Treasure hunt in an amoeba: Non-coding RNAs in Dictyostelium discoideum

ARTICLE *in* CURRENT GENETICS · APRIL 2007

Impact Factor: 2.68 · DOI: 10.1007/s00294-006-0112-z · Source: PubMed

2 AUTHORS:

Andrea Hinas
Uppsala University
9 PUBLICATIONS **182** CITATIONS
SEE PROFILE

Fredrik Söderbom
Uppsala University
33 PUBLICATIONS **793** CITATIONS
SEE PROFILE

Curr Genet (2007) 51:141–159
DOI 10.1007/s00294-006-0112-z

REVIEW ARTICLE

Treasure hunt in an amoeba: non-coding RNAs in *Dictyostelium discoideum*

Andrea Hinas · Fredrik Söderbom

Received: 16 October 2006 / Revised: 22 November 2006 / Accepted: 23 November 2006 / Published online: 15 December 2006

Abstract The traditional view of RNA being merely an intermediate in the transfer of genetic information, as mRNA, spliceosomal RNA, tRNA, and rRNA, has become outdated. The recent discovery of numerous regulatory RNAs with a plethora of functions in biological processes has truly revolutionized our understanding of gene regulation. Tiny RNAs such as microRNAs and small interfering RNAs play vital roles at different levels of gene control. Small nucleolar RNAs are much more abundant than previously recognized, and new functions beyond processing and modification of rRNA have recently emerged. Longer non-coding RNAs (ncRNAs) can also have important regulatory roles in the cell, e.g., antisense RNAs that control their target mRNAs. The majority of these important findings arose from analyses in various model organisms. In this review, we focus on ncRNAs in the social amoeba *Dictyostelium discoideum.* This important genetically tractable model organism has recently received renewed attention in terms of discovery, regulation and functional studies of ncRNAs. Old and recent findings are discussed and put in context of what we today know about ncRNAs in other organisms.

Communicated by S. Hohmann.

A. Hinas · F. Söderbom (✉)
Department of Molecular Biology,
Biomedical Center, Swedish University of Agricultural Sciences, Box 590, 75124 Uppsala, Sweden
e-mail: fredde@xray.bmc.uu.se

Keywords *Dictyostelium discoideum* · Protist · Non-coding RNA · RNAi · Small RNA · Antisense RNA

Introduction

In a review article from 1981, Gary W. Zieve made an inventory of the stable and abundant small RNAs that had been discovered in mammalian cells—at that time about ten RNA species [not counting transfer RNA (tRNA) and ribosomal RNA (rRNA)]—which he divided into two groups (Zieve 1981). In his concluding remarks he stated, "The elucidation of the role of these two groups of small RNAs will require the efforts of both cellular and molecular biologists and should be a fruitful area of research over the next decade". Subsequently, most of these RNAs were defined as spliceosomal RNAs. Today, we know that small RNAs were not only an interesting finding that fascinated biologists in the 1980s, but will also continue to attract researchers from many fields for many years to come.

Non-coding RNAs (ncRNAs) have recently emerged as a broad and very important group of molecules with crucial functions in a wide variety of cellular processes. The epithet "non-coding" implies that these RNAs do not encode proteins, but are functional as RNAs per se. Classical and well-known examples of ncRNAs are those involved in different levels of protein synthesis, such as rRNAs, tRNAs, and spliceosomal RNAs—also referred to as small nuclear RNAs (snRNAs). However, recent experimental and bioinformatics-based searches for new ncRNAs in a number of model organisms, from *Escherichia coli* to human, have uncovered an unexpectedly great number of ncR-

NAs with a plethora of functions (Gottesman 2005; Griffiths-Jones et al. 2005; Hüttenhofer et al. 2005; Storz et al. 2005; Zamore and Haley 2005; Hüttenhofer and Schattner 2006). Examples of biological processes regulated by ncRNAs are apoptosis, stress responses, development, chromosome maintenance, and virulence, just to mention a few (reviewed in Gottesman 2005; Matzke and Birchler 2005; Esquela-Kerscher and Slack 2006; Romby et al. 2006). In order to control these functions, ncRNAs act on, e.g., chromatin remodeling, transcription, RNA stability and biogenesis, and protein transport (Blum et al. 1990; Bachellerie et al. 2002; Halic and Beckmann 2005; Matzke and Birchler 2005; Stuart et al. 2005; Goodrich and Kugel 2006; Söderbom 2006; Valencia-Sanchez et al. 2006). Furthermore, misregulation of certain ncRNAs has been implicated in diseases, e.g., cancer, Prader-Willi syndrome, and cartilage-hair hypoplasia (Cavaillé et al. 2000; Ridanpää et al. 2001; Vitali et al. 2003; Kishore and Stamm 2006; Esquela-Kerscher and Slack 2006).

The great majority of the identified ncRNAs is between 20 and 500 nucleotides long and therefore termed small ncRNAs. Commonly, these RNA molecules do not work alone; instead, they form ribonucleoprotein (RNP) complexes in which each class of ncRNA is associated with a specific set of proteins. The RNA moiety often confers *specificity* by base-pairing to complementary nucleotide sequences within the target RNA, e.g., mRNAs and rRNAs. Thus, the associated proteins are guided by the ncRNAs to the specific target RNAs where they exert the *function* of the RNP, such as cleavage or chemical modification (Hüttenhofer and Schattner 2006).

Important classes of ncRNAs where hundreds of new members have recently been discovered are small nucleolar RNAs (snoRNAs), microRNAs (miRNAs) and small interfering RNAs (siRNAs). SnoRNPs are involved in modification of rRNAs, snRNAs and tRNAs, while si/miRNPs target mRNAs for degradation and/or translational inhibition (Bachellerie et al. 2002; Zamore and Haley 2005; Söderbom 2006). Although most of the identified ncRNAs with known functions are small, much longer ncRNAs, up to several kilobases long, are also expressed in eukaryotic cells. One well-studied example is the mammalian Xist RNA, which controls X-chromosome inactivation and dosage compensation (Plath et al. 2002; Heard 2005). Other long ncRNAs act as antisense RNAs to regulate genes; these RNAs are often fully complementary to their target (sense) RNAs, which can be either mRNA or other ncRNAs, and are commonly transcribed in *cis* from the opposite DNA strand. Although genome-scale microarray analysis and mining of expressed sequence tag (EST) libraries indicate the presence of great numbers of antisense transcripts in eukaryotes, so far only a few have known biological roles and/or effects on sense RNAs (Werner and Berdal 2005). Examples are antisense regulation of (1) the prespore gene *psvA* in *Dictyostelium discoideum*, (2) the gene for the ncRNA Xist in mammals, and (3) fibroblast growth factor-2 (FGF-2) in vertebrates (Kimelman and Kirschner 1989; Hildebrandt and Nellen 1992; Asa et al. 2001; Plath et al. 2002; Heard 2005). In the latter case, however, the antisense RNA carries an open reading frame thus making its role more complex.

Although the functions of many of the identified ncRNAs are now known, the last few years' large-scale isolation of new small RNAs have revealed many RNAs that still await functional assignment (Hüttenhofer et al. 2005; Mattick and Makunin 2006). This is illustrated by the newly identified classes of RNAs of unknown function in *Caenorhabditis elegans* and *D. discoideum* (Aspegren et al. 2004; Deng et al. 2006).

Besides tRNA and rRNA, knowledge concerning ncRNAs in *D. discoideum* has lagged behind compared to that in many other model organisms. For example, it was only very recently that the major spliceosomal RNAs were identified (Hinas et al. 2006). However, the discovery and understanding of small RNAs in *D. discoideum* have taken a leap forward during the last few years due to extensive computational and experimental searches. These findings have also established *D. discoideum* as a relevant model organism for studies of ncRNAs during growth and development.

The social amoeba *D. discoideum* has for many years proven useful as a model organism to analyze different aspects of cellular biology. More recently, this intriguing organism has also been developed as a host model for a number of pathogens (Steinert and Heuner 2005). *D. discoideum* has the high experimental power of other genetically tractable model organisms such as yeast, but with the added advantage of sharing many features with multicellular organisms (Kessin 2001; Eichinger and Rivero 2006). In nature, *D. discoideum* lives in the soil as single cell amoebae, feeding on bacteria. When challenged by starvation, the cells start secreting cAMP, which acts as a chemoattractant for other *D. discoideum* cells nearby. Up to 100,000 cells aggregate, followed by a series of well-defined developmental stages during which they differentiate into two main cell types, prestalk and prespore cells. The final structure is a stalk topped with a ball of spores (Kessin 2001). The full developmental cycle is completed within 24 h in the laboratory (Fig. 1a). The fashion in which *D. discoideum* forms a multicellular-like organism, separating growth and differentiation, differs substantially

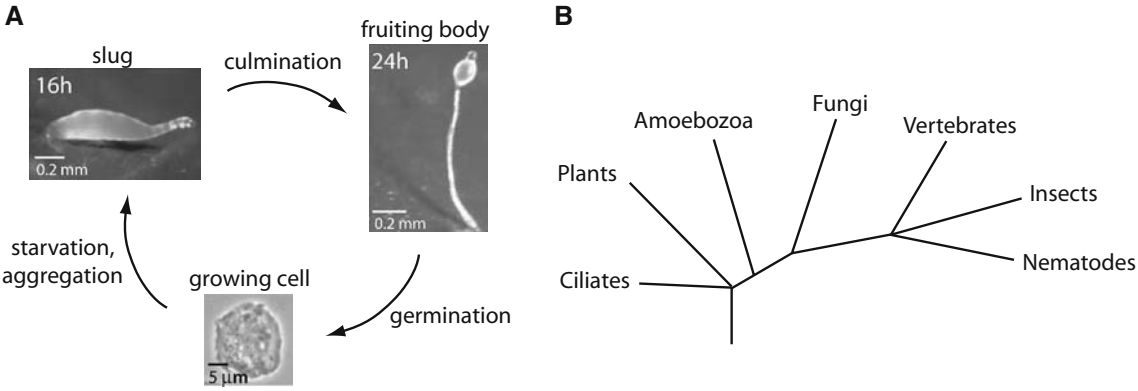

Fig. 1 Life cycle and evolutionary position of *D. discoideum*. **a** Starvation triggers aggregation of single cells into multicellular structures. Time points (16 and 24 h) indicate hours post-starvation under laboratory conditions. At 16 h of development, *D. discoideum* can either form a finger-like structure or a motile slug. The wt strain AX4 was used to illustrate the different developmental stages. **b** Schematic evolutionary tree based on protein sequence comparisons depicting the evolutionary position of *D. discoideum* (Amoebozoa) (Bapteste et al. 2002; Eichinger et al. 2005). Other phylogenetic groups are included based on relevance for this review

from how multicellularity is achieved in animals and plants. Nevertheless, certain processes such as signal transduction are, in many respects, very similar between animals and *D. discoideum* (Williams et al. 2005; Bagorda et al. 2006). Furthermore, the 34 Mbp haploid genome sequence was recently completed and revealed a greater number of protein-encoding genes shared with animals than with plants and fungi (Eichinger et al. 2005). This positions *D. discoideum* in a very interesting place from an evolutionary point of view; based on comparisons of large numbers of protein sequences, the lineage leading to this organism is believed to have branched out before animals and fungi but after plants (Fig. 1b, Bapteste et al. 2002; Eichinger et al. 2005). Another important virtue of *D. discoideum* as a model organism, its genetic tractability is demonstrated by the many molecular and biochemical tools available to mine the functions of its genes. Not only can genes easily be disrupted by homologous recombination, and new genes identified by insertional mutagenesis, but also RNA interference (RNAi), an invaluable tool for gene knock-down in animals and plants, can be used to investigate gene function in *D. discoideum* (Martens et al. 2002). Information about the *D. discoideum* genome, gene annotations, literature and much more can be found at http://www.dictybase.org (Chisholm et al. 2006).

In this review, we aim to give a comprehensive overview of ncRNAs in *D. discoideum*, excluding rRNA and tRNAs (for information about tRNA and rRNA, see Eichinger et al. 2005 and references therein). The identification and functional studies of ncRNAs will be described from a historical perspective, starting with the first handful of small RNAs identified in the late 1970s and continuing with today's large-scale isolation of numerous RNA molecules. The *D. discoideum* ncRNAs will also be discussed in the context of what is known about these RNAs from other organisms. Furthermore, the methods used to identify new ncRNAs, both experimental and computational, will be described.

