Most eukaryotic mRNAs depend upon precise removal of introns by the spliceosome, a complex of RNAs and proteins. Splicing of pre-mRNA is known to take place in *Dictyostelium discoideum*, and we previously isolated the U2 spliceosomal RNA experimentally. In this study, we identified the remaining major spliceosomal RNAs in *Dictyostelium* by a bioinformatical approach. Expression was verified from 17 small nuclear RNA (snRNA) genes. All these genes are preceded by a putative noncoding RNA gene promoter. Immunoprecipitation showed that snRNAs U1, U2, U4, and U5, but not U6, carry the conserved trimethylated 5′ cap structure. A number of divergent U2 species are expressed in *Dictyostelium*. These RNAs carry the U2 RNA hallmark sequence and structure motifs but have an additional predicted stem-loop structure at the 5′ end. Surprisingly, and in contrast to the other spliceosomal RNAs in this study, the new U2 variants were enriched in the cytoplasm and were developmentally regulated. Furthermore, all of the snRNAs could also be detected as polyadenylated species, and polyadenylated U1 RNA was demonstrated to be located in the cytoplasm.

Removal of introns by splicing is an essential step in the maturation of most eukaryotic mRNAs. This reaction is carried out by the spliceosome, a complex of the spliceosomal small nuclear RNAs (snRNAs) U1, U2, U4, U5, and U6 and their associated proteins. The spliceosomal RNAs interact with each other and the pre-mRNA by base pairing, thus keeping the pre-mRNA–spliceosome complex in position for the splicing reaction to occur (for reviews, see references 46 and 63). Most likely, the snRNAs are also responsible for the catalysis of the reaction (52).

The U1, U2, U4, and U5 snRNAs are transcribed by RNA polymerase II and acquire a monomethylated 5′ cap structure (m7G) in the nucleus (for reviews of small nuclear ribonucleoprotein [snRNP] biogenesis and snRNA modifications, see references 34 and 58). Following transport to the cytoplasm, the snRNAs are subjected to further maturation. Here, the 5′ cap is trimethylated (m2,2,7G), and specific proteins are assembled upon the different snRNAs to form snRNPs, which are then transported back into the nucleus. The biogenesis pathway of U6 snRNA is different in that this RNA is transcribed by RNA polymerase III, the 5′-end triphosphate is γ-monomethylated, and the RNA is thought never to leave the nucleus (58).

During the past few years, polyadenylation of noncoding RNA (ncRNA) such as snRNA, small nuclear RNA (snRNA), and...
spliced gene. The introns harbor the canonical GT and AG nucleotides as part of their 5′ and 3′ splice sites, respectively, and are typically short, i.e., ~150 nucleotides (nt) (12, 51). Earlier studies with *Dictyostelium* showed the presence of small nuclear RNAs of lengths similar to those of splicosomal RNAs from other organisms, but the sequences and further characteristics of these putative snRNAs were not determined (60).

We recently reported the identification of the U2 spliceosomal RNA, which was represented in a *Dictyostelium* cDNA library constructed from small RNAs (6). In the present study, we bioinformatically identified the remaining major spliceosomal RNAs (U1, U4, U5, and U6) by using a combination of search criteria: sequence homology, RNA-RNA interaction, and secondary structure. All the predicted RNAs, expressed from multiple loci in the genome, were sequenced and investigated with respect to expression, putative promoter elements, and 5′ cap status. In addition to the nuclear U2 RNA previously identified, we also found 5′-extended U2-like RNAs which are enriched in the cytoplasm. Interestingly, these RNAs are down-regulated during development. Furthermore, all of the spliceosomal RNAs could be found as polyadenylated species, and polyadenylated U1 RNA was shown to accumulate in the cytoplasm.

**MATERIALS AND METHODS**

**Spliceosomal snRNA sequences used for homology searches.** Spliceosomal snRNAs used for homology searches were obtained from GenBank. For the following snRNAs, the DBDB/EMBL/GenBank accession numbers are given in parentheses: human U1 snRNA (M14386), mouse U1 snRNA (M14121), *Caenorhabditis elegans* U1 snRNA (X51371), *Arabidopsis thaliana* U1 snRNA (X53175), U5 snRNA (X31012), and U6 snRNA (X25257).

**Bioinformatic analysis.** The version 2 *Dictyostelium* genome sequence was downloaded from the dictyBase website (P. Fey, P. Gaudet, E. M. Just, S. N. Merchant, K. E. Filcher, W. A. Kibbe, and R. L. Chisholm, June 2004 [http://dictybase.org/]). Secondary structures were predicted by minimum free-energy optimization using the RNAfold routine of the Vienna RNA package v1.4 (19), available at [http://www.tbi.univie.ac.at/~ivo/RNA/](http://www.tbi.univie.ac.at/~ivo/RNA/) as of April 2005. BLAST analysis was carried out using a local installation of Standalone BLAST v2.2.9 (4). Multiple sequence alignments were constructed using the T-Coffee web server at [http://www.ch.embnet.org/software/TCoffee.html](http://www.ch.embnet.org/software/TCoffee.html) (38) and edited using the BioEdit sequence alignment editor, v7.0.4.1 (14). All algorithms described were implemented with Java 2 SDK v1.4.2 and run on a conventional benchtop computer.

**Oligonucleotides.** Sequences of the DNA oligonucleotides (Invitrogen) used in this study are presented elsewhere (see Table S1 in the supplemental material). Sequences of the RNA oligonucleotides (Dharmacon) are described in reference 6.

**Growth conditions.** *Dictyostelium discoideum* strain AX4 (26) was grown axenically in HL5 medium and developed on nitrocellulose membranes (50).

