by cse3 or cas6. Additionally, precrRNA is exonucleolytic processed into mature crRNA consisting of a repeat tag (folding into a single hairpin) and the spacer sequence.

(CRISPRs). Transcripts from these loci target foreign nucleic acids in by specific-sequence detection. In addition to other well-described defense strategies, such as prevention of adsorption, blocking of injection and abortive infection,[74] the CRISPR clusters store invading genetic information in preparation for an "immune response" during a second infection. Immunization is established in three phases:[75,76] (1) incorporation of new spacers into CRISPR arrays, (2) expression and processing of CRISPR RNAs (crRNA) and (3) CRISPR interference.

Various *cas* proteins are located upstream and optionally downstream of CRISPR cluster (Fig. 8). Six core genes *cas1* to *cas6*, located <1kb around CRISPR have been identified to interact beside repeat-associated mysterious proteins (RAMP) with mature processed CRISPR (crRNA). *cas1* acts as endonuclease within immunization process and *cas2 and cas6* are a sequence specific endoribonuclease. The CRISPR leader located upstream of repeat tags is defined by a low-complexity A-T-rich region. This area acts as promoter for the CRISPR transcript. The spacer addition provides novel phage resistance, whereas spacer deletion was shown to result in a loss of phage resistance.[77] With the CRISPR/Cas system, horizontal gene transfer between distant organisms was shown by resistance of bacteria that were never invaded by that phage before.[78]

Recently, a possible relationship between CRISPR interference and eukaryotic RNA interference (RNAi) by siRNA, miRNA and piRNA has been discussed,[76] since there are many obvious similarities: (1) RNA guiding effector apparatus to the target, (2) the gene function is affected in a programmable and sequence directed manner, (3) adaptive and heritable components used to establish recoverable genomic records of past invasions (e.g., piRNA of eukaryotic silencing). On the other hand, the two systems clearly are not homologous: both the protein machinery is completely different and crRNAs are not amplified posttranscriptionally from single stranded precursors. For more details,

we refer to the reviews in references 74, 76 and 79 (see Chapter 13, pages 213-218 for more detail on the CRISPR system).

Ribozymes

A ribonucleic acid enzyme (ribozyme) is an RNA molecule that, like the much more familiar protein-based enzymes, catalyzes a chemical react ion. Although most of them function together with protein components in vivo, experiments in in vitro prove that the catalytic activity resides in the RNA component. Natural ribozymes include, for example, the peptidyl transferase activity of the 23S rRNA, the hammerhead and hairpin ribozyme, and RNase P RNA. Similar to the ribosome, the spliceosome is in essence also a ribozyme.[80] The efficiency of the catalysis, however, depends to a certain extent on secondary structure features of the mRNA to be catalyzed.[81]

Ribonuclease P (RNase P) and mitochondrial RNA processing (RNase MRP) are ribozymes acting in processing tRNA and rRNA, respectively. The RNA subunits are evolutionarily related. RNase MRP is eukaryotic specific whereas RNase P is present in all kingdoms of life.[82] The high similarity of P and MRP RNA secondary structures[83] and similarity of the protein contents and interactions of RNase P and MRP[84,85] suggest that P and MRP RNAs are paralogues. Both enzymes consist of an RNA molecule and 9-12 protein subunits, most of which are part of both enzymes. An *in silico* relationship was shown for Pop8 and Rpp14/Pop5 as well as Rpp25 and Pop6.[86] The RNA subunit and its interacting proteins build a coevoluting network.

Beyond the relatively small collection of natural ribozymes, it is quite easy to "breed" catalytically active RNAs (and even single-stranded DNAs) in in vitro selection (SELEX) experiments, see for example a recent review.[87] In fact, most of the chemical transformations that are catalyzed by protein enzymes in extant organism can be implemented in moderate-size RNAs. This lends further credibility not only to the RNA World Hypothesis but opens the possibility that at least some ncRNAs might also be catalytically active. There is, however, no indication that RNA catalysis is wide-spread phenomenon beyond some classes of "selfish elements" such as viroids.

ncRNA-PROTEIN INTERACTIONS

Ribonucleoparticles

Ribonucleoparticles (RNPs) are composed of both RNA and protein components. In some cases, such as RNase P, where the RNA is catalytically active, one speaks of ribonucleoproteins. In the following we briefly explore the huge variety of RNA-protein interactions using 7SK RNA, SRP RNA, Telomerase RNA, spliceosomal RNAs and vault RNA as examples.

