



Review

Functional roles of non-coding Y RNAs



Madzia P. Kowalski¹, Torsten Krude*

Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, United Kingdom

ARTICLE INFO

Article history:

Received 18 May 2015

Received in revised form 3 July 2015

Accepted 4 July 2015

Available online 6 July 2015

Keywords:

Non-coding RNA

Y RNA

DNA replication

RNA stability

RNA domains

ABSTRACT

Non-coding RNAs are involved in a multitude of cellular processes but the biochemical function of many small non-coding RNAs remains unclear. The family of small non-coding Y RNAs is conserved in vertebrates and related RNAs are present in some prokaryotic species. Y RNAs are also homologous to the newly identified family of non-coding stem-bulge RNAs (sbRNAs) in nematodes, for which potential physiological functions are only now emerging. Y RNAs are essential for the initiation of chromosomal DNA replication in vertebrates and, when bound to the Ro60 protein, they are involved in RNA stability and cellular responses to stress in several eukaryotic and prokaryotic species. Additionally, short fragments of Y RNAs have recently been identified as abundant components in the blood and tissues of humans and other mammals, with potential diagnostic value. While the number of functional roles of Y RNAs is growing, it is becoming increasingly clear that the conserved structural domains of Y RNAs are essential for distinct cellular functions. Here, we review the biochemical functions associated with these structural RNA domains, as well as the functional conservation of Y RNAs in different species. The existing biochemical and structural evidence supports a domain model for these small non-coding RNAs that has direct implications for the modular evolution of functional non-coding RNAs.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Contents

1. Introduction and historical overview	20
2. Y RNAs form RoRNPs that are involved in the regulation of RNA stability and cellular stress responses	22
3. Y RNAs are essential factors for the initiation of chromosomal DNA replication	23
4. Y RNA localisation	25
5. Y RNA loop domain-binding proteins	25
6. Y RNA-derived small RNAs	26
7. Conclusions	27
Acknowledgements	27
References	27

1. Introduction and historical overview

Small non-coding Y RNAs were first discovered in 1981 as components of ribonucleoproteins (RNPs) complexed with Ro60 and La proteins, autoantigens which are targets of the immune system in patients suffering from the autoimmune diseases systemic lupus erythematosus (SLE) and Sjögren's syndrome (Hendrick et al., 1981; Lerner et al., 1981). These non-coding RNAs were initially found in the cytoplasm of mammalian cells (human, mouse and monkey) and were therefore given the prefix 'Y', for cytoplasmic Y RNAs, to distinguish them from nuclear U RNAs (Lerner et al., 1981).

There are four non-coding Y RNAs in humans (hY1, hY3, hY4 and hY5 RNA; an hY2 RNA was also originally described, but was later

Abbreviations: sbRNA, stem-bulge RNA; RNP, ribonucleoprotein; PTB, polypyrimidine tract-binding protein; ZBP1, zipcode binding protein 1; Yr1 RNA, Y RNA-like RNA; snRNA, small nuclear RNA; UV, ultraviolet; PNPase, polynucleotide phosphorylase; RYPER, Ro60/Y RNA/PNPase Exoribonuclease RNP; NMR, nuclear magnetic resonance; CD, circular dichroism; RNAi, RNA interference; MO, antisense morpholino oligonucleotide; MBT, mid-blastula transition; ORC, origin recognition complex; RoBPI, RoRNP binding protein I; YsRNAs, Y RNA-derived small RNAs.

* Corresponding author.

E-mail address: tk218@cam.ac.uk (T. Krude).

¹ Present address: Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA 94305, USA.

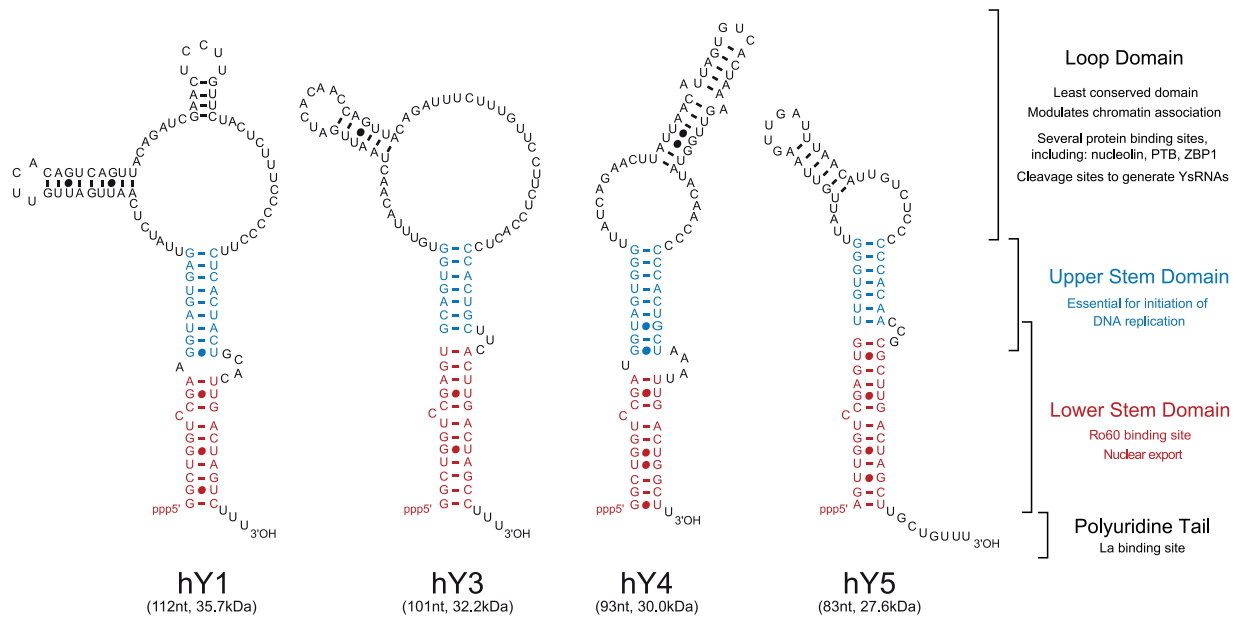


Fig. 1. The non-coding human Y RNAs. The nucleotide sequences and secondary structures of hY RNAs are derived from sequence alignment and enzymatic and chemical probing (Teunissen et al., 2000; van Gelder et al., 1994). The conserved structural RNA domains and their associated functions are highlighted for each hY RNA. The size in nucleotides (nt) and molecular weight (kDa) of each RNA is indicated. See main text for references.

removed from the list as it was found to be a degradation product of hY1 RNA). Y RNAs are present in all vertebrate species investigated so far, with between one and four different genes per species reflecting gene loss and duplication events during vertebrate evolution (Mosig et al., 2007; Perreault et al., 2007). In humans, the four Y RNA genes are clustered together at a single chromosomal locus on chromosome 7q36 (Maraia et al., 1994, 1996). A similar syntenic arrangement of Y RNA genes has been described in other vertebrates (Farris et al., 1996; O'Brien et al., 1993). Individual Y RNA genes are transcribed by RNA polymerase III from distinct promoters (Hendrick et al., 1981; Wolin and Steitz, 1983). Y RNAs are relatively small at 100 ± 20 nucleotides in size, and fold into characteristic stem-loop secondary structures (Fig. 1). Chemical and enzymatic structure probing experiments have revealed that the 5' and 3' RNA ends hybridise to form predominantly double-stranded upper and lower stem domains with an internal loop (Teunissen et al., 2000; van Gelder et al., 1994). The nucleotide sequences of the lower and upper stems are highly conserved, whereas the sequences – and hence predicted structures – of the internal loop vary greatly between individual Y RNAs.

The existence of Y RNAs is not restricted to vertebrates. The family of small non-coding stem-bulge RNAs (sbRNAs) in nematodes has recently been shown to be homologous in structure and function to vertebrate Y RNAs (Boria et al., 2010; Kowalski et al., 2015). Small non-coding RNAs bearing similarities to vertebrate Y RNAs or nematode sbRNAs have also been reported in other eukaryotes, including the insects *Anopheles gambiae* (Perreault et al., 2007) and *Bombyx mori* (Duarte et al., 2015), and the lancet *Branchiostoma floridae* (Mosig et al., 2007). Similar RNAs have also been described in some prokaryotes, including *Deinococcus radiodurans* (Chen et al., 2000), *Salmonella enterica* serovar Typhimurium (Chen et al., 2013), *Mycobacterium smegmatis* and possibly many more (Chen et al., 2014). These bacterial RNAs are highly divergent from, and not homologous with, the four vertebrate Y RNA clades (Perreault et al., 2007).

Since their discovery in 1981, the Y RNA field has grown considerably, concomitant with the rise in number of independent cellular roles associated with Y RNAs (Fig. 2). Biochemical functions of Y RNAs have been mapped to distinct structural domains of the Y

and sbRNAs. Therefore, to reconcile this diverse range of functions, a concept of modular structure and evolution of these RNAs is now becoming apparent (Fig. 3).

Initially, investigations into Y RNA function focussed on the conserved binding sites for Ro60 and La proteins, which are required for Y RNAs to associate with these proteins to form RoRNPs (Chen and Wolin, 2004). RoRNPs are currently implicated in RNA processing and quality control (Hall et al., 2013; Wolin et al., 2012). The precise cellular function of RoRNPs is not yet clear and furthermore they are not essential as deletion mutants of Ro60 are viable. The highly conserved binding sites for Ro60 and La proteins are present in the lower stem and polyuridine tail domains of Y RNAs, respectively (Fig. 1).

