The small RNA repertoire of *Dictyostelium discoideum* and its regulation by components of the RNAi pathway

Andrea Hinas[1], Johan Reimegård[2], E. Gerhart H. Wagner[2], Wolfgang Nellen[3], Victor R. Ambros[4] and Fredrik Söderbom[1,*]

[1]Department of Molecular Biology, Biomedical Center, Swedish University of Agricultural Sciences, Box 590, SE-75124 Uppsala, Sweden, [2]Department of Cell and Molecular Biology, Biomedical Center, Uppsala University, Box 596, SE-75124 Uppsala, Sweden, [3]Department of Genetics, Kassel University, Heinrich Plett Strasse 40, 34132 Kassel, Germany and [4]Department of Genetics, Dartmouth Medical School, Hanover, NH 03755, USA

ABSTRACT

Small RNAs play crucial roles in regulation of gene expression in many eukaryotes. Here, we report the cloning and characterization of 18–26 nt RNAs in the social amoeba *Dictyostelium discoideum*. This survey uncovered developmentally regulated microRNA candidates whose biogenesis, at least in one case, is dependent on a Dicer homolog, DrnB. Furthermore, we identified a large number of 21 nt RNAs originating from the DIRS-1 retrotransposon, clusters of which have been suggested to constitute centromeres. Small RNAs from another retrotransposon, Skipper, were significantly up-regulated in strains depleted of the second Dicer-like protein, DrnA, and a putative RNA-dependent RNA polymerase, RrpC. In contrast, the expression of DIRS-1 small RNAs was not altered in any of the analyzed strains. This suggests the presence of multiple RNAi pathways in *D. discoideum*. In addition, we isolated several small RNAs with antisense complementarity to mRNAs. Three of these mRNAs are developmentally regulated. Interestingly, all three corresponding genes express longer antisense RNAs from which the small RNAs may originate. In at least one case, the longer antisense RNA is complementary to the spliced but not the unspliced pre-mRNA, indicating synthesis by an RNA-dependent RNA polymerase.

INTRODUCTION

Since their initial discovery in worms, 18–26 nt small RNAs have now been identified in many eukaryotes (1,2).

By antisense complementarity, they confer specificity to associated protein complexes and can thereby regulate expression of their target genes at the level of transcription, mRNA stability or translation (3–5). These small RNAs can be divided into two main classes; microRNAs (miRNAs), which are processed from imperfectly base-paired hairpin transcripts, and small interfering RNAs (siRNAs), which are derived from long double-stranded RNAs (dsRNAs) (5–7). The dsRNA precursors of siRNAs originate e.g. from viruses, repetitive elements or are synthesized by an RNA-dependent RNA polymerase (RdRP) using a single-stranded RNA as a template (2,8–10). Despite their different origin, miRNA and siRNA pathways share many similarities. In both cases, these small RNAs are processed from precursors into mature small RNAs by the RNase III-type Dicer proteins, and are subsequently incorporated in an effector complex that contains an Argonaute–Piwi family protein (5,6). The miRNA/siRNA guides the protein complex to its complementary target RNA and induces cleavage of the target if the small RNA and target RNA form a perfectly or close to perfectly base-paired duplex. In plants, most miRNAs exert their effect in this way. If base pairing is only partial, as is the common feature of animal miRNA–target interactions, the main effect seems to be inhibition of translation, although some mRNA degradation is also frequently observed (4).

miRNAs are common in multicellular organisms and animal viruses and were only recently discovered in a unicellular organism, the green alga *Chlamydomonas reinhardtii* (6,11–13). In animals and plants, miRNAs play important cellular roles by modulating the expression of endogenous genes and it has been estimated that up to one-third of all human genes may be regulated by miRNAs (5,14). The physiological role of miRNAs in *C. reinhardtii*

*To whom correspondence should be addressed. Tel: +46 18 471 4901; Fax: +46 18 536971; Email: fredde@xray.bmc.uu.se
Present address:
Andrea Hinas, Department of Molecular and Cellular Biology, Harvard University, 16 Divinity Ave, Cambridge, MA 02138, USA.

is still not known, however, verified targets include genes encoding flagellum-associated proteins (11,13).

siRNAs are present in eukaryotes from all major phylogenetic branches including plants, animals and fungi where they act e.g. in a defense mechanism against viral RNAs and mobilization of transposons (2,15). Repetitive elements, such as transposons and retrotransposons, constitute substantial parts of the centromeres in many eukaryotes, and the RNAi machinery is required for silencing of centromeric repeats in e.g. fission yeast (16,17). Although most studies on natural siRNAs have focused on their roles in defense against viruses and repetitive elements, an increasing number of reports point to an additional role in regulation of non-transposon genes. For example, large-scale cloning of small RNAs from *Caenorhabditis elegans* and *Arabidopsis thaliana* has identified many small RNAs with antisense complementarity to genes other than repetitive elements (18–20). Recently, such small RNAs were demonstrated to be involved in regulation of overlapping *A. thaliana* genes during salt stress and bacterial infection (20,21). Furthermore, whole-genome microarray analyses and expressed sequence tags (ESTs) indicate substantial expression of longer RNAs with antisense complementarity to mRNAs in organisms ranging from protozoa to mammals (22,23). Also, many mRNAs are transcribed such that they may overlap with other mRNAs, thus potentially forming dsRNAs.

The genetically tractable social amoeba *Dictyostelium discoideum* has proven a valuable model organism in many different research areas, from cell differentiation to host–pathogen interactions (24,25). The recently sequenced 34 Mb genome has been annotated and revealed ~12 500 genes (26). Under normal conditions, *D. discoideum* grows and divides as single cells, but when challenged by starvation, cells stream together to form aggregates containing ~100 000 cells. Together, the cells go through development as a multicellular organism, ultimately forming a ball of spores on top of a stalk (24). The intermediate evolutionary position of *D. discoideum*, branching after plants but before the division between animals and fungi, further contributes to its significance as a model organism (26).

Transgenic RNA interference (RNAi) is a standard method to knock down gene expression in *D. discoideum* (27–29). Several genes predicted to encode homologs to RNAi machinery proteins are present in the genome, including those encoding two Dicer-like proteins, *drnA* and *drnB*, and three RdRPs, *rrpA*, *rrpB* and *rrpC* (29,30). Of these, only *rrpA* is required for transgenic RNAi (29). In addition, deletion of the gene encoding a putative RNA helicase, *helF*, renders cells hypersensitive to RNAi (31). Five genes (and one apparently truncated gene, our unpublished data) predicted to encode Piwi-like proteins are present in the genome (15).

Here we report the first large-scale isolation of small RNAs (18–26 nt) from the unicellular amoeba *D. discoideum*, and the involvement of RNAi-machinery homologs in endogenous small RNA biogenesis, indicating multiple RNAi pathways. We found developmentally regulated miRNA candidates of which at least one is dependent on a Dicer homolog for its biogenesis,

and isolated developmentally regulated RNAs derived from retrotransposons, one of which has been suggested to constitute centromeres. In addition, a number of small RNAs antisense to mRNAs were identified, possibly generated from longer antisense RNAs involving RdRP-directed synthesis. These results establish *D. discoideum* as a model for functional analysis of endogenous small RNAs and allows for simultaneous studies of their role in single cell growth and multicellular development.

MATERIALS AND METHODS

Accession numbers

Isolated small RNA sequences in this study have been deposited in Gene Expression Omnibus: platform GPL5734, and in GenBank: EU113320–EU115992. The micro RNA candidates mica1190 and mica1198 are named ddi-mir-1176 and ddi-mir-1177, respectively, in the miRBase database (41).

Growth and development

All *D. discoideum* strains were grown axenically in HL5 medium and synchronously developed on nitrocellulose membranes (32).

Oligonucleotides

DNA oligonucleotides (Invitrogen) used in this study are listed in the Supplementary Data (Table S1).

Cloning of small RNAs

cDNA libraries of 18–26 nt *D. discoideum* RNAs were constructed according to two different protocols (33,34). Briefly, total RNA was isolated by the TRIzol method (Invitrogen) from growing *D. discoideum* AX4 strain cells as well as from cells developed for 16 and 24 h, and the fractions were subsequently pooled. After size fractionation and ligation of a 3′ linker, the RNA was divided into two fractions, one of which was directly ligated to a 5′ linker, thus selecting for small RNAs with 5′ monophosphates (34). Following RT-PCR, the PCR fragments were cloned and sequenced. The second fraction was subjected to reverse transcription directly after 3′ ligation (33). Apart from synthesizing a complementary DNA strand, the reverse transcriptase adds a few non-templated C residues at the 3′ end. These C:s were then hybridized to a DNA oligo with three 3′ G:s followed by PCR, cloning and sequencing. This approach is insensitive of the nature of the 5′ end of the small RNA.

Computational analysis

The genome sequence (v 2.5) and annotation of mRNAs (dated 2006-01-04) in *D. discoideum* was downloaded from dictybase. We divided the ~34.2 Mb genome into open reading frames (ORFs), intergenic regions, repeats and genes for non-coding RNAs, roughly sized 21.7 Mb, 9.2 Mb, 3.3 Mb and 43 kb, respectively. Annotation of complex repeats and non-coding RNAs was carried out as described elsewhere (Larsson *et al.*, manuscript in preparation). BLASTN (NCBI BLAST 2.2.14) was used

to determine the location of the cDNA sequences in the genome sequence of *D. discoideum*. The cutoff for BLASTN was empirically set to of 5.0×10^{-3} in order to distinguish hits that had a length of at least 18 nt. This approach also allowed us to find cDNAs that matched the genome sequence perfectly as well as those that were either post-transcriptionally modified, missequenced, or had cloning artifacts at the ends of the cDNA. Of the final sequences that matched the criteria stated above a minor fraction, 74 sequences (4%) in the DYNL library and 92 sequences (10%) in the DYSM library, had one mismatch within the aligned region. Importantly, none of the in detail analyzed small RNAs in this report had any mismatches compared to the genomic sequence.