The 7SK RNA has been described in a diverse set of animals.[88-92] It negatively controls transcription by regulating the activity of the Positive Transcription Elongation Factor b (P-TEFb) for Polymerase II.[93,94] In this process HEXIM1/2 proteins dissolve from P-TEFb and bind to the polymerase III transcript 7SK RNA.[95-97] Furthermore the La-related protein 7 (LARP7) was proven to regulate the stability of 7SK RNA.[88,98,99] The network of 7SK RNA and its specific protein partners is a metazoan invention.[90-92,100] The well characterized LARP7 homologs[101] are found in most metazoan clades, HEXIM

was found in all clades of metazoans excluding Platyhelminthes. Two copies are known for eutheria.[92,102] In contrast, another protein known to interact with 7SK RNA, MePCE/BCDIN3 has a much broader phylogenetic distribution, indicating that it has other important functions beyond its interaction with 7SK RNA.

The signal recognition particle (SRP) is present in all living organisms. This ribonucleoparticle targets ribosomes to the endoplasmatic reticulum (ER) in order to translate proteins at the ER. The SRP complex consists of a highly conserved 300 nt SRP RNA (or 7S RNA) and six proteins (9, 14, 19, 54, 68 and 72 SRPs) in eukaryotes. In archaeas four protein particles are absent and SRP19 and SRP54 exists compared to one Ffh protein (homolog to SRP54) and a RNA molecule (also known as 4.5S RNA or 6S RNA in *Bacillus*).[103] Although the network of SRP RNA and corresponding proteins is highly conserved, substantial variations between the three main kingdoms of life are clearly visible.[104]

Eukaryotic cells have been shown to decrease protein synthesis and increasing the expression of protein quality control mechanisms, such as chaperones and proteases. The bacterial SRP receptor, FtsY, inhibits the translation of both SRP-dependent and SRP-independent proteins.[105] In higher plants two different SRP-dependent mechanisms are known: once proteins are posttranslational transferred to chloroplasts and on the other hand proteins are cotranslational encoded by the plastid genome.[106] Although for the first system in general no RNA component seems to be necessary. However in single plastids from red algae and chlorophyta a reminiscent eubacterial SRP was identified. The cotranslational protein requires the SRP RNA, which accelerates the interaction between the SRP and SRP receptor 200-fold.[107] This SRP RNA is missing in the chloroplast SRP (cpSRP) pathway. Instead, the cpSRP and cpSRP receptor (cpFtsY) by themselves can interact 200-fold faster than their bacterial homologues.[107]

Linear chromosomal ends of most eukaryotes are replicated by a telomerase enzyme which consists of a catalytic protein component, the telomerase reverse transcriptase (TERT) and the core functional unit telomerase RNA.[66] Telomerase dates back to the origin of eukaryotes.[108] Notable exceptions are diptera including *Anopheles* and *Drosophila*, which use retrotransposons or unequal recombination instead of a telomerase enzyme. Although TERT proteins are experimentally validated for a wide range of eukaryotes, the knowledge of RNA:protein interacting networks implies that telomerase RNA exists for more organisms than the experimentally validated small phylogenetic groups within vertebrates, yeasts, ciliates and plasmodia.[109,110] Telomerase is normally active only early in development and in stem cells and the germline. It extends the ends of linear chromosomes and thus counteracts the gradual shortening of the ends caused by intrisic limitations of the replication machinery. In cancer research, telomerase has received a lot of interest because of this crucial role in cell proliferation and the observation that telomerase is also active in many cancer cells.[111]

Several small ncRNAs are involved in splicing processes. Small nuclear RNAs (snRNAs) act with up to 200 proteins as a large RNP (spliceosome) in eukaryotes to splice exons of protein-coding genes together.[112] We find two homologous systems for splicing in eukaryotes: The major spliceosome involves U1, U2, U4, U5 and U6 snRNA and splices more than 90% of (human) protein-coding genes. In contrast, the minor spliceosome utilizes the alternative U11, U12, U4atac, U5 and U6atac RNAs and plays an ancillary role in the nucleus. Although the splicing reactions seem to date back until the last unknown common ancestor (LUCA), its extensive use in protein expression seems to be an eukaryotic innovation.[113] Contrary to *cis*-splicing, in eight eukaryotic phyla

(kinetoplastids, nematodes, platyhelminthes, cnidaria, arthropods, tunicates, rotifera, dinoflagellates) a short leader sequence derived from small SL RNA is added to the 5′ end of a mRNA by *trans*-splicing.[114,115] The Sm protein binds to SL RNAs in order to act in this spliceosome-catalyzed process. In nematoda SmY RNAs are hypothesized to recycle the spliceosome proteins after SL RNAs are consumed in the *trans*-splicing reaction.[116] Direct interactions are not known.