Non-coding RNAs in *D. discoideum*: where it all started

The discovery of ncRNAs in the 1970s, termed small nuclear RNAs, present in immunoprecipitations using serum from patients with the autoimmune disease systemic lupus erythematosus (SLE) was a major breakthrough for RNA research (Lerner and Steitz 1979, 1981). Until then, ncRNAs had been considered to be more or less restricted to tRNAs and rRNAs but now the door to a whole new world of RNA had been opened. The idea that at least some of the identified snRNAs were involved in pre-mRNA splicing (Lerner et al. 1980; Rogers and Wall 1980) lead to extensive efforts to find homologous RNAs in other organisms, *D. discoideum* being no exception. The first indication of such RNAs in *D. discoideum* came in 1979, when Takeishi and Kaneda (1979) demonstrated the presence of nuclear RNAs in the expected size range by gel electrophoresis. In the next few years, these RNAs were further characterized and a few additional nuclear RNAs were discovered (Wise and Weiner 1980, 1981; Takeishi and Kaneda 1981; Kaneda et al. 1983). One of these, D2/Dd9, turned out to have a high

degree of sequence similarity to the previously identified mammalian U3 RNA present in the nucleolus (Nakamura et al. 1968; Weinberg and Penman 1968; Prestayko et al. 1970; Wise and Weiner 1980; Takeishi and Kaneda 1981). U3 RNA was later recognized to play a role in rRNA processing (reviewed in Culver 2002; Nazar 2004). The only other *D. discoideum* nuclear RNA that was sequenced, D1/Dd8, revealed a 188 nt RNA with no apparent homology to RNAs from other species (Kaneda et al. 1983).

Non-coding RNAs as regulators of gene expression: the discovery of antisense RNAs

A few years later, Firtel and colleagues introduced longer artificial antisense RNAs as an important genetic tool in *D. discoideum* research (Crowley et al. 1985). They demonstrated that, upon introduction of specific antisense RNAs, it was possible to decrease the level of the complementary target mRNA and hence that of the associated protein. These results were in line with previous observations in other organisms, from bacteria to mammalian cells (Green et al. 1986). This discovery paved the way for studies in which antisense RNA was used as a fast alternative/complement to gene knockouts in order to mine the functions of specific genes in *D. discoideum* (Nellen et al. 1992; Kuspa et al. 1995; Gomer 1999). However, it was not until almost a decade later that an *endogenous* sense-antisense transcript pair in *D. discoideum* was reported by Nellen and co-workers (Hildebrandt and Nellen 1992). They showed that the developmentally important 2.2 kbp prespore gene *EB4* (later termed *psvA*) was regulated by a 1.8 kb antisense transcript which was complementary to a large part of the mRNA. The mRNA and the antisense RNA showed opposite expression patterns, where the mRNA was up-regulated during development, and the antisense RNA primarily expressed in growing cells. However, nuclear run-on experiments demonstrated that *psvA* mRNA transcription was more or less equal in growing and developing cells. Thus, the authors concluded that the difference in mRNA levels observed by Northern blot analysis must be due to post-transcriptional regulation. On the other hand, the differential expression of the antisense RNA seemed to be almost entirely dependent on transcriptional control. Furthermore, inhibition of transcription resulted in stabilization of psvA mRNA. Taken together, these results indicate that the antisense RNA developmentally down-regulates *psvA* mRNA levels. To our knowledge, this was the first reported example of an endogenous antisense regulated system in a eukaryotic organism. Another *D. discoideum* gene suggested to be regulated by an antisense-mediated mechanism is *nxnA*, encoding annexin VII (recently classified as annexin C1) (Okafuji et al. 1997; Moss and Morgan 2004). The evidence for antisense regulation is so far not as strong as for *psvA*. Whether or not there are additional *D. discoideum* genes that are post-transcriptionally controlled by antisense interactions is presently unresolved, but a recent in situ hybridization experiment gives us a hint (Maruo et al. 2004). In this study, hybridization with fluorescently labeled probes targeting the mRNA and antisense strands of a developmentally regulated prespore gene (designated SSH704) of unknown function yielded signals in developing *D. discoideum*. Interestingly, the mRNA and antisense RNA seem to accumulate in different cell types, suggestive of antisense-mediated mRNA control.

An mRNA-like non-coding RNA in *D. discoideum*

Another class of long ncRNAs has mRNA-like features such as spliced-out introns and a poly(A) tail (Erdmann et al. 2000). Interestingly, the expression of several such mRNA-like ncRNAs in, e.g., the fruit fly and mouse have been demonstrated to be spatially as well as temporally restricted (Mehler and Mattick 2006). An mRNA-like ncRNA in *D. discoideum* was reported in 1994 (Yoshida et al. 1994). The corresponding gene, at the time called *DC6*, had previously been identified as developmentally regulated; its expression was induced during aggregation early in *D. discoideum* development (Yoshida et al. 1991). The longest open reading frame of the 1,322 nt RNA could only give rise to a protein of 42 amino acids and subcellular fractionation indicated that the RNA was mainly cytoplasmic but not associated with ribosomes. The gene was therefore renamed *dutA* for development-specific but untranslatable RNA. The *dutA* gene has an unusually high A/T content (83%), even for the A/T rich genome of *D. discoideum* (72% A/T in open reading frames) (Eichinger et al. 2005). When this gene was disrupted by homologous recombination, no obvious developmental phenotype was obtained (Yoshida et al. 1994). In a follow-up study, Okamoto and colleagues demonstrated that the dutA RNA is transcribed by RNA pol II, and that developmental regulation was primarily due to changes in transcription (Kumimoto et al. 1995). The function of the dutA RNA is still unknown, but an intriguing clue came from a recent in situ hybridization analysis of developing *D. discoideum* (Maeda et al. 2003). In this study, dutA (referred to as

clone SSJ314) exhibits a well-defined expression pattern in two prestalk subtypes: PstA and PstO. This may suggest that *dutA* plays a role in *D. discoideum* development, but that its function may be redundant as indicated by gene disruption experiments. Furthermore, a microarray analysis in which *D. discoideum* cells were used as a model for infection with *Legionella pneumophila* indicated down-regulation of dutA mRNA during infection. A possible function of *dutA* in host–pathogen interaction remains to be elucidated (Farbrother et al. 2006).

A novel mitochondrial small RNA

Thus far, all described *D. discoideum* ncRNAs originated from the nuclear chromosome (except for mitochondrial rRNAs and tRNAs). However, a 129 nt mitochondrial small RNA (msRNA) was reported in 1998 (Pi et al. 1998). This RNA shows some 5S rRNA-like features but lacks extensive sequence similarity to the rRNA. Moreover, its apparent absence from mitochondrial ribosomes suggests an alternative role. Recently, several new mitochondrial ncRNAs have been isolated from mouse, perhaps indicating a more complex picture of the mitochondrial transcriptome than previously anticipated (Lung et al. 2006).

Non-coding RNAs in the twenty-first century

The last few years' advances in experimental as well as computational methodology, to isolate and identify small ncRNAs, have dramatically changed our view on how different processes are regulated in the cell (Hüttenhofer et al. 2005; Hüttenhofer and Vogel 2006). Large-scale cloning efforts and tiling microarray analyses have revealed a huge number of small RNAs in organisms from all kingdoms of life. Moreover, the power of today's computers have enabled whole-genome bioinformatics searches for small RNAs, using sequence and/or structure similarities to previously identified RNAs as well as other search criteria, such as G/C content. Before these new experimental and computational methods were developed, the identification of ncRNAs was restricted to highly abundant RNAs, which had to be purified from cell extract and sequenced by time-consuming methods.

Until just a few years ago, only a handful of small RNAs had been reported in *D. discoideum* (see above). The new advances in ncRNA biology encouraged us to take advantage of both experimental and bioinformatical tools in order to investigate the small RNA population in *D. discoideum*. These analyses have exposed classes of RNAs previously identified in other organisms as well as RNAs that may be unique to

Table 1 Non-coding RNAs (except for rRNA and tRNA) in *D. discoideum*

RNA or class of RNAs	Assigned/putative function	Identification method	Reference(s)
RNAs identified in other organisms			
Spliceosomal RNAs	Pre-mRNA splicing	cDNA, bioinformatics (experimentally verified)	Aspegren et al. (2004) Hinas et al. (2006)
snoRNAs			
D2/Dd9 (U3)	rRNA processing	RNA sequencing	Wise and Weiner (1981) Takeishi and Kaneda (1981)
Box C/D	RNA methylation	cDNA	Aspegren et al. (2004)
Box H/ACA	RNA pseudouridylation	cDNA	Aspegren et al. (2004)
SRP RNA	Protein localization	cDNA	Aspegren et al. (2004)
RNase P RNA	tRNA processing	Bioinformatics[a], genomic PCR	Piccinelli et al. (2005) Marquez et al. (2005)
RNase MRP RNA	rRNA processing	Bioinformatics[a]	Piccinelli et al. (2005)
Antisense RNA	Post-transcriptional regulation of gene expression	Northern blot analysis	Hildebrandt and Nellen (1992)
siRNAs	Transcriptional and post-transcriptional regulation of gene expression	cDNA	Kuhlmann et al. (2005)
RNAs unique (?) to *D. discoideum*			
D1/Dd8		RNA sequencing	Kaneda et al. (1983)
dutA		cDNA	Yoshida et al. (1994)
msRNA		RNA sequencing	Pi et al. (1998)
Class I RNAs		cDNA	Aspegren et al. (2004)
Class II RNAs		cDNA	Aspegren et al. (2004)

RNA classes identified in other organisms as well as RNA classes so far only found in *D. discoideum* are listed, along with their function (assigned or putative) and means of identification

[a] Indicate computationally predicted ncRNAs that have subsequently been experimentally verified by us (Larsson et al., manuscript in preparation)

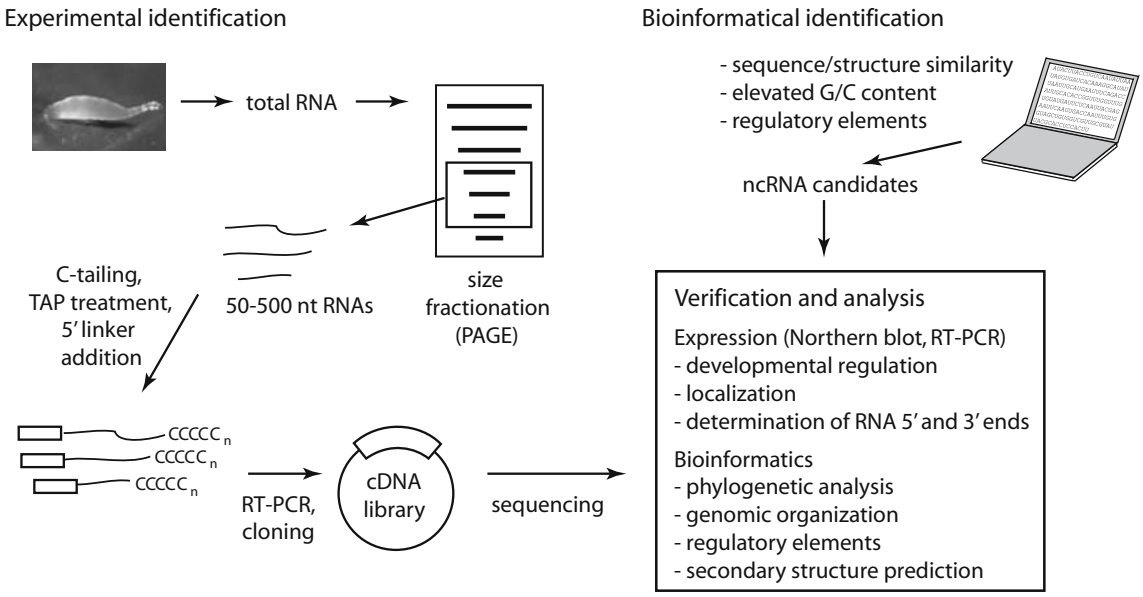

Fig. 2 Experimental and bioinformatical approaches used to identify ncRNAs (>50 nt) in *D. discoideum*. In the experimental approach, total RNA was isolated from cells undergoing development, size-fractionated on polyacrylamide gels, and small RNAs ranging in size from 50 to 500 nt were extracted. A C-tail was added to their 3′ ends followed by treatment with tobacco acid pyrophosphatase (*TAP*) to remove cap structures from the 5′ ends. Subsequently, an RNA oligo was ligated to the 5′ end of the C-tailed RNA. The RNA, now with known sequences at both ends, was converted into cDNA by reverse transcription PCR (*RT-PCR*) and ligated into plasmids. The generated cDNA libraries represented full-length small RNAs in the size range of 50–500 nt of which 36 unique cDNAs were derived from small ncRNAs. A similar experimental method was used to isolate 18–25 nt RNAs (Ambros and Lee 2004)

D. discoideum. The methods used and the identified small RNAs are described in the following sections and summarized in Table 1 and Fig. 2.