More recently Y RNAs were shown to be essential factors for the initiation step of chromosomal DNA replication in human cell nuclei (Christov et al., 2006; Krude et al., 2009). Strikingly, the essential cellular function of Y RNAs in DNA replication is conserved in vertebrates and nematodes, since functional inactivation of Y RNAs in *Xenopus laevis* and *Danio rerio* embryos, or of sbRNAs in *Caenorhabditis elegans* leads to abrogation of DNA replication, cell cycle arrest and embryonic lethality (Collart et al., 2011; Kowalski et al., 2015). In contrast to Ro60 and La binding, this essential function resides in the upper stem domain of these RNAs (Gardiner et al., 2009; Kowalski et al., 2015; Wang et al., 2014).

The loop domain of Y RNAs is diverse in sequence and has been reported to bind several different proteins, including nucleolin, polypyrimidine tract-binding protein (PTB) and zipcode binding protein 1 (ZBP1) (Köhn et al., 2013) (Fig. 1). The roles of the interactions with these proteins are unclear, but it has been suggested that they could modulate the subcellular localisation of Ro60 (Sim and Wolin, 2011), and also confer specialised cellular functions by binding preferentially to specific Y RNAs (Hogg and Collins, 2007; Langley et al., 2010). Consistent with its wide array of binding proteins, the loop domain has been implicated in modulating the association of Y RNAs with subnuclear chromatin domains (Zhang et al., 2011).

Finally, small RNA fragments derived from Y RNAs become enriched in apoptotic cells, possibly as result of apoptotic degradation processes (Rutjes et al., 1999). However, recent

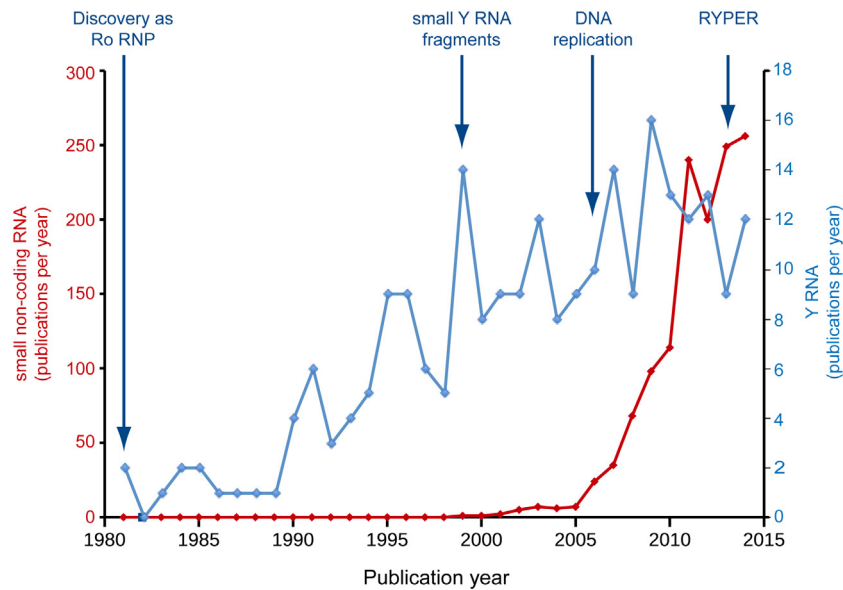


Fig. 2. Approximate number of publications per year relating to small non-coding RNA (left y-axis) and Y RNA (right y-axis), respectively. Data are the number of hits on a Thomson Reuters Web of Science topic search for the phrases 'small non-coding RNA' and 'Y RNA' (including 'YRNA', 'RoRNP', 'Ro RNP' and 'Y5 RNA' with white space characters).

high-throughput RNA sequencing approaches have now established that small Y RNA fragments are also highly abundant in cells, tissues and body fluids of humans and mammals, as well as in a range of tumours (Dhahbi et al., 2014; Meiri et al., 2010; Nicolas et al., 2012; Vojtech et al., 2014). Therefore, these Y RNA-derived fragments are now of clinical interest and have attracted much recent attention as potential biomarkers for disease.

In this review, we discuss the biochemical and cellular functions of Y RNAs, as well as their evolutionary conservation. The reader is referred to earlier reviews that have focused on individual aspects of Y RNA and RoRNP biology (Chen and Wolin, 2004; Hall et al., 2013; Köhn et al., 2013; Krude, 2010; Pruijn et al., 1997; Verhagen and Pruijn, 2011; Wolin et al., 2012, 2013). Here, we present a current integrated view of Y RNA function, focussing on the modular domain structure of Y RNAs, which can mediate the

assembly of distinct Y RNPs under different cellular conditions and contexts.

2. Y RNAs form RoRNPs that are involved in the regulation of RNA stability and cellular stress responses

Y RNAs were first identified as non-coding RNAs bound by the Ro60 protein, a common 60 kDa antigen detected by antibodies from patients with the autoimmune diseases Systemic Lupus Erythematosus and Sjögren's syndrome (Lerner et al., 1981). Ro60 is conserved in vertebrates and homologues have been identified in most metazoa, and also in ~5% of sequenced bacterial genomes, including *D. radiodurans* and *Salmonella* (Sim and Wolin, 2011; Wolin et al., 2013).

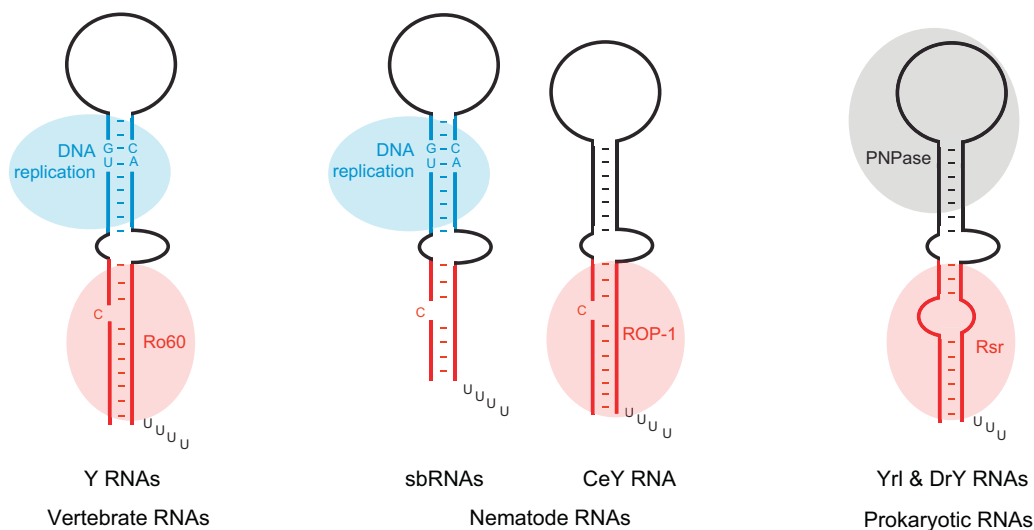


Fig. 3. Modular structure of prokaryotic and eukaryotic Y RNAs. The overall secondary structures of the indicated RNAs are shown schematically and are based on consensus structures. Secondary structures in the divergent loop domains are omitted for clarity. RNA motifs acting as binding sites for Ro60 protein orthologues are shown in red. RNA motifs required for DNA replication are highlighted in blue, and interaction sites for the prokaryotic exonuclease PNPase are highlighted by a grey circle. See main text for references. (For interpretation of the references to color in this legend, the reader is referred to the web version of the article.)

In all organisms studied, orthologues of Ro60 protein bind to Y RNAs, or Y RNA-like (Yrl) non-coding RNAs to form RoRNPs (Chen et al., 2013, 2014) (Fig. 3). In vertebrates, the Ro60 binding site on the lower stem domain of Y RNAs is well characterised and comprises a seven-base-pair helix, a single bulged cytidine and a three-nucleotide bulge on the opposite strand (Pruijn et al., 1991; Wolin and Steitz, 1984) (Fig. 1). In vertebrates, Ro60-binding to the Y RNA lower stem is dependent on both RNA sequence-specific interactions and shape complementarity (Stein et al., 2005). The two bulges in the lower stem of Y RNAs distort its helical structure, making the major groove of the RNA accessible to the amino side chains of Ro60 (Green et al., 1998). Mutations in the Y RNA lower stem that remove either bulge, or change the conserved nucleotide sequence, abolish Ro60 binding (Green et al., 1998; Pruijn et al., 1991). Finally, X-ray crystallography studies show that Ro60 is toroidal in shape, binds Y RNAs on its outer surface and contains a positively charged central channel that can accommodate single-stranded, but not double-stranded RNA (Stein et al., 2005). Immunoprecipitation experiments have shown that in the nematode *C. elegans* the Ro60 orthologue protein, ROP-1, binds only a single major RNA, termed CeY RNA (Van Horn et al., 1995). Although a recent study suggested that the sbrRNA CeN72 interacts with ROP-1 in a gel shift assay *in vitro* (Xiao et al., 2012), neither CeN72 nor any of the other 17 *C. elegans* sbrRNAs were identified in ROP-1 immunoprecipitates from worm extracts (Van Horn et al., 1995). Consistent with these observations, it was suggested that the CeY RNA is an outlier of the sbrRNA family and that it may have undergone a functional specialisation towards RoRNP-related functions in nematodes (Boria et al., 2010). In prokaryotes, the Ro60 orthologue protein Rsr binds to the lower stem of DrY RNA in *D. radiodurans* and Yrl RNAs in *Salmonella* and *M. smegmatis* (Chen et al., 2013, 2014).

La protein is a 50 kDa auto-antigen found complexed with a subset of RoRNPs (Hendrick et al., 1981). La is required for accurate and efficient termination of RNA polymerase III transcription, and binds to the 3' polyuridine tail of newly synthesised RNAs in the nucleus (Stefano, 1984). While most mature RNA transcripts lose their polyuridine tail, Y RNAs retain theirs (Fig. 1), and so can maintain association with La. La is implicated in the nuclear retention of Y RNAs and protecting RNAs from exonucleolytic cleavage (Wolin and Cedervall, 2002).