**Expression analysis: 5′ RACE, 3′ RACE, RT-PCR, and Northern blot analysis.** 5′ rapid amplification of cDNA ends (5′ RACE), 3′ RACE, and Northern blot analysis were carried out as described previously (6). Reverse transcription-PCR (RT-PCR) to detect polyadenylated snRNA species was performed as described previously (6) with the exception that the RNA was reverse transcribed using the GeneRacer oligo(dT) primer (Invitrogen), a procedure followed by PCR with GeneRacer 3′ primer and gene-specific primers. 5′ RACE PCR products of U5A/USB and U6 could be obtained only by nested PCR. In this case, reverse transcription was carried out initially as before, and subsequently an internal gene-specific primer was added to the PCR amplification to enhance the specificity.

**RNase H digestion of poly(A) tails.** RNase H digestion of poly(A) tails was carried out essentially as described by Lie and Maclonald (30). Briefly, 40 μg total RNA or 5.2 μg nuclear RNA (amount adjusted to give the same Northern blot U1 signal intensity as 40 μg of total RNA) dissolved in water was denatured at 85°C for 5 min together with 800 pmol oligo(dT)20 oligonucleotide and, in control reactions, with 0.1 pmol polyadenylated in vitro transcript L3(A20) (33) (kind gift from P. Nilsson). The buffer conditions were adjusted to 40 mM Tris-Cl (pH 8.0), 4 mM MgCl2, 1 mM diethiothreitol, and 30 ng/ml bovine serum albumin, and the reaction mixtures were incubated at 42°C for 10 min. Finally, 1 unit of RNase H (Invitrogen) was added to each reaction mixture, and the mixtures were subsequently incubated at 37°C for 30 min. After phenol-chloroform extraction and ethanol precipitation, half of each reaction mixture was subjected to Northern blot analysis as described above.

**Genomic PCR.** The U6 genomic locus was PCR amplified using *Tag* polymerase (Amersham Biosciences) and 25 pmol each of primers U6 5′ RT and U6 3′ genomic. Cycling conditions were as follows: 95°C for 5 min and 30 cycles of 95°C for 30 s, 50°C for 30 s, and 60°C for 2 min followed by 60°C for 7 min. The PCR product was gel purified using the QiAEX II gel extraction kit (QIAGEN), cloned into the pCR2.1 TOPO vector (Invitrogen), and sequenced on an ABI Prism 3700 DNA analyzer (Applied Biosystems).

**Immunoprecipitation of trimethyl-capped RNAs.** Trimethyl-capped RNAs were immunoprecipitated using a rabbit polyclonal anti-m7G-cap antibody (Synaptic Systems) formerly referred to as R1131 (31). Twenty microliters of serum was coupled to 200 μl of 10% protein G-Sepharose 4 fast-flow beads (Amersham Biosciences; kind gift from G. Akusjarvi) in 50 mM Tris-Cl (pH 7.5), 150 mM NaCl, and 0.1% NP-40 by rotation at 4°C for 20 h. The immobilized antibody was subsequently incubated with 200 μg total RNA prepared from growing *Dictyostelium* cells and 330 units of RNaseguard (Amersham Biosciences) in a total volume of 500 μl (same buffer used as described above) for 2 hours on a rotator at 4°C. After being washed five times with 1 ml buffer (5-min rotation at 4°C followed by sedimentation at 500 × g for 5 min), proteins were digested with 100 μg proteinase K (Amersham Biosciences) at 37°C for 20 min. The released RNA was finally extracted by use of the TRIzol method (Invitrogen), separated on a 10% polyacrylamide gel (7 M urea, 1× Tris-borate-EDTA), and subjected to Northern blot analysis as described previously (6). As a control, the whole procedure was carried out in parallel using serum from a nonimmunized rabbit (generous gift from S. Svard).

**Nucleotide sequence accession numbers.** Newly identified *Dictyostelium* spliceosomal snRNA accession numbers are as follows: U1A and U1B, AY953940 and AY953941; U1C and U1D, DQ012950 and DQ012951; U2C to U2G, DQ012952 to DQ012956; U4A to U4C, AY918063 to AY918065; U5A and USB, DQ01173 and DQ001174; and U6, AY953942. Previously reported U2A/U2B (DdR-19) has accession number J093580 (6).

**RESULTS**

**Identification and characterization of the *Dictyostelium* spliceosomal RNAs.** (i) U1 snRNA. U1 snRNA genes were identified by searching the *Dictyostelium* genome for the conserved and functionally important 5′ splice site recognition motif ACUUAUCCG, located 2 nucleotides from the 5′ end of U1 snRNAs (32, 63). At every occurrence of the canonical recognition motif in the *Dictyostelium* genome, a 165-nt-long sequence was extracted, starting 2 nucleotides upstream of the motif. The secondary structures were predicted by minimum free-energy optimization with the constraint that the 5′ splice site recognition motif must be single stranded. The minimum free energy was compared to a threshold value generated from the minimum free energy of the predicted secondary structures of U1 snRNAs from human, mouse, *C. elegans*, and *A. thaliana*. *Dictyostelium* sequences with minimum free-energy values above this threshold value were discarded. After the remaining sequences were analyzed, two identical candidates fulfilled the criteria of being predicted to fold into the highly conserved cloverleaf structure and containing other conserved features, e.g., Sm-binding site and sequence/structure motifs required for interaction with the U1-specific proteins U1 70K and U1A (Fig. 1A) (15, 41, 44). These genomic sequences were named *U1A* and *U1B*. The *U1A* and *U1B* sequences were used to perform a BLASTN (4) search against the *Dictyostelium* genome; this search identified three additional U1 candidate
genes, which we named \textit{U1C}, \textit{U1D}, and \textit{U1E} (see Fig. S1A in the supplemental material).