Experimental isolation of *D. discoideum* small RNAs

In a shotgun cloning approach, 36 new small RNAs were identified in *D. discoideum* cells undergoing development (for details, see Fig. 2 and Aspegren et al. 2004). These were further analyzed regarding (1) sequence, (2) similarity to small ncRNAs present in other organisms, (3) putative structure, (4) developmental expression, (5) subcellular localization (nucleus/cytoplasm), (6) genomic location, and (7) putative regulatory sequences, e.g., promoter motifs (Aspegren et al. 2004). In a similar approach, we cloned 18–25 nt RNAs in order to investigate the presence of siRNAs and possibly miRNAs in *D. discoideum* (Kuhlmann et al. 2005; Hinas et al., manuscript in preparation). The isolated RNAs will be described below.

Bioinformatics approaches to identify *D. discoideum* ncRNAs

Computational searches usually require additional information such as sequence elements that constitute hallmark motifs for specific classes of ncRNAs, often combined with comparative genomics. Our group has successfully used bioinformatics to search for ncRNAs based on conserved primary and secondary structural motifs and thereby identified the complete set of spliceosomal RNAs in *D. discoideum* (Hinas et al. 2006). In addition, other groups have identified ribonuclease P (RNase P) and MRP RNAs by similar methods (Marquez et al. 2005; Piccinelli et al. 2005). The characteristics of these RNAs are described in later sections.

Bioinformatical predictions of ncRNAs by sequence similarity are, naturally, limited to the discovery of already known classes of RNAs. Less biased bioinformatical methods are thus highly attractive. One such method, which is regularly used in combination with other techniques, is comparative genomics. In this approach, the genomes of closely related species are aligned to search for conserved sequences. In the case of *D. discoideum*, no close relative with a sequenced genome is presently available, although this will soon change since the genomes of several species closely related to *D. discoideum* are currently being sequenced (L. Eichinger, personal communication). However, certain features of *D. discoideum* have enabled us to take shortcuts in the quest for novel small RNAs, such as the high A/T content (overall 78%) of the *D. discoideum* genome (Eichinger et al. 2005).

In our analysis of the *D. discoideum* small RNAs identified by others and by us, we noted that, (1) all the RNAs are derived from intergenic regions and (2) their average G/C content (45%) is considerably higher than that of the surrounding intergenic sequences (14%). Based on these observations, we designed and experimentally validated an algorithm to identify novel RNA genes in the genome (Larsson et al., manuscript in preparation). Similar computational screens have been employed to predict ncRNA genes in other organisms with A/T rich genomes, such as hyperthermophiles and the malaria-causing parasite *Plasmodium falciparum* (Klein et al. 2002; Upadhyay et al. 2005). Furthermore, we identified a putative ncRNA promoter (Aspegren et al. 2004; Hinas et al. 2006), present upstream of many regions encoding ncRNAs in *D. discoideum* which may be instrumental to ncRNA gene discovery.

Small nucleolar RNAs

Large-scale isolation of small RNAs from a variety of model organisms has revealed a surprisingly large number of snoRNAs (Hüttenhofer et al. 2001; Marker et al. 2002; Tang et al. 2002; Yuan et al. 2003; Aspegren et al. 2004; Deng et al. 2006; Hüttenhofer and Vogel 2006; Zemann et al. 2006). These RNAs (60–300 nt) are present in the eukaryotic nucleus, mainly in the nucleolus, where some have roles in rRNA processing but the majority target specific nucleotides in rRNA for modification (Kiss 2002). However, snoRNAs can also modify other small RNAs, i.e., spliceosomal RNAs (Tycowski et al. 1998; Ganot et al. 1999; Jády and Kiss 2001; Söderbom 2006), possibly spliced leader (SL) RNA in trypanosomatids (Liang et al. 2002; Barth et al. 2005; Hastings 2005) and tRNA in archaea (Clouet d'Orval et al. 2001; Omer et al. 2003; Singh et al. 2004). The snoRNAs can be divided into two main families, box C/D and box H/ACA snoRNAs, which guide 2′-*O*-ribose methylation and pseudouridylation, respectively. The modifications are introduced by specific proteins associated with each family of snoRNAs (Bachellerie et al. 2002).

BLASTP search of the *D. discoideum* genome revealed a homolog with more than 60% identity to the box C/D snoRNA associated methyltransferase fibrillarin/Nop1 from human, *Arabidopsis thaliana*, and *Saccharomyces cerevisiae* (our unpublished observation and Söderbom 2006). In addition, the gene for the box H/ACA snoRNA associated pseudouridine synthase Cbf5 has been identified in the *D. discoideum* genome and the protein is expressed (Reinders et al. 2006).

U3 RNA, a snoRNA involved in rRNA processing in other eukaryotes, was discovered in *D. discoideum* in the late 1970s. It took another 25 years before larger scale isolation and cloning of ncRNAs in *D. discoideum* identified several other members of this class of ncRNAs (Aspegren et al. 2004). These, however, are predicted to belong to the modifying type, i.e., 17 box C/D and one H/ACA snoRNAs. The *D. discoideum* snoRNAs carry the hallmark features essential for snoRNA function, i.e., conserved sequence elements and predicted structures. In addition, *D. discoideum* snoRNAs seem to reside in the nucleus [only one snoRNA tested (Hinas and Söderbom, unpublished)] which is their expected location. Taken together, this strongly indicates that the isolated *D. discoideum* ncRNAs are *bona fide* snoRNAs. Expression analysis of the isolated snoRNAs demonstrated that several of them are developmentally regulated, suggesting roles at specific stages of growth or development (Aspegren et al. 2004). Furthermore, snoRNAs in other organisms contain ~10–21 nt sequence motifs that can base pair with complementary target sequences and thereby specifically mark the nucleotide to be modified (Bachellerie et al. 1995; Cavaillé et al. 1996; Kiss-László et al. 1996; Nicoloso et al. 1996; Tycowski et al. 1996; Ganot et al. 1997; Ni et al. 1997). This also seems to be true for the *D. discoideum* snoRNAs. About two thirds of the snoRNAs have ~10 nt sequence elements complementary to rRNA, and 2′-*O*-ribose methylation of one predicted nucleotide was experimentally demonstrated (Aspegren et al. 2004). What about the remaining "orphan" one-third? It seems unlikely that they would introduce modification in the canonical target ncRNAs, since bioinformatics searches for complementary target sequences failed to identify any good candidates (Aspegren et al. 2004). Interestingly, new targets for snoRNAs have recently been suggested, namely mRNAs. The best-studied example is the mammalian brain specific box C/D snoRNA HBII-52 which shows sequence complementarity to the serotonin receptor 5-HT_{2C}R mRNA and may influence alternative splicing of the pre-mRNA (Cavaillé et al. 2000; Vitali et al. 2005; Kishore and Stamm 2006). Loss of expression of HBII-52 snoRNA has been suggested to contribute to Prader–Willi syndrome, a severe human disease, possibly by affecting mRNA processing (Cavaillé et al. 2000; Kishore and Stamm 2006). Bioinformatic searches based on sequence complementarity to the *D. discoideum* snoRNAs with unassigned targets have so far failed to identify specific mRNA target candidates, and depletion of one of the orphan snoRNAs by gene disruption did not yield any obvious phenotype (Schumacher and Söderbom, unpublished). Thus, targets and

functions for the snoRNAs lacking sequence complementarity to canonical target RNAs in *D. discoideum* remain elusive.

The *D. discoideum* spliceosomal RNAs: expected and unexpected findings

The spliceosomal RNAs (snRNAs) U1, U2, U4, U5, and U6 associate with a large number of proteins to form snRNPs which recognize and remove introns from pre-mRNA transcripts (Will and Lührmann 2001). A growing number of studies also indicate a role for snRNAs per se in catalysis of the splicing reaction (Valadkhan 2005).

D. discoideum genes have few introns. About 30% of the predicted genes lack introns and the remaining genes have an average of 1.9 introns (Eichinger et al. 2005). The introns carry the canonical GT and AG dinucleotides as parts of their 5′ and 3′ splice site, respectively, and are generally short and very A/T rich (Eichinger et al. 2005; Szafranski et al. 2005). In spite of this low occurrence of introns, alternative splicing has been reported for a number of genes in *D. discoideum* (e.g., Grant et al. 1990; Louis et al. 1993; Escalante et al. 2003).

The first *D. discoideum* spliceosomal RNA, U2 snRNA, was represented in one of our cDNA libraries (Aspegren et al. 2004), and the remaining U snRNAs were identified bioinformatically (Hinas et al. 2006). Altogether, 18 spliceosomal RNA genes were predicted, and 17 of these were shown to be expressed. The *D. discoideum* snRNAs can be folded into the well-conserved secondary structures known from other organisms and have all the sequence motifs required for splicing (Staley and Guthrie 1998; Yu et al. 1999; Hinas et al. 2006). In accordance with other eukaryotes, U1, U2, U4, and U5, but not U6, carry a trimethylated 5′ cap (Hinas et al. 2006). A second less abundant spliceosome, the U12-dependent spliceosome, with a distinct set of snRNAs has been discovered in some organisms (reviewed in Will and Lührmann 2005). Currently, this minor spliceosome and its unique RNAs have not been identified in *D. discoideum*.

At a first glance, the *D. discoideum* snRNAs conformed to the snRNAs from other organisms. However, upon closer inspection, we noticed the presence of a subclass of four divergent *D. discoideum* U2 RNAs among the seven predicted U2 genes. These RNAs have conserved motifs typical for U2 RNAs, such as branch point recognition sequences and sequence motifs predicted to form interactions with U6 RNA (Yu et al. 1999; Hinas et al. 2006), but carry an additional sequence at their 5′ ends that potentially could fold into a stem-loop structure (Hinas et al. 2006). Interestingly, and in sharp contrast to other spliceosomal RNAs, these U2-like RNAs are enriched in the cytoplasm and down-regulated during development. To our knowledge, such atypical U2-like RNAs have not been observed in any other species.

Another unexpected finding was that a fraction of all *D. discoideum* snRNAs is polyadenylated (Hinas et al. 2006). Polyadenylation in eukaryotes has long been thought to solely be associated with mRNAs, where poly(A) tails increase mRNA stability and translation (Colgan and Manley 1997). On the other hand, polyadenylation in bacteria stimulates RNA degradation (Deutscher 2006). In the last few years, polyadenylation has been shown to destabilize ncRNA also in yeast, and a number of studies have directly or indirectly revealed polyadenylated ncRNAs in many eukaryotes, including humans (Stolc et al. 2004, 2005; Anderson 2005; Slomovic et al. 2006). Thus, polyadenylation in eukaryotes may also function as part of a surveillance machinery in which aberrant RNA is tagged with an A-tail and subsequently degraded (Anderson 2005). The proteins required for polyadenylation and degradation of ncRNAs in yeast are located in the nucleus (Allmang et al. 1999; Huh et al. 2003). However, the A-tailed snRNAs in *D. discoideum* are enriched in the cytoplasm, at least in the case of U1 RNA, indicating either the presence of poly(A) polymerase activity in the cytoplasm or, alternatively, nuclear polyadenylation followed by export to the cytoplasm. Whether polyadenylation of snRNA in *D. discoideum* is linked to degradation of aberrant RNAs, or is part of normal snRNA biogenesis, is presently unknown.