RoRNPs are currently implicated in non-coding RNA quality control, RNA stability and in cellular responses to stress in several organisms (Sim and Wolin, 2011). Ro60 binds aberrant non-coding RNAs such as misfolded 5S rRNA or U2 snRNA in a range of species, including *X. laevis*, *C. elegans* and *Mus musculus* (Chen et al., 2003; Labbe et al., 1999a,b; O'Brien and Wolin, 1994). *C. elegans* lacking ROP-1 are viable, but these deletion strains show defects in dauer larvae formation, an alternative developmental stage induced by starvation or stress that allows them to survive unfavourable environmental conditions (Labbe et al., 1999b, 2000). Furthermore, mouse cells and *D. radiodurans* upregulate and accumulate RoRNPs in response to ultraviolet (UV) irradiation and both cell types lacking Ro60 have decreased survival following UV irradiation (Chen et al., 2000, 2003; Xue et al., 2003). In *D. radiodurans*, the Ro60 orthologue protein Rsr also has a role in heat-stress-induced rRNA maturation and starvation-induced rRNA decay (Chen et al., 2007; Wurtmann and Wolin, 2010). Ro60 binding to misfolded non-coding RNAs is not largely sequence specific, indicating that Ro60 could potentially bind a wide range of RNAs (Fuchs et al., 2006). It has therefore been proposed that RoRNPs function as cellular stress sensors, which scavenge and process aberrant non-coding RNAs that fail to associate with their cognate RNA-binding proteins (Fuchs et al., 2006; Hogg and Collins, 2007).

Genetic deletion studies have established that Ro60 protein and its orthologues ROP-1 and Rsr in nematodes and prokaryotes,

respectively, are not essential for cell proliferation, or the viability and development of the unperturbed organism (Chen et al., 2000, 2003; Labbe et al., 1999b; Xue et al., 2003). However, these deletions of Ro60 and its orthologues resulted in significant reductions in the levels of soluble eukaryotic or prokaryotic Y RNAs. It has therefore been concluded from these studies that Ro60 proteins play a functional role in the stability of their associated Y RNAs.

The precise role of Y RNAs in vertebrate RoRNPs has proved controversial. Structural and biochemical studies have shown that misfolded RNAs insert through the Ro60 cavity and also bind to the Ro60 outer surface at a region that partially overlaps with the Y RNA-binding domain (Fuchs et al., 2006; Stein et al., 2005). Since Y RNAs bind Ro60 in a sequence-specific manner and with higher affinity than misfolded RNAs, it has been suggested that a bound Y RNA could sterically block misfolded RNA binding to Ro60 (Fuchs et al., 2006; Stein et al., 2005). However, all four human Y RNAs bind to Ro60 and La, but only hY5 RNA co-purified with a common target of RNA quality control, 5S rRNA, via ribosomal protein L5 (Hogg and Collins, 2007). This work suggested that hY5-RoRNPs interact with target 5S-L5 RNPs, so that specific Y RNAs might themselves modulate the recruitment of misfolded or variant non-coding RNAs to RoRNPs. A unified model has been proposed in which Y RNAs can both positively and negatively regulate the target specificity of non-coding RNA quality control mediated by Ro60 (Hogg and Collins, 2007).

Interestingly, recent studies in *D. radiodurans* have demonstrated that prokaryotic Y RNAs regulate both access of the Ro60 orthologue protein Rsr to RNA substrates and also recruit exonucleases involved in their maturation or degradation (Chen et al., 2007, 2013). In *D. radiodurans*, the prokaryotic DrY RNA tethers Rsr to the exoribonuclease polynucleotide phosphorylase (PNPase), forming RYPER (Ro60/Y RNA/PNPase exoribonuclease RNP), an RNA degradation complex that cleaves structured RNAs (Chen et al., 2013; Wolin et al., 2013). In this specialised RoRNP, the Y RNA acts as a scaffold linking Rsr with PNPase. It also serves as a gate mediating the entry of single-stranded RNA substrates into the PNPase cavity, thereby modulating the substrate specificity of the enzyme and increasing the effectiveness of RYPER (Chen et al., 2013). It remains to be seen whether metazoan RoRNPs with their Y RNAs are also involved in nucleolytic degradation of target RNAs.

Taken together, in these past three-and-a-half decades of research on the RoRNP, a substantial body of structural and functional data has accumulated that supports a functional role for Y RNAs in RNA stability and quality control. This allocation of Y RNA function comes from a sequence- and structure-specific association of pro- and eukaryotic Y RNAs with members of the Ro protein family. In all cases, this interaction, and thus Y RNA involvement in a functional role of the resulting RNP, occurs *via* the evolutionarily conserved lower stem of the Y RNAs (Fig. 3).

3. Y RNAs are essential factors for the initiation of chromosomal DNA replication

The first direct and essential cellular function that has been experimentally demonstrated for Y RNAs is their involvement in the initiation of chromosomal DNA replication (Christov et al., 2006, 2008; Collart et al., 2011; Gardiner et al., 2009; Krude et al., 2009). In an unbiased approach, Y RNAs were purified by biochemical fractionation of a human cell extract as an activity that is essential for the reconstitution of chromosomal DNA replication in a cell-free system (Christov et al., 2006). This *in vitro* system uses nuclei that are prepared from late G1 phase human cells. Semi-conservative DNA replication initiates and subsequently elongates in these nuclei upon the addition of cytosolic extract from proliferating human cells (Krude, 2000; Krude et al., 1997). During its

step-wise purification, one particular cytosolic replication factor maintained DNA replication activity over several biochemical steps that enriched for poly-anionic factors. Surprisingly, this approach led to the purification of non-coding Y RNAs as the relevant factor and not of a protein, which was expected at the time (Christov et al., 2006). In support of these gain-of-function experiments, specific degradation of Y RNAs from unfractionated cytosolic cell extract abrogates the initiation step of DNA replication (Christov et al., 2006; Gardiner et al., 2009; Krude et al., 2009). DNA replication can then be restored by addition of non-targeted individual human or vertebrate Y RNAs, but not of other small non-coding RNAs such as 5S ribosomal RNA or U2 RNA. Therefore, vertebrate Y RNAs are required specifically for DNA replication, and they function redundantly with each other in this system. Single molecule analysis of Y RNA depletion and reconstitution experiments provided detailed and direct evidence that Y RNAs are required for the initiation step of DNA replication, leading to the establishment of new DNA replication forks on human chromosomal DNA (Krude et al., 2009). In contrast, Y RNAs are not required for the elongation of existing DNA replication forks, and are thus not involved in the DNA copying mechanism as such (Krude et al., 2009). From these experiments, it has also become clear that Y RNAs do not fulfil this important function in isolation, but require interaction with other DNA replication proteins that are also present in the extract (Christov et al., 2006).

Systematic mutagenesis of vertebrate Y RNAs identified that the upper stem domain is necessary and also sufficient for Y RNA function in the initiation of DNA replication (Gardiner et al., 2009). The upper stem domain is present in all vertebrate Y RNAs and can thus explain the functional redundancy of vertebrate Y RNAs. This domain contains a highly conserved central GUG–CAC nucleotide sequence motif (Gardiner et al., 2009) (Fig. 1). A recent structure analysis of the upper stem by nuclear magnetic resonance (NMR) and far-UV circular dichroism (CD) spectroscopy provided evidence that this domain adopts a locally destabilised A-form helix under physiological conditions in solution (Wang et al., 2014). The helix is stabilised by two flanking G–C base pairs, but the central section around the highly conserved G–C base pair (*i.e.* the upper one of the GUG–CAC motif) is unstable and the accessible bases may thus be involved in specific interactions of this domain with as yet unknown proteins (Wang et al., 2014). Mutations in this sequence motif abrogate the initiation activity of the Y RNA, concomitant with structural perturbation of the upper stem domain (Gardiner et al., 2009; Wang et al., 2014). Conversely, an insertion of the upper stem domain into a similarly folded, but inactive, backbone of a synthetic RNA results in the full activation of this previously inert RNA as a DNA replication initiation factor (Gardiner et al., 2009). Furthermore, the lower stem and loop domains of Y RNAs are dispensable as they can be entirely removed from the RNA without loss of DNA replication initiation function (Gardiner et al., 2009).

Importantly, neither Ro60 or La proteins, nor their binding sites in vertebrate Y RNAs, are required for the initiation of DNA replication. Immunoprecipitation experiments indicate that ~50% of Y RNAs in human cell extracts are present outside Ro60 and La RNPs (Langley et al., 2010). Immunodepletion of Ro60 and La RNPs from human cytosolic extracts does not inhibit DNA replication in human cell nuclei (Langley et al., 2010). Furthermore, addition of recombinant purified Ro60 or La proteins has no effect on DNA replication *in vitro* (Langley et al., 2010). Deletion of Ro60 and La binding sites on the lower stem domain of vertebrate Y RNAs does not inhibit the DNA replication initiation activity of the mutant Y RNAs (Christov et al., 2006; Gardiner et al., 2009). These findings indicate that Y RNAs mediate the initiation of DNA replication independently of RoRNPs. Consistent with these *in vitro* findings, genetic knockout of Ro60 in various organisms has no effect on DNA replication or viability (Chen et al., 2000; Labbe et al., 1999b; Xue et al., 2003).

Vertebrate Y RNAs are also essential for initiation of DNA replication *in vivo*. Disruption of Y RNAs by RNAi in proliferating vertebrate cells in culture blocks DNA replication and cell proliferation (Christov et al., 2006, 2008; Collart et al., 2011). Interestingly, transfection of a synthetic small double-stranded RNA derived from the upper stem of hY1 RNA overcomes this inhibition (Gardiner et al., 2009), demonstrating that its replication function is indeed responsible for the *in vivo* phenotype. Furthermore, functional inactivation of Y RNAs by microinjection of antisense morpholino oligonucleotides (MOs) into *D. rerio* or *X. laevis* embryos leads to DNA replication inhibition, arrested development and early embryonic death (Collart et al., 2011), which occurs right after the mid-blastula transition (MBT) (Langley et al., 2014). Consistent with a functional role in DNA replication and cell proliferation, Y RNAs are over-expressed in human solid tumours, when compared with the corresponding healthy tissues (Christov et al., 2008). Taken together, these observations establish that Y RNA function is built on a modular structure of the overall full-length RNA. The essential function for chromosomal DNA replication in vertebrates can be ascribed to the short upper stem domain of the vertebrate Y RNAs.