In cases where one cDNA gave multiple hits, all were kept for further analysis. Subsequently, the cDNAs were annotated as part of a complex repeat, non-coding RNA, coding RNA, or intergenic region. cDNAs that matched protein-coding RNAs were further subdivided as located in the untranslated regions (UTRs), exons, introns or in antisense orientation to the mRNA. miRNA candidates were predicted using cDNAs that matched intergenic, intron or antisense (to mRNAs) regions, followed by extraction of the genomic sequences 150 nt upstream and downstream of each cDNA. The extracted sequences were analyzed by MiRfold (35) with standard settings to identify putative pre-miRNA structures. The predicted hairpin structures were further analyzed by miR-*abela* (36). All remaining sequences that had a score above 0.0 were regarded as candidate miRNAs. Secondary RNA structures were predicted using the Mfold software v.2.3 at folding temperature 22°C (37).

RT-PCR ANALYSIS

For RT-PCR analysis, *D. discoideum* AX4 strain RNA pooled from growing cells and cells developed for 16 and 24 h (equal amounts) was DNase-treated twice. Next, 2 μg of RNA was reverse transcribed at 55°C, 60°C and 65°C (20 min at each temperature) using 15 U of Thermoscript reverse transcriptase (Invitrogen) and gene-specific primers. Primers used for reverse transcription of asRNA and mRNA, respectively, were oligos 241 and 242 (*rsmF*), 243 and 245 (*hatA*), 274 and 275 (*hatB*), 276 and 277 (*hatC*) and 246 and 248 (*DDB0230011*). Negative controls were treated identically except that no reverse transcriptase was added to these reactions. After RNase H treatment (Invitrogen), PCR was carried out using 0.5 μl reverse transcription mix, 20 pmol of gene-specific primers and 0.6 U AmpliTaq Gold polymerase (Roche). The same primers were used for amplification of cDNA corresponding to both mRNA and asRNA of each gene. Primers for each gene were the same as in the reverse transcription, except for DDB0230011 mRNA and asRNA, which were amplified in a nested PCR reaction using oligos 247 and 251. Cycling conditions were 95°C for 9 min; 5 cycles of 95°C for 30 s, 55–50°C (*hatA*), 60–55°C (*hatB, hatC*), 62–57°C (*rsmF*) or 59–54°C (*DDB0230011*) for 40 s, 72°C for 40 s; 35 cycles of 95°C for 30 s, 50°C (*hatA*), 55°C (*hatB, hatC*), 57°C (*rsmF*) or 54°C (*DDB0230011*) for 40 s, 72°C for 40s; followed by 72° for 10 min. For *hatA*,

genomic PCR was carried out using oligos 243 and 245. Due to the high A/T content of the *hatA* intron, cycling conditions were 95°C for 5 min; 35 cycles of 95°C for 30 s, 47°C for 30 s, 60°C for 2 min; followed by 60° for 7 min. The PCR products were cloned into the pCR2.1-TOPO vector (Invitrogen) and sequenced using an ABI PRISM 3700 DNA Analyzer (Applied Biosystems).

Northern blot analysis

For northern blots of small RNAs, total RNA was extracted by the TRIzol method and 20 μg RNA was separated on a 12% polyacrylamide/7 M urea/1 × TBE gel and electroblotted to a Hybond-N$^+$ membrane (GE Healthcare) at 20 V at 4°C for 16 h. The RNA was immobilized by UV cross-linking (150 mJ) and radioactively 5′ labeled DNA oligonucleotides were hybridized in Church buffer (7% SDS, 0.5 M NaPO$_4$ pH 7.2, 1 mM EDTA and 1% BSA) at 42°C overnight. The membranes were washed at 42°C; 2×5 min with $2 \times$ SSC/0.1% SDS, 2×10 min with $1 \times$ SSC/0.1% SDS, 2×5 min with $0.5 \times$ SSC/0.1% SDS and hybridization signals were analyzed using a Phosphorimager (Molecular Dynamics). Northern blot of mRNAs and antisense RNAs was performed essentially as described above, but using a 6% polyacrylamide/7 M urea/1 × TBE gel. Probes were *in vitro* transcribed from purified PCR products with ^{32}P-labeled UTP using T7 polymerase (Ambion). Templates for *in vitro* transcription were PCR-amplified from pCR2.1-TOPO plasmids (see above) containing the *rsmF*, *hatA* or *DDB0230011* sequences by using primers 257 or 258 in combination with gene-specific primers 241 or 242 (*rsmF*), 244 or 252 (*hatA*), and 247 or 251 (*DDB0230011*). Hybridizations were carried out in Church buffer at 65°C overnight. The membranes were washed under the same conditions as above, but at 65°C. The SRP RNA probe was a 5′-labeled DNA oligonucleotide (38). The DNA size marker, pUC Mix Marker 8 (Fermentas) and the RNA decade marker (Ambion) were 5′-labeled with ^{32}P.

RESULTS

Two cDNA libraries representing 18–26 nt *D. discoideum* RNAs

The aim of this study was to investigate the small RNA (18–26 nt) profile of *D. discoideum* during growth and development. For this reason, we cloned and sequenced pooled small RNAs from growing single cells and from two different multicellular stages (16 and 24 h of development). At 16 h of development, cells have differentiated into two main cell types, pre-spore and pre-stalk cells, and after 24 h development has reached the final stage; a ball of spores on top of a stalk (24).

The specific characteristics of different small RNA classes determine their accessibility to different cloning methods (2). One important feature is the 5′ end of the RNA, e.g. a mono- or a triphosphate as a result of different biogenesis pathways, or a hydroxyl group in the case of hydrolyzed degradation products. Therefore, we used two different methods for the addition of the 5′ linker
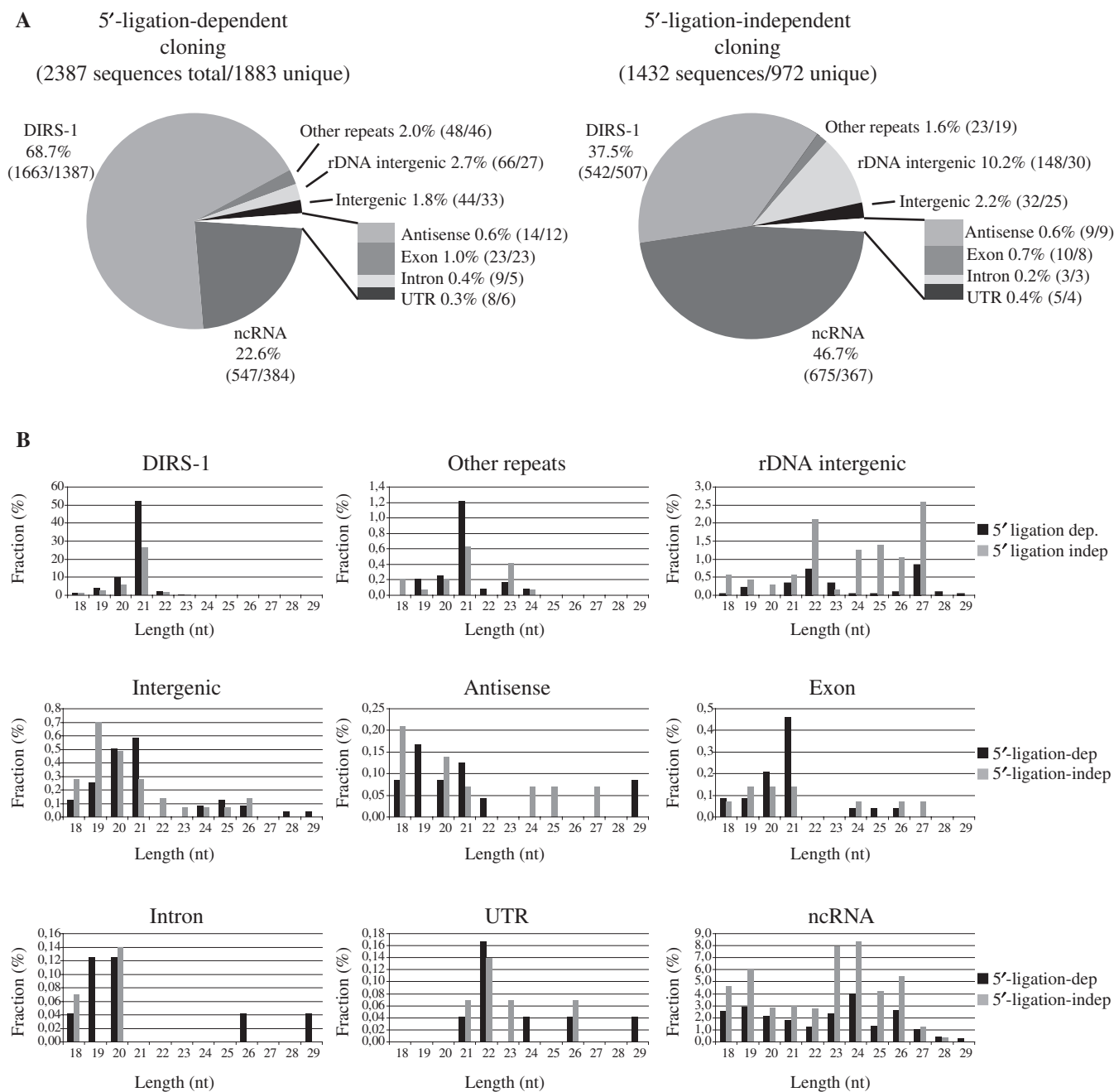
A

5′-ligation-dependent cloning
(2387 sequences total/1883 unique)

5′-ligation-independent cloning
(1432 sequences/972 unique)

B

Figure 1. Two cDNA libraries representing 18–26 nt RNAs from *D. discoideum*. (**A**) Classification of small RNAs represented in the cDNA libraries constructed by 5′-ligation-dependent cloning (left) and 5′-ligation-independent cloning (right). Numbers of sequences (total as well as unique) are indicated in brackets. (**B**) Size distribution of small RNAs represented in the cDNA libraries. Sequences (percentage of each library) from the 5′-ligation-dependent and 5′-ligation-independent libraries are represented by black and gray bars, respectively.

prior to cloning (see Materials and Methods section). In the first approach, referred to as the 5′-ligation-dependent method, an RNA oligo was ligated to the 5′ end of the size-fractionated RNA without prior treatment of the RNA. This selects for RNAs with 5′ monophosphates, a characteristic feature of Dicer cleavage products (34). In the second approach, the 5′-ligation-independent method, we took advantage of the ability of the reverse transcriptase to add a few untemplated terminal C residues. This allows for a hybridization-based

addition of the 5′ linker and thereby isolation of RNAs regardless of the nature of their 5′ ends (33).