Non-coding RNAs in RNA processing and protein localization: ribonuclease P, ribonuclease MRP, and signal recognition particle RNAs

In eukaryotic cells, two ribonucleoprotein (RNP) enzymes have important roles in RNA processing. RNase P, which is also present in bacteria and archaea, cleaves off 5′-leader sequences of tRNAs to produce mature 5′ ends (Hartmann and Hartmann 2003). The RNase P RNA subunit alone from bacteria and archaea can perform the catalytic activity in vitro and hence act as a ribozyme (Guerrier-Takada et al. 1983; Pannucci et al. 1999). RNase MRP is exclusively present in eukaryotes and promotes rRNA processing/maturation (Schmitt and Clayton 1993; Chu et al. 1994; Lygerou et al. 1994, 1996). Mutations within the RNase

MRP RNA gene are associated with human diseases, such as cartilage-hair hypoplasia and anauxetic dysplasia (Ridanpää et al. 2001, 2002; Thiel et al. 2005).

D. discoideum RNase P activity was first demonstrated in 1995 by Drainas and colleagues (Stathopoulos et al. 1995). The cellular enzyme requires both protein and RNA subunits for catalytic activity as in other eukaryotes. However, the identity of the subunits was not determined. The presence of RNase MRP was only recently suggested in *D. discoideum*. In 2005, two independent reports shed some light on the missing components of the two RNP complexes. By experimental and bioinformatic methods, the putative genes for RNase P and MRP RNAs from a great number of organisms were identified, including *D. discoideum* (Marquez et al. 2005; Piccinelli et al. 2005). We recently verified the expression of these genes throughout development of *D. discoideum* (Avesson and Söderbom, unpublished).

Another RNA–protein complex present in all kingdoms of life is the signal recognition particle (SRP) (Halic and Beckmann 2005). In eukaryotes, the SRP binds to the signal peptide of nascent proteins as it emerges from the ribosome and causes translational arrest, whereafter the SRP guides the protein–ribosome complex to the endoplasmic reticulum (ER). Here, SRP is released, translation resumes, and the nascent peptide chain is co-translationally inserted into the lumen of ER for subsequent transport to the appropriate membrane. The *D. discoideum* SRP RNA was represented in our cDNA library and the ~280 nt RNA is expressed throughout development (Aspegren et al. 2004). The RNA is predicted to fold into the Y-like structure characteristic of SRP RNAs in archaea and eukaryotes, and the conserved sequence motifs are present (Nagai et al. 2003; Rosenblad et al. 2004). Not surprisingly, a recent bioinformatic analysis of the *D. discoideum* genome identified the protein homologs shown to interact with SRP RNA in other organisms (Andersen et al. 2006). An additional copy of the SRP RNA gene with 88% sequence identity is present in *D. discoideum*. Whether this copy is transcribed or constitutes a pseudogene is presently not known.

Novel RNA classes: abundant junk or hidden jewels?

Possibly the most exciting result from the *D. discoideum* small RNA cDNA libraries was the presence of two novel classes of RNAs (Aspegren et al. 2004). These RNAs (55–65 nt) have no apparent sequence homology to RNAs isolated from any other organisms. Based on shared sequence and structure motifs, they were denoted Class I RNAs (14 unique sequences) and Class II RNAs (two unique sequences) (Fig. 3). The Class I RNAs have conserved sequence elements located at their 5′ and 3′ ends, 16 and 8 nt, respectively. Parts of these sequence motifs are complementary and form a six-nucleotide stem structure keeping the ends of the RNAs together (Avesson and Söderbom, unpublished). Interestingly, Class I RNA is down-regulated ~twofold during development (Aspegren et al. 2004). Experimental data suggest that members of this class of RNA are transcribed as longer precursors which are subsequently processed to their mature size (Schumacher and Söderbom, unpublished). The 14 unique Class I RNAs are predicted to be encoded from 17 loci, where two and three identical Class I RNA genes, respectively, are present on chromosome 4 (see below). Furthermore, bioinformatical analysis of the *D. discoideum* genome revealed an additional 24 unique regions expected to be transcribed into Class I RNAs (Aspegren et al. 2004). In a later study, Stadler and colleagues found a similar number of genes when using the Class I RNAs to validate a search algorithm for small RNAs (Mosig et al. 2006). The second class of novel RNAs, Class II RNA (two almost identical sequences), is also predicted to a form a short RNA duplex by base-pairing between conserved sequences in the 5′ and 3′ ends (Fig. 3). In contrast to the Class I RNAs, Class II RNAs show no apparent developmental regulation (Aspegren et al. 2004).

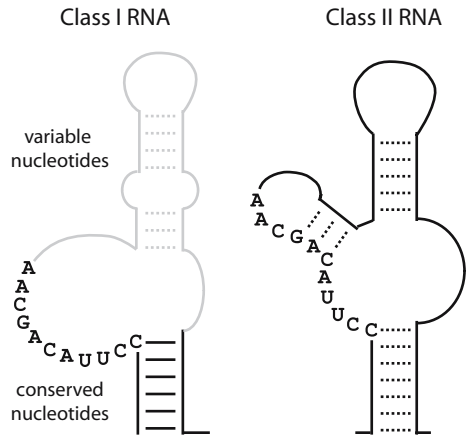

Fig. 3 Predicted secondary structure (Mfold) of Class I and Class II RNAs. Conserved and variable nucleotides within each class are depicted in *black* and *gray*, respectively (the two nucleotides differing between the Class II RNAs have been left out for simplicity). In one unique Class I RNA sequence, an A to G substitution in the stem is predicted to maintain base-pairing, corroborating the predicted stem structure. The 11 nt sequence motif shared between the two classes is indicated. Experimentally verified base pairs (Avesson and Söderbom, unpublished) and predicted base pairs are represented by *solid* and *dashed lines*, respectively

Class I and II RNAs have several features in common. They are abundant, expressed at all tested stages of development, predominantly localized in the cytoplasm, their ends have the potential to form a short stem, and the regions encoding these RNAs are preceded by a conserved putative promoter element (Aspegren et al. 2004). Intriguingly, the two classes also share an 11 nt sequence element predicted to be located at similar positions within their structures. All these features suggest that these RNAs play biological role(s) in the cell as opposed to being merely nonfunctional transcriptional noise. Eukaryotic small RNAs of unknown function have been isolated in similar cloning approaches but no extensive classes based on sequence and structural motifs were reported (Hüttenhofer et al. 2001; Marker et al. 2002; Yuan et al. 2003; Zemann et al. 2006). However, in a recent report, two new classes of small RNAs were isolated from the nematode *C. elegans* (Deng et al. 2006). One of these classes resembles Class I and II from *D. discoideum* in that the RNAs have conserved 5′ and 3′ sequences with the potential to form a duplex, interspersed by sequences of variable length and nucleotide composition. So far no function has been assigned to these RNAs.

It is tempting to speculate that the conserved sequence and/or structural elements present in *D. discoideum* Class I and II RNAs constitute binding motifs for specific proteins. The variable sequences in the RNAs could bind to their specific target RNAs via antisense interactions, guiding the associated proteins to their site of action. Hence, the RNA would provide specificity for the modifying or regulatory protein components. This situation would be analogous to how many other small ncRNAs function, e.g., snoRNAs and si/miRNAs. Moreover, the failure to detect homologs to Class I and II RNAs based on sequence searches of genome databases does not necessarily mean that these RNAs are specific to *D. discoideum*. Secondary structures are often more conserved than primary sequences in ncRNAs, but much harder to find by computational methods (Eddy 2002). At this point, the possible biological roles of the novel classes of RNAs in *D. discoideum* remain elusive. Experiments to understand their function, by gene disruptions and RNA–protein pull-downs, are underway.

Very small RNAs and their protein partners: RNA interference

The technique to use antisense RNA to knock down specific genes was established several decades ago and has been extensively used to study gene function in many organisms, including *D. discoideum*. However, in 1998, Fire, Mello, and colleagues reported that injection of double stranded RNA (dsRNA) into the nematode *C. elegans* resulted in a far more potent silencing of target genes than antisense (or sense) RNA alone. This sequence-dependent dsRNA-induced gene silencing was termed RNA interference (RNAi) (Fire et al. 1998).

RNAi is not a phenomenon invented by clever scientists—it has an endogenous role in most eukaryotic cells. In the last few years, RNAi has been demonstrated to affect gene expression at a number of different levels, from RNA degradation to transcriptional silencing through DNA and/or histone methylation and even DNA elimination in certain organisms (for reviews, see Birchler and Matzke 2005; Mochizuki and Gorovsky 2004; Wassenegger 2005; Zamore and Haley 2005). It is now known that the observed gene silencing is mediated by the so-called small interfering RNAs (siRNAs), which are only ~20–26 nt long and derived from longer dsRNA molecules (Hamilton and Baulcombe 1999; Hammond et al. 2000; Zamore et al. 2000). Homologs to proteins involved in RNAi, as well as the ability to knock down gene expression by RNAi have been demonstrated in organisms representing all eukaryotic supergroups (Cerutti and Casas-Mollano 2006 and references therein). It seems likely that the RNAi machinery is an ancient feature of eukaryotes that has subsequently been lost in some lineages, such as *Saccharomyces cerevisiae* (Cerutti and Casas-Mollano 2006). In *D. discoideum*, knock-down of gene expression by RNAi was first achieved by transgenic expression of an RNA hairpin construct, now a standard procedure for knock-down of *D. discoideum* genes (Martens et al. 2002; Lee et al. 2005; Boeckeler et al. 2006).

Core proteins in the RNAi pathway are *Dicers*, *Argonaute-Piwi* proteins and, in some organisms, *RNA-dependent RNA polymerases*. These protein classes and their counterparts in *D. discoideum* are described in the following sections. *Dicers* belong to a class of RNase III-type enzymes and process longer dsRNA molecules into siRNAs (Zamore et al. 2000; Bernstein et al. 2001). In *D. discoideum*, two genes encoding Dicer-like proteins, *drnA* and *drnB*, have been identified (Martens et al. 2002). Gene disruption of either *drnA* or *drnB* alone does not affect RNAi efficiency or the accumulation of silencing-associated siRNAs, indicating a functional redundancy as described for Dicers in *A. thaliana* (Martens et al. 2002; Gasciolli et al. 2005).

Argonute-Piwi proteins are commonly divided into two classes, Ago-like and Piwi-like, based on similarity to the *A. thaliana* Argonaute and *D. melanogaster* Piwi proteins, respectively (Carmell et al. 2002). Some Ago-

like proteins, often referred to as "slicers", introduce cleavage of the target RNA when guided by cognate siRNAs (Liu et al. 2004; Meister et al. 2004; Okamura et al. 2004; Baumberger and Baulcombe 2005). Piwi-like proteins have been implicated in germ line development in mammals and *D. melanogaster*, but their mechanism of action remains largely unknown (Carmell et al. 2002; Jaronczyk et al. 2005). An exception to this comes from a very recent study in *D. melanogaster*, which for the first time demonstrated slicer activity for a Piwi-like protein (Saito et al. 2006). Also, an abundant novel class of 26–30 nt Piwi-interacting RNAs (piRNAs) was recently discovered in mammalian testes (Aravin et al. 2006; Girard et al. 2006; Lau et al. 2006; Watanabe et al. 2006).

The distribution of Argonaute-Piwi proteins differs substantially between organisms. Animals possess both Ago-like and Piwi-like proteins, while only the Ago-like type is present in plants and fungi. In Amoebozoa (including *D. discoideum*) and ciliates only Piwi-like proteins have been identified (Cerutti and Casas-Mollano 2006). The *D. discoideum* genome contains five genes predicted to encode putative Piwi-like proteins (Cerutti and Casas-Mollano 2006). Interestingly, they all contain the potentially catalytic Asp-Asp-His motif at conserved positions (our unpublished observation). This catalytic triad (or a similar motif) has been demonstrated to be necessary, although not sufficient, for RNA cleavage by both Ago-like and Piwi-like proteins (Parker and Barford 2006). It is thus possible that the *D. discoideum* Piwi-like proteins are capable of cleaving RNA; this remains, however, to be demonstrated.