While the spliceosome and its RNAs are well understood, there are several much more enigmatic ncRNA species. One example is the vault ribonucleoprotein complex, one of the largest RNPs. It consists of 8-16 noncoding vault RNAs, a major vault protein (MVP) and two minor vault proteins (VPARP and TEP1).[117] It is involved in transport processes to and from the nuclear membrane. Vault RNAs are polymerase III transcripts with length of about 100nt which are known throughout the deuterostomes[118] and in a few protostomes.[119] The RNA components are not necessary for the structural assembly of vault particles.[120] They strongly bind to a variety of organic molecules including chemotherapy drugs, which might explain the relation of the vault complex with chemoresistance.

Long ncRNA as Coat-Hangers?

A handful of long noncoding RNAs appear to play a crucial role in the organization of complex RNPs, acting like a "coat-hanger" to bring together different proteins. An example is the huge telomerase RNA of yeasts.[121] Another example is the heat-shock ncRNAs hsrω of *Drosophila*. Different isoforms of hsrω are expressed nearly ubiquitously in a developmentally regulated pattern. Upon heat shock, the nuclear-retained longest isoform sequesters a variety of nuclear RNA processing factors.[122] Similarly, NEAT1 (also known as MENε) may act as organizing factor of the paraspeckles,[123,124] which are ribonucleoprotein bodies found in the interchromatin space of mammalian cell nuclei. The hsrω-n RNA of *Drosophila* probably plays a similar function in the organization of the ω speckles.[122] It is a crucial component of the stress response.

At present, it is unknown to what extent other long ncRNAs function as stabilizing backbones of large RNPs, bringing together specific combinations of proteins and what role such complexes might play.

CONCLUSION

There is compelling evidence, that ncRNAs and especially long ncRNAs do not constitute a homogeneous class of transcripts but rather differ dramatically in structure, function, mechanisms and processing. While some are retained in the nucleus, others are predominantly cytoplasmic. As in the case of *Drosophila* hsrω, different splice variants and different poly-adenylation sites can be associated with differences in cellular localization.[122] Some mRNA-like noncoding RNs (mlncRNAs) are primarily involved in shaping spatial structures. Others, such as the noncoding host genes of snoRNAs and microRNAs[125] are expressed to deliver their pay-load of small structured RNAs to the appropriate processing machinery. Yet another group is processed into a large collection of short RNAs such as the precursors of piRNAs,[126] and nuclear ncRNAs interact directly with the transcription machinery.[127]

It may not come as a surprise that functionalities are not strictly separated. Many of the evolutionary ancient and well-conserved "housekeeping RNAs", including tRNAs,

snoRNAs and snRNAs, give rise to smaller processing products[118,128-130] that at least in some cases have their own function in gene regulation.[131-133] Similarly, MALAT1 and MENβ are prolific sources of small RNAs.[134] We may conclude therefore, that the network of a cell's RNA infrastructure is not only shaped by complex patterns of interactions among RNAs and between RNA and protein and DNA, but also by an intricate tangle of processing pathways that extract multiple functions from different "developmental stages" of the same primary transcript.

ACKNOWLEDGEMENTS

This work was supported in part by the State of Saxony through a "Landesstipendium".