Of the 5009 cloned sequences from the 5′-ligation-dependent cDNA library, 48% (2387 sequences) matched the published *D. discoideum* genome sequence (26) whereas the corresponding number for the 5′-ligation-independent library was 36% (1432 sequences) of a total of 4002 sequenced clones. The relative representation of different RNA classes in the two libraries is shown in Figure 1A, along with the numbers of unique sequences

identified. Most clones represented small RNAs derived from the DIRS-1 retrotransposon (68.7% and 37.5% in the 5′-ligation-dependent and 5′-ligation-independent libraries, respectively). Another large fraction was derived from putative degradation products of rRNAs, tRNAs, small nucleolar RNAs, small nuclear RNAs and other non-coding RNAs (22.6% and 46.7%). A substantial portion (2.7% and 10.2%) matched the intergenic region between the 26S rRNA and 5S rRNA genes, which in *D. discoideum* is present on the 88 kb extrachromosomal rDNA palindrome (39). Out of these clones, 90% were derived from a 27 nt region situated ∼450 bp downstream and ∼1500 nt upstream of the 26S rRNA and 5S rRNA genes, respectively. This region has previously been referred to as non-transcribed (39). Northern blot analysis of this region yielded several different hybridization signals of which none was shorter than ∼40 nt in length (data not shown).

Besides the DIRS-1 retrotransposon, other repetitive elements—retrotransposons as well as DNA transposons—were also represented, albeit at a low frequency (2.0% and 1.6% in the 5′-ligation-dependent and 5′-ligation-independent libraries, respectively). These include the Skipper retrotransposon, which is further described below. Other isolated small RNAs originated from intergenic regions (1.8% and 2.2%), exons (1.0% and 0.7%), introns (0.4% and 0.2%), UTRs of protein-coding genes (0.3% and 0.4%) and small RNAs with antisense complementarity to protein-coding genes (0.6% of each library). Further analysis of these small RNAs is described below.

The size distribution of the small RNAs represented in the libraries revealed a strong bias towards 21 nt RNAs. This was particularly evident for RNAs derived from DIRS-1 (Figure 1B). Notably, the frequency of DIRS-1 small RNAs was substantially higher in the 5′-ligation-dependent compared to the 5′-ligation-independent library. On the other hand, the reverse situation was seen for clones representing fragments of non-coding RNAs (ncRNAs), e.g. rRNAs and tRNAs. These sequences were much more abundant in the 5′-ligation-independent library than in the 5′-ligation-dependent library (Figure 1A). This indicates that the majority of the DIRS-1 small RNAs carry 5′ monophosphates, characteristic of Dicer products.

miRNA candidates and the expression of one candidate in different RNAi knockout strains

Until very recently, miRNAs had only been identified in multicellular organisms and their viruses (11,13). To explore the possibility that miRNAs were represented in our cDNA libraries, we computationally predicted secondary structures of RNA sequences flanking intergenic, intronic and antisense small RNA loci (see Materials and Methods section). Five out of 167 folded sequences fulfilled standard criteria for miRNA precursors; four were derived from intergenic regions, and one was in antisense orientation to an intron of a predicted gene. By northern blot analysis, expression of ∼22 nt RNAs was detected for two of the intergenic small

RNAs, mica1190 and mica1198. Furthermore, both are up-regulated during development (16 h slugs and 24 h fruiting bodies) (Figure 2A). The predicted precursors of the miRNA candidates are shown in Figure 2B. No signals corresponding to pre-miRNAs could be detected in the northern blot experiment; however, this is also a common situation in plants where miRNA precursors are rarely seen (40).

A third predicted candidate miRNA gave a strong hybridization signal of ∼60 nt, which also increased during development, but no signal corresponding to the potential mature miRNA was observed (data not shown). Notably, the corresponding clone was derived from the 5′-ligation-independent cDNA library, which has a relatively high proportion of ncRNA fragments. Thus, this small RNA may be a degradation product of a previously unknown longer ncRNA. The other four predicted miRNAs were found in the 5′-ligation-dependent library, which contains candidate Dicer cleavage products. In order to address possible conservation of the *D. discoideum* miRNA candidates, we searched for homologs of known miRNAs in the miRBase database (41). No miRNAs with significant sequence similarity were identified.

To investigate the biogenesis pathway of the miRNA candidates, the expression of one of them (mica1198) was analyzed in *D. discoideum* strains lacking various RNAi-related proteins. The disrupted genes encode two Dicer-like proteins (DrnA and DrnB), three RdRP homologs (RrpA, RrpB and RrpC), and a putative RNA helicase (HelF) (29–31). In most of the mutant strains, the expression of mica1198 was essentially unaffected when compared to wild-type cells (Figure 2C). However, in the strain lacking the Dicer-like protein DrnB, the miRNA candidate could not be detected. This suggests that DrnB is responsible for processing of mica1198 from its precursor. No hybridization signals from the putative precursor RNAs could be detected, which may indicate that the same Dicer is responsible for generating the pre-miRNA as is the case in plants (42). In contrast, expression of mica1198 was up-regulated in the strain lacking the RdRP RrpC. As a control, the same membrane was probed for a small DIRS-1-derived RNA, whose level was unchanged in all knockout strains (Figure 2C and see below).

Small RNAs derived from the DIRS-1 retrotransposon

DIRS-1 is the most abundant retrotransposon in the *D. discoideum* genome and DIRS-1-rich regions are found at one end of each chromosome (26). These regions cluster during mitosis and have therefore been suggested to function as centromeres. The complete DIRS-1 retrotransposon encodes a 4.5 kb mRNA and a 900 nt antisense RNA and is flanked by inverted long terminal repeats (LTRs) (43–45). We identified small RNAs along almost the entire retrotransposon, not only from the segments that would be predicted to form dsRNAs, e.g. mRNA-antisense RNA and LTRs (Figure 3A).

DIRS-1 (also referred to as Tdd-1) expression is up-regulated during development (44). Therefore, the developmental expression of a randomly chosen DIRS-1
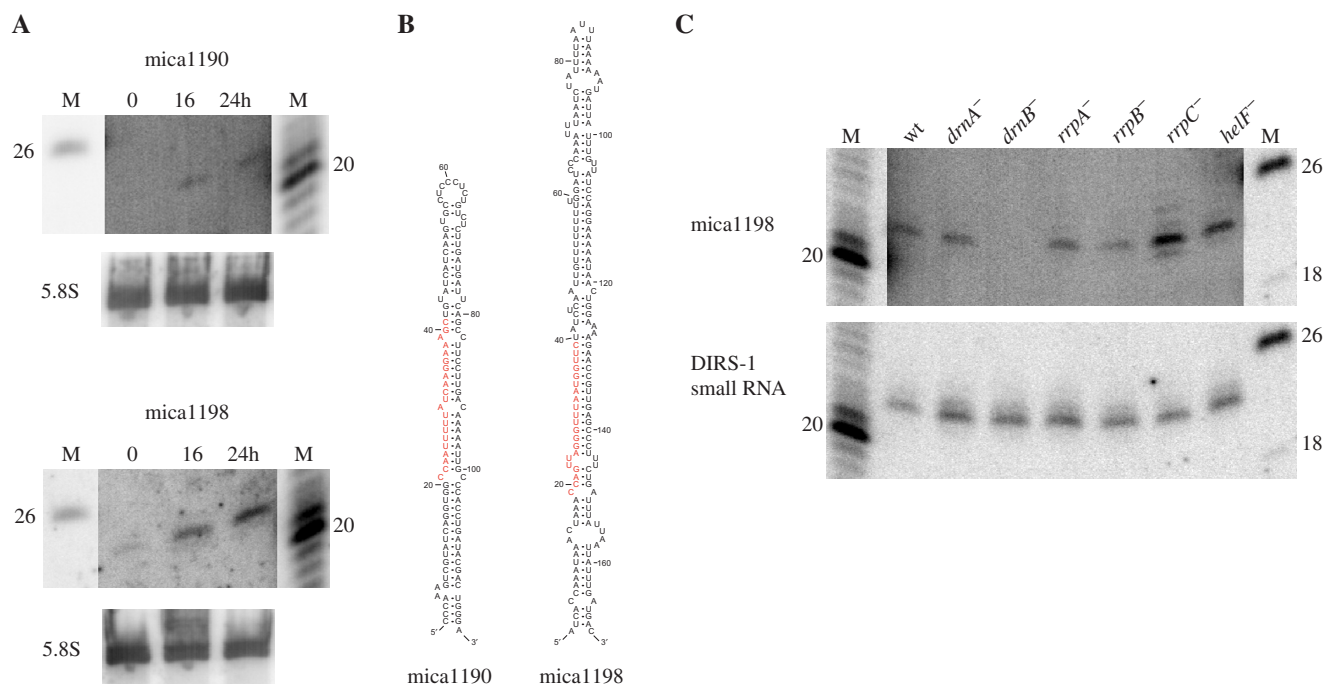
Figure 2. miRNA candidates in *D. discoideum* (**A**) northern blot analysis of mica1190 and mica1198 using RNA isolated from growing cells (0 h) and developed cells (16 and 24 h). The 5.8S rRNA is used as a loading control and radioactively labeled size markers (M) are shown. (**B**) Predicted secondary structure (Mfold) of sequences surrounding miRNA candidates mica1190 and mica1198. (**C**) Northern blot analysis showing the expression of mica1198 in strains where genes encoding putative Dicer homologs (*drnA⁻* and *drnB⁻*), RdRp (*rrpA⁻*, *rrpB⁻* and *rrpC⁻*) and RNA helicase (*helF⁻*) have been disrupted. Wild-type (wt) represents *D. discoideum* AX2 strain, which is the background strain of all knockout strains except *rrpC⁻*. *rrpC⁻* was constructed in the AX4 wild-type strain, which displayed similar expression levels as AX2 for all RNAs analyzed in this study (data not shown). The same membrane was stripped and reprobed with the DIRS-1 small RNA which is expressed approximately equally in all strains analyzed and was therefore used as a loading control.

small RNA was investigated by northern blot analysis. This RNA, identified in the 5′-ligation-dependent library, originates from the antisense strand immediately downstream of the reported 900 nt antisense RNA (Figure 3A). Similar to the DIRS-1 mRNA, the small DIRS-1 RNA is up-regulated at 16 and 24 h of development compared to the levels in growing cells (Figure 3B).