Recently, a family of related small non-coding RNAs, termed stem-bulge RNAs (sbrRNAs), was identified in nematode worms (Aftab et al., 2008; Boria et al., 2010; Deng et al., 2006), and an sbrRNA was also described for the silkworm, *B. mori* (Duarte et al., 2015). The genome of *C. elegans* contains at least 18 sbrRNA genes, in addition to the related aforementioned CeY RNA gene, each with a putative RNA polymerase III promoter, with many sbrRNA genes also present in other nematode species (Boria et al., 2010). A computational analysis based on nucleotide sequence and structural motifs suggested that sbrRNAs might be homologues of vertebrate Y RNAs, the previously described CeY RNA being an outlier of this group because of a lower sequence conservation compared to the other sbrRNAs (Boria et al., 2010). Y RNAs and sbrRNAs share an overall stem-loop structure containing double-stranded upper and lower stem domains, as well as a single-stranded internal loop (Boria et al., 2010). The upper stem domain of sbrRNAs, like vertebrate Y RNAs, contains a highly conserved A/GUG–CAC/U motif (Boria et al., 2010; Kowalski et al., 2015) (Fig. 3). We have recently shown that sbrRNAs from several nematode species can functionally substitute for vertebrate Y RNAs and support the initiation of chromosomal DNA replication *in vitro*, whereas CeY RNA does not (Kowalski et al., 2015). Importantly, the initiation activity of full-length sbrRNAs was dependent on the upper stem domain and intriguingly, also on a conserved UUAUC motif in the loop domain, which is also present in vertebrate Y RNAs (Kowalski et al., 2015). Furthermore, functional inhibition of sbrRNAs in *C. elegans* resulted in DNA replication defects and lethality during early embryogenesis (Kowalski et al., 2015). Collectively, these findings indicate that sbrRNAs are functional homologues of vertebrate Y RNAs. To date, candidate Y RNAs or sbrRNAs have not been identified in plants or fungi and it therefore remains to be seen to what extent the regulation of DNA replication by small stem-loop RNAs has been conserved during eukaryotic evolution.

Y RNAs have also been reported in prokaryotes and some other isolated eukaryotic species (Chen et al., 2000, 2013, 2014; Mosig et al., 2007; Perreault et al., 2007; Van Horn et al., 1995). The Y RNAs from *C. elegans* (CeY RNA), *B. floridae* (BFY RNA) and *D. radiodurans* (DrY RNA) do not have sequence similarity to vertebrate Y RNAs in the upper stem domain and they are unable to substitute for vertebrate Y RNAs in DNA replication assays *in vitro* (Gardiner et al., 2009). These results indicate that these non-vertebrate Y RNAs do not fulfil the role of vertebrate Y RNAs in DNA replication. Furthermore, *D. radiodurans* or *C. elegans* with a deletion in their respective genes coding for DrY and CeY RNA are viable, so these RNAs are not essential for DNA replication and viability of the organism (Boria et al., 2010; Chen et al., 2007). Thus, in *C. elegans* a large family of

sbRNAs is found, at least some of which function in DNA replication and do not appear to bind Ro60, whilst the divergent CeY RNA binds to Ro60 and does not function in DNA replication (Fig. 3).

The mechanism of Y RNA function in the initiation step of DNA replication in vertebrates is not yet clear, although several key features are emerging. Y RNAs interact biochemically with several DNA replication initiation proteins, including the origin recognition complex ORC, and initiation proteins Cdc6, Cdt1 and DUE-B (Collart et al., 2011; Zhang et al., 2011). In contrast, hY RNAs do not interact biochemically with DNA replication fork proteins including the DNA helicase subunits MCM2–7, GINS complex, primase, or DNA polymerases (Zhang et al., 2011). These biochemical interactions would therefore suggest a functional interaction between Y RNAs and the protein machinery of the DNA replication initiation complex. Using fluorescently-labelled hY RNAs, it was shown that hY RNAs also associate dynamically with unreplicated chromatin in G1 phase nuclei *in vitro*, where they co-localise with several DNA replication proteins on chromatin before the initiation of DNA replication, including ORC, Cdt1, MCM2 and Cdc45 (Zhang et al., 2011). In *X. laevis*, Y RNA binding to chromatin occurs only after the MBT and is ORC-dependent (Collart et al., 2011). Once DNA replication initiates in a Y RNA-dependent manner, Y RNAs are locally displaced from these initiation sites, and they are consequently absent from the sites of ongoing DNA synthesis in these nuclei (Zhang et al., 2011). It has therefore been suggested that Y RNAs could function in a ‘catch-and-release’ mechanism on chromatin in human cells (Zhang et al., 2011), which is consistent with the original ‘licensing factor’ model of Blow and Laskey (Blow et al., 1987; Laskey et al., 1981). Future experiments are needed to test this hypothesis and resolve the underlying molecular mechanism of Y RNA function during the initiation of chromosomal DNA replication. Questions to address will include whether or not the upper stem domain of Y RNAs hybridises with other nucleic acids in order to execute its essential function; which are the functionally essential interacting DNA replication proteins and whether Y RNAs activate these proteins or inactivate any potential repressors.

4. Y RNA localisation

In eukaryotes the biogenesis of Y RNAs begins in the nucleus, as RNA polymerase III transcription is terminated. As with other RNA polymerase III transcripts such as tRNAs or pre-miRNAs, Y RNAs can be exported to the cytoplasm, or like U snRNAs, they can remain in the nucleus after transcription.

There are conflicting reports on the relative distribution of Y RNAs in the nucleus and cytoplasm of vertebrate cells, likely in part due to different methodologies used (Hall et al., 2013; Pruijn et al., 1997; Zhang et al., 2011). Early enucleation and cell fractionation experiments revealed that Y RNAs were predominantly, or even exclusively, cytoplasmic in cultured mammalian cells and *X. laevis* oocytes (O’Brien et al., 1993; Peek et al., 1993; Simons et al., 1994). One study reported more recently that in human and mouse cells, h/mY1, h/mY3 and hY4 RNAs are found in the cytoplasm, whereas hY5 RNA localises to the nucleus (Gendron et al., 2001). *In situ* hybridisation and ultrastructural analysis by electron microscopy, however, showed that Y RNAs are present at discrete sites in both the nucleus and cytoplasm of cultured human cells (Farris et al., 1997; Matera et al., 1995). Furthermore, in proliferating human cells hY1, hY3 and hY5 RNAs also localise to the edge of nucleoli (the perinucleolar compartment) and co-localise with PTB at these sites in the cell nucleus (Matera et al., 1995). Using fluorescently-labelled hY RNAs it was shown that all four hY RNAs bind chromatin in G1 phase nuclei dynamically from a soluble pool as the nuclei enter S phase *in vitro* (Zhang et al., 2011). While hY1, hY3 and hY4 co-localise with each other and associate mostly with

early-replicating euchromatin, hY5 is enriched in nucleoli (Zhang et al., 2011). The loop domain of hY RNAs modulates this differential association with chromatin as mutant hY RNAs lacking this domain bind to chromatin indiscriminately (Zhang et al., 2011). It remains to be seen if this dynamic chromatin association of Y RNAs is important for their function in the initiation of DNA replication, and in how far it is regulated during the cell cycle. In any case, there is now accumulating evidence that Y RNAs are present in both the nuclear and cytoplasmic fractions of eukaryotic cells, and that their relative abundance in these cell compartments most likely reflects the methodologies used for study and/or the physiological state of the cells.

The export pathways used by Y RNAs are also becoming increasingly well understood. Y RNA export is dependent on the small GTPase Ran, indicating that exportins likely serve as transport receptor proteins for Y RNAs (Rutjes et al., 2001). The lower stem of Y RNAs resembles a dsRNA mini-helix present in other exportin-5 substrates and exportin-5 was shown to associate in a complex with hY1 RNA and RanGTP (Gwizdek et al., 2001, 2003). Furthermore, deletion of the lower stem of hY1 RNA results in defective nuclear export of the mutant RNA in *Xenopus* oocytes (Rutjes et al., 2001). Therefore, it seems likely that Y RNAs are exported from the nucleus in an exportin-5-dependent manner. Direct mechanistic evidence of a re-import pathway of Y RNAs back into the nucleus has not been identified to date.

Ro60 binding protects Y RNAs from exonucleolytic degradation and is required for stable accumulation of Y RNAs in a range of species (Chen and Wolin, 2004). It has been proposed that one function of mammalian Y RNAs is to regulate the subcellular localisation of Ro60 (Sim and Wolin, 2011). Ro60 is present in both the nucleus and cytoplasm of cells and it has been shown that Y RNAs can influence this distribution (Sim and Wolin, 2011). In mouse cells, Y RNA binding to Ro60 occludes a nuclear localisation signal on the Ro60 surface, thereby retaining RoRNPs in the cytoplasm (Sim et al., 2009). Another study has demonstrated that Ro60 binding is a prerequisite for efficient nuclear export of Y RNAs in *X. laevis* oocytes (Simons et al., 1996).

The intracellular localisation of Y RNAs changes under conditions of cellular stress (Chen and Wolin, 2004). Both Ro60 and Y RNAs accumulate in the nucleus after UV irradiation or oxidative stress in several species (Chen et al., 2000, 2003; Sim et al., 2009, 2012). This is consistent with a role for nuclear Ro60-Y RNA complexes in cellular stress responses. However, it could also arise from stress-induced inhibition of the RanGTP gradient, resulting in defective nuclear export and thus nuclear accumulation (Köhn et al., 2013).