REFERENCES

1. Amaral PP, Dinger ME, Mercer TR et al. The eukaryotic genome as an RNA machine. Science 2008; 319:1787-1789.
2. Collins LC, Penny D. The RNA infrastructure: dark matter of the eukaryotic cell? Trends Genet 2009; 25:120-128.
3. The Athanasius F. Bompfunewerer RNA Consortium: RNA's everywhere: genome-wide annotation of structured RNA's. J Exp Zool B: Mol Dev Evol 2007; 308B:1-25.
4. Bompfünewerer AF, Flamm C, Fried C et al. Evolutionary patterns of noncoding RNAs. Theory Biosci 2005; 123:301-369.
5. Leontis NB, Westhof E. Analysis of RNA motifs. Curr Opin Struct Biol 2003; 13:300-308.
6. Komissarova N, Becker J, Solter S et al. Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. Molecular Cell 2002; 10:1151-1162.
7. MacNeill SA. DNA replication: partners in the Okazaki two-step. Curr Biol 2001; 11(20):R842-R844.
8. Nadal A, Eritja R, Esteve T et al. "Parallel" and "antiparallel tail-clamps" increase the efficiency of triplex formation with structured DNA and RNA targets. Chembiochem 2005; 6:1034-1042.
9. Suzuki T. Targeted gene modification by oligonucleotides and small DNA fragments in eukaryotes. Front Biosci 2008; 13:737-744.
10. The ENCODE Project Consortium. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 2007; 447:799-816.
11. The FANTOM Consortium. The Transcriptional Landscape of the Mammalian Genome. Science 2005; 309:1159-1563.
12. Khalil AM, Guttman M, Huarte M et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci USA 2009; 106:11675-11680.
13. Sheik MJ, Gaughwin PM, Lim B et al. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. RNA 2010; 16:324-337.
14. Mattick JS, Amaral PP, Dinger ME et al. RNA regulation of epigenetic processes. Bioessays 2009; 31:51-59.
15. Mondal T, Rasmussen M, Pandey GK et al. Characterization of the RNA content of chromatin. Nat Methods. 2010;7(8):582
16. Harrison BR, Yazgan O, Krebs JE. Life without RNAi: noncoding RNAs and their functions in Saccharomyces cerevisiae. Biochem Cell Biol 2009; 87:767-779.
17. Matzke M, Kanno T, Daxinger L et al. RNA-mediated chromatin-based silencing in plants. Curr Opin Cell Biol 2009; 21:367-376.
18. Swiezewski S, Liu F, Magusin A et al. Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. Nature 2009; 462:799-802.
19. Marques AC, Ponting CP. Catalogues of mammalian long noncoding RNAs: modest conservation and incompleteness. Genome Biol 2009; 10:R124.
20. Hiller M, Findeiß S, Lein S et al. Conserved introns reveal novel transcripts in Drosophila melanogaster. Genome Res 2009; 19:1289-1300.
21. Hekimoglu B, Ringrose L. Noncoding RNAs in polycomb/trithorax regulation. RNA Biol 2009; 6:129-137.
22. Stein AJ, Fuchs G, Fu C et al. Structural insights into RNA quality control: the Ro autoantigen binds misfolded RNAs via its central cavity. Cell 2005; 121:529-539.