Small RNAs originating from the Skipper retrotransposon

Skipper is the second most abundant retrotransposon in the *D. discoideum* genome (26). In contrast to DIRS-1, Skipper is flanked by direct LTRs, and no antisense RNA has yet been reported to be transcribed from this retrotransposon (46). In spite of its rather high copy number in the genome, only eight small RNAs originating from this repetitive element were identified by cDNA cloning and all were represented in the ligation-dependent library only. Interestingly, they all matched a partial Skipper fragment on chromosome 2, although a few of them could also originate from a number of complete or partial Skipper sequences present at different locations in the genome. Except for a number of ESTs of 270–320 nt that include one of the isolated small RNAs (Figure 4A), no expression for this partial Skipper fragment has been demonstrated prior to this study. A genomic region with the potential to give rise to a long hairpin RNA, with a ~320 bp stem and ~900 nt loop, could be the source for all eight small RNAs (Figure 4A) Notably, this places two of

the isolated small RNAs on opposite strands in the same 21 bp region. However, six out of the eight small RNAs could potentially originate from an alternative stem-loop structure that resembles a miRNA precursor (Figure 4B). If so, these small RNAs would originate from the predicted loop sequence (three staggered sequences) and the 3′ part (three staggered sequences) of the hairpin. Although it is likely that all eight small RNAs are derived from one or several longer RNA molecules transcribed from the same DNA strand, two of the three small RNAs from the loop sequence in Figure 4B also match a second site at the opposite DNA strand elsewhere in the Skipper fragment.

The expression of the Skipper small RNAs originating from the predicted loop and 3′ stem structure during *D. discoideum* growth and development was analyzed by northern blot. Since the hybridization patterns for both sets of small RNAs were identical, only the result for one set is shown in Figure 4C. The sequence recognized by the oligonucleotide probe is indicated in Figure 4A and B. The result demonstrated up-regulation of the Skipper small RNA during development, much like the DIRS-1 small RNA (Figure 3B). However, in contrast to DIRS-1, the Skipper mRNA is expressed at low levels both during growth and development (30). It should be noted that the two oligonucleotides used most likely will hybridize to all three small RNAs from each region, respectively. In addition, northern blot also verified expression of the
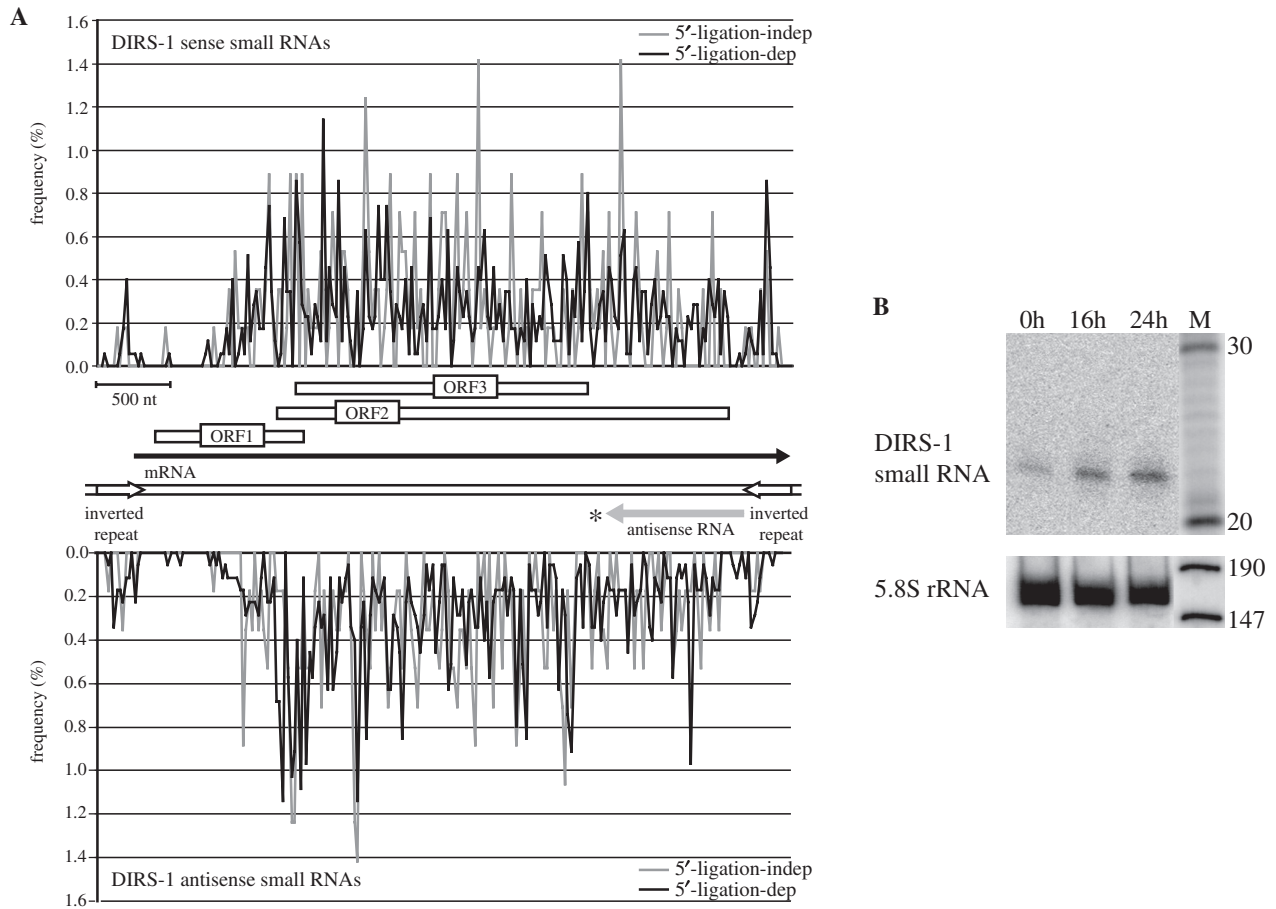
Figure 3. Small RNAs derived from the DIRS-1 retrotransposon (**A**) Distribution of cloned DIRS-1 sequences in the two cDNA libraries in sense (top) and antisense (bottom) orientation relative to the DIRS-1 retrotransposon (middle), which is drawn to scale. The asterisk indicates the location of the small RNA for which expression was analyzed in Figure 3B. (**B**) Northern blot analysis of one DIRS-1 small RNA during growth (0 h) and development (16 and 24 h) of *D. discoideum*. Approximately equal loading was visualized by use of a probe against 5.8S rRNA. Radioactively labeled size markers (M) are shown for reference.
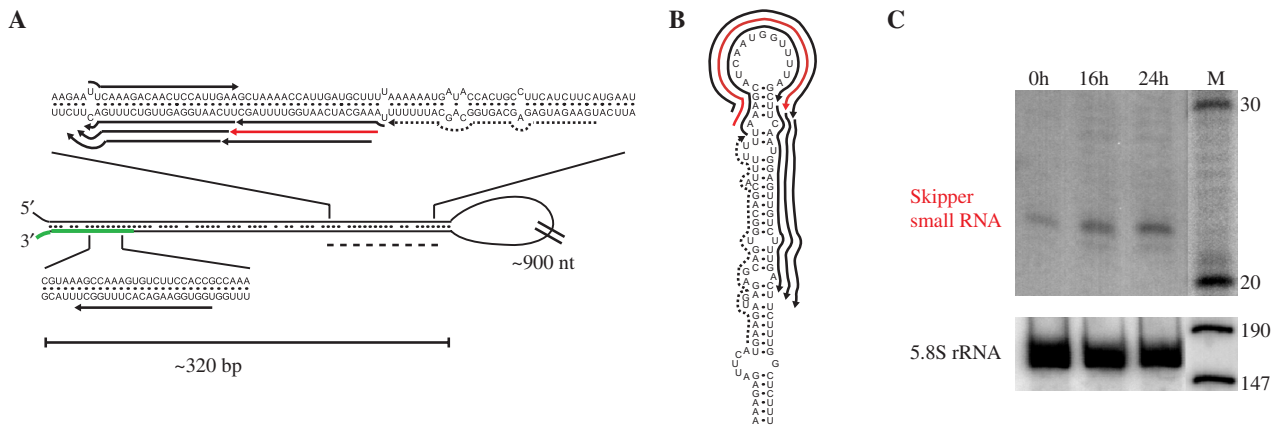