In many organisms, such as nematodes, plants and fungi, *RNA-dependent RNA polymerases* (RdRPs) are also required for RNAi. RdRPs synthesize a complementary RNA strand using a single stranded RNA as a template and are believed to amplify the initial RNAi signal, leading to more potent silencing (Wassenegger and Krczal 2006). *D. discoideum* encodes three RdRP homologs, *rrpA, rrpB* and *rrpC*, and *rrpA* is required for RNAi-mediated knock-down of gene expression; *rrpB* and *rrpC* are dispensable (Martens et al. 2002).

A nuclear protein in *D. discoideum*, HelF, with homology to the RNA helicase domain of Dicer proteins, counteracts RNAi-mediated silencing through an unknown mechanism (Popova et al. 2006). Suppression of RNAi has also been shown for certain proteins in *C. elegans*, e.g., a putative RdRP and an exonuclease (Simmer et al. 2002; Kennedy et al. 2004; Duchaine et al. 2006). Interestingly, these proteins seem to be required for the generation of endogenous small RNAs (Duchaine et al. 2006; Lee et al. 2006). In line with this, it is possible that the *D. discoideum* HelF protein affects the biogenesis of endogenous small RNAs, although at this point there is no supporting evidence.

Why does the cell need an RNAi machinery? In order to assess the function of endogenous RNAi, small RNAs have been cloned from many organisms (reviewed in Aravin and Tuschl 2005). The emerging pattern of the cloned siRNAs points towards two major roles of RNAi—fighting off viral infections and preventing mobilization of genomic retrotransposons (Aravin and Tuschl 2005; Li and Ding 2005; Voinnet 2005).

In *D. discoideum*, large-scale cloning of 18–25 nt RNAs revealed a disproportionately large fraction of 21 nt RNAs derived from the retrotransposon DIRS-1 (Fig. 4) (Kuhlmann et al. 2005; Hinas et al., manuscript in preparation). DIRS-1 is the most abundant retrotransposon in the *D. discoideum* genome and DIRS-1 clusters, which are located at one end of each chromosome, have been suggested to constitute centromeres (Eichinger et al. 2005). An antisense transcript of ~900 nt covers part of the DIRS-1 mRNA, hence 21 nt RNAs could be generated by cleavage of arising dsRNA by Dicer (Cappello et al. 1985; Kuhlmann et al. 2005). However, the isolated small RNAs cover both strands of the entire retrotransposon, including parts not engaged in putative base-pairing between mRNA and antisense RNA. Presently, we do not know by which mechanism these RNAs are generated. Furthermore, DIRS-1 sequences have been reported to be subject to DNA methylation by the Dnmt2-type DNA methylase DnmA, but so far there is no evidence linking the DNA methylation to the DIRS-1 siRNAs (Kuhlmann et al. 2005; Katoh et al. 2006). In addition to the DIRS-1 derived siRNAs, a number of small RNAs antisense to mRNAs were represented in our cDNA library (Hinas et al, manuscript in preparation). Such small antisense RNAs have also been identified in, e.g., *C. elegans* and *Tetrahymena thermophila* (Ambros et al. 2003; Lee and Collins 2006).

Besides siRNAs, a great number of a related class of small RNAs, miRNAs, has been identified in plants, animals and animal viruses (Lee et al. 1993; Lee and Ambros 2001; Nair and Zavolan 2006). RNAs of this class have been assigned crucial roles in gene regulation in a wide variety of cellular processes, ranging from apoptosis to development and stress responses (Carthew 2006; Esquela-Kerscher and Slack 2006; Mallory and Vaucheret 2006). The miRNAs are similar in size to siRNAs and require the same or similar proteins for their biogenesis, but are processed from imperfect hairpin precursors instead of perfectly base-paired dsRNA (Du and Zamore 2005). To date, miRNAs have not been reported in any unicellular organism. In our cDNA library, we isolated a small number of inter-

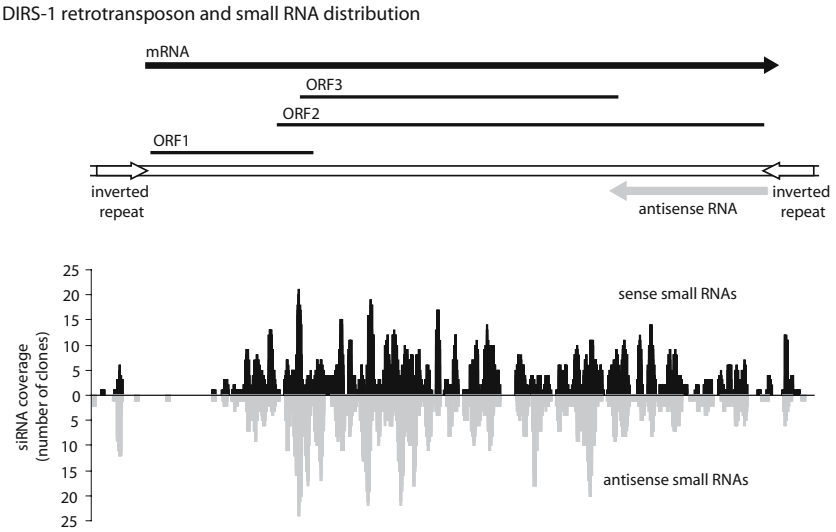

Fig. 4 Schematic diagram of the DIRS-1 retrotransposon showing the positions and orientations of the inverted repeats (*open arrows*), mRNA (*black arrow*) with three overlapping open reading frames (ORFs *solid lines*), and the ~900 nt antisense RNA (*gray arrow*). The cloned small RNAs are relatively evenly distributed along both strands

genic small RNAs and we are presently investigating the possibility that these may represent miRNAs.

Genomic organization

So far, all *D. discoideum* small ncRNA genes were found in intergenic regions, i.e., transcribed from sequences between protein-coding genes (Fig. 5a), as are most ncRNA in other organisms with snoRNAs and miRNAs as notable exceptions.

The genomic organization of modification snoRNA genes varies between different organisms. The snoRNAs in *D. discoideum* appear to be transcribed from independent genes, or transcribed as polycistronic units (Aspegren et al. 2004). This organization resembles that in *A. thaliana* (Brown et al. 2003), budding yeast (Lowe and Eddy 1999; Schattner et al. 2004) and protists (Uliel et al. 2004; Yang et al. 2005). In contrast, vertebrate and *D. melanogaster* snoRNAs seem to be mainly derived from intronic sequences (Tycowski and Steitz 2001; Yuan et al.

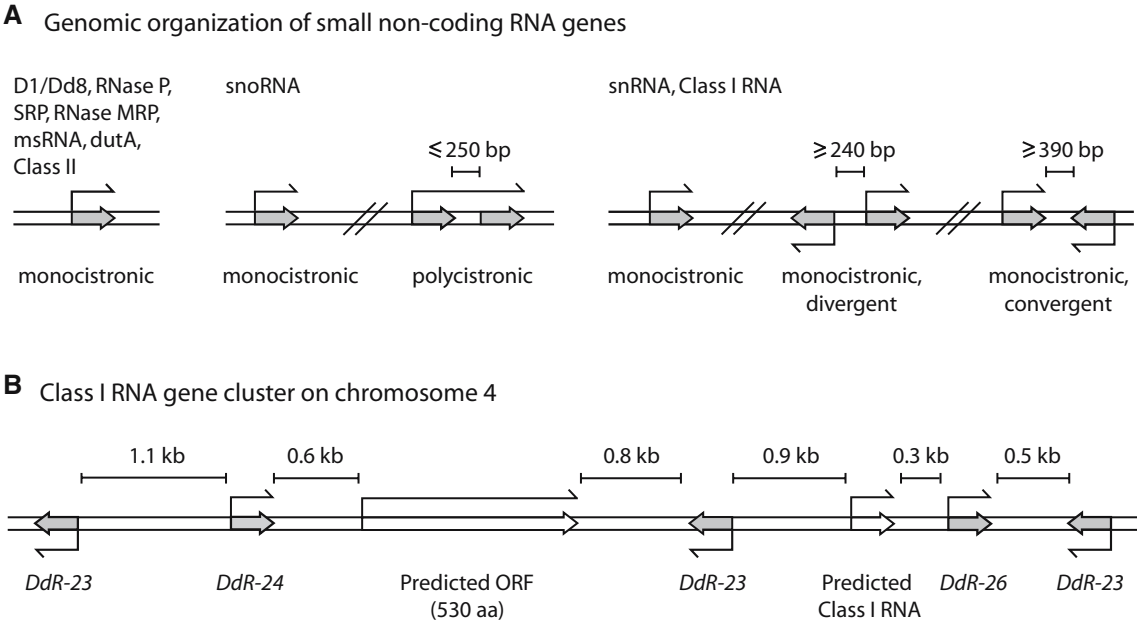

Fig. 5 Gene organization of *D. discoideum* ncRNAs. **a** The majority of the ncRNA classes are monocistronic, with the exception of some snoRNAs that are cotranscribed. Several snRNA and Class I RNA genes are situated closely together in divergent or convergent pairs. **b** Example of a Class I RNA gene cluster on chromosome 4. Three copies of Class I RNA gene *DdR-23* are positioned closely together in the same orientation, interspersed by other Class I RNA genes as well as a predicted ORF. Predicted genes and genes with demonstrated expression are represented by *open* and *gray arrows*, respectively

2003; Accardo et al. 2004; Huang et al. 2005; Mattick and Makunin 2005).

The 17 expressed genes encoding spliceosomal RNAs are located on four of the six chromosomes in *D. discoideum* (Hinas et al. 2006). The majority is present as multiple copies, four U1, seven U2, three U4, two U5 snRNA genes, with only minor sequences variations within each class. The exception is the U6 snRNA which appears to be derived from a single gene. This gene organisation, with a variable number of copies for each spliceosomal RNA gene, is common in other eukaryotes as well (e.g., Goffeau et al. 1996; Adams et al. 2000; The Arabidopsis Genome 2000; Stein et al. 2003). Moreover, most *D. discoideum* snRNA genes are clustered in pairs, where two copies encoding homologous genes are transcribed either divergently, convergently, or in tandem. Gene pairs are more similar to each other than to genes within the same gene class placed further away on the same or on other chromosomes. This suggests that many of the snRNA genes arose via gene duplication. Interestingly, genes within three of these clusters are positioned very close together with only 242–344 nt between the experimentally determined start or end of transcription (Hinas et al. 2006).

The Class I RNA genes (Aspegren et al. 2004) show a somewhat different genomic localization. Although copies of these genes can be found on all chromosomes except chromosome six, a larger cluster is present on chromosome four. A peculiar observation is that genes encoding (or predicted to encode) the same Class I RNA species (DdR-23) are repeated, i.e., three identical copies are present as closely spaced direct repeats (Fig. 5b). Interspersed between the DdR-23 genes are other Class I RNA genes, as well as a predicted protein-coding gene, transcribed in opposite directions as compared to DdR-23. The regions between the RNA genes are relatively short (0.3–3 kb).

In accordance with the genomic organization of Class I and snRNA genes, many tRNA genes in *D. discoideum* are also arranged in clusters where pairs with identical anticodons are commonly found in tandem, diverging or converging orientation (Eichinger et al. 2005). However, while only a few hundred nucleotides can separate members of the Class I and snRNA genes, respectively, the tRNA gene pairs are positioned further apart.

How are ncRNA genes transcribed?

In eukaryotes, RNA polymerases I, II, and III all have distinct roles in transcription of ncRNA genes (reviewed in Paule and White 2000; Hernandez 2001; Schramm and Hernandez 2002; Smale and Kadonaga 2003). rRNAs are transcribed by RNA pol I and III. In addition, RNA pol III is responsible for transcription of genes encoding SRP, RNase P, RNase MRP, tRNA and U6 RNAs. The remaining major spliceosomal RNAs, U1, U2, U4, and U5, are transcribed by RNA pol II. A fourth RNA pol, recently identified in plants, is involved in the generation of small RNAs (~22 nt) that have been proposed to function in chromatin silencing (Herr et al. 2005; Onodera et al. 2005; Vaucheret 2005).