Northern blot experiments showed that \textit{U1A}/\textit{U1B} RNA was expressed in both growing and differentiated cells (data not shown). The two RNAs are identical, and we could therefore not determine whether both of them are expressed. Additional Northern blot analysis of nuclear and total RNA demonstrated that the \textit{U1A}/\textit{U1B} RNA is located in the nucleus (Fig. 2). In order to determine the exact ends of the \textit{U1A}/\textit{U1B} RNA, 5'- and 3' RACE experiments were carried out following sequencing of isolated RACE products (see Table S1 and Fig. S1 in the supplemental material). The \textit{U1C} and \textit{U1D} genes differ by only one nucleotide. The expression of both RNAs was demonstrated by 5' RACE followed by sequencing, where the

FIG. 1. Predicted secondary structures and snRNA-snRNA interactions of the \textit{Dictyostelium} spliceosomal RNAs. Gray boxes represent Sm-binding motifs; regions predicted to be involved in snRNA-snRNA interactions are highlighted in orange, red, and yellow for U2 to U6 helices Ia and Ib, II, and III, respectively. The highly conserved U6 motif ACAGAG is indicated by a solid black line. (A) U1A. The 5' splice site recognition motif is indicated in blue. (B) U2A. The pre-mRNA branch site recognition motif is boxed in green. (C) U4A-U6 interaction. (D) U5A. The purple box represents the U5 motif interacting with the exon sequences at the 5' and 3' splice sites. (E) U2A-U6 interaction. Only the first 55 nucleotides of U2A are shown. Adapted from reference 63 with permission of the publisher.
U1C and U1D RNAs were represented by three and eight clones, respectively (data not shown). No clones representing U1E RNA were present among the 11 5’/H11032 RACE and 9 3’/H11032 RACE clones sequenced, nor was any PCR product obtained in a 3’/H11032 RACE using a primer specific for U1E. We conclude that either the U1E gene is not expressed at all or the expression level is below the detection limit of the RACE analysis. This is in agreement with the observation that the putative promoter motif in front of the U1E gene is degenerated (see below).

Taken together, these results show that at least three similar U1 snRNAs are expressed in Dictyostelium and transcribed from three or four genes. The U1A/U1B RNAs are predominantly located in the nucleus and carry the conserved trimethylated 5’ cap structure.

(ii) U2 snRNA. In a previous study, we experimentally isolated U2 RNA (previously referred to as DdR-19), which was shown to be a primary transcript, expressed throughout development, and localized to the nucleus and which was predicted to be transcribed from two genomic loci, here named U2A and U2B. Furthermore, five additional putative U2 genes were demonstrated to be present in the Dictyostelium genome (6); in this report, the genes are named U2C to U2G (see Fig. S1B in the supplemental material). In the present study, immunoprecipitation in combination with Northern blot analysis showed that the U2 snRNA carries a trimethylated 5’ cap—the probe used for Northern blotting was designed to hybridize to U2A to U2C (Fig. 3B).

Expression of the remaining five U2 candidates, U2C to U2G, was verified by 3’/H11032 RACE using primers specific for each RNA. The genes U2D to U2G differed from U2A to U2C in that the putative promoter sequence was located further upstream from the predicted start of transcription and were thus further investigated (see below).

Hence, U2 snRNA in Dictyostelium carries the conserved cap structure, is transcribed from six or seven loci, and is located mainly in the nucleus.

(iii) U4 snRNA. In order to identify the Dictyostelium U4 snRNA, we took advantage of the conserved and extensive base pairing between the U4 and U6 snRNAs as well as the important base-pairing interactions within the U4 RNA (reviewed in references 32 and 46). The sequence of Dictyostelium U4 RNA demonstrates nuclear localization of all snRNAs. Each panel shows the same membrane, which was successively probed for the different snRNAs. The U1A/U1B panel is derived from the Northern blot hybridization presented in Fig. 5. U2A to U2C and DdR-21 (indicated by asterisks) were previously shown to be nuclear and cytoplasmic, respectively (6), and are shown for reference. Closed triangles indicate the longer U4 transcript, predicted to form an additional 3’ stem-loop structure (two different exposures are shown). A radioactively labeled size marker is indicated by M; numbers to the left indicate sizes in bp. Loading of RNA was chosen so that U1A/U1B signals were equal in the nuclear and total fractions, thus facilitating an evaluation of the relative intensities of the RNA bands.

U1C and U1D RNAs were represented by three and eight clones, respectively (data not shown). No clones representing U1E RNA were present among the 11 5’/RACE and 9 3’/RACE clones sequenced, nor was any PCR product obtained in a 3’/RACE using a primer specific for U1E. We conclude that either the U1E gene is not expressed at all or the expression level is below the detection limit of the RACE analysis. This is in agreement with the observation that the putative promoter motif in front of the U1E gene is degenerated (see below).

Taken together, these results show that at least three similar U1 snRNAs are expressed in Dictyostelium and transcribed from three or four genes. The U1A/U1B RNAs are predominantly located in the nucleus and carry the conserved trimethylated 5’ cap structure.

(ii) U2 snRNA. In a previous study, we experimentally isolated U2 RNA (previously referred to as DdR-19), which was shown to be a primary transcript, expressed throughout development, and localized to the nucleus and which was predicted to be transcribed from two genomic loci, here named U2A and U2B. Furthermore, five additional putative U2 genes were demonstrated to be present in the Dictyostelium genome (6); in this report, the genes are named U2C to U2G (see Fig. S1B in the supplemental material). In the present study, immunoprecipitation in combination with Northern blot analysis showed that the U2 snRNA carries a trimethylated 5’ cap—the probe used for Northern blotting was designed to hybridize to U2A to U2C (Fig. 3B).

Expression of the remaining five U2 candidates, U2C to U2G, was verified by 3’ RACE using primers specific for each RNA. The genes U2D to U2G differed from U2A to U2C in that the putative promoter sequence was located further upstream from the predicted start of transcription and were thus further investigated (see below).

Hence, U2 snRNA in Dictyostelium carries the conserved cap structure, is transcribed from six or seven loci, and is located mainly in the nucleus.