23. Christov CP, Gardiner TJ, Szüts D et al. Functional requirement of noncoding Y RNAs for human chromosomal DNA replication. Mol Cell Biol 2006; 26:6993-7004.
24. Krude T. Non-coding RNAs: New players in the field of eukaryotic DNA replication. In: Nasheuer HP, ed. Genome Stability and Human Diseases. New York: Springer, 2010:105-118.
25. Szüts D, Christov C, Kitching L et al. Distinct populations of human PCNA are required for initiation of chromosomal DNA replication and concurrent DNA repair. Exp Cell Res 2005; 311:240-250.
26. Szüts D, Kitching L, Christov C et al. RPA is an initiation factor for human chromosomal DNA replication. Nucleic Acids Res 2003; 31:1725-1734.
27. Mosig A, Guofeng M, Stadler BMR et al. Evolution of the vertebrate Y RNA cluster. Theory Biosci 2007; 126:9-14.
28. Perreault J, Perreault JP, Boire G. Ro-associated Y RNAs in metazoans: evolution and diversification. Mol Biol Evol 2007; 24:1678-1689.
29. Boria I, Gruber AR, Tanzer A et al. Nematode sbRNAs: homologs of vertebrate Y RNAs. J Mol Evol 2010.
30. Aitken CE, Petrov A, Puglisi JD. Single ribosome dynamics and the mechanism of translation. Annu Rev Biophys 2010; 39.
31. Valadkhan S. snRNAs as the catalysts of pre-mRNA splicing. Curr1 Op1 Chem1 Biol 2005; 9:603-608.
32. Brunel C, Romby P. Probing RNA structure and RNA-ligand complexes with chemical probes. Methods Enzymol 2000; 318:3-21.
33. Mathews D, Sabina J, Zuker M et al. Expanded sequence dependence of thermodynamic parameters improves prediction of scRNA secondary structure. J Mol Biol 1999; 288:911-940.
34. Alkan C, Karakoc E, Nadeau JH et al. RNA-RNA interaction prediction and antisense RNA target search. J Comput Biol 2006; 13:267-282.
35. Hofacker IL, Fontana W, Stadler PF et al. Fast folding and comparison of RNA secondary structures. Monatsh Chem 1994; 125:167-188.
36. Rivas E, Eddy SR. A Dynamic programming algorithms for RNA structure prediction including pseudoknots. J Mol Biol 1999; 285:2053-2068.
37. Rehmsmeier M, Steffen P, Höchsmann M et al. Fast and effective prediction of microRNA/target duplexes. Gene 2004; 10:1507-1517.
38. Mückstein U, Tafer H, Hackermüller J et al. Thermodynamics of RNA-RNA binding. Bioinformatics 2006; 22:1177-1182.
39. Busch A, Richter AS, Backofen R. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. Bioinformatics 2008; 24:2849-2856.
40. Kim D, Rossi J. RNAi mechanisms and applications. Biotechniques 2008; 44:613-616.
41. Carthew RW, Sontheimer EJ. Origins and mechanisms of miRNAs and siRNAs. Cell 2009; 136:642-655.
42. Verdel A, Vavasseur A, Le Gorrec M et al. Common themes in siRNA-mediated epigenetic silencing pathways. Int J Dev Biol 2009; 53:245-257.
43. Hertel J, Lindemeyer M, Missal K et al. The expansion of the metazoan MicroRNA repertoire. BMC Genomics 2006; 7:15.
44. Brodersen P, Voinnet O. Revisiting the principles of microRNA target recognition and mode of action. Nat Rev Mol Cell Biol 2009; 10:141-148.
45. Song JJ, Smith SK, Hannon GJ et al. Crystal structure of Argonaute and its implications for RISC slicer activity. Science 2004; 305:1434-1437.
46. Vogel J, Sharma CM. How to find small noncoding RNAs in bacteria. Biol Chem 2005; 386:1219-1238.
47. Waters LS, Storz G. Regulatory RNAs in bacteria. Cell 2009; 136:615-628.
48. Sharma CM, Hoffmann S, Darfeuille F et al. The primary transcriptome of the major human pathogen Helicobacter pylori. Nature 2010; 464:250-255.
49. Sharma CM, Darfeuille F, Plantinga TH et al. A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. Genes Dev 2007; 21:2804-2817.
50. Fröhlich KS, Vogel J. Activation of gene expression by small RNA. Curr Opin Microbiol 2009; 12:674-682.
51. Meisner NC, Hackermüller J, Uhl V et al. mRNA openers and closers: modulating AU-rich element-controlled mRNA stability by a molecular switch in mRNA secondary structure. Chembiochem 2004; 5:1432-1447.
52. Hackermüller J, Meisner NC, Auer M et al. The effect of RNA secondary structures on RNA-ligand binding and the modifier RNA mechanism: a quantitative model. Gene 2005; 345:3-12.
53. Henkin TM. Riboswitch RNAs: using RNA to sense cellular metabolism. Genes Dev 2008; 22:3383-3390.
54. Kaempfer R. RNA sensors: novel regulators of gene expression. EMBO Rep 2003; 4:1043-1047.
55. Ray PS, Jia J, Yao P et al. A stress-responsive RNA switch regulates VEGFA expression. Nature 2009; 457:915-919.
56. Limbach PA, Crain PF, McCloskey JA. Summary: the modified nucleosides of RNA. Nucleic Acids Res 1994; 22:2183-2196.
57. Dunin-Horkawicz S, Czerwoniec A, Gajda MJ et al. MODOMICS: a database of RNA modification pathways. Nucleic Acids Res 2006; 34:D145-D149.