To approach transcriptional regulation of ncRNA genes in *D*, discoideum, we searched for putative promoter elements in the DNA sequences preceding the start of transcription. This analysis revealed a highly conserved sequence motif of eight base pairs ([A/T]CCCA[C/T]AA) ~63 nt upstream of the transcriptional start site (Aspegren et al. 2004; Hinas et al. 2006). We named this putative promoter sequence DUSE (*D. discoideum* Upstream Element). DUSE is present in front of ncRNA-encoding sequences that, in analogy to the situation in other eukaryotes, are expected to be transcribed by RNA pol II (U1–U5 snRNAs) as well as those predicted to be RNA pol III dependent (SRP, RNase P, RNase MRP and U6 RNAs). In the cases where RNA pol II and III recognize the same promoter element within the same organism, other features such as additional promoter elements (e.g., TATA box), their spacing and distance from the transcription start, define which genes are transcribed by a certain polymerase (Schramm and Hernandez 2002). So far, such features are not found near the ncRNA genes in *D. discoideum*. However, TATA boxes are hard to spot due to the high A/T content surrounding the transcribed regions. Nevertheless, since the *D. discoideum* U1–U5, but not U6, snRNAs carry the characteristic trimethylated cap structure (Hinas et al. 2006), this strongly indicates that U1–U5 snRNAs are transcribed by RNA pol II and U6 snRNA by RNA pol III, as in other eukaryotes (Will and Lührmann 2001).

The DUSE consensus sequence, and its distance from the start of transcription, is similar to the USE (Upstream Sequence Element) reported for promoters of snRNA genes in *T. thermophila* (Orum et al. 1992) and *A. thaliana* (Vankan et al. 1988; Waibel and Filipowicz 1990), and for SRP RNA genes (Heard et al. 1995; Yukawa et al. 2005), and RNase MRP RNA genes in plants (Kiss et al. 1992).

DUSE is also present in front of Class I and II RNA genes (Aspegren et al. 2004), and those encoding the previously identified ncRNAs D1/Dd8, D2/Dd9, and

dutA (Larsson et al., manuscript in preparation). Further evidence for DUSE constituting a remarkably conserved promoter element for ncRNA genes in *D. discoideum* comes from our recent bioinformatic screen of the genome. By combining solely two search criteria, (1) sequences with elevated G/C content and (2) presence of DUSE, we identified a number of putative ncRNAs. Northern blot analysis of a handful of these predicted ncRNAs showed that they were all expressed as small RNAs (Larsson et al., manuscript in preparation).

Conclusion

Progress in bioinformatics and experimental techniques, combined with the growing number of sequenced genomes, have recently revealed unexpected layers of information. One of the most astonishing revelations is the large fraction, maybe the majority, of the eukaryotic genome that is transcribed into RNAs (Frith et al. 2005). However, e.g., in humans, mRNAs represent only 2% of the euchromatic genome (Human Genome Sequencing 2004). These findings point to an intricate and unexpected complexity of the eukaryotic transcriptome—most of which we to date do not understand. In the upcoming years, we need to address several questions such as do the majority of these non-protein coding RNAs have a function, or are they merely transcriptional noise without any physiological roles? Today, we know that RNA sequences such as introns that historically have been regarded as by-products from pre-mRNA processing have important cellular functions. Often, they are further processed to mature regulatory ncRNAs, e.g., snoRNAs and miRNAs, with central roles in a variety of biological processes. Also, many regulatory RNAs are derived from genomic sequences in between protein encoding regions (hence challenging the widely used term "intergenic"). Although the biological roles of some of the regulatory RNAs have been elucidated, the majority still awaits functional assignment. In this context, the importance of studying different model organisms cannot be overrated since the regulatory RNAs, their associated proteins, and often their targets are conserved. This has already given us enough information to generalize some of the cellular roles for certain classes of ncRNAs.

D. discoideum has recently emerged as a model to discover and analyze the function of ncRNAs. Many fundamental biological processes in *D. discoideum,* such as chemotaxis, signal transduction and cell differentiation, just to mention a few, have been investigated in great detail and share many features with those of multicellular organisms, e.g., mammals. Complex biological pathways can be investigated by utilizing the molecular and biochemical toolbox available for this organism. Yet another important advantage is the ease with which this amoeba grows and synchronously develops under laboratory conditions. Taken together, these features make *D. discoideum* an amenable model to investigate functional aspects of ncRNAs. This has already been demonstrated in a number of studies reaching back to the late 1970s, the era when studies of small RNAs took off. Not surprisingly, *D. discoideum* encodes the classes of ncRNAs present in other eukaryotes, i.e., snoRNAs, snRNAs, SRP, RNase P and MRP RNAs, siRNAs, and longer antisense RNAs. However, distinct novel classes of RNAs are present that may or may not be organism-specific. Other recently uncovered surprises are cytoplasmic polyadenylated snRNAs and the presence of developmentally regulated spliceosomal-like U2 RNAs that are mainly present in the cytoplasm. Issues to be addressed in the upcoming years are if these new findings are *D. discoideum*-specific "oddities" or more general phenomenons, and the biological significance of their presence.

Acknowledgments We thank the members of our research group, Pontus Larsson, and Johan Reimegård for supplying unpublished data. Gerhart Wagner, Victor Ambros, and Jan Andersson are gratefully acknowledged for comments on the manuscript. This work was supported by grants from the European Community (FOSRAK, EC005120) and The Swedish Research Council (Linneaus grant).

References

Accardo MC et al (2004) A computational search for box C/D snoRNA genes in the *D. melanogaster* genome. Bioinformatics 20:3293–3301

Adams MD et al (2000) The genome sequence of *Drosophila melanogaster*. Science 287:2185–2195

Allmang C, Petfalski E, Podtelejnikov A, Mann M, Tollervey D, Mitchell P (1999) The yeast exosome and human PM-Scl are related complexes of $3' \rightarrow 5'$ exonucleases. Genes Dev 13:2148–2158

Ambros V, Lee RC (2004) Identification of microRNAs and other tiny noncoding RNAs by cDNA clothing. In: Gott JM (ed) RNA interferance, editing, and modification. Human Press, Totowa, pp 131–158

Ambros V, Lee RC, Lavanway A, Williams PT, Jewell D (2003) MicroRNAs and other tiny endogenous RNAs in *C. elegans*. Curr Biol 13:807–818

Andersen ES et al (2006) The tmRDB and SRPDB resources. Nucleic Acids Res 34:D163–D168

Anderson JT (2005) RNA turnover: unexpected consequences of being tailed. Curr Biol 15:R635–638

Aravin A, Tuschl T (2005) Identification and characterization of small RNAs involved in RNA silencing. FEBS Lett 579:5830–5840

Aravin A et al (2006) A novel class of small RNAs bind to MILI protein in mouse testes. Nature 442:203–207
Asa SL, Ramyar L, Murphy PR, Li AW, Ezzat S (2001) The endogenous fibroblast growth factor-2 antisense gene product regulates pituitary cell growth and hormone production. Mol Endocrinol 15:589–599
Aspegren A, Hinas A, Larsson P, Larsson A, Söderbom F (2004) Novel non-coding RNAs in *Dictyostelium discoideum* and their expression during development. Nucleic Acids Res 32:4646–4656
Bachellerie JP, Michot B, Nicoloso M, Balakin A, Ni J, Fournier MJ (1995) Antisense snoRNAs: a family of nucleolar RNAs with long complementarities to rRNA. Trends Biochem Sci 20:261–264
Bachellerie JP, Cavaillé J, Hüttenhofer A (2002) The expanding snoRNA world. Biochimie 84:775–790
Bagorda A, Mihaylov VA, Parent CA (2006) Chemotaxis: moving forward and holding on to the past. Thromb Haemost 95:12–21
Bapteste E et al (2002) The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium*, *Entamoeba*, and *Mastigamoeba*. Proc Natl Acad Sci USA 99:1414–1419
Barth S, Hury A, Liang XH, Michaeli S (2005) Elucidating the role of H/ACA-like RNAs in trans-splicing and rRNA processing via RNA interference silencing of the *Trypanosoma brucei* CBF5 pseudouridine synthase. J Biol Chem 280:34558–34568
Baumberger N, Baulcombe DC (2005) *Arabidopsis* ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. Proc Natl Acad Sci USA 102:11928–11933
Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409:363–366
Birchler JA, Matzke MA (2005) RNAi-mediated pathways in the nucleus. Nat Rev Genet 6:24–35
Blum B, Bakalara N, Simpson L (1990) A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information. Cell 60:189–198
Boeckeler K, Tischendorf G, Mutzel R, Weissenmayer B (2006) Aberrant stalk development and breakdown of tip dominance in *Dictyostelium* cell lines with RNAi-silenced expression of calcineurin B. BMC Dev Biol 6:12
Brown JW, Echeverria M, Qu LH (2003) Plant snoRNAs: functional evolution and new modes of gene expression. Trends Plant Sci 8:42–49
Cappello J, Handelsman K, Lodish HF (1985) Sequence of *Dictyostelium* DIRS-1: an apparent retrotransposon with inverted terminal repeats and an internal circle junction sequence. Cell 43:105–115
Carmell MA, Xuan Z, Zhang MQ, Hannon GJ (2002) The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev. 16:2733
Carthew RW (2006) Gene regulation by microRNAs. Curr Opin Genet Dev 16:203–208
Cavaillé J, Nicoloso M, Bachellerie JP (1996) Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides. Nature 383:732–735
Cavaillé J et al (2000) Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. Proc Natl Acad Sci USA 97:14311–14316
Cerutti H, Casas-Mollano JA (2006) On the origin and functions of RNA-mediated silencing: from protists to man. Curr Genet 50:81–99
Chisholm RL et al (2006) dictyBase, the model organism database for *Dictyostelium discoideum*. Nucleic Acids Res 34:D423–427
Chu S, Archer RH, Zengel JM, Lindahl L (1994) The RNA of RNase MRP is required for normal processing of ribosomal RNA. Proc Natl Acad Sci USA 91:659–663
Clouet d'Orval B, Bortolin ML, Gaspin C, Bachellerie JP (2001) Box C/D RNA guides for the ribose methylation of archaeal tRNAs. The tRNATrp intron guides the formation of two ribose-methylated nucleosides in the mature tRNATrp. Nucleic Acids Res 29:4518–4529
Colgan DF, Manley JL (1997) Mechanism and regulation of mRNA polyadenylation. Genes Dev 11:2755–2766
Crowley TE, Nellen W, Gomer RH, Firtel RA (1985) Phenocopy of discoidin I-minus mutants by antisense transformation in *Dictyostelium*. Cell 43:633–641
Culver GM (2002) Sno-capped: 5′ ends of preribosomal RNAs are decorated with a U3 SnoRNP. Chem Biol 9:777
Deng W et al (2006) Organization of the *Caenorhabditis elegans* small non-coding transcriptome: genomic features, biogenesis, and expression. Genome Res 16:20–29
Deutscher MP (2006) Degradation of RNA in bacteria: comparison of mRNA and stable RNA. Nucleic Acids Res 34:659–666
Du T, Zamore PD (2005) Microprimer: the biogenesis and function of microRNA. Development 132:4645–4652
Duchaine TF et al (2006) Functional proteomics reveals the biochemical niche of C. *elegans* DCR-1 in multiple small-RNA-mediated pathways. Cell 124:343–354
Eddy SR (2002) Computational genomics of noncoding RNA genes. Cell 109:137–140
Eichinger L, Rivero F (2006) *Dictyostelium discoideum* protocols. Humana Press, Totowa, NJ
Eichinger L et al (2005) The genome of the social amoeba *Dictyostelium discoideum*. Nature 435:43–57
Erdmann VA, Szymanski M, Hochberg A, Groot N, Barciszewski J (2000) Non-coding, mRNA-like RNAs database Y2K. Nucleic Acids Res 28:197–200
Escalante R, Moreno N, Sastre L (2003) *Dictyostelium discoideum* developmentally regulated genes whose expression is dependent on MADS box transcription factor SrfA. Eukaryot Cell 2:1327–1335
Esquela-Kerscher A, Slack FJ (2006) Oncomirs—microRNAs with a role in cancer. Nat Rev Cancer 6:259–269
Farbrother P et al (2006) *Dictyostelium* transcriptional host cell response upon infection with *Legionella*. Cellular Microbiol 8:438–456
Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391:806
Frith MC, Pheasant M, Mattick JS (2005) Genomics: the amazing complexity of the human transcriptome. Eur J Hum Genet 13:894
Ganot P, Bortolin ML, Kiss T (1997) Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. Cell 89:799–809
Ganot P, Jády BE, Bortolin ML, Darzacq X, Kiss T (1999) Nucleolar factors direct the 2′-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. Mol Cell Biol 19:6906–6917
Gasciolli V, Mallory AC, Bartel DP, Vaucheret H (2005) Partially redundant functions of *Arabidopsis* DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. Curr Biol 15:1494–1500
Girard A, Sachidanandam R, Hannon GJ, Carmell MA (2006) A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature 442:199–202