(iii) U4 snRNA. In order to identify the Dictyostelium U4 snRNA, we took advantage of the conserved and extensive base pairing between the U4 and U6 snRNAs as well as the important base-pairing interactions within the U4 RNA (reviewed in references 32 and 46). The sequence of Dictyostelium
U6 snRNA (see below) was used to search the *Dictyostelium* genome for U4 candidate sequences with the possibility of interacting with U6 snRNA to form the conserved stem structures I and II (Fig. 1C). Identified sequences were subjected to secondary-structure predictions in order to investigate if the putative U4 RNAs could form the conserved internal 5′ and central stem-loops. These analyses left us with three U4 snRNA candidates which, in addition, all carried the K-turn motif described by Nottrott et al. (39) as well as the Sm-binding motif (Fig. 1C). These genomic sequences were named U4A, U4B, and U4C (see Fig. S1C in the supplemental material).

Northern blot and RACE experiments utilizing a single probe or primer to detect U4A to U4C RNAs were used to analyze expression. The RNAs were found to be expressed throughout development (data not shown) and located predominantly in the nucleus (Fig. 2). The 5′RACE experiment confirmed that the U4 RNA is a primary transcript with a 5′ triphosphate (Fig. 3A) and the RNA could be immunoprecipitated with an antibody specific for the m^2,2,7G cap (Fig. 3B). Only the sequence of U4A was present among the 14 clones that were sequenced from the 5′RACE experiment. U4A was also represented by 14 clones from the 3′RACE experiment, whereas only 1 and 2 clones representing U4B and U4C RNAs, respectively, were isolated (see Fig. S2 in the supplemental material). Taken together, these data indicate an expression bias towards the U4A variant. In the 3′RACE, most of the U4A 3′ ends (12 out of 14 clones) varied in length by only 2 nucleotides, but two clones showed slightly longer extensions of 12 or 13 nt compared to the shortest, most abundant species. The single clone representing the U4B RNA showed a 3′ end similar to those of the longer U4A RNAs (see Fig. S2 in the supplemental material). The longer extension is matched by an additional band from the Northern blot analysis (Fig. 2) which does not disappear after treatment with RNase H in combination with an oligo(dT) oligonucleotide (data not shown). This extended 3′ end possibly could form a stem-loop similar to the one observed in U4 RNAs from, for example, plants and mammals (data not shown). Interestingly, this 3′ stem is absent in organisms such as *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, and the slime mold *Physarum polycephalum* (13, 35).

Thus, the U4 spliceosomal RNAs in *Dictyostelium* are expressed from three genes, are located in the nucleus, and have 5′ ends that carry the conserved m^2,2,7G cap.

(iv) U5 snRNA. The U5 snRNA genes were identified by performing a BLASTN search against the *Dictyostelium* genome with the *A. thaliana* U5 snRNA sequence as a query. This search resulted in two sequences with perfect matches to the conserved pre-mRNA-interacting loop motif (36, 37, 61). Upon closer inspection, the sequences flanking the matching regions were found to contain Sm-binding motifs and could be folded into the conserved secondary structure of U5 snRNAs (13, 63) (Fig. 1D). The genomic sequences, which differed by only one nucleotide, were named U5A and U5B (see Fig. S1D in the supplemental material).

The expression of both genes was verified by 3′RACE experiments in which U5A and U5B snRNAs were represented by six and two clones, respectively (see Fig. S2 in the supplemental material). Northern blot analysis showed expression in growing cells as well as in differentiated cells (data not shown) and a nuclear localization (Fig. 2). By 5′RACE, the U5 snRNA was determined to be a primary transcript (Fig. 3A), and immunoprecipitation indicated the presence of a trimethylated 5′ cap (Fig. 3B). The exact 3′ ends were determined by 3′RACE and were only slightly different (see Fig. S2 in the supplemental material).

Hence, two U5 snRNA genes are expressed in *Dictyostelium*, and the RNA has the conserved 5′ cap structure and is localized to the nucleus.

(v) U6 snRNA. The *Dictyostelium* U6 snRNA was found by conducting a BLASTN search against the *Dictyostelium* genome with the U6 snRNA from *A. thaliana* as a query. The identified U6 snRNA candidate contains the phylogenetically highly conserved motif ACAGAG, the 5′ terminal stem-loop, and the intramolecular U6 helix (Fig. 1E) (13, 32). U6 snRNA is known to form extensive intramolecular base-pairing interactions with the U2 snRNA (helices Ia, Ib, II, and III) as well as with the U4 snRNA (stems I and II) (32, 46, 49, 63). The predicted U2A-U6 and U4A-U6 interactions are depicted in Fig. 1C and E, respectively.

The U6 snRNA was shown to be expressed throughout development (data not shown) and located in the nucleus (Fig. 2). The 5′RACE experiment demonstrated the presence of two 5′ variants of U6 RNA (Fig. 3A), and sequencing identified the longer fragment as the predicted full-length 5′ end of the *Dictyostelium* U6 RNA (Fig. 1C and E). Furthermore, this fragment was more abundant when TAP had been used to process 5′ triphosphates to monophosphates, indicating that this U6 snRNA is a primary transcript. The shorter 5′RACE product was more abundant in the absence of TAP treatment, suggesting that it represented a monophosphorylated processed U6 species. U6 snRNA, which in most organisms lacks the trimethyl guanosine cap (34), could not be immunoprecipitated with an antibody specific for the trimethylated cap structure (Fig. 3B). Furthermore, one clone out of the five sequenced clones from a 3′RACE experiment represented a U6 snRNA with a short poly(A) tail (see Fig. S2 in the supplemental material). This led to further analysis by oligo(dT) RT-PCR (see below).

As determined by sequencing of clones from the 3′RACE experiment, the U6 RNA sequences analyzed all deviated from the publicly available genome sequence by three nucleotides but were identical to *Dictyostelium* expressed sequence tag sequences (all sequences were obtained from dictyBase; see Materials and Methods). Therefore, we PCR amplified and sequenced the U6 region of the *Dictyostelium* genome and found the genomic sequence to be identical to the RNA and expressed sequence tag sequences. The deviations in the publicly available genome sequence are thus most likely due to sequencing errors.