Figure 4. Small RNAs derived from the Skipper retrotransposon. (**A**) Predicted secondary structure (Mfold) of a partial Skipper retrotransposon encompassing all eight cloned small RNAs. Black arrows indicate sequences represented in the cDNA libraries. The region shown in B is indicated by a dashed line. The sequence complementary to the probe used in the northern blot analysis is represented by a red arrow. The dotted arrow indicates a sequence analyzed by northern blot but not present in any of the cDNA libraries. A region covered by EST sequences (http://www.dictybase.org) is shown in green (these sequences, 270–320 nt long, extend further 3′ of the drawn structure). (**B**) Alternative prediction of secondary structure of a region of the Skipper fragment from which six of the eight cloned small RNAs may be derived. The arrows represent the same features as in Figure 4A. (**C**) Northern blot analysis of one Skipper small RNA during growth (0 h) and development (16 and 24 h) of *D. discoideum*. Oligonucleotide probes targeting sequences immediately upstream and downstream of this sequence gave the same results (data not shown). The 5.8S rRNA signals indicate equal loading. Radioactively labeled size markers (M) are shown for reference.
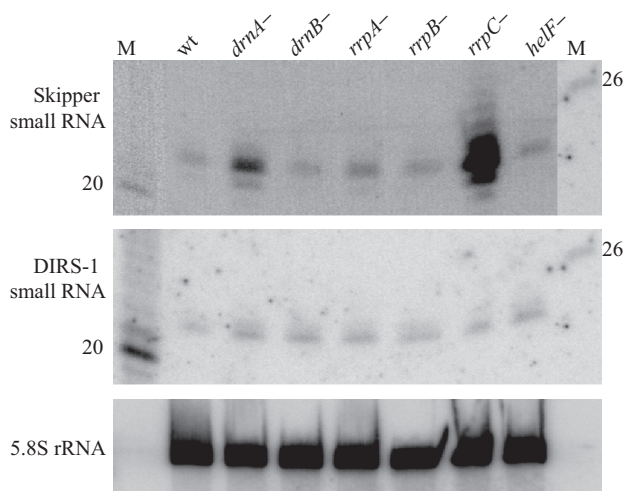
Figure 5. Expression of repeat-associated small RNAs in RNAi knockout strains. Northern blot analysis of Skipper and DIRS-1 small RNAs in a number of different RNAi knockout strains (Figure 2C). Equal loading is visualized by 5.8S rRNA hybridization signals.

same sized RNA from the 5′-part of the predicted stem, showing the same developmental regulation. The rational for analyzing RNAs derived from this region, albeit no matching small RNAs were represent in the cDNA libraries, was the predicted pre-miRNA-like structure which, if cleaved by Dicer, should release such a small RNA.

The isolated Skipper small RNAs demonstrate partial or complete sequence identity or complementarity to Skipper sequences throughout the genome. Interestingly, however, we observed that some of the small RNAs also displayed near-perfect complementarity to mRNAs unrelated to Skipper. For the predicted glycosyltransferase gene, *ggtA*, one of the small RNAs can potentially form a duplex with the mRNA with only one mismatch 6 nt from the 5′end of the small RNA (data not shown).

Different RNAi pathways involved in generation of DIRS-1 and Skipper small RNAs

In order to investigate the possible role of the RNAi machinery in the accumulation of repeat-associated small RNAs, northern blot analysis of DIRS-1 and Skipper small RNAs was carried out using RNA isolated from strains in which RNAi-related genes had been disrupted (see above). The DIRS-1 small RNA accumulated to approximately equal levels in all RNAi knockout strains (Figure 5). The small RNA derived from the predicted loop and 5′-stem of the partial Skipper copy was, however, significantly up-regulated in the strain lacking the *rrpC* and *drnA* genes (only the result for the loop-derived RNA is shown). This indicates that DIRS-1 and Skipper small RNAs are generated by different pathways.

Expression of longer antisense RNAs may generate small RNAs with antisense complementarity to mRNA

A number of clones in the cDNA libraries represented small RNAs with antisense complementarity to predicted or validated protein-coding genes. Three of these genes were selected for further analysis; *hatA*, *rsmF* and *DDB0230011* (http://www.dictybase.org).

hatA encodes the actin-binding protein hisactophilin I. A closely related gene, encoding hisactophilin II, (*hatB*), is located ∼600 bp downstream of *hatA*. Gene disruption of both *hatA* and *hatB* renders *D. discoideum* more sensitive to hyperosmotic stress than wild-type cells (47). A third, highly similar gene, *hatC*, is located on another chromosome although there is no previously reported evidence for its expression. *hatA* and *hatB* each contain one intron, whereas *hatC* lacks introns (Figure 6A). Although these three genes are very similar, the small RNA identified in our cDNA library is fully complementary only to *hatA*.

rsmF (Ras superfamily member F) is predicted to encode a small GTPase (26), although no expression has previously been shown for this gene.

DDB0230011 is predicted to encode a cysteine-rich protein with sequence similarity to the Fraser syndrome protein, FRAS1, as well as the *Xenopus laevis* dorsalizing factor kielin (48–50). In an earlier study, the *D. discoideum DDB0230011* mRNA was demonstrated to be expressed in prestalk cells during development (51).

In order to investigate if the isolated small RNAs with antisense complementarity to mRNAs could originate from longer antisense transcripts, RT-PCR analysis followed by cloning and sequencing of the PCR product was carried out for both strands of *hatA, hatB, hatC, rsmF* and *DDB0230011* (Figure 6A). This experiment verified the expression of both mRNA and longer antisense RNAs for all genes except *hatC*, for which only mRNA was detected. In the case of *hatA*, the primers used for RT-PCR were located in exons one and two. As expected, the PCR product corresponding to the mRNA strand represented the mature, spliced, mRNA without intron. Intriguingly, RT-PCR of the antisense strand gave rise to a product of the same size, and sequencing confirmed that the antisense RNA was perfectly complementary to the spliced mRNA and not to the intron-containing pre-mRNA (Figure 6A). This suggests that at least one RdRP is involved in the production of the antisense RNA. For the remaining genes, the primers were located within one exon. It could therefore not be determined whether the corresponding antisense RNAs were complementary to unspliced or spliced mRNAs.

Developmental regulation of mRNAs and antisense RNAs

Analysis of mRNA and antisense expression for *hatA*, *rsmF* and *DDB0230011* in growing cells and 16 h developed cells, respectively, was also carried out by northern blot using strand-specific probes. It should be noted that the *in vitro* transcribed RNA probe most likely would not distinguish between *hatA*, *hatB* and *hatC* due to their high degree of sequence similarity. *hatA*/*hatB* mRNA has previously been demonstrated to be expressed during early development but is down-regulated
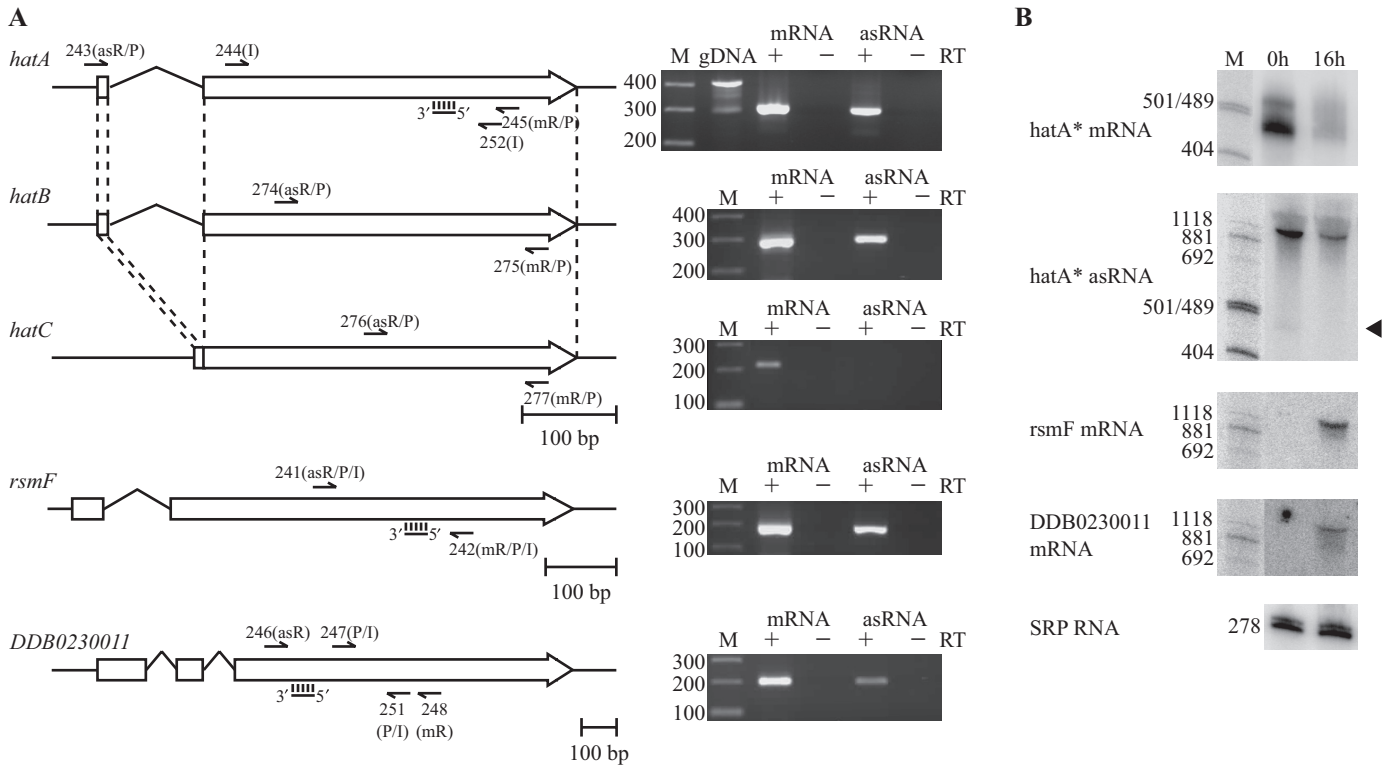
Figure 6. Small antisense RNAs may originate from longer transcripts. (**A**) Schematic drawing (left) of structures of genes analyzed by RT-PCR (right). The positions of sequences represented in the cDNA libraries are indicated for *hatA*, *rsmF* and *DDB0230011*. Primers used in the study are represented by small arrows. Letters in brackets following the oligo numbers indicate whether the oligo was used for reverse transcription of mRNA (mR) or asRNA (asR), PCR (P), or to construct templates for *in vitro* transcription (I). +/− RT indicates the presence or absence of reverse transcriptase in the reaction, respectively. The faint band at ∼300 bp for the PCR reaction performed with genomic DNA (gDNA) and *hatA* primers is probably due to cross-hybridization of primers to *hatC*. (**B**) Northern blot analysis during growth (0 h) and development (16 h) using *in vitro* transcribed, strand-specific probes. The probes constructed against *hatA* will most likely also hybridize to *hatB* and *hatC* and are therefore designated *hatA**. The arrowhead next to the *hatA** asRNA panel indicates a signal corresponding to the size of the *hatA** mRNA.

during culmination (52). Our results confirmed this observation; the mRNA was down-regulated 5-fold in 16 h cells compared to growing cells (Figure 6B). Hybridization of the same membrane with a probe against the antisense strand gave a signal with a similar developmental expression pattern as the mRNA. Interestingly, this signal indicated that the antisense RNA is more than twice as long (∼1100 nt) as the mRNA (∼470 nt). A faint, additional antisense RNA with a size similar to that of the mRNA was also detected in RNA isolated from growing cells.