Mammalian Y RNAs can also be selectively packaged into viruses. This has been demonstrated for the human immunodeficiency virus type 1 (HIV-1) and Moloney murine leukaemia virus (Garcia et al., 2009; Wang et al., 2007). This process does not require Ro60 binding and likely occurs during early stages of Y RNA biogenesis when nascent Y RNAs are present in the nucleus (Wang et al., 2007). It is currently unknown whether Y RNAs are involved in retroviral function.

Unlike vertebrate Y RNAs, the intracellular localisations of nematode sbRNAs and CeY RNA have not yet been investigated. It therefore remains to be seen whether the functional homology between vertebrate Y RNAs and nematode sbRNAs is also reflected in a similar distribution of these nematode sbRNAs in the nuclear and cytoplasmic fractions of the cell.

5. Y RNA loop domain-binding proteins

Recently, several novel Y RNA-binding proteins have been identified in eukaryotic cells. It was shown using gel filtration that Y

RNPs present in human cell extracts range in size from 150 to 550 kDa, indicating that Y RNAs can associate with multiple proteins simultaneously (Fabini et al., 2000). However, apart from the canonical Y RNA-binding proteins Ro60 and La, the interactions between other cellular proteins and Y RNAs are currently poorly characterised.

Several proteins interact with Y RNAs through the loop domain, which is the least conserved domain in Y RNAs; it is heterogeneous in structure and varies in nucleotide sequence between individual Y RNAs (Fig. 1) (Farris et al., 1999; Teunissen et al., 2000). The loop domains of hY1, hY3 and hY5 RNAs are pyrimidine-rich, and in hY1 and hY3 RNAs, contain stretches of poly-pyrimidine sequences. The loop domains of Y RNAs can interact with a different set of proteins to form distinct Y RNPs (Bouffard et al., 2000; Fabini et al., 2001; Hogg and Collins, 2007). The loop domain may therefore specialise individual Y RNAs for specific cellular functions (Hogg and Collins, 2007).

In human cells, several Y RNA-loop-binding proteins have been identified, including nucleolin, PTB/hnRNP I, hnRNP K and ZBP1 (Köhn et al., 2013). Nucleolin binds pyrimidine-rich stretches in the loop domain of Y RNAs in human cells, and while it preferentially associates with hY1 and hY3, it is present in stable cytosolic RNPs with all four hY RNAs (Fabini et al., 2001; Langley et al., 2010). Nucleolin is involved in many metabolic processes, including rRNA processing, ribosome biogenesis and nucleo-cytoplasmic transport (Ginisty et al., 1999). Cytosolic nucleolin RNPs are distinct from Ro60 and La RNPs and immunodepletion of nucleolin RNPs from cytosolic extracts does not inhibit DNA replication initiation in human cell nuclei (Langley et al., 2010). Therefore the interaction between Y RNAs and soluble nucleolin is not required for Y RNA function in this process (Langley et al., 2010).

PTB and hnRNP K bind preferentially to poly-pyrimidine tracts in the loop domains of hY1 and hY3 RNAs, like nucleolin, but their association with hY4 and hY5 RNAs has not been detected (Fabini et al., 2001; Fouraux et al., 2002). Efficient binding of PTB and hnRNP K to Y RNAs also likely requires La, because deletion of the La binding site significantly decreases binding of both proteins to the RNA (Fabini et al., 2001). PTB and hnRNP K are both involved in several aspects of RNA processing and are thought to shuttle between the nucleus and the cytoplasm (Krecic and Swanson, 1999). PTB, hnRNP K and La function as RNA chaperones *in vitro* and mediate RNA folding without a requirement for ATP consumption (Belisova et al., 2005). Since Y RNA binding to these proteins inhibits their RNA chaperone activity, it has been proposed that Y RNAs mediate the transport of hnRNP K, PTB and La to specific targets, before releasing the proteins to execute their function (Belisova et al., 2005). ZBP1 interacts with mouse Y1 and Y3 RNAs *via* their loop domains (Köhn et al., 2013; Sim et al., 2012). Depletion of ZBP1 results in accumulation of Y3 RNA in mouse cell nuclei, indicating that ZBP1 is involved in nuclear export of Y RNPs (Sim et al., 2012).

All four human Y RNAs associate with the antiviral cytidine deaminase APOBEC3G, which is also a component of Ro60 and La RNPs (Chiu et al., 2006; Gallois-Montbrun et al., 2008). Although the function of these Y RNPs is unknown, it has been speculated that APOBEC3G could be involved in RNA editing of Y RNAs to modulate their function (Chiu et al., 2006; Köhn et al., 2013).

There are also several proteins that bind preferentially or exclusively to Y5 RNAs, indicating a specialised role for this Y RNA. RoRNP binding protein I (RoBPI) mainly associates with hY5 RNA in human cells (Bouffard et al., 2000) but also binds hY1 and hY3 RNAs (Hogg and Collins, 2007). RoBPI is a DNA- and RNA-binding protein involved in several nuclear processes, such as transcription and RNA splicing (Page-McCaw et al., 1999). In addition, Interferon-induced protein with tetratricopeptide repeats 5 (IFIT5) only interacts with hY5 RNA (Hogg and Collins, 2007). Ribosomal protein L5 also interacts specifically with hY5 RNA (Hogg

and Collins, 2007). The L5 protein forms complexes with 5S rRNA (Steitz et al., 1988) and Y5 RNA also associates with 5S rRNA, with a strong preference for a misfolded variant (Hogg and Collins, 2007). Together with the observed enrichment of hY5 in nucleoli (Zhang et al., 2011), these results indicate that hY5 RNA could be involved in rRNAs biogenesis (Hogg and Collins, 2007).

The identification of these non-canonical Y RNA-binding proteins, many of which show preferential binding to specific Y RNAs, indicates that Y RNAs are likely incorporated into multiple, distinct RNPs to carry out specialised functions (Bouffard et al., 2000; Fabini et al., 2001; Hogg and Collins, 2007; Langley et al., 2010). It is also possible that Y RNAs sequester multiple cellular proteins until they are needed, for example in response to stress (Köhn et al., 2013). In addition, the localisation of Y RNAs is also influenced by these interacting proteins (Köhn et al., 2013; Zhang et al., 2011). In conclusion, the loop domains of eukaryotic Y RNAs attract many divergent binding proteins and further work is required to understand mechanistically any potential functional roles for the resulting different Y RNP complexes.

6. Y RNA-derived small RNAs

Recently, a plethora of deep sequencing studies in eukaryotes have identified small RNA fragments derived from longer RNAs (Rother and Meister, 2011; Tuck and Tollervey, 2011). Accumulating evidence suggests that these RNA fragments, while derived from pre-existing small non-coding RNAs, are themselves physiologically relevant in both healthy and diseased cells (Dhahbi, 2014; Hall and Dalmay, 2013).

High levels of Y RNA-derived small RNAs (YsRNAs) of 22–36 nucleotides are produced in apoptotic cells (Rutjes et al., 1999). Immunoprecipitation experiments have revealed that these YsRNAs are bound to Ro60 and La proteins, suggesting that the binding sites of these proteins in the lower stem domain of Y RNAs are protected from the nucleolytic degradation process (Rutjes et al., 1999). It remains to be seen whether the upper stem domain of Y RNAs, which is required for Y RNA functionality in DNA replication and cell proliferation, is actively targeted during apoptosis.

YsRNAs are also detected in proliferating cells, both cancerous and non-cancerous, at levels similar to that of known miRNAs (Nicolas et al., 2012). YsRNAs are also found in the brain, retina and other healthy mammalian tissues, as well as in a range of tumours (Chen and Heard, 2013; Meiri et al., 2010; Verhagen and Puijn, 2011; Yamazaki et al., 2014). Some of these YsRNAs were initially mis-annotated as a novel type of miRNA generated by the processing of full-length Y RNAs (Meiri et al., 2010; Verhagen and Puijn, 2011). However, it has subsequently been shown that YsRNA biogenesis is independent of the canonical miRNA biogenesis pathway. YsRNAs do not associate with Argonaute proteins (Chen and Heard, 2013; Nicolas et al., 2012), and the generation of YsRNAs appears to be independent of Dicer (Langenberger et al., 2013). Furthermore, in contrast to miRNAs, YsRNAs do not have gene silencing activity in the luciferase reporter assay (Meiri et al., 2010). The role of these intracellular YsRNAs is currently unknown.

YsRNAs, along with tRNA fragments, have been recently identified as highly abundant small RNAs circulating in the blood of humans and other mammals (Dhahbi, 2014). Circulating YsRNAs of 25–33 nucleotides are present in human blood in multiple forms, including within vesicles and as cell-free RNP complexes of 100–300 kDa (Dhahbi, 2014; Dhahbi et al., 2013, 2014). These YsRNAs are derived from the 5' and 3' termini of full-length Y RNAs by cleavage within the internal loop domain (Dhahbi et al., 2013, 2014). The levels of these 5' and 3'-Y RNA and 5'-tRNA-derived fragments were found to be significantly different in a group of breast cancer patients compared to healthy individuals,

suggesting that these fragments may have some diagnostic value as cancer biomarkers (Dhahbi et al., 2014). Interestingly, many of the Y RNAs from which the fragments are derived have previously been annotated as Y RNA pseudogenes, and therefore not thought to be expressed or processed (Dhahbi et al., 2013, 2014). Since YsRNAs have no known functions, the significance of the pseudogene expression is unclear.