In conclusion, U6 snRNA in *Dictyostelium* is expressed from a single gene, is located in the nucleus and, as expected, does not have a trimethylated cap.

Developmentally regulated U2 RNA variants are enriched in the cytoplasm. U2D, U2E, U2F, and U2G were the only RNAs in this study that had the putative promoter element (see below) located further upstream, i.e., 92 to 102 nt instead of the consensus ~63 nt, from the predicted 5′ end. Therefore, these four RNAs were subjected to 5′RACE analysis. In this experiment, an internal oligonucleotide primer that could hy-
bridize to all four RNAs was used. Interestingly, the experiment generated distinct products of lengths corresponding only to extended 5′/H11032 ends (Fig. 4A). Sequencing of two clones from the 5′/H11032 RACE experiment showed that the U2D and U2F RNAs are primary transcripts with 5′ extensions of 30 and 39 nt, respectively, compared to the U2A/U2B RNA. This places the start of their transcription ~63 nt (the consensus distance) from the putative promoter element. Furthermore, the extended 5′ ends have the possibility to form stem structures (Fig. 4B). Although only two clones were sequenced from the 5′RACE analysis, the U2E and U2G RNAs are expected to also have these longer 5′ ends, based on the locations of the upstream sequence element (USE) and the absence of shorter PCR products (Fig. 4A). Surprisingly, and in sharp contrast to results for the other snRNAs in this study, Northern blot experiments using an oligonucleotide probe hybridizing to the extended 5′ ends of U2E to U2G showed that these RNAs are down-regulated during development (Fig. 4C) and enriched in the cytoplasm (Fig. 4D). Furthermore, the hybridization signals for the U2E to U2G RNAs were about 200-fold weaker than those for the U2A to U2C RNAs, indicating a lower expression level (data not shown). Whether the U2E to U2G RNAs have an m2,2,7G cap structure or not was not possible to determine due to their low abundance.

These results demonstrate cytoplasmic enrichment and developmental regulation of U2 RNA variants with extended 5′ ends.

Polyadenylation of snRNAs. The 3′RACE analysis of U1A/U1B and U6 identified a few cDNA clones representing polyadenylated RNAs (see above and see Fig. S2 in the supplemental material). Northern blot analysis of U1A/U1B showed, in addition to the mature U1 RNA, the presence of low-abundance longer transcripts which disappeared after treatment with RNase H and oligo(dT). Interestingly, these polyadenylated U1 RNA species could be detected only in the total RNA and not in the nuclear RNA fraction (Fig. 5). Except for the U1A/U1B RNAs, Northern blot analysis combined with RNase H and oligo(dT) treatment failed to detect any polyadenylated species for any of the other spliceosomal RNAs (data not shown). However, RT-PCR assays with an oligo(dT) primer in combination with primers specific for the different snRNAs (U1A to U1D, U2A to U2G, U4A to U4C,....

VOL. 5, 2006

DICTYOSTELIUM SPLICEOSOMAL RNAs 929

FIG. 4. Analysis of U2D to U2G RNAs. (A) 5′RACE of U2D to U2G. One primer hybridizing internally to all four RNAs, U2D to U2G, was used to perform 5′RACE with and without prior treatment with TAP. The upper and lower black triangles correspond to primary transcripts of the expected lengths of U2E to U2G and of U2D, respectively. Two clones were sequenced and found to represent U2D and U2F. The upper open triangle probably represents an unspecific PCR product, and the fragment indicated by the lower open triangle was sequenced and found to originate from primer dimers. (B) Predicted secondary structures of 5′-extended U2 variants U2D and U2E to U2G. Dots indicate positions where variants differ in terms of sequence. The corresponding nucleotides are indicated in black boxes for U2E, U2F, and U2G. The pre-mRNA branch site recognition motif is boxed in green. Regions predicted to be involved in interactions with U6 are highlighted in orange, red, and yellow for U2 to U6 helices I, II, and III, respectively. (C) Northern blot analysis showing down-regulation of U2E to U2G (normalized to U2A to U2C signals) during development. Time points represent hours of development, with 0 h designating axenically growing cells. Both panels show the same membrane, which was successively probed for the different U2 RNAs. (D) Northern blot analysis showing cytoplasmic enrichment of U2E to U2G, compared and normalized to U6 signals from the same membrane. Radioactively labeled size markers are indicated by M; numbers to the left indicate sizes in bp.
US5A and US5B, and U6) yielded products for all snRNAs. The PCR products were cloned (1 to 4 clones/snRNA) and sequenced. The observed lengths of the poly(A) tails varied between 17 and 65 nt, with the majority being between 20 and 30 nt (see Fig. S2 in the supplemental material). This experimentally determined length of the poly(A) tail does not necessarily reflect the true tail length, since the oligo(dT) primer is expected to hybridize randomly within the A-tail. In a minority of the cases, the amplified fragments could originate from A-rich genomic sequences downstream of the predicted RNA gene. Nevertheless, the majority of the product-yielding RNAs either are transcribed from genes that are not immediately followed by A-rich sequences, have poly(A) tails considerably longer than the 18-nt oligo(dT) primer used in the reverse transcription reaction, or were shown to be polyadenylated in vitro. The polyadenylated U1A/U1B RNAs are approximately 2,500-fold less abundant than the U1A/U1B RNAs.

FIG. 5. Northern blot analysis of RNA before and after digestion of poly(A) tails with RNase H and oligo(dT) shows that a fraction of the U1A/U1B transcripts is polyadenylated. The polyadenylated U1A/U1B RNAs are excluded from the nucleus. The upper and middle panels are different exposures from the same hybridization. The lower panel is derived from the same membrane but shows hybridization with a labeled oligonucleotide probe detecting a polyadenylated in vitro transcript [in vitro poly(A)], included to ensure proper conditions for RNaseH/oligo(dT) digestion. A radioactively labeled size marker is indicated by M; numbers to the left indicate sizes in bp. The polyadenylated U1A/U1B RNAs are approximately 2,500-fold less abundant than the U1A/U1B RNAs.