For *rsmF* and *DDB0230011*, northern blot analysis failed to detect expression from the antisense strands. However, expression of both mRNAs was strongly developmentally regulated and could only be detected in 16 h cells (Figure 6B). Interestingly, the *rsmF* mRNA is substantially longer (∼1100 nt) than anticipated. The predicted open reading frame of *rsmF* is 627 nt and 5′ and 3′ UTRs are in general short in *D. discoideum* (26).

DISCUSSION

Small RNAs are known to be important regulators of gene expression in most eukaryotes. In the model organism *D. discoideum*, RNAi has for several years been used as a tool to knock down gene expression (27–29), but in spite of this, little was known about endogenous small RNAs in this organism (30). This prompted us to investigate the small RNA repertoire of *D. discoideum* by cloning and sequencing 18–26 nt RNAs.

In other organisms, multiple classes of small RNAs of different, distinct, lengths have been identified (2). This is exemplified by miRNAs, which are usually 21–22 nt long in animals and plants, and 26–28 nt repeat-associated siRNAs (ra-siRNAs) isolated from *Schizosaccharomyces pombe* and *Drosophila melanogaster* (6,16,53). In plants, different Dicer-like proteins specifically produce 21-, 22- or 24-nt small RNAs (42). In our cDNA libraries from *D. discoideum*, the size distribution of the cloned small RNAs derived from repetitive elements was clearly biased towards 21 nt RNAs. For the other groups of small RNAs, the number of sequences was too low to confidently address this question. However, for some of the groups, e.g. small RNAs derived from intergenic and exonic regions, most isolated RNAs were between 19 and 21 nt long. This size distribution is in contrast to that of the clones derived from longer non-coding RNAs, presumably degradation products, which demonstrated a more random size distribution from 18 to 26 nt.

When cDNA libraries of small RNAs have been constructed for other organisms, e.g. *C. elegans*, cloning

efficiency for different small RNA classes depends on the method used (2). In an attempt to avoid this bias and to recover as many different small RNAs as possible, we employed two different cloning protocols; one 5′-ligation-dependent and one 5′-ligation-independent. The most prominent difference between the two cDNA libraries was the much higher number of ncRNA fragments, putative degradation products, in the 5′-ligation-independent library, while the DIRS-1 small RNAs were significantly enriched in the 5′-ligation-dependent library. This suggests that the DIRS-1 small RNAs are predominantly Dicer products, since generation of 5′ monophosphate-carrying RNA is characteristic of this enzyme's activity. The unambiguous length bias towards 21 nt for the DIRS-1 RNAs further supports this.

The cDNA libraries were searched for putative miRNAs and five top scoring candidates were identified based on their genomic location and predicted precursor structure. Two of these RNAs, both 21 nt long, could be detected by northern blot analysis, however only after several days of exposure. Hence, the other miRNA candidates could be present at even lower levels. Four of the candidates were present only in the 5′-ligation-dependent library, indicating processing by Dicer. The involvement of a Dicer enzyme was further substantiated by the absence of one of the miRNA candidates in a strain depleted of the Dicer homolog DrnB. However, the abundance of this small RNA was unaffected in a strain in which the gene for the other Dicer homolog, *drnA*, had been disrupted. This result is reminiscent of the situation in *A. thaliana* (42) and *D. melanogaster* (54–56) where miRNA biogenesis is strictly dependent on one of their Dicer-like enzymes, DCL1 and Dicer-1, respectively. Furthermore, the miRNA candidate that is absent in the *D. discoideum drnB* mutant is up-regulated in a strain in which one of the RdRP genes has been disrupted. Again, this parallels observations in *A. thaliana*, where the majority of known miRNAs is enriched when one of the RdRPs (RDR2) is mutated (57).

Notably, the two Dicer-like proteins in *D. discoideum* do not contain the N-terminal helicase domain, which is a conserved part of Dicers in most organisms, including animals, fungi and plants (15,29). In *D. discoideum*, this motif is present on the RdRPs, suggesting domain swapping between the Dicers and the polymerases (29). Nevertheless, Dicer B is still required for the biogenesis of at least one of the *D. discoideum* miRNA candidates. Dicer-like sequences from *Tetrahymena thermophila*, *Trypanosoma brucei* and *Giardia lamblia* also lack the helicase domain but take part in RNAi-related mechanisms *in vivo* and/or *in vitro* (cleavage of dsRNAs) (58–60). It should be noted that none of these Dicer-mediated processes include miRNA biogenesis.

Several targets for the miRNA candidates were predicted, both with near-perfect complementarity to ORFs and imperfect base pairing within 3′UTRs (data not shown). Due to the intermediate evolutionary position of *D. discoideum*, indicated by branching out after plants but before the fungi and animal lineages, we cannot predict whether the miRNA candidates would interact with targets in a plant or animal-like fashion, to primarily promote target cleavage or translational repression, respectively. Experiments are under way in order to validate predicted targets. In the unicellular alga *C. reinhardtii*, miRNAs were shown to induce cleavage of target mRNAs in accordance to its evolutionary position, i.e. close to plants (11,13). Although *D. discoideum* is a unicellular organism, in contrast to *C. reinhardtii*, multicellular development is a natural process initiated by starvation. Since expression of the miRNA candidates is up-regulated during development, this could indicate a role during the multicellular life cycle. However, since expression is also evident during growth, a function during the single cell stage is plausible as well.

The by far most abundant class of small RNAs was derived from the DIRS-1 retrotransposon (30). A randomly chosen DIRS-1 small RNA displayed an expression pattern similar to the one reported for DIRS-1 mRNA, i.e. up-regulation during development (44). It is likely that these results reflect the action of the RNAi machinery to prevent mobilization of the transposon, as reported for several other eukaryotes (61). In addition, DIRS-1-rich clusters are located at one end of each chromosome, and have therefore been suggested to constitute centromeres (26). The large number of DIRS-1 small RNAs may thus be analogous to the small RNAs derived from centromeric repeats in *S. pombe* and *A. thaliana* (16,17,62). siRNAs has been demonstrated to direct DNA methylation of both transposons and centromeric repeats (3). This is in line with the observed DNA methyltransferase-dependent methylation of DIRS-1, however no clear connection between DNA methylation and the RNAi machinery in *D. discoideum* has yet been established (30).

Small RNAs derived from the Skipper retrotransposon were also isolated, although at a much lower frequency. These small RNAs were only represented in the ligation-dependent cDNA library and they all matched a particular partial Skipper fragment. A longer putative RNA expressed from this fragment has the potential to form a long foldback structure from which the small RNAs could be processed. A predicted similar foldback structure turned out to be a rich source of small RNAs in *A. thaliana* when RDR2 had been mutated (57). Alternatively, six of the eight small RNAs could originate from a shorter stem-loop structure, similar to a miRNA precursor. Which, if any, of these conformations is the true precursor of the small RNAs remains to be determined. The analyzed Skipper small RNAs demonstrated a developmentally up-regulated expression, much like the DIRS-1 small RNA. In contrast, no developmental regulation has been reported for the Skipper mRNA (30). Interestingly, some of the Skipper small RNAs were almost fully complementary to non-retrotransposon mRNAs. The question of whether these mRNAs are *bona fide* targets for the Skipper small RNAs is presently unresolved. Some small RNAs derived from repetitive elements have been implicated in miRNA pathways in mammals and plants (63–65).

In a previous study, differential accumulation of the Skipper and DIRS-1 mRNAs was observed in strains in which genes encoding the putative RNAi machinery

components *drnA* (Dicer homolog), *rrpA*, *rrpB* and *rrpC* (RdRP homologs), and *helF* (putative RNA helicase) had been disrupted by homologous recombination (30). This report demonstrated up-regulation of the Skipper mRNA in all knockout strains except for *helF⁻*, whereas the DIRS-1 mRNA exhibited normal expression in all strains tested, except for *rrpC⁻*, in which expression was significantly up-regulated. To compare the expression pattern of the retrotransposon mRNAs to the corresponding small RNAs, we analyzed the levels of the small RNAs derived from Skipper and DIRS-1 in the same knockout strains, as well as in a strain where *drnB* had been disrupted. In the case of DIRS-1, the small RNA expression levels remained unchanged in all the mutant backgrounds. This may indicate that the functions of the two Dicer-like proteins are redundant, as observed for siRNA biogenesis in plants (42). Alternatively, it is conceivable that the component responsible for the processing of the DIRS-1 small RNAs has not yet been identified. In contrast to the DIRS-1 small RNA, the Skipper small RNA was significantly up-regulated in two knockout strains, *rrpC⁻* (putative RdRP) and *drnA⁻*, (one of the two Dicer homologs). Notably, the small RNA derived from the *A. thaliana* foldback structure mentioned above is strongly expressed in the RDR2 mutant (57). One possible explanation for the up-regulation of Skipper small RNA in *drnA⁻* and *rrpC⁻* strains is that the components of the RNAi machinery are competing for common ligands. For example, RrpC may normally be engaged in the production of several different classes of small RNAs and thereby tether some of the ligands also used to generate Skipper small RNAs. When *rrpC* is knocked out, factors normally associated with RrpC are released and can be recruited to the processing machinery required for the biogenesis of Skipper small RNAs (and to a lesser extent for mica1198).