YsRNAs and full-length Y RNAs have also been detected in vesicles released by mouse immune cells (Nolte-t Hoen et al., 2012) and YsRNAs comprise a substantial fraction of the RNA component of exosome vesicles present in human semen (Vojtech et al., 2014). The function of these extracellular YsRNAs is currently unknown and whilst it has been speculated that YsRNAs are specifically processed and secreted as part of an as yet undefined signalling process (Dhahbi, 2014), these RNA fragments could alternatively just be passive, stable degradation products of highly abundant cellular Y RNAs. However, it is an intriguing possibility that small RNAs could mediate intercellular physiological signals (Chen et al., 2012; Hoy and Buck, 2012; Sarkies and Miska, 2014; Valadi et al., 2007) and therefore the clinical potential of YsRNAs as diagnostic biomarkers or blood-delivered therapy targets remains an open and exciting possibility.

7. Conclusions

Y RNAs are small non-coding RNAs involved in a range of cellular processes, including DNA replication, RNA stability and cellular stress responses. The modular domains of Y RNAs mediate their distinct cellular roles (Figs. 1–3). The upper stem domain of vertebrate Y RNAs and homologous nematode sbRNAs is essential for the initiation of chromosomal DNA replication. The lower stem is required for Ro60 binding and hence is involved in stress responses, RNA surveillance, stability control and RoRNP-mediated RNA degradation across several eukaryotic and prokaryotic species. Furthermore, the evolutionary conservation of the structure and nucleotide sequence of the upper and lower stem domains is tightly correlated to the conservation of Y RNA functions across species, indicative of high selective pressure on the Y RNA domains. The loop domain, although part of the conserved overall secondary structure of Y RNAs, is highly varied in its nucleotide sequence. This might facilitate functional sub-specialisation of different Y RNA molecules within the same species. This RNA module-based strategy might provide a way of regulating and separating important cellular functions of the small non-coding Y RNAs. With the recent emergence of an expansive landscape of pervasive transcription and non-coding RNAs in mammalian cells (Clark et al., 2011; Iyer et al., 2015), this principle also has direct implications for non-coding RNA

evolution. As seen with Y and sbRNAs across eukaryotes and prokaryotes, each modular RNA domain can evolve separately and thus pave the way for functional divergence and specialisation of non-coding RNAs (Fig. 3). Since the discovery of Y RNAs in 1981, much progress has been made in elucidating their binding proteins, evolutionarily conservation and important cellular roles. However, many key questions remain, which provide opportunity for exciting further growth in the field of non-coding Y RNAs (Box 1).

Acknowledgements

We thank Paulo Amaral, Howard Baylis and Alexander Langley for critical reading of the manuscript. Research in the authors' laboratory has been funded by grants from Cancer Research UK (CRUK), the Association for International Cancer Research (AICR) and the Biotechnology and Biological Sciences Research Council (BBSRC). The funders had no involvement in the design and preparation of this article.

References

- Aftab, M.N., He, H., Skogerbo, G., Chen, R., 2008. Microarray analysis of ncRNA expression patterns in *Caenorhabditis elegans* after RNAi against snoRNA associated proteins. *BMC Genomics* 9, 278.
- Belisova, A., Semrad, K., Mayer, O., Kocian, G., Waigmann, E., Schroeder, R., Steiner, G., 2005. RNA chaperone activity of protein components of human Ro RNPs. *RNA* 11, 1084–1094.
- Blow, J.J., Dilworth, S.M., Dingwall, C., Mills, A.D., Laskey, R.A., 1987. Chromosome replication in cell-free systems from *Xenopus* eggs. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 317, 483–494.
- Boria, I., Gruber, A.R., Tanzer, A., Bernhart, S.H., Lorenz, R., Mueller, M.M., Hofacker, I.L., Stadler, P.F., 2010. Nematode sbRNAs: homologs of vertebrate Y RNAs. *J. Mol. Evol.* 70, 346–358.
- Bouffard, P., Barbar, E., Briere, F., Boire, G., 2000. Interaction cloning and characterization of RoBPI, a novel protein binding to human Ro ribonucleoproteins. *RNA* 6, 66–78.
- Chen, C.J., Heard, E., 2013. Small RNAs derived from structural non-coding RNAs. *Methods* 63, 76–84.
- Chen, X., Liang, H., Zhang, J., Zen, K., Zhang, C.Y., 2012. Horizontal transfer of microRNAs: molecular mechanisms and clinical applications. *Protein Cell* 3, 28–37.
- Chen, X., Quinn, A.M., Wolin, S.L., 2000. Ro ribonucleoproteins contribute to the resistance of *Deinococcus radiodurans* to ultraviolet irradiation. *Genes Dev.* 14, 777–782.
- Chen, X., Sim, S., Wurtmann, E.J., Feke, A., Wolin, S.L., 2014. Bacterial noncoding Y RNAs are widespread and mimic tRNAs. *RNA* 20, 1715–1724.
- Chen, X., Smith, J.D., Shi, H., Yang, D.D., Flavell, R.A., Wolin, S.L., 2003. The Ro autoantigen binds misfolded U2 small nuclear RNAs and assists mammalian cell survival after UV irradiation. *Curr. Biol.* 13, 2206–2211.
- Chen, X., Taylor, D.W., Fowler, C.C., Galan, J.E., Wang, H.W., Wolin, S.L., 2013. An RNA degradation machine sculpted by Ro autoantigen and noncoding RNA. *Cell* 153, 166–177.
- Chen, X., Wolin, S.L., 2004. The Ro 60 kDa autoantigen: insights into cellular function and role in autoimmunity. *J. Mol. Med. (Berl.)* 82, 232–239.
- Chen, X., Wurtmann, E.J., Van Batavia, J., Zybailov, B., Washburn, M.P., Wolin, S.L., 2007. An ortholog of the Ro autoantigen functions in 23S rRNA maturation in *D. radiodurans*. *Genes Dev.* 21, 1328–1339.
- Chiu, Y.L., Witkowska, H.E., Hall, S.C., Santiago, M., Soros, V.B., Esnault, C., Heidmann, T., Greene, W.C., 2006. High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition. *Proc. Natl. Acad. Sci. U. S. A.* 103, 15588–15593.
- Christov, C.P., Gardiner, T.J., Szüts, D., Krude, T., 2006. Functional requirement of noncoding Y RNAs for human chromosomal DNA replication. *Mol. Cell. Biol.* 26, 6993–7004.
- Christov, C.P., Trivier, E., Krude, T., 2008. Noncoding human Y RNAs are overexpressed in tumours and required for cell proliferation. *Br. J. Cancer* 98, 981–988.
- Clark, M.B., Amaral, P.P., Schlesinger, F.J., Dinger, M.E., Taft, R.J., Rinn, J.L., Ponting, C.P., Stadler, P.F., Morris, K.V., Morillon, A., Rozowsky, J.S., Gerstein, M.B., Wahlestedt, C., Hayashizaki, Y., Carninci, P., Gingeras, T.R., Mattick, J.S., 2011. The reality of pervasive transcription. *PLoS Biol.* 9, e1000625, discussion e1001102.
- Collart, C., Christov, C.P., Smith, J.C., Krude, T., 2011. The midblastula transition defines the onset of Y RNA-dependent DNA replication in *Xenopus laevis*. *Mol. Cell. Biol.* 31, 3857–3870.
- Deng, W., Zhu, X., Skogerbo, G., Zhao, Y., Fu, Z., Wang, Y., He, H., Cai, L., Sun, H., Liu, C., Li, B., Bai, B., Wang, J., Jia, D., Sun, S., He, H., Cui, Y., Wang, Y., Bu, D., Chen, R., 2006. Organization of the *Caenorhabditis elegans* small non-coding transcriptome: genomic features, biogenesis, and expression. *Genome Res.* 16, 20–29.

Box 1

Outstanding questions

- How do Y RNAs regulate the initiation of DNA replication in vertebrates?
- Is the requirement for small stem-loop RNAs in the initiation of chromosomal DNA replication conserved beyond vertebrates and nematodes?
- Are Y RNAs in RoRNPs involved in RNA degradation in vertebrates?
- What are the functions of non-canonical Y RNA-binding proteins and their associated RNPs?
- Are small RNA fragments derived from mammalian Y RNAs of functional significance and could they be used therapeutically?
- As a general outlook, why have so many different biological functions been compressed into such a small RNA molecule?