58. Bachellerie JP, Cavaillé J, Hüttenhofer A. The expanding snoRNA world. Biochimie 2002; 84:775-790.
59. Henras AK, Capeyrou R, Henry Y et al. Cbf5p, the putative pseudouridine synthase of H/ACA-type snoRNPs, can form a complex with Gar1p and Nop10p in absence of Nhp2p and box H/ACA snoRNAs. RNA 2004; 10:1704-1712.
60. Singh SK, Gurha P, Gupta R. Dynamic guide-target interactions contribute to sequential 2′-O-methylation by a unique archaeal dual guide box C/D sRNP. RNA 2008; 14:1411-1423.
61. 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; 169:745-753.
62. Rogelj B. Brain-specific small nucleolar RNAs. J Mol Neurosci 2006; 28:103-109.
63. Proudfoot N. New perspectives on connecting messenger RNA 3′ end formation to transcription. Curr Opin Cell Biol 2004; 16:272-278.
64. Marzluff WF. Metazoan replication-dependent histone mRNAs: a distinct set of RNA polymerase II transcripts. Curr Opin Cell Biol 2005; 17:274-280.
65. Baillat D, Hakimi MA, Näär AM et al. Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. Cell 2005; 123:265-276.
66. Box JA, Bunch JT, Tang W et al. Spliceosomal cleavage generates the 3′ end of telomerase RNA. Nature 2008; 456:910-914.
67. Wilusz JE, Spector DL. An unexpected ending: noncanonical 3′ end processing mechanisms. RNA 2010; 16:259-266.
68. Marz M, Stadler PF. Comparative analysis of eukaryotic U3 snoRNA. RNA Biol 2009; 6:503-507.
69. Atzorn V, Fragapane P, Kiss T. U17/snR30 Is a Ubiquitous snoRNA with Two Conserved Sequence Motifs Essential for 18S rRNA Production. Mol Cell Biol 2004; 24:1769-1778.
70. Shikanai T. RNA editing in plant organelles: machinery, physiological function and evolution. Cell Mol Life Sci 2006; 63:698-708.
71. Simpson L, Aphasizhev R, Gao G et al. Mitochondrial proteins and complexes in Leishmania and Trypanosoma involved in U-insertion/deletion RNA editing. RNA 2004; 10:159-170.
72. Ochsenreiter T, Cipriano M, Hajduk SL. Alternative mRNA editing in Trypanosomes is extensive and may contribute to mitochondrial protein diversity. PLoS One 2008; 3:e1566.
73. Madej MJ, Niemann M, Hüttenhofer A et al. Identification of novel guide RNAs from the mitochondria of Trypanosoma brucei. RNA Biology 2008; 5:84-91.
74. Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. Science 2010; 327:167-170.
75. van der Oost J, Jore MM, Westra ER et al. CRISPR-based adaptive and heritable immunity in prokaryotes. Trends Biochem Sci 2009; 34:401-407.
76. Marraffini LA, Sontheimer EJ. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nat Rev Genet 2010; 11:181-190.
77. Barrangou R, Fremaux C, Deveau H et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science 2007; 315:1709-1712.
78. Godde JS, Bickerton A. The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. J Mol Evol 2006; 62:718-729.
79. Karginov FV, Hannon GJ. The CRISPR system: small RNA-guided defense in bacteria and archaea. Mol Cell 2010; 37:7-19.
80. Valadkhan S. The spliceosome: a ribozyme at heart? Biol Chem 2007; 388:693-697.
81. Warf MB, Berglund JA. Role of RNA structure in regulating pre-mRNA splicing. Trends Biochem Sci 2010; 35:169-178.
82. Piccinelli P, Rosenblad MA, Samuelsson T. Identification and analysis of ribonuclease P and MRP RNA in a broad range of eukaryotes. Nucleic Acids Res 2005; 33:4485-4495.
83. Collins LJ, Moulton V, Penny D. Use of RNA secondary structure for studying the evolution of RNase P and RNase MRP. J Mol Evol 2000; 51:194-204.
84. Aspinall TV, Gordon JMB, Bennet HJ et al. Interactions between subunits of Saccharomyces cerevisiae RNase MRP support a conserved eukaryotic RNase P/MRP architecture. Nucleic Acids Res 2007; 35:6439-6450.
85. Walker SC, Engelke DR. Ribonuclease P: the evolution of an ancient RNA enzyme. Crit Rev Biochem Mol Biol 2006; 41:77-102.
86. Rosenblad MA, López MD, Piccinelli P et al. Inventory and analysis of the protein subunits of the ribonucleases P and MRP provides further evidence of homology between the yeast and human enzymes. Nucleic Acids Res 2006; 34:5145-5156.
87. Talini G, Gallori E, Maurel MC. Natural and unnatural ribozymes: back to the primordial RNA world. Res Microbiol 2009; 160:457-465.
88. Krueger BJ, Jeronimo C, Roy BB et al. LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. Nucleic Acids Res 2008; 36:2219-2229.
89. Murphy S, Di Liegro C, Melli M. The in vitro transcription of the 7SK RNA gene by RNA polymerase III is dependent only on the presence of an upstream promoter. Cell 1987; 51:81-87.