In conclusion, most of the spliceosomal RNAs in *Dictyostelium* cells are also present as polyadenylated species. These species are, at least in the case of U1A/U1B, mainly cytoplasmic.

**Genomic organization.** The present computational and experimental search for spliceosomal RNAs in *Dictyostelium* demonstrated that these RNAs, except for U6, are transcribed from multiple genes which in most cases generate distinct snRNA variants. Eighteen snRNA genes were predicted (see Fig. S3 in the supplemental material), and RNAs with sequences corresponding to 17 of these were experimentally detected (see above). However, the U1A and U1B genes, as well as the U2A and U2B genes, are predicted to express identical transcripts. At this point it is unresolved whether or not all four genes are transcriptionally active. The spliceosomal RNA genes are distributed on four of the six chromosomes in *Dictyostelium* (see Fig. S3 in the supplemental material). The majority of these genes are clustered in pairs; each gene pair belongs to the same snRNA gene class, and most often the RNAs are derived from opposite DNA strands. The largest gene class, i.e., the seven genes expressing U2 snRNA-like sequences, is distributed on three of the chromosomes. The U2B and U2C genes are situated on opposite strands, with the transcription start sites separated by only 344 bp. In a similar way, the U4A and U4B genes and the U5A and U5B genes are located on chromosomes 3 and 5 and separated by 392 and 242 bp, respectively. Furthermore, sequences within each class of snRNAs that are situated close together in pairs show a level of similarity generally higher than that of those located further away, e.g., on other chromosomes.

**Flanking sequence elements.** We have previously reported a sequence element situated upstream of regions expressing snRNAs in *Dictyostelium* (6). This short element is located approximately 63 nt upstream of the transcription start site and has the consensus sequence (A/T)CCCCAC/T)AA. Here we named this consensus motif the “*Dictyostelium* upstream sequence element” (DUSE). With some minor sequence variations, this element is present 62 to 66 nt upstream of the transcription start site in all spliceosomal RNA genes reported in this study (Table 1). The U1A and U1B, U2A to U2C, U4A, U5A and U5B, and U6 genes have the same DUSE sequence, ACCCATAA. The DUSEs of the remaining genes derive from the consensus sequence by only one nucleotide, with the exception that the DUSEs of U1E and U2F have two nucleotides substituted. However, in U1E, the two deviating nucleotides are located within the highly conserved CCC triplet, and interestingly, the expression of the U1E gene could not be detected, indicating the importance of an intact DUSE motif. Furthermore, Northern blot experiments as well as RACE analysis indicate that RNA genes, with DUSE sequences in which one nucleotide deviates from the consensus sequence, are less expressed, corroborating the importance of the conserved sequence of the upstream motif for transcription. In addition, the regions expressing the developmentally regulated U2D to U2G are all preceded by a DUSE with two (three for
TABLE 1. Conserved genomic elements flanking the Dictyostelium spliceosomal RNA genes

| Gene | DUSEa | Position (DUSE)b | DMc | Position (DM)d |
|------|-------|------------------|-----|---------------|
| U1A  | ACCCATAA | 63 | TTTATGAAAAA | 19 |
| U1B  | ACCCATAA | 63 | TTTATGAAAAA | 19 |
| U1C  | AaCCATAA | 63 | TTTATGAAAAA | 18 |
| U1D  | AaCCATAA | 64 | TTTATGAAAAA | 18 |
| U1E  | AaCTAAT | 63 | TTTATGAAAAA | 18 |
| U2A  | ACCCATAA | 62 | TTTAGAAAAAA | 19 |
| U2B  | ACCCATAA | 64 | TTTAGAAAAAA | 20 |
| U2C  | ACCCATAA | 62 | TTTAGAAAAAA | 19 |
| U2D  | TCCAAaAA | 62 | TTTAGAAAAAA | 15 |
| U2E  | TCCAAaAA | 64 | TTTAGAAAAAA | 12 |
| U2F  | TCCAAaAL | 63 | TTTAGAAAAAA | 12 |
| U2G  | TCCAAaA | 64 | TTTAGAAAAAA | 12 |
| U4A  | ACCCATAA | 63 | TTTAGAAAAAA | 26 |
| U4B  | cCCCATAA | 64 | TTTAGAAAAAA | 24 |
| U4C  | gCCCATAA | 64 | TTTAGAAAAAA | 23 |
| U5A  | ACCCATAA | 66 | TTTAGAAAAAA | 13 |
| U5B  | ACCCATAA | 64 | TTTAGAAAAAA | 13 |
| U6   | ACCCATAA | 64 | TTTAGAAAAAA | 13 |

a Lowercase letters in DUSE motifs indicate deviations from consensus.
b Position (DUSE) depicts the distance, in nucleotides, between the 5'-most nucleotide of the DUSE motif and the transcription start site.
c DM, downstream motif.
d Position (DM) indicates the distance from the 3' end of the snRNA-expressing region to the first 5' nucleotide of the DM.

U2F) specific nucleotide substitutions, whereas U2A to U2C are constitutively expressed throughout development (Table 1) (see Discussion).

It has been well established in research with, e.g., vertebrates and plants that formation of mature 3' ends of the spliceosomal RNAs transcribed by RNA polymerase II is dependent upon a short regulatory element, the 3' box, which is located just downstream of the mature RNAs (9, 17, 64). Upon closer inspection, we could identify a conserved purine-rich element 12 to 26 nt downstream of the observed or predicted 3' ends of all Dictyostelium snRNAs except for U6 (Table 1). Downstream of the U6 locus, we instead observe a stretch of thymidine residues which corresponds to the terminator motif for RNA polymerase III, the polymerase associated with U6 snRNA (58).