- Dhabhi, J.M., 2014. Circulating small noncoding RNAs as biomarkers of aging. *Ageing Res. Rev.* 17, 86–98.
- Dhabhi, J.M., Spindler, S.R., Atamna, H., Boffelli, D., Martin, D.I., 2014. Deep sequencing of serum small RNAs identifies patterns of 5' tRNA half and YRNA fragment expression associated with breast cancer. *Biomark. Cancer* 6, 37–47.
- Dhabhi, J.M., Spindler, S.R., Atamna, H., Boffelli, D., Mote, P., Martin, D.I., 2013. 5'-YRNA fragments derived by processing of transcripts from specific YRNA genes and pseudogenes are abundant in human serum and plasma. *Physiol. Genomics* 45, 990–998.
- Duarte Jr., F.F., de Lima Neto, Q.A., Rando Fdos, S., de Freitas, D.V., Pattaro Junior, J.R., Polizelli, L.G., Munhoz, R.E., Seixas, F.A., Fernandez, M.A., 2015. Identification and molecular structure analysis of a new noncoding RNA, a sbRNA homolog, in the silkworm *Bombyx mori* genome. *Mol. Biosyst.* 11, 801–808.
- Fabini, G., Rajmakers, R., Hayer, S., Fouraux, M.A., Pruijn, G.J., Steiner, G., 2001. The heterogeneous nuclear ribonucleoproteins I and K interact with a subset of the Ro ribonucleoprotein-associated Y RNAs in vitro and in vivo. *J. Biol. Chem.* 276, 20711–20718.
- Fabini, G., Rutjes, S.A., Zimmermann, C., Pruijn, G.J., Steiner, G., 2000. Analysis of the molecular composition of Ro ribonucleoprotein complexes. Identification of novel Y RNA-binding proteins. *Eur. J. Biochem.* 267, 2778–2789.
- Farris, A.D., Gross, J.K., Hanas, J.S., Harley, J.B., 1996. Genes for murine Y1 and Y3 Ro RNAs have class 3 RNA polymerase III promoter structures and are unlinked on mouse chromosome 6. *Gene* 174, 35–42.
- Farris, A.D., Koelsch, G., Pruijn, G.J., van Venrooij, W.J., Harley, J.B., 1999. Conserved features of Y RNAs revealed by automated phylogenetic secondary structure analysis. *Nucleic Acids Res.* 27, 1070–1078.
- Farris, A.D., Puvion-Dutilleul, F., Puvion, E., Harley, J.B., Lee, L.A., 1997. The ultrastructural localization of 60-kDa Ro protein and human cytoplasmic RNAs: association with novel electron-dense bodies. *Proc. Natl. Acad. Sci. U. S. A.* 94, 3040–3045.
- Fouraux, M.A., Bouvet, P., Verkaar, S., van Venrooij, W.J., Pruijn, G.J., 2002. Nucleolin associates with a subset of the human Ro ribonucleoprotein complexes. *J. Mol. Biol.* 320, 475–488.
- Fuchs, G., Stein, A.J., Fu, C., Reinisch, K.M., Wolin, S.L., 2006. Structural and biochemical basis for misfolded RNA recognition by the Ro autoantigen. *Nat. Struct. Mol. Biol.* 13, 1002–1009.
- Gallois-Montbrun, S., Holmes, R.K., Swanson, C.M., Fernandez-Ocana, M., Byers, H.L., Ward, M.A., Malim, M.H., 2008. Comparison of cellular ribonucleoprotein complexes associated with the APOBEC3F and APOBEC3G antiviral proteins. *J. Virol.* 82, 5636–5642.
- Garcia, E.L., Onafuwa-Nuga, A., Sim, S., King, S.R., Wolin, S.L., Telesnitsky, A., 2009. Packaging of host my RNAs by murine leukemia virus may occur early in Y RNA biogenesis. *J. Virol.* 83, 12526–12534.
- Gardiner, T.J., Christov, C.P., Langley, A.R., Krude, T., 2009. A conserved motif of vertebrate Y RNAs essential for chromosomal DNA replication. *RNA* 15, 1375–1385.
- Gendron, M., Roberge, D., Boire, G., 2001. Heterogeneity of human Ro ribonucleoproteins (RNPs): nuclear retention of Ro RNPs containing the human hY5 RNA in human and mouse cells. *Clin. Exp. Immunol.* 125, 162–168.
- Ginisty, H., Sicard, H., Roger, B., Bouvet, P., 1999. Structure and functions of nucleolin. *J. Cell Sci.* 112 (Pt 6), 761–772.
- Green, C.D., Long, K.S., Shi, H., Wolin, S.L., 1998. Binding of the 60-kDa Ro autoantigen to Y RNAs: evidence for recognition in the major groove of a conserved helix. *RNA* 4, 750–765.
- Gwizdek, C., Bertrand, E., Dargemont, C., Lefebvre, J.C., Blanchard, J.M., Singer, R.H., Doglio, A., 2001. Terminal minihelix, a novel RNA motif that directs polymerase III transcripts to the cell cytoplasm. Terminal minihelix and RNA export. *J. Biol. Chem.* 276, 25910–25918.
- Gwizdek, C., Ossareh-Nazari, B., Brownawell, A.M., Doglio, A., Bertrand, E., Macara, I.G., Dargemont, C., 2003. Exportin-5 mediates nuclear export of minihelix-containing RNAs. *J. Biol. Chem.* 278, 5505–5508.
- Hall, A.E., Dalmay, T., 2013. Discovery of novel small RNAs in the quest to unravel genome complexity. *Biochem. Soc. Trans.* 41, 866–870.
- Hall, A.E., Turnbull, C., Dalmay, T., 2013. Y RNAs: recent developments. *Biomol. Concepts* 4, 103–110.
- Hendrick, J.P., Wolin, S.L., Rinke, J., Lerner, M.R., Steitz, J.A., 1981. Ro small cytoplasmic ribonucleoproteins are a subclass of La ribonucleoproteins: further characterization of the Ro and La small ribonucleoproteins from uninfected mammalian cells. *Mol. Cell. Biol.* 1, 1138–1149.
- Hogg, J.R., Collins, K., 2007. Human Y5 RNA specializes a Ro ribonucleoprotein for 5S ribosomal RNA quality control. *Genes Dev.* 21, 3067–3072.
- Hoy, A.M., Buck, A.H., 2012. Extracellular small RNAs: what, where, why? *Biochem. Soc. Trans.* 40, 886–890.
- Iyer, M.K., Niknafs, Y.S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., Barrette, T.R., Prensner, J.R., Evans, J.R., Zhao, S., Poliakov, A., Cao, X., Dhanasekaran, S.M., Wu, Y.M., Robinson, D.R., Beer, D.G., Feng, F.Y., Iyer, H.K., Chinnaiyan, A.M., 2015. The landscape of long noncoding RNAs in the human transcriptome. *Nat. Genet.* 47, 199–208.
- Köhn, M., Pazaitis, N., Huttelmaier, S., 2013. Why YRNAs? About versatile RNAs and their functions. *Biomolecules* 3, 143–156.
- Kowalski, M.P., Baylis, H.A., Krude, T., 2015. Non-coding stem-bulge RNAs are required for cell proliferation and embryonic development in *C. elegans*. *J. Cell Sci.* 128, 2118–2129.
- Krecic, A.M., Swanson, M.S., 1999. hnRNP complexes: composition, structure, and function. *Curr. Opin. Cell Biol.* 11, 363–371.
- Krude, T., 2000. Initiation of human DNA replication in vitro using nuclei from cells arrested at an initiation-competent state. *J. Biol. Chem.* 275, 13699–13707.
- Krude, T., 2010. Non-coding RNAs: new players in the field of eukaryotic DNA replication. *Subcell Biochem.* 50, 105–118.
- Krude, T., Christov, C.P., Hyrien, O., Marheineke, K., 2009. Y RNA functions at the initiation step of mammalian chromosomal DNA replication. *J. Cell Sci.* 122, 2836–2845.
- Krude, T., Jackman, M., Pines, J., Laskey, R.A., 1997. Cyclin/Cdk-dependent initiation of DNA replication in a human cell-free system. *Cell* 88, 109–119.
- Labbe, J.C., Burgess, J., Rokeach, L.A., Hekimi, S., 2000. ROP-1, an RNA quality-control pathway component, affects *Caenorhabditis elegans* dauer formation. *Proc. Natl. Acad. Sci. U. S. A.* 97, 13233–13238.
- Labbe, J.C., Hekimi, S., Rokeach, L.A., 1999a. Assessing the function of the Ro ribonucleoprotein complex using *Caenorhabditis elegans* as a biological tool. *Biochem. Cell Biol.* 77, 349–354.
- Labbe, J.C., Hekimi, S., Rokeach, L.A., 1999b. The levels of the RoRNP-associated Y RNA are dependent upon the presence of ROP-1, the *Caenorhabditis elegans* Ro60 protein. *Genetics* 151, 143–150.
- Langenberger, D., Cakir, M.V., Hoffmann, S., Stadler, P.F., 2013. Dicer-processed small RNAs: rules and exceptions. *J. Exp. Zool. B: Mol. Dev. Evol.* 320, 35–46.
- Langley, A.R., Chambers, H., Christov, C.P., Krude, T., 2010. Ribonucleoprotein particles containing non-coding Y RNAs, Ro60, La and nucleolin are not required for Y RNA function in DNA replication. *PLoS ONE* 5, e13673.
- Langley, A.R., Smith, J.C., Stemple, D.L., Harvey, S.A., 2014. New insights into the maternal to zygotic transition. *Development* 141, 3834–3841.
- Laskey, R.A., Harland, R.M., Earnshaw, W.C., Dingwall, C., 1981. Chromatin assembly and the co-ordination of DNA replication in the eukaryotic chromosome. In: Schweiger, H.G. (Ed.), *International Cell Biology 1980–1981*. Springer Verlag, Berlin, pp. 162–167.
- Lerner, M.R., Boyle, J.A., Hardin, J.A., Steitz, J.A., 1981. Two novel classes of small ribonucleoproteins detected by antibodies associated with lupus erythematosus. *Science* 211, 400–402.
- Maraia, R., Sakulich, A.L., Brinkmann, E., Green, E.D., 1996. Gene encoding human Ro-associated autoantigen Y5 RNA. *Nucleic Acids Res.* 24, 3552–3559.
- Maraia, R.J., Sasaki-Tozawa, N., Driscoll, C.T., Green, E.D., Darlington, G.J., 1994. The human Y4 small cytoplasmic RNA gene is controlled by upstream elements and resides on chromosome 7 with all other hY scRNA genes. *Nucleic Acids Res.* 22, 3045–3052.
- Matera, A.G., Frey, M.R., Margelot, K., Wolin, S.L., 1995. A perinucleolar compartment contains several RNA polymerase III transcripts as well as the polypyrimidine tract-binding protein, hnRNP I. *J. Cell Biol.* 129, 1181–1193.
- Meiri, E., Levy, A., Benjamin, H., Ben-David, M., Cohen, L., Dov, A., Dromi, N., Elyakim, E., Yerushalmi, N., Zion, O., Lithwick-Yanai, G., Sitbon, E., 2010. Discovery of microRNAs and other small RNAs in solid tumors. *Nucleic Acids Res.* 38, 6234–6246.
- Mosig, A., Guofeng, M., Stadler, B.M.R., Stadler, P.F., 2007. Evolution of the vertebrate Y RNA cluster. *Theory Biosci.* 129, 9–14.
- Nicolas, F.E., Hall, A.E., Csorba, T., Turnbull, C., Dalmay, T., 2012. Biogenesis of Y RNA-derived small RNAs is independent of the microRNA pathway. *FEBS Lett.* 586, 1226–1230.
- Noite-t Hoen, E.N., Buermans, H.P., Waasdorp, M., Stoorvogel, W., Wauben, M.H., 't Hoen, P.A., 2012. Deep sequencing of RNA from immune cell-derived vesicles uncovers the selective incorporation of small non-coding RNA biotypes with potential regulatory functions. *Nucleic Acids Res.* 40, 9272–9285.
- O'Brien, C.A., Margelot, K., Wolin, S.L., 1993. Xenopus Ro ribonucleoproteins: members of an evolutionarily conserved class of cytoplasmic ribonucleoproteins. *Proc. Natl. Acad. Sci. U. S. A.* 90, 7250–7254.
- O'Brien, C.A., Wolin, S.L., 1994. A possible role for the 60-kD Ro autoantigen in a discard pathway for defective 5S rRNA precursors. *Genes Dev.* 8, 2891–2903.
- Page-McCaw, P.S., Amonlirdviman, K., Sharp, P.A., 1999. PUF60: a novel U2AF65-related splicing activity. *RNA* 5, 1548–1560.
- Peek, R., Pruijn, G.J., van der Kemp, A.J., van Venrooij, W.J., 1993. Subcellular distribution of Ro ribonucleoprotein complexes and their constituents. *J. Cell Sci.* 106 (Pt 3), 929–935.
- Perreault, J., Perreault, J.P., Boire, G., 2007. The Ro associated Y RNAs in metazoans: evolution and diversification. *Mol. Biol. Evol.* 24, 1678–1689.
- Pruijn, G.J., Simons, F.H., van Venrooij, W.J., 1997. Intracellular localization and nucleocytoplasmic transport of Ro RNP components. *Eur. J. Cell Biol.* 74, 123–132.
- Pruijn, G.J., Slobbe, R.L., van Venrooij, W.J., 1991. Analysis of protein-RNA interactions within Ro ribonucleoprotein complexes. *Nucleic Acids Res.* 19, 5173–5180.
- Rother, S., Meister, G., 2011. Small RNAs derived from longer non-coding RNAs. *Biochimie* 93, 1905–1915.
- Rutjes, S.A., Lund, E., van der Heijden, A., Grimm, C., van Venrooij, W.J., Pruijn, G.J., 2001. Identification of a novel cis-acting RNA element involved in nuclear export of hY RNAs. *RNA* 7, 741–752.
- Rutjes, S.A., van der Heijden, A., Utz, P.J., van Venrooij, W.J., Pruijn, G.J., 1999. Rapid nucleolytic degradation of the small cytoplasmic Y RNAs during apoptosis. *J. Biol. Chem.* 274, 24799–24807.
- Sarkies, P., Miska, E.A., 2014. Small RNAs break out: the molecular cell biology of mobile small RNAs. *Nat Rev. Mol. Cell Biol.* 15, 525–535.
- Sim, S., Weinberg, D.E., Fuchs, G., Choi, K., Chung, J., Wolin, S.L., 2009. The subcellular distribution of an RNA quality control protein, the Ro autoantigen, is regulated by noncoding Y RNA binding. *Mol. Biol. Cell* 20, 1555–1564.