90. Gruber AR, Kilgus C, Mosig A et al. Arthropod 7SK RNA. Mol Biol Evol 2008; 25:1923-1930.
91. Gruber AR, Koper-Emde D, Marz M et al. Invertebrate 7SK snRNAs. J Mol Evol 2008; 66:107-115.
92. Marz M, Donath A, Verstraete N et al. Evolution of 7SK RNA and its protein partners in metazoa. Mol Biol Evol 2009; 26:2821-2830.
93. Peterlin BM, Price DH. Controlling the elongation phase of transcription with P-TEFb. Mol1 Cell 2006; 23:297-305.
94. Egloff S, Van Herreweghe E, Kiss T. Regulation of polymerase II transcription by 7SK snRNA: two distinct RNA elements direct P-TEFb and HEXIM1 binding. Mol1 Cell1 Biol 2006; 26:630-642.
95. Yik JH, Chen R, Nishimura R et al. Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. Mol Cell 2003; 12:971-982.
96. Michels AA, Nguyen VT, Fraldi A et al. MAQ1 and 7SK RNA interact with CDK9/cyclin T complexes in a transcription-dependent manner. Mol Cell Biol 2003; 23:4859-4869.
97. Blazek D, Barboric M, Kohoutek J et al. Oligomerization of HEXIM1 via 7SK snRNA and coiled-coil region directs the inhibition of P-TEFb. Nucleic Acids Res 2005; 33:7000-7010.
98. Markert A, Grimm M, Martinez J et al. The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes. EMBO Rep 2008; 9:569-575.
99. He S, Su H, Liu C et al. MicroRNA-encoding long noncoding RNAs. BMC Genomics 2008; 21:236.
100. Gürsoy HC, Koper D, Benecke BJ. The vertebrate 7S K RNA separates hagfish (Myxine glutinosa) and lamprey (Lampetra fluviatilis). J1 Mol1 Evol 2000; 50:456-464.
101. Bousquet-Antonelli C, Deragon JM. A comprehensive analysis of the La-motif protein superfamily. RNA 2009; 15:750-764.
102. Byers SA, Price JP, Cooper JJ et al. HEXIM2:a HEXIM1-related protein, regulates positive transcription elongation factor b through association with 7SK. J Biol Chem 2005; 280:16360-16377.
103. Rosenblad MA, Gorodkin J, Knudsen B et al. SRPDB: signal recognition particle database. Nucleic Acids Res 2003; 31:363-364.
104. Andersen ES, Rosenblad MA, Larsen N et al. The tmRDB and SRPDB resources. Nucleic Acids Res 2006; 34:163-168.
105. Bürk J, Weiche B, Wenk M et al. Depletion of the signal recognition particle receptor inactivates ribosomes in Escherichia coli. J Bacteriol 2009; 191:7017-7026.
106. Rosenblad MA, Samuelsson T. Identification of chloroplast signal recognition particle RNA genes. Plant Cell Physiol 2004; 45:1633-1639.
107. Jaru-Ampornpan P, Nguyen TX, Shan SO. A distinct mechanism to achieve efficient signal recognition particle (SRP)-SRP receptor interaction by the chloroplast srp pathway. Mol Biol Cell 2009; 20:3965-3973.
108. Chen JLWGC. An emerging consensus for telomerase RNA structure. Proc Natl Acad Sci USA 2004; 101:14683-14684.
109. Podlevsky JD, Bley CJ, Omana RV et al. The telomerase database. Nucleic Acids Res 2008; 36:D339-D343.
110. Xie M, Mosig A, Qi X et al. Size variation and structural conservation of vertebrate telomerase RNA. J Biol Chem 2008; 283:2049-2059.
111. Artandi SE, DePinho RA. Telomeres and telomerase in cancer. Carcinogenesis 2010; 31:9-18.
112. Nilsen TW. The spliceosome: the most complex macromolecular machine in the cell? Bioessays 2003; 25:1147-1149.
113. Dávila López M, Rosenblad MA, Samuelsson T. Computational screen for spliceosomal RNA genes aids in defining the phylogenetic distribution of major and minor spliceosomal components. Nucleic Acids Res 2008; 36:3001-10.
114. Hastings KE. SL trans-splicing: easy come or easy go? Trends Genet 2005; 21:240-247.
115. Nilsen TW. Evolutionary origin of SL-addition trans-splicing: still an enigma. Trends Genet 2001; 17:678-680.
116. MacMorris M, Kumar M, Lasda E et al. A novel family of C. elegans snRNPs contains proteins associated with trans-splicing. RNA 2007; 13:511-520.
117. Smith S. The world according to PARP. Trends Biochem Sci 2001; 26:174-179.
118. Stadler PF, Chen JJL, Hackermüller J et al. Evolution of vault RNAs. Mol Biol Evol 2009; 26:1975-1991.
119. Mosig A, Zhu L, Stadler PF. Customized strategies for discovering distant ncRNA homologs. Brief Funct Genomics Proteomics 2009; 8:451-460.
120. van Zon A, Mossink MH, Scheper RJ et al. The vault complex. Cell Mol Life Sci 2003; 60:1828-1837.
121. Zappulla DC, Cech TR. RNA as a flexible scaffold for proteins: yeast telomerase and beyond. Cold Spring Harb Symp Quant Biol 2006; 71:217-224.
122. Jolly C, Lakhotia SC. Human sat III and Drosophila hsrω transcripts: a common paradigm for regulation of nuclear RNA processing in stressed cells. Nucleic Acids Res 2006; 34:5508-5514.
123. Bond A, VanGompel MJW, Sametsky E et al. Balanced gene regulation by an embryonic brain ncRNA is critical for adult hippocampal GABA circuitry. Nature Neuroscience 2009; 12:1020-1027.