DISCUSSION

In this study, we have identified the RNA components of the major spliceosome in Dictyostelium. The sequences were predicted from the available draft genome by use of a combination of sequence homology, snRNA-snRNA interactions, snRNA-pre-mRNA interactions, and secondary structure predictions. This analysis also shows that the spliceosomal RNAs of Dictyostelium are capable of folding into the evolutionarily highly conserved secondary structures and contain the sequence and structure motifs required for splicing activity. All the predicted snRNA genes, with the exception of U1E, were shown to be expressed. In agreement with observations from other organisms, all the snRNAs except for U6 could be immunoprecipitated using an antibody specific for the trimethylated cap characteristic of RNA polymerase II-transcribed snRNAs (34).

Intriguingly, our results show that all the Dictyostelium spliceosomal RNAs are polyadenylated to some extent. Polyadenylation in eukaryotes has been considered to be restricted to mRNA, where one of its assigned functions is to increase RNA stability (57). Recently, however, ncRNA polyadenylation by poly(A) polymerases Trf4 (reference 5 and references therein) and Trf5 (16, 20) has been shown to activate the nuclear exosome in budding yeast, leading to RNA degradation. Furthermore, rRNA polyadenylation has been reported for, e.g., Leishmania species (10) and Schizosaccharomyces pombe (59). In the case of S. pombe, at least part of the polyadenylation was dependent on the nuclear Trf4/Trf5 homolog Cid14 (59). As for spliceosomal RNAs, polyadenylation has been demonstrated, but only in strains carrying mutations either in the nuclear exosome component Rrp6p (2, 54) or in RNase III (1). To our knowledge, our study provides the first report of direct evidence for polyadenylated spliceosomal RNAs in an organism where no mutations have been introduced in genes affecting RNA processing.

Thus far, most evidence for polyadenylated ncRNA has come from budding yeast. However, a growing number of microarray surveys of different organisms, in which oligo(dT) oligonucleotides have been used to amplify the cDNA probes, have also yielded hybridization signals from ncRNA targets. Examples of this are whole-genome microarray experiments using RNA from humans (24), Drosophila melanogaster (47), and A. thaliana (48). Even though this method does not provide any direct evidence for polyadenylation, it indicates that the phenomenon of polyadenylated ncRNAs is more widespread than previously anticipated.

The role of snRNA polyadenylation in Dictyostelium, however, is unclear. It is tempting to speculate that the polyadenylation promotes the degradation of aberrant snRNAs, as shown for tRNAs in budding yeast (23, 53). In budding and fission yeast species, the polyadenylation complex, as well as the exosome responsible for degradation of the polyadenylated ncRNA, is nuclear (3, 21, 59). For Dictyostelium, on the other hand, Northern blot hybridizations show that the polyadenylated U1A/U1B species are located in the cytoplasm (this study). The question of whether the polyadenylation of Dictyostelium snRNAs takes place in the nucleus or in the cytoplasm is presently unresolved. However, cytoplasmic homologs of the Trf4/Trf5 poly(A) polymerases have been identified in other eukaryotes, such as S. pombe (42, 43) and C. elegans (56), although their possible role in ncRNA polyadenylation remains to be investigated.

The seven U2 snRNA-like sequences in the Dictyostelium genome can be divided into two distinct groups based on sequence similarity, U2A to U2C and U2D to U2G. The U2A to U2C RNAs closely resemble the canonical U2 snRNAs of other organisms (46, 63), whereas U2D to U2G appear to have a slightly diverged primary sequence. In addition, the U2D to U2G RNAs have 30- to 39-nt-long extensions in the 5' end. Nevertheless, the RNAs of both groups can be folded into the conserved secondary structure of U2 snRNA, preserving the functionally important sequence motifs as well as nucleotides important for U2-U6 interactions (32, 63). Thus, the differences in primary sequence are explained mainly by compensatory base-pair changes in stems, changes in presumably non-functional single-stranded regions, and a substantial variation in stem-loop III. This stem-loop has been reported to be non-essential for splicing activity in yeast and is absent in trypanosomes (22, 45). The fact that the U2-like RNAs, U2D to U2G,
exist in parallel with the canonical U2 snRNAs, U2A to U2C, and exhibit conserved secondary structures and interaction motifs despite the apparently diverged primary sequence is interesting. This could imply that the U2-like RNAs are under selection pressure and are important for the fitness of the organism. The same tendency towards variation in primary sequence but conservation in secondary structure can be observed, although to a lesser extent, in the U1C, U1D, U4B, and U4C RNAs.

The additional nucleotides in the 5’ region of the U2D to U2G RNAs are predicted to form an extra stem-loop structure. Remarkably, these variant U2 RNAs were found to be enriched in the cytoplasm. Moreover, they were shown to be developmentally regulated and overall much less abundant than the U2A to U2C snRNAs. Together, these data indicate an alternative function of these RNAs compared to the conventional U2 snRNAs. The altered structure of the U2D to U2G variants, in particular their additional 5’ stem-loop, could function as a cytoplasmic localization signal and/or could be important for the observed developmental down-regulation of the RNAs. However, developmental regulation of these RNAs could also be due to transcriptional control (see below). The small sizes of these RNAs, as well as their sequence and structural similarities to nuclear U2 RNAs, which are not developmentally controlled, give us an excellent tool to investigate RNA localization signals as well as features important for developmental control.