- Sim, S., Wolin, S.L., 2011. Emerging roles for the Ro 60-kDa autoantigen in noncoding RNA metabolism. *Wiley Interdiscip. Rev. RNA* 2, 686–699.
- Sim, S., Yao, J., Weinberg, D.E., Niessen, S., Yates 3rd, J.R., Wolin, S.L., 2012. The zipcode-binding protein ZBP1 influences the subcellular location of the Ro 60-kDa autoantigen and the noncoding Y3 RNA. *RNA* 18, 100–110.
- Simons, F.H., Pruijn, G.J., van Venrooij, W.J., 1994. Analysis of the intracellular localization and assembly of Ro ribonucleoprotein particles by microinjection into *Xenopus laevis* oocytes. *J. Cell Biol.* 125, 981–988.
- Simons, F.H., Rutjes, S.A., van Venrooij, W.J., Pruijn, G.J., 1996. The interactions with Ro60 and La differentially affect nuclear export of hY1 RNA. *RNA* 2, 264–273.
- Stefano, J.E., 1984. Purified lupus antigen La recognizes an oligouridylylate stretch common to the 3' termini of RNA polymerase III transcripts. *Cell* 36, 145–154.
- Stein, A.J., Fuchs, G., Fu, C., Wolin, S.L., Reinisch, K.M., 2005. Structural insights into RNA quality control: the Ro autoantigen binds misfolded RNAs via its central cavity. *Cell* 121, 529–539.
- Steitz, J.A., Berg, C., Hendrick, J.P., La Branche-Chabot, H., Metspalu, A., Rinke, J., Yario, T., 1988. A 5S rRNA/L5 complex is a precursor to ribosome assembly in mammalian cells. *J. Cell Biol.* 106, 545–556.
- Teunissen, S.W., Kruithof, M.J., Farris, A.D., Harley, J.B., van Venrooij, W.J., Pruijn, G.J., 2000. Conserved features of Y RNAs: a comparison of experimentally derived secondary structures. *Nucleic Acids Res.* 28, 610–619.
- Tuck, A.C., Tollervey, D., 2011. RNA in pieces. *Trends Genet.* 27, 422–432.
- Valadi, H., Ekstrom, K., Bossios, A., Sjostrand, M., Lee, J.J., Lotvall, J.O., 2007. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* 9, 654–659.
- van Gelder, C.W., Thijssen, J.P., Klaassen, E.C., Sturchler, C., Krol, A., van Venrooij, W.J., Pruijn, G.J., 1994. Common structural features of the Ro RNP associated hY1 and hY5 RNAs. *Nucleic Acids Res.* 22, 2498–2506.
- Van Horn, D.J., Eisenberg, D., O'Brien, C.A., Wolin, S.L., 1995. *Caenorhabditis elegans* embryos contain only one major species of Ro RNP. *RNA* 1, 293–303.
- Verhagen, A.P., Pruijn, G.J., 2011. Are the Ro RNP-associated Y RNAs concealing microRNAs? Y RNA-derived miRNAs may be involved in autoimmunity. *Bioessays* 33, 674–682.
- Vojtech, L., Woo, S., Hughes, S., Levy, C., Ballweber, L., Sauteraud, R.P., Strobl, J., Westerberg, K., Gottardo, R., Tewari, M., Hladik, F., 2014. Exosomes in human semen carry a distinctive repertoire of small non-coding RNAs with potential regulatory functions. *Nucleic Acids Res.* 42, 7290–7304.
- Wang, I., Kowalski, M.P., Langley, A.R., Rodriguez, R., Balasubramanian, S., Hsu, S.T., Krude, T., 2014. Nucleotide contributions to the structural integrity and DNA replication initiation activity of noncoding y RNA. *Biochemistry* 53, 5848–5863.
- Wang, T., Tian, C., Zhang, W., Luo, K., Sarkis, P.T., Yu, L., Liu, B., Yu, Y., Yu, X.F., 2007. 7SL RNA mediates virion packaging of the antiviral cytidine deaminase APOBEC3G. *J. Virol.* 81, 13112–13124.
- Wolin, S.L., Belair, C., Boccitto, M., Chen, X., Sim, S., Taylor, D.W., Wang, H.W., 2013. Non-coding Y RNAs as tethers and gates: insights from bacteria. *RNA Biol.* 10, 1602–1608.
- Wolin, S.L., Cedervall, T., 2002. The La protein. *Annu. Rev. Biochem.* 71, 375–403.
- Wolin, S.L., Sim, S., Chen, X., 2012. Nuclear noncoding RNA surveillance: is the end in sight? *Trends Genet.* 28, 306–313.
- Wolin, S.L., Steitz, J.A., 1983. Genes for two small cytoplasmic Ro RNAs are adjacent and appear to be single-copy in the human genome. *Cell* 32, 735–744.
- Wolin, S.L., Steitz, J.A., 1984. The Ro small cytoplasmic ribonucleoproteins: identification of the antigenic protein and its binding site on the Ro RNAs. *Proc. Natl. Acad. Sci. U. S. A.* 81, 1996–2000.
- Wurtmann, E.J., Wolin, S.L., 2010. A role for a bacterial ortholog of the Ro autoantigen in starvation-induced rRNA degradation. *Proc. Natl. Acad. Sci. U. S. A.* 107, 4022–4027.
- Xiao, T., Wang, Y., Luo, H., Liu, L., Wei, G., Chen, X., Sun, Y., Chen, X., Skogerbo, G., Chen, R., 2012. A differential sequencing-based analysis of the *C. elegans* noncoding transcriptome. *RNA* 18, 626–639.
- Xue, D., Shi, H., Smith, J.D., Chen, X., Noe, D.A., Cedervall, T., Yang, D.D., Eynon, E., Brash, D.E., Kashgarian, M., Flavell, R.A., Wolin, S.L., 2003. A lupus-like syndrome develops in mice lacking the Ro 60-kDa protein, a major lupus autoantigen. *Proc. Natl. Acad. Sci. U. S. A.* 100, 7503–7508.
- Yamazaki, F., Kim, H.H., Lau, P., Hwang, C.K., Iuvone, P.M., Klein, D., Clokie, S.J., 2014. pY RNA1-s2: a highly retina-enriched small RNA that selectively binds to Matrin 3 (Matr3). *PLoS ONE* 9, e88217.
- Zhang, A.T., Langley, A.R., Christov, C.P., Kheir, E., Shafee, T., Gardiner, T.J., Krude, T., 2011. Dynamic interaction of Y RNAs with chromatin and initiation proteins during human DNA replication. *J. Cell Sci.* 124, 2058–2069.