124. Sasaki YTF, Ideue T, Sano M et al. MENεβ noncoding RNAs are essential for structural integrity of nuclear paraspeckles. Proc Natl Acad Sci USA 2009; 106:2525-2530.
125. Mourtada-Maarabouni M, Pickard MR, Hedge VL et al. GAS5:a nonprotein-coding RNA, controls apoptosis and is downregulated in breast cancer. Oncogene 2009; 28:195-208.
126. Betel D, Sheridan R, Marks DS et al. Computational analysis of mouse piRNA sequence and biogenesis. PLoS Comput Biol 2007; 3:e222.
127. Bond CS, Fox AH. Paraspeckles: nuclear bodies built on long noncoding RNA. J Cell Biol 2009; 186:637-644.
128. Cole C, Sobala A, Lu C et al. Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. RNA 2009; 15:2147-2160.
129. Langenberger D, Bermudez-Santana C, Stadler PF et al. Identification and classification of small RNAs in transcriptome sequence data. Pac1 Symp1 Biocomput 2010; 15:80-87.
130. Taft RJ, Glazov EA, Lassmann T et al. Small RNAs derived from snoRNAs. RNA 2009; 15:1233-1240.
131. Ender C, Krek A, Friedländer MR et al. A human snoRNA with microRNA-like functions. Mol Cell 2008; 32:519-528.
132. Lee YS, Shibata Y, Malhotra A et al. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). Genes Dev 2009; 23:2639-2649.
133. Persson H, Kvist A, Vallon-Christersson J et al. The noncoding RNA of the multidrug resistance-linked vault particle encodes multiple regulatory small RNAs. Nat Cell Biol 2009; 11:1268-1271.
134. Stadler P. Evolution of the long noncoding RNAs MALAT1 and MENβε. In: Ferreira CE, Satoru M, Stadler P eds. Proceedings of the Brazilian Symposium on Bioinformatics. Lecture Notes in Computer Science. New York:Springer, 2010.
135. Huang FWD, Qin J, Reidys CM et al. Target prediction and a statistical sampling algorithm for RNA-RNA interaction. Bioinformatics 2010; 26:175-181.
136. Ni J, Tien AL, Fournier MJ. Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. Cell 1997; 89:565-573.
137. Yu LE, Koslowsky DJ. Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: Structure probing of a gRNA bound to its cognate mRNA. RNA 2006; 12:1050-1060.