The different expression patterns of the DIRS-1 and Skipper small RNAs, as well as the tested miRNA candidate in the mutant backgrounds, argues for three, at least partly distinct biogenesis pathways for small RNAs in *D. discoideum*. Similarly, multiple RNAi pathways have been demonstrated in e.g. *C. elegans*, *A. thaliana* and *T. thermophila* (19,42,66–68).

The relatively large fraction of small RNAs derived from a 27 nt region in-between the 26S and 5S rRNA genes on the rDNA palindrome is puzzling. At this point we do not know what, if any, their function is. A similar situation is present in mouse cells where RNA transcribed from the intergenic region located ∼2 kb from the transcriptional start site of pre-rRNA recently have been suggested to affect rDNA transcription. In this case, a longer transcript is processed into shorter intermediates (150–300 nt) which promote heterochromatin formation (69).

In addition to the repeat-associated small RNAs, putative miRNAs, and small RNAs derived from the rDNA palindrome, we isolated a small number of RNAs with complete antisense complementarity to annotated protein-coding mRNAs. Such small RNAs have previously been identified in e.g. *C. elegans*, *T. thermophila* and *A. thaliana* (18–20,66,70). In *T. thermophila*, these

RNAs are believed to play a role in the repression of pseudogene expression (66). In *A. thaliana*, a 24 nt small RNA is produced from the overlapping region of two convergent mRNAs, one of which is up-regulated upon salt stress (20). This RNA in turn triggers the synthesis of 21 nt secondary siRNAs which subsequently down-regulate the second, complementary, mRNA.

We were unable to detect any of the three tested *D. discoideum* small antisense RNAs by northern blot analysis. This is analogous to the small antisense RNAs in *C. elegans*, of which only a few could be detected by this method (19). It should be noted, however, that we used total RNA for northern blot analyses, and it is possible that enrichment for small RNAs would enable detection of these RNAs. Furthermore, in *A. thaliana*, the stress-induced small RNA could only be detected under salt stress conditions (20). This may suggest that we yet need to find proper conditions for accumulation of the *D. discoideum* small antisense RNAs to detectable levels.

Only a few longer antisense RNAs had previously been described in *D. discoideum* (71). Interestingly, we demonstrate here the presence of longer antisense transcripts from all three tested genes for which we had identified small antisense RNAs. For one of these genes, *hatA*, both the antisense RNA and the mRNA were abundant enough to be detected by northern blot analysis. In this case, the mRNA and antisense RNA demonstrated similar patterns of down-regulation at 16 h of development compared to growing cells. The regulation of expression between these two time points appeared to be exerted mainly at the level of transcription as indicated by nuclear run-on transcription experiment (data not shown). However, RT-PCR analysis of the *hatA* antisense RNA suggested a post-transcriptional mechanism, since the cloned and sequenced antisense RNA was fully complementary to the spliced mRNA and not to the intron-containing pre-mRNA. This suggests synthesis of the antisense RNA by an RdRP. A similar unusual antisense RNA corresponding to the mature mRNA has previously been described for annexin VII (72). On the other hand, the antisense RNA detected by our northern blot analysis was approximately twice as long as the mRNA (although a faint band corresponding to the mRNA size could also be seen), and thus it is difficult to visualize how the antisense RNA could be synthesized by an RdRP using the mature mRNA as a template. It is conceivable that the hatA gene is regulated both transcriptionally and post-transcriptionally, possibly involving both DNA-dependent and RNA-dependent synthesis of antisense RNAs.

In summary, our survey of small RNAs in the unicellular model organism *D. discoideum* revealed several classes of RNAs. Strong candidates for miRNAs were identified of which at least one is dependent on Dicer B. The majority of the small RNAs were derived from the DIRS-1 retrotransposon, suggesting a role for RNAi in the control of transposon mobility and/or establishment of functional centromeres. In addition, small RNAs corresponding to other repetitive elements were also isolated. Multiple pathways seem to be involved in the biogenesis of these different classes of small RNAs, demonstrating an endogenous role for the RNAi

machinery in *D. discoideum*. Furthermore, small RNAs with antisense complementarity to mRNAs were identified. Intriguingly, expression of longer antisense RNAs was detected from these and other loci. Antisense transcription thus seems to be widespread in *D. discoideum*, as has been observed in other organisms. The possible role of these antisense RNAs in regulation of their complementary mRNAs, as well as the question of whether or not this regulation is RNAi-mediated, will require further investigation. Taken together, these results point to the power of *D. discoideum* as a model for complex RNAi-related processes. Not only will the many genetic and biochemical tools available be of great importance, but also the intriguing evolutionary position of this organism, between plants and animals, and between uni- and multicellularity.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

REFERENCES

1. Lee,R.C., Feinbaum,R.L. and Ambros,V. (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, **75**, 843–854.
2. Aravin,A. and Tuschl,T. (2005) Identification and characterization of small RNAs involved in RNA silencing. *FEBS Lett.*, **579**, 5830–5840.
3. Matzke,M.A. and Birchler,J.A. (2005) RNAi-mediated pathways in the nucleus. *Nat. Rev. Genet.*, **6**, 24–35.
4. Valencia-Sanchez,M.A., Liu,J., Hannon,G.J. and Parker,R. (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev.*, **20**, 515–524.
5. Zamore,P.D. and Haley,B. (2005) Ribo-gnome: the big world of small RNAs. *Science*, **309**, 1519–1524.
6. Bartel,D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.
7. Kim,V.N. (2005) MicroRNA biogenesis: coordinated cropping and dicing. *Nat. Rev. Mol. Cell Biol.*, **6**, 376–385.
8. Li,H.W. and Ding,S.W. (2005) Antiviral silencing in animals. *FEBS Lett.*, **579**, 5965–5973.
9. Voinnet,O. (2005) Induction and suppression of RNA silencing: insights from viral infections. *Nat. Rev. Genet.*, **6**, 206–220.
10. Wassenegger,M. and Krczal,G. (2006) Nomenclature and functions of RNA-directed RNA polymerases. *Trends Plant Sci.*, **11**, 142–151.
11. Molnar,A., Schwach,F., Studholme,D.J., Thuenemann,E.C. and Baulcombe,D.C. (2007) miRNAs control gene expression in the single-cell alga *Chlamydomonas reinhardtii. Nature.*, **477**, 1126–1129.
12. Nair,V. and Zavolan,M. (2006) Virus-encoded microRNAs: novel regulators of gene expression. *Trends Microbiol.*, **14**, 169–175.
13. Zhao,T., Li,G., Mi,S., Li,S., Hannon,G.J., Wang,X.J. and Qi,Y. (2007) A complex system of small RNAs in the unicellular green alga *Chlamydomonas reinhardtii. Genes Dev.*, **21**, 1190–1203.
14. Lewis,B.P., Burge,C.B. and Bartel,D.P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, **120**, 15–20.
15. Cerutti,H. and Casas-Mollano,J.A. (2006) On the origin and functions of RNA-mediated silencing: from protists to man. *Curr. Genet.*, **50**, 81–99.
16. Reinhart,B.J. and Bartel,D.P. (2002) Small RNAs correspond to centromere heterochromatic repeats. *Science*, **297**, 1831.
17. Volpe,T.A., Kidner,C., Hall,I.M., Teng,G., Grewal,S.I. and Martienssen,R.A. (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science*, **297**, 1833–1837.
18. Ambros,V., Lee,R.C., Lavanway,A., Williams,P.T. and Jewell,D. (2003) MicroRNAs and other tiny endogenous RNAs in *C. elegans. Curr. Biol.*, **13**, 807–818.
19. Lee,R.C., Hammell,C.M. and Ambros,V. (2006) Interacting endogenous and exogenous RNAi pathways in *Caenorhabditis elegans. RNA*, **12**, 589–597.
20. Borsani,O., Zhu,J., Verslues,P.E., Sunkar,R. and Zhu,J.K. (2005) Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in *Arabidopsis. Cell*, **123**, 1279–1291.
21. Katiyar-Agarwal,S., Morgan,R., Dahlbeck,D., Borsani,O., Villegas,, Zhu,J.K., Staskawicz,B.J. and Jin,H. (2006) A pathogen-inducible endogenous siRNA in plant immunity. *Proc. Natl Acad. Sci. USA*, **103**, 18002–18007.
22. Werner,A. and Berdal,A. (2005) Natural antisense transcripts: sound or silence? *Physiol. Genomics*, **23**, 125–131.
23. Munroe,S.H. and Zhu,J. (2006) Overlapping transcripts, double-stranded RNA and antisense regulation: a genomic perspective. *Cell. Mol. Life Sci.*, **63**, 2102–2118.
24. Kessin,R.H. (2001) *Dictyostelium—Evolution, Cell Biology, and the Development of Multicellularity*. Cambridge University Press, Cambridge, UK.
25. Williams,R.S., Boeckeler,K., Gräf,R., Muller-Taubenberger,A., Li,Z., Isberg,R.R., Wessels,D., Soll,D.R., Alexander,H. *et al.* (2006) Towards a molecular understanding of human diseases using *Dictyostelium discoideum. Trends Mol. Med.*, **12**, 415–424.
26. Eichinger,L., Pachebat,J.A., Glöckner,G., Rajandream,M.-A., Sucgang,R., Berriman,M., Song,J., Olsen,R., Szafranski,K. *et al.* (2005) The genome of the social amoeba *Dictyostelium discoideum. Nature*, **435**, 43–57.
27. Boeckeler,K., Tischendorf,G., Mutzel,R. and 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.
28. Lee,C.H., Jeong,S.Y., Kim,B.J., Choi,C.H., Kim,J.S., Koo,B.M., Seok,Y.J., Yim,H.S. and Kang,S.O. (2005) *Dictyostelium* CBP3 associates with actin cytoskeleton and is related to slug migration. *Biochim. Biophys. Acta*, **1743**, 281–290.
29. Martens,H., Novotny,J., Oberstrass,J., Steck,T.L., Postlethwait,P. and Nellen,W. (2002) RNAi in *Dictyostelium*: the role of RNA-directed RNA polymerases and double-stranded RNase. *Mol. Biol. Cell*, **13**, 445–453.
30. Kuhlmann,M., Borisova,B.E., Kaller,M., Larsson,P., Stach,D., Na,J., Eichinger,L., Lyko,F., Ambros,V. *et al.* (2005) Silencing of retrotransposons in *Dictyostelium* by DNA methylation and RNAi. *Nucleic Acids Res.*, **33**, 6405–6417.
31. Popova,B., Kuhlmann,M., Hinas,A., Söderbom,F. and 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.
32. Sussman,M. (1987) Cultivation and synchronous morphogenesis of *Dictyostelium* under controlled experimental conditions. *Methods Cell Biol.*, **28**, 9–29.
33. Lee,R.C. and Ambros,V. (2001) An extensive class of small RNAs in *Caenorhabditis elegans. Science*, **294**, 862–864.
34. Lau,N.C., Lim,L.P., Weinstein,E.G. and Bartel,D.P. (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans. Science*, **294**, 858–862.