Goffeau A et al (1996) Life with 6000 genes. Science 274:546, 563–547
Gomer RH (1999) Gene identification by shotgun antisense. Methods 18:311
Goodrich JA, Kugel JF (2006) Non-coding-RNA regulators of RNA polymerase II transcription. Nat Rev Mol Cell Biol 7:612–616
Gottesman S (2005) Micros for microbes: non-coding regulatory RNAs in bacteria. Trends Genet 21:399–404
Grant CE, Bain G, Tsang A (1990) The molecular basis for alternative splicing of the CABP1 transcripts in *Dictyostelium discoideum*. Nucleic Acids Res 18:5457–5463
Green PJ, Pines O, Inouye M (1986) The role of antisense RNA in gene regulation. Ann Rev Biochem 55:569–597
Griffiths-Jones S, Moxon S, Marshall M, Khanna A, Eddy SR, Bateman A (2005) Rfam: annotating non-coding RNAs in complete genomes. Nucleic Acids Res 33:D121–D124 (database issue)
Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell 35:849–857
Halic M, Beckmann R (2005) The signal recognition particle and its interactions during protein targeting. Curr Opin Struct Biol 15:116
Hamilton AJ, Baulcombe DC (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. Science 286:950–952
Hammond SM, Bernstein E, Beach D, Hannon GJ (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. Nature 404:293–296
Hartmann E, Hartmann RK (2003) The enigma of ribonuclease P evolution. Trends Genet 19:561
Hastings KEM (2005) SL trans-splicing: easy come or easy go? Trends Genet 21:240–247
Heard E (2005) Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome. Curr Opin Genet Dev 15:482
Heard DJ, Filipowicz W, Marques JP, Palme K, Gualberto JM (1995) An upstream U-snRNA gene-like promoter is required for transcription of the *Arabidopsis thaliana* 7SL RNA gene. Nucleic Acids Res 23:1970–1976
Hernandez N (2001) Small nuclear RNA genes: a model system to study fundamental mechanisms of transcription. J Biol Chem 276:26733–26736
Herr AJ, Jensen MB, Dalmay T, Baulcombe DC (2005) RNA polymerase IV directs silencing of endogenous DNA. Science 308:118–120
Hildebrandt M, Nellen W (1992) Differential antisense transcription from the *Dictyostelium EB4* gene locus: implications on antisense-mediated regulation of mRNA stability. Cell 69:197–204
Hinas A, Larsson P, Avesson L, Kirsebom LA, Virtanen A, Söderbom F (2006) Identification of the major spliceosomal RNAs in *Dictyostelium discoideum* reveals developmentally regulated U2 variants and polyadenylated snRNAs. Eukaryot Cell 5:924–934
Huang ZP, Zhou H, He HL, Chen CL, Liang D, Qu LH (2005) Genome-wide analyses of two families of snoRNA genes from *Drosophila melanogaster*, demonstrating the extensive utilization of introns for coding of snoRNAs. RNA 11:1303–1316
Huh WK et al (2003) Global analysis of protein localization in budding yeast. Nature 425:686–691
Human Genome Sequencing C (2004) Finishing the euchromatic sequence of the human genome. Nature 431:931
Hüttenhofer A, Schattner P (2006) The principles of guiding by RNA: chimeric RNA-protein enzymes. Nat Rev Genet 7:475–482
Hüttenhofer A, Vogel J (2006) Experimental approaches to identify non-coding RNAs. Nucleic Acids Res 34:635
Hüttenhofer A et al (2001) RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. EMBO J 20:2943–2953
Hüttenhofer A, Schattner P, Polacek N (2005) Non-coding RNAs: hope or hype? Trends Genet 21:289–297
Jády BE, Kiss T (2001) A small nucleolar guide RNA functions both in 2′-*O*-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. EMBO J 20:541–551
Jaronczyk K, Carmichael JB, Hobman TC (2005) Exploring the functions of RNA interference pathway proteins: some functions are more RISCy than others? Biochem J 387:561–571
Kaneda S, Gotoh O, Seno T, Takeishi K (1983) Nucleotide sequence of *Dictyostelium* small nuclear RNA Dd8 not homologous to any other sequenced small nuclear RNA. J Biol Chem 258:10606–10613
Katoh M, Curk T, Xu Q, Zupan B, Kuspa A, Shaulsky G (2006) Developmentally regulated DNA methylation in *Dictyostelium discoideum*. Eukaryot Cell 5:18–25
Kennedy S, Wang D, Ruvkun G (2004) A conserved siRNA-degrading RNase negatively regulates RNA interference in *C.* elegans. Nature 427:645–649
Kessin RH (2001) *Dictyostelium*—evolution, cell biology, and the development of multicellularity. Cambridge University Press, Cambridge, UK
Kimelman D, Kirschner MW (1989) An antisense mRNA directs the covalent modification of the transcript encoding fibroblast growth factor in *Xenopus* oocytes. Cell 59:687
Kishore S, Stamm S (2006) The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. Science 311:230–232
Kiss T (2002) Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. Cell 109:145–148
Kiss T, Marshallsay C, Filipowicz W (1992) 7–2/MRP RNAs in plant and mammalian cells: association with higher order structures in the nucleolus. EMBO J 11:3737–3746
Kiss-László Z, Henry Y, Bachellerie JP, Caizergues-Ferrer M, Kiss T (1996) Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. Cell 85:1077–1088
Klein RJ, Misulovin Z, Eddy SR (2002) Noncoding RNA genes identified in AT-rich hyperthermophiles. Proc Natl Acad Sci U S A 99:7542–7547
Kuhlmann M et al (2005) Silencing of retrotransposons in *Dictyostelium* by DNA methylation and RNAi. Nucleic Acids Res 33:6405–6417
Kumimoto H, Yoshida H, Okamoto K (1995) RNA polymerase II transcribes *Dictyostelium* untranslatable gene, *dutA*, specifically in the developmental phase. Biochem Biophys Res Commun 216:273–278
Kuspa A, Dingermann T, Nellen W (1995) Analysis of gene function in *Dictyostelium*. Experientia 51:1116–1123
Lau NC et al (2006) Characterization of the piRNA complex from rat testes. Science 313:363–367
Lee RC, Ambros V (2001) An extensive class of small RNAs in *Caenorhabditis elegans*. Science 294:862–864
Lee SR, Collins K (2006) Two classes of endogenous small RNAs in *Tetrahymena thermophila*. Genes Dev 20:28–33
Lee RC, Feinbaum RL, Ambros V (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75:843–854
Lee CH et al (2005) *Dictyostelium* CBP3 associates with actin cytoskeleton and is related to slug migration. Biochim Biophys Acta 1743:281–290

Lee RC, Hammell CM, Ambros V (2006) Interacting endogenous and exogenous RNAi pathways in *Caenorhabditis elegans*. RNA 12:589–597

Lerner MR, Steitz JA (1979) Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. Proc Natl Acad Sci USA 76:5495–5499

Lerner MR, Steitz JA (1981) Snurps and scyrps. Cell 25:298–300

Lerner MR, Boyle JA, Mount SM, Wolin SL, Steitz JA (1980) Are snRNPs involved in splicing? Nature 283:220–224

Li HW, Ding SW (2005) Antiviral silencing in animals. FEBS Lett 579:5965–5973

Liang XH, Xu YX, Michaeli S (2002) The spliced leader-associated RNA is a trypanosome-specific sn(o) RNA that has the potential to guide pseudouridine formation on the SL RNA. RNA 8:237–246

Liu J et al (2004) Argonaute2 is the catalytic engine of mammalian RNAi. Science 305:1437–1441

Louis JM, Saxe CL 3rd, Kimmel AR (1993) Two transmembrane signaling mechanisms control expression of the cAMP receptor gene CAR1 during *Dictyostelium* development. Proc Natl Acad Sci USA 90:5969–5973

Lowe TM, Eddy SR (1999) A computational screen for methylation guide snoRNAs in yeast. Science 283:1168–1171

Lung B et al (2006) Identification of small non-coding RNAs from mitochondria and chloroplasts. Nucleic Acids Res 34:3842–3852

Lygerou Z, Mitchell P, Petfalski E, Seraphin B, Tollervey D (1994) The POP1 gene encodes a protein component common to the RNase MRP and RNase P ribonucleoproteins. Genes Dev 8:1423–1433

Lygerou Z, Allmang C, Tollervey D, Seraphin B (1996) Accurate processing of a eukaryotic precursor ribosomal RNA by ribonuclease MRP in vitro. Science 272:268–270

Maeda M et al (2003) Changing patterns of gene expression in *Dictyostelium* prestalk cell subtypes recognized by in situ hybridization with genes from microarray analyses. Eukaryot Cell 2:627–637

Mallory AC, Vaucheret H (2006) Functions of microRNAs and related small RNAs in plants. Nat Genet 38(Suppl):S31–36

Marker C et al (2002) Experimental RNomics. Identification of 140 Candidates for Small Non- Messenger RNAs in the Plant *Arabidopsis thaliana*. Curr Biol 12:2002–2013

Marquez SM et al (2005) Structural implications of novel diversity in eucaryal RNase P RNA. RNA 11:739–751

Martens H, Novotny J, Oberstrass J, Steck TL, Postlethwait P, Nellen W (2002) RNAi in *Dictyostelium*: the role of RNA-directed RNA polymerases and double-stranded RNase. Mol Biol Cell 13:445–453

Maruo T et al (2004) Control of cell type proportioning in *Dictyostelium discoideum* by differentiation-inducing factor as determined by in situ hybridization. Eukaryot Cell 3:1241–1248

Mattick JS, Makunin IV (2005) Small regulatory RNAs in mammals. Hum Mol Genet 14(1):R121–R132

Mattick JS, Makunin IV (2006) Non-coding RNA. Hum Mol Genet 15(1):R17–R29

Matzke MA, Birchler JA (2005) RNAi-mediated pathways in the nucleus. Nat Rev Genet 6:24–35

Mehler MF, Mattick JS (2006) Non-coding RNAs in the nervous system. J Physiol 575:333–341

Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G, Tuschl T (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol Cell 15:185–197

Mochizuki K, Gorovsky MA (2004) Small RNAs in genome rearrangement in *Tetrahymena*. Curr Opin Genet Dev 14:181–187

Mosig A, Sameith K, Stadler P (2006) Fragrep: an efficient search tool for fragmented patterns in genomic sequences. Genomics Proteomics Bioinformatics 4:56–60

Moss S, Morgan R (2004) The annexins. Genome Biol 5:219

Nagai K, Oubridge C, Kuglstatter A, Menichelli E, Isel C, Jovine L (2003) Structure, function and evolution of the signal recognition particle. EMBO J 22:3479–3485

Nair V, Zavolan M (2006) Virus-encoded microRNAs: novel regulators of gene expression. Trends Microbiol 14:169–175

Nakamura T, Prestayko AW, Busch H (1968) Studies on nucleolar 4 to 6 S ribonucleic acid of Novikoff hepatoma cells. J Biol Chem 243:1368–1375

Nazar RN (2004) Ribosomal RNA processing and ribosome biogenesis in eukaryotes. IUBMB Life 56:457–465

Nellen W et al (1992) Mechanisms of gene regulation by endogenous and artificially introduced antisense RNA. Biochem Soc Trans 20:750–754

Ni J, Tien AL, Fournier MJ (1997) Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. Cell 89:565–573

Nicoloso M, Qu LH, Michot B, Bachellerie JP (1996) Intron-encoded, antisense small nucleolar RNAs: the characterization of nine novel species points to their direct role as guides for the 2′-O-ribose methylation of rRNAs. J Mol Biol 260:178–195

Okafuji T, Abe F, Maeda Y (1997) Antisense-mediated regulation of Annexin VII gene expression during the transition from growth to differentiation in *Dictyostelium discoideum*. Gene 189:49

Okamura K, Ishizuka A, Siomi H, Siomi MC (2004) Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. Genes Dev 18:1655–1666

Omer AD, Ziesche S, Decatur WA, Fournier MJ, Dennis PP (2003) RNA-modifying machines in archaea. Mol Microbiol 48:617–629

Onodera Y, Haag JR, Ream T, Nunes PC, Pontes O, Pikaard CS (2005) Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. Cell 120:613

Orum H, Nielsen H, Engberg J (1992) Structural organization of the genes encoding the small nuclear RNAs U1 to U6 of *Tetrahymena thermophila* is very similar to that of plant small nuclear RNA genes. J Mol Biol 227:114–121

Pannucci JA, Haas ES, Hall TA, Harris JK, Brown JW (1999) RNase P RNAs from some Archaea are catalytically active. Proc Natl Acad Sci USA 96:7803–7808

Parker JS, Barford D (2006) Argonaute: a scaffold for the function of short regulatory RNAs. Trends Biochem Sci 31:622–630

Paule MR, White RJ (2000) Survey and summary transcription by RNA polymerases I and III. Nucleic Acids Res 28:1283–1298

Pi M, Morio T, Urushihara H, Tanaka Y (1998) Characterization of a novel small RNA encoded by *Dictyostelium discoideum* mitochondrial DNA. Mol Gen Genet 257:124–131

Piccinelli P, Rosenblad MA, Samuelsson T (2005) Identification and analysis of ribonuclease P and MRP RNA in a broad range of eukaryotes. Nucleic Acids Res 33:4485–4495

Plath K, Mlynarczyk-Evans S, Nusinow DA, Panning B (2002) Xist RNA and the mechanism of X chromosome inactivation. Ann Rev Genet 36:233–278

Popova B, Kuhlmann M, Hinas A, Söderbom F, Nellen W (2006) HelF, a putative RNA helicase acts as a nuclear suppressor of RNAi but not antisense mediated gene silencing. Nucleic Acids Res 34:773–784

Prestayko AW, Tonato M, Busch H (1970) Low molecular weight RNA associated with 28 s nucleolar RNA. J Mol Biol 47:505–515

Reinders Y, Schulz I, Gräf R, Sickmann A (2006) Identification of novel Proteins in *Dictyostelium discoideum* by comparative proteomic approaches. J Proteome Res 5:589–598

Ridanpää M, Sistonen P, Rockas S, Rimoin DL, Makitie O, Kaitila I (2002) Worldwide mutation spectrum in cartilage-hair hypoplasia: ancient founder origin of the major70A → G mutation of the untranslated RMRP. Eur J Hum Genet 10:439–447

Ridanpää M et al (2001) Mutations in the RNA component of RNase MRP cause a pleiotropic human disease, cartilage-hair hypoplasia. Cell 104:195–203

Rogers J, Wall R (1980) A mechanism for RNA splicing. Proc Natl Acad Sci USA 77:1877–1879

Romby P, Vandenesch F, Wagner EGH (2006) The role of RNAs in the regulation of virulence-gene expression. Curr Opin Microbiol 9:229

Rosenblad MA, Zwieb C, Samuelsson T (2004) Identification and comparative analysis of components from the signal recognition particle in protozoa and fungi. BMC Genomics 5:5

Saito K et al (2006) Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. Genes Dev 20:2214–2222

Schattner P, Decatur WA, Davis CA, Ares M Jr, Fournier MJ, Lowe TM (2004) Genome-wide searching for pseudouridylation guide snoRNAs: analysis of the *Saccharomyces cerevisiae* genome. Nucleic Acids Res 32:4281–4296

Schmitt ME, Clayton DA (1993) Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in *Saccharomyces cerevisiae*. Mol Cell Biol 13:7935–7941

Schramm L, Hernandez N (2002) Recruitment of RNA polymerase III to its target promoters. Genes Dev 16:2593–2620

Simmer F et al (2002) Loss of the putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi. Curr Biol 12:1317–1319

Singh SK, Gurha P, Tran EJ, Maxwell ES, Gupta R (2004) Sequential 2′-*O*-methylation of archaeal pre-tRNATrp nucleotides is guided by the intron-encoded but trans-acting box C/D ribonucleoprotein of pre-tRNA. J Biol Chem 279:47661–47671

Slomovic S, Laufer D, Geiger D, Schuster G (2006) Polyadenylation of ribosomal RNA in human cells. Nucleic Acids Res 34:2966–2975

Smale ST, Kadonaga JT (2003) The RNA polymerase II core promoter. Ann Rev Biochem 72:449–479

Söderbom F (2006) Small nucleolar RNAs: identification, structure, and function. In: Nellen W, Hammann C (eds) Small RNAs: analysis and regulatory functions. Springer Berlin Heidelberg New York 31–55

Staley JP, Guthrie C (1998) Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell 92:315–326

Stathopoulos C, Kalpaxis DL, Drainas D (1995) Partial purification and characterization of RNase P from *Dictyostelium discoideum*. Eur J Biochem 228:976–980

Stein LD et al (2003) The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. PLoS Biol 1:E45

Steinert M, Heuner K (2005) *Dictyostelium* as host model for pathogenesis. Cell Microbiol 7:307–314

Stolc V et al (2004) A gene expression map for the euchromatic genome of *Drosophila melanogaster*. Science 306:655–660

Stolc V et al (2005) Identification of transcribed sequences in *Arabidopsis thaliana* by using high-resolution genome tiling arrays. Proc Natl Acad Sci USA 102:4453–4458

Storz G, Altuvia S, Wassarman KM (2005) An abundance of RNA regulators. Ann Rev Biochem 74:199–217

Stuart KD, Schnaufer A, Ernst NL, Panigrahi AK (2005) Complex management: RNA editing in trypanosomes. Trends Biochem Sci 30:97–105

Szafranski K, Lehmann R, Parra G, Guigo R, Glöckner G (2005) Gene organization features in A/T-rich organisms. J Mol Evol 60:90–98

Takeishi K, Kaneda S (1979) Low molecular weight nuclear RNA species in *Dictyostelium discoideum*. Nucleic Acids Symp Ser s125–127

Takeishi K, Kaneda S (1981) Isolation and characterization of small nuclear RNAs from *Dictyostelium discoideum*. J Biochem. 90:299–308

Tang TH et al (2002) Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. Proc Natl Acad Sci USA 99:7536–7541

The Arabidopsis Genome I (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796

Thiel CT et al (2005) Severely incapacitating mutations in patients with extreme short stature identify RNA-processing endoribonuclease RMRP as an essential cell growth regulator. Am J Hum Genet 77:795–806

Tycowski KT, Steitz JA (2001) Non-coding snoRNA host genes in *Drosophila*: expression strategies for modification guide snoRNAs. Eur J Cell Biol 80:119–125

Tycowski KT, Smith CM, Shu MD, Steitz JA (1996) A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in *Xenopus*. Proc Natl Acad Sci USA 93:14480–14485

Tycowski KT, You ZH, Graham PJ, Steitz JA (1998) Modification of U6 spliceosomal RNA is guided by other small RNAs. Mol Cell 2:629–638

Uliel S, Liang XH, Unger R, Michaeli S (2004) Small nucleolar RNAs that guide modification in trypanosomatids: repertoire, targets, genome organisation, and unique functions. Int J Parasitol 34:445–454

Upadhyay R, Bawankar P, Malhotra D, Patankar S (2005) A screen for conserved sequences with biased base composition identifies noncoding RNAs in the A-T rich genome of *Plasmodium falciparum*. Mol Biochem Parasitol 144:149–158

Valadkhan S (2005) snRNAs as the catalysts of pre-mRNA splicing. Curr Opin Chem Biol 9:603

Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev 20:515–524

Vankan P, Edoh D, Filipowicz W (1988) Structure and expression of the U5 snRNA gene of *Arabidopsis thaliana*. Conserved upstream sequence elements in plant U-RNA genes. Nucleic Acids Res 16:10425–10440

Vaucheret H (2005) RNA polymerase IV and transcriptional silencing. Nat Genet 37:659–660

Vitali P, Royo H, Seitz H, Bachellerie J-P, Hüttenhofer A, Cavaillé J (2003) Identification of 13 novel human modification guide RNAs. Nucleic Acids Res 31:6543–6551

Vitali P et al (2005) ADAR2-mediated editing of RNA substrates in the nucleolus is inhibited by C/D small nucleolar RNAs. J Cell Biol 169:745–753

Voinnet O (2005) Induction and suppression of RNA silencing: insights from viral infections. Nat Rev Genet 6:206–220

Waibel F, Filipowicz W (1990) RNA-polymerase specificity of transcription of *Arabidopsis* U snRNA genes determined by promoter element spacing. Nature 346:199

Wassenegger M (2005) The role of the RNAi machinery in heterochromatin formation. Cell 122:13–16

Wassenegger M, Krczal G (2006) Nomenclature and functions of RNA-directed RNA polymerases. Trends Plant Sci 11:142–151

Watanabe T et al (2006) Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev 20:1732–1743

Weinberg RA, Penman S (1968) Small molecular weight monodisperse nuclear RNA. J Mol Biol 38:289–304

Werner A, Berdal A (2005) Natural antisense transcripts: sound or silence? Physiol Genomics 23:125–131

Will CL, Lührmann R (2001) Spliceosomal UsnRNP biogenesis, structure and function. Curr Opin Cell Biol 13:290–301

Will CL, Lührmann R (2005) Splicing of a rare class of introns by the U12-dependent spliceosome. Biol Chem 386:713–724

Williams JG, Noegel AA, Eichinger L (2005) Manifestations of multicellularity: Dictyostelium reports in. Trends Genet 21:392–398

Wise JA, Weiner AM (1980) *Dictyostelium* small nuclear RNA D2 is homologous to rat nucleolar RNA U3 and is encoded by a dispersed multigene family. Cell 22:109–118

Wise JA, Weiner AM (1981) The small nuclear RNAs of the cellular slime mold *Dictyostelium discoideum*. J Biol Chem 256:956–963

Yang C-Y, Zhou H, Luo J, Qu L-H (2005) Identification of 20 snoRNA-like RNAs from the primitive eukaryote, *Giardia lamblia*. Biochem Biophys Res Commun 328:1224–1231

Yoshida H, Yamada Y, Okamoto K (1991) DC6, a novel type of *Dictyostelium discoideum* gene regulated by secreted factors but not by cAMP. Differentiation 46:161–166

Yoshida H, Kumimoto H, Okamoto K (1994) dutA RNA functions as an untranslatable RNA in the development of *Dictyostelium discoideum*. Nucleic Acids Res 22:41–46

Yu Y-T, Scharl EC, Smith CM, Steitz JA (1999) The growing World of small nuclear ribonucleoproteins. In: Gesteland RF, Cech TR, Atkins JF (eds) The RNA World, 2nd edn. Cold Spring Harbor Laboratory Press, New York, pp 487–524

Yuan G, Klambt C, Bachellerie JP, Brosius J, Hüttenhofer A (2003) RNomics in *Drosophila melanogaster*: identification of 66 candidates for novel non-messenger RNAs. Nucleic Acids Res 31:2495–2507

Yukawa Y et al (2005) Plant 7SL RNA genes belong to type 4 of RNA polymerase III- dependent genes that are composed of mixed promoters. Plant J 43:97–106

Zamore PD, Haley B (2005) Ribo-gnome: the big world of small RNAs. Science 309:1519–1524

Zamore PD, Tuschl T, Sharp PA, Bartel DP (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell 101:25–33

Zemann A, op de Bekke A, Kiefmann M, Brosius J, Schmitz J (2006) Evolution of small nucleolar RNAs in nematodes. Nucl. Acids Res 34:2676–2685

Zieve GW (1981) Two groups of small stable RNAs. Cell 25:296