35. Billoud,B., De Paepe,R., Baulcombe,D. and Boccara,M. (2005) Identification of new small non-coding RNAs from tobacco and *Arabidopsis*. *Biochimie*, **87**, 905.

36. Sewer,A., Paul,N., Landgraf,P., Aravin,A., Pfeffer,S., Brownstein,M.J., Tuschl,T., van Nimwegen,E. and Zavolan,M. (2005) Identification of clustered microRNAs using an *ab initio* prediction method. *BMC Bioinformatics*, **6**, 267.

37. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.

38. Aspegren,A., Hinas,A., Larsson,P., Larsson,A. and Söderbom,F. (2004) Novel non-coding RNAs in *Dictyostelium discoideum* and their expression during development. *Nucleic Acids Res.*, **32**, 4646–4656.

39. Sucgang,R., Chen,G., Liu,W., Lindsay,R., Lu,J., Muzny,D., Shaulsky,G., Loomis,W., Gibbs,R. *et al.* (2003) Sequence and structure of the extrachromosomal palindrome encoding the ribosomal RNA genes in *Dictyostelium*. *Nucleic Acids Res.*, **31**, 2361–2368.

40. Jones-Rhoades,M.W., Bartel,D.P. and Bartel,B. (2006) MicroRNAs and their regulatory roles in plants. *Annu. Rev. Plant Biol.*, **57**, 19–53.

41. Griffiths-Jones,S., Grocock,R.J., van Dongen,S., Bateman,A. and Enright,A.J. (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.*, **34**, D140–D144.

42. Vazquez,F. (2006) *Arabidopsis* endogenous small RNAs: highways and byways. *Trends Plant Sci.*, **11**, 460–468.

43. Glöckner,G., Szafranski,K., Winckler,T., Dingermann,T., Quail,M.A., Cox,E., Eichinger,L., Noegel,A.A. and Rosenthal,A. (2001) The complex repeats of *Dictyostelium discoideum*. *Genome Res.*, **11**, 585–594.

44. Rosen,E., Sivertsen,A. and Firtel,R.A. (1983) An unusual transposon encoding heat shock inducible and developmentally regulated transcripts in *Dictyostelium*. *Cell*, **35**, 243–251.

45. Zuker,C., Cappello,J., Lodish,H.F., George,P. and Chung,S. (1984) *Dictyostelium* transposable element DIRS-1 has 350-base-pair inverted terminal repeats that contain a heat shock promoter. *Proc. Natl Acad. Sci. USA*, **81**, 2660–2664.

46. Leng,P., Klatte,D.H., Schumann,G., Boeke,J.D. and Steck,T.L. (1998) Skipper, an LTR retrotransposon of *Dictyostelium*. *Nucleic Acids Res.*, **26**, 2008–2015.

47. Pintsch,T., Zischka,H. and Schuster,S.C. (2002) Hisactophilin is involved in osmoprotection in *Dictyostelium*. *BMC Biochem.*, **3**, 10.

48. Matsui,M., Mizuseki,K., Nakatani,J., Nakanishi,S. and Sasai,Y. (2000) *Xenopus kielin*: a dorsalizing factor containing multiple chordin-type repeats secreted from the embryonic midline. *Proc. Natl Acad. Sci. USA*, **97**, 5291–5296.

49. McGregor,L., Makela,V., Darling,S.M., Vrontou,S., Chalepakis,G., Roberts,C., Smart,N., Rutland,P., Prescott,N. *et al.* (2003) Fraser syndrome and mouse blebbed phenotype caused by mutations in FRAS1/Fras1 encoding a putative extracellular matrix protein. *Nat. Genet.*, **34**, 203–208.

50. Vrontou,S., Petrou,P., Meyer,B.I., Galanopoulos,V.K., Imai,K., Yanagi,M., Chowdhury,K., Scambler,P.J. and Chalepakis,G. (2003) Fras1 deficiency results in cryptophthalmos, renal agenesis and blebbed phenotype in mice. *Nat. Genet.*, **34**, 209–214.

51. Maeda,M., Sakamoto,H., Iranfar,N., Fuller,D., Maruo,T., Ogihara,S., Morio,T., Urushihara,H., Tanaka,Y. *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.

52. Stoeckelhuber,M., Noegel,A.A., Eckerskorn,C., Kohler,J., Rieger,D. and Schleicher,M. (1996) Structure/function studies on the pH-dependent actin-binding protein hisactophilin in *Dictyostelium* mutants. *J. Cell Sci.*, **109**, 1825–1835.

53. Aravin,A.A. (2003) The small RNA profile during *Drosophila melanogaster* development. *Dev. Cell*, **5**, 337.

54. Jiang,F., Ye,X., Liu,X., Fincher,L., McKearin,D. and Liu,Q. (2005) Dicer-1 and R3D1-L catalyze microRNA maturation in *Drosophila*. *Genes Dev.*, **19**, 1674–1679.

55. Lee,Y.S., Nakahara,K., Pham,J.W., Kim,K., He,Z., Sontheimer,E.J. and Carthew,R.W. (2004) Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell*, **117**, 69–81.

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

57. Lu,C., Kulkarni,K., Souret,F.F., MuthuValliappan,R., Tej,S.S., Poethig,R.S., Henderson,I.R., Jacobsen,S.E., Wang,W. *et al.* (2006) MicroRNAs and other small RNAs enriched in the *Arabidopsis* RNA-dependent RNA polymerase-2 mutant. *Genome Res.*, **16**, 1276–1288.

58. MacRae,I.J., Li,F., Zhou,K., Cande,W.Z. and Doudna,J.A. (2006) Structure of Dicer and mechanistic implications for RNAi. *Cold Spring Harb. Symp. Quant. Biol.*, **71**, 73–80.

59. Mochizuki,K. and Gorovsky,M.A. (2005) A Dicer-like protein in *Tetrahymena* has distinct functions in genome rearrangement, chromosome segregation, and meiotic prophase. *Genes Dev.*, **19**, 77–89.

60. Shi,H., Tschudi,C. and Ullu,E. (2006) An unusual Dicer-like1 protein fuels the RNA interference pathway in *Trypanosoma brucei*. *RNA*, **12**, 2063–2072.

61. Slotkin,R.K. and Martienssen,R. (2007) Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.*, **8**, 272–285.

62. May,B.P., Lippman,Z.B., Fang,Y., Spector,D.L. and Martienssen,R.A. (2005) Differential regulation of strand-specific transcripts from Arabidopsis centromeric satellite repeats. *PLoS Genet.*, **1**, e79.

63. Smalheiser,N.R. and Torvik,V.I. (2006) Alu elements within human mRNAs are probable microRNA targets. *Trends Genet.*, **22**, 532–536.

64. Smalheiser,N.R. and Torvik,V.I. (2005) Mammalian microRNAs derived from genomic repeats. *Trends Genet.*, **21**, 322–326.

65. Arteaga-Vazquez,M., Caballero-Perez,J. and Vielle-Calzada,J.P. (2006) A Family of MicroRNAs present in plants and animals. *Plant Cell*, **18**, 3355–3369.

66. Lee,S.R. and Collins,K. (2006) Two classes of endogenous small RNAs in *Tetrahymena thermophila*. *Genes Dev.*, **20**, 28–33.

67. Duchaine,T.F., Wohlschlegel,J.A., Kennedy,S., Bei,Y., Conte,J.D., Pang,K., Brownell,D.R., Harding,S. and Mitani,S. (2006) Functional proteomics reveals the biochemical niche of *C. elegans* DCR-1 in multiple small-RNA-mediated pathways. *Cell*, **124**, 343.

68. Yigit,E., Batista,P.J., Bei,Y., Pang,K.M., Chen,C.C., Tolia,N.H., Joshua-Tor,L., Mitani,S., Simard,M.J. *et al.* (2006) Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell*, **127**, 747–757.

69. Mayer,C., Schmitz,K.M., Li,J., Grummt,I. and Santoro,R. (2006) Intergenic transcripts regulate the epigenetic state of rRNA genes. *Mol. Cell*, **22**, 351–361.

70. Ruby,J.G., Jan,C., Player,C., Axtell,M.J., Lee,W., Nusbaum,C., Ge,H. and Bartel,D.P. (2006) Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell*, **127**, 1193–1207.

71. Hinas,A. and Söderbom,F. (2007) Treasure hunt in an amoeba: non-coding RNAs in *Dictyostelium discoideum*. *Curr. Genet.*, **51**, 141–159.

72. Okafuji,T., Abe,F. and Maeda,Y. (1997) Antisense-mediated regulation of Annexin VII gene expression during the transition from growth to differentiation in *Dictyostelium discoideum*. *Gene*, **189**, 49–56.