Core promoters of small nuclear RNAs generally consist of an 11- to 21-bp sequence element located 50 to 70 nt upstream of the transcription start site. The upstream elements vary in sequence and are known as, e.g., “proximal sequence elements” or “upstream sequence elements” and are believed to be essential in the transcription of snRNA genes by both RNA polymerase II and RNA polymerase III (for a review, see reference 18). Analysis of the 5’ flanking sequence of the Dictyostelium snRNA genes uncovered a highly conserved motif positioned 62 to 66 nt upstream of the transcription start site. This motif is similar to an 8-bp putative promoter element previously reported by us (6), here named DUSE [consensus sequence, (A/T)CCCA(C/T)AA]. The DUSE is similar to USE sequences reported for snRNA genes in Tetrahymena thermophila and A. thaliana, particularly in the conserved CCC triplet (40, 52). In previous phylogenetic comparisons of promoter elements, a strict conservation of two C residues was identified in snRNA proximal sequence element sequences of Xenopus laevis, human, chicken, and mouse (29). Results from RT-PCR experiments indicate that the expression levels of the different Dictyostelium snRNAs are affected by the DUSE. When performing 5’- and 3’RACE experiments with internal primers hybridizing to all three U4 copies, the great majority of the sequenced clones represented U4A, expressed from a gene which has the consensus DUSE sequence. Only a few clones were found to represent U4B and U4C, where the first nucleotide in the DUSE differs from that found in the consensus sequence. Furthermore, the upstream element of the U1E gene is apparently degenerated with a disrupted CCC triplet and, in agreement with the results given above, transcription from this locus was undetectable. In addition, Northern blot experiments to detect U2A to U2C (consensus DUSE) and U2E to U2G (two or three specific substitutions compared to the DUSE for U2A to U2C) transcripts indicated lower expression of the U2E to U2G RNAs, corroborating the importance of the upstream element. Taken together, these data suggest that the DUSE is a promoter element. Furthermore, the developmentally regulated U2 variants, which are enriched in the cytoplasm, are derived from regions that are preceded by highly similar DUSES deviating by two nucleotides (or three nucleotides, in the case of U2F) from the U2A to U2C DUSE sequence. This nucleotide variation is unique for U2D to U2G, and since only the corresponding RNAs are developmentally regulated, it suggests that this motif could be involved in control of transcription during development.

In other organisms, U1 to U5 snRNA genes have been shown to be transcribed by RNA polymerase II. In contrast, U6 snRNA genes are transcribed by RNA polymerase III. Whether this is also the case in Dictyostelium is presently not known. However, the immunoprecipitation experiment showed the presence of a trimethylated 5’ cap structure for U1 to U5 snRNAs but not for U6, strongly indicating similar transcriptional organization. If Dictyostelium U6 snRNA is transcribed by RNA polymerase III, the presence of the upstream DUSE and the absence of apparent box A and box B promoter elements (data not shown) important for transcription of yeast U6 snRNA genes indicate that the mode of transcription is more similar to that seen in plants and animals (18). Due to the high A/T content in regions surrounding the snRNA genes, we were unable to determine the presence of TATA boxes, which constitute important promoter elements preceding many snRNA genes in other organisms (18).

In conclusion, the identification of the Dictyostelium spliceosomal RNAs revealed both expected and unexpected features. All the major spliceosomal RNAs were identified, and they appear to be expressed and processed in a way similar to what has been described for metazoans and plants. Interestingly, most of the snRNAs could also be found as polyadenylated species, of which at least polyadenylated U1 seems to be cytoplasmic. The discovery of polyadenylated snRNAs in Dictyostelium strongly supports the emerging picture of ncRNA polyadenylation as a conserved evolutionrary feature. Remarkably, the analysis of the spliceosomal RNAs in Dictyostelium also identified a class of U2 variants with extended 5’ ends which are developmentally regulated and enriched in the cytoplasm. The implications of these new findings remain to be investigated.

ACKNOWLEDGMENTS

We are grateful to Göran Akusjärvi, Gerhart Wagner, and Dan Larhammar for helpful discussion and to Gerhart Wagner for critically reading the manuscript. We also thank dictyBase (http://dictybase.org) and the Dictyostelium Sequencing Center Consortium. This work was supported by grants from the Swedish Research Council and Wallenberg Consortium North.

REFERENCES

1. Abou Elela, S., and M. Ares, Jr. 1998. Depletion of yeast RNase III blocks correct U2 3’ end formation and results in polyadenylated but functional U2 snRNA. EMBO J. 17:3753–3764.
2. Allmang, C., J. Kufel, G. Chanfreau, P. Mitchell, E. Petfalksi, and D. Tollervey. 1999. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. EMBO J. 18:5399–5410.
3. Allmang, C., E. Petfalksi, A. Podtelejnikov, M. Mann, D. Tollervey, and P. Mitchell. 1999. The yeast exosome and human PM-Scl are related complexes of 5’ → 5’ exonucleases. Genes Dev. 13:2148–2158.

Allmang, C., E. Petfalksi, A. Podtelejnikov, M. Mann, D. Tollervey, and P. Mitchell. 1999. The yeast exosome and human PM-Scl are related complexes of 5’ → 5’ exonucleases. Genes Dev. 13:2148–2158.
4. Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402.

5. Anderson, J. T. 2005. RNA turnover: unexpected consequences of being tailor-made.Curr. Biol. 15:R45–R51.

6. Aspégren, A., A. Hinas, P. Larsson, A. Larsson, and F. Söderbom. 2004. Novel non-coding RNAs in Dicyostelium discoideum and their expression during development. Nucleic Acids Res. 32:4646–4656.

7. Baldauf, S. L., A. J. Roger, I. Wenk-Siefert, and W. F. Doolittle. 2000. A kingdom-level phylogeny of eukaryotes based on combined protein data. Science 290:972–977.

8. Colgan, D. F., and J. L. Manley. 1997. Mechanism and regulation of mRNA nuclear export. Curr. Opin. Genet. Dev. 11:2775–2786.

9. Connolly, S., and W. Filipowicz. 1993. Activity of chimeric U small nuclear RNA (snRNA)/mRNA genes in transcribed protoplasts of Nicotiana plumaginifolia. U snRNA 3’ end formation and transcription initiation can occur independently in plants. Mol. Cell. Biol. 13:6403–6415.

10. Decuypere, S., J. Vandesompele, V. Yardley, S. De Doncker, T. Laurent, S. Delpeyrat, and F. Doolittle. 2000. Identification of novel non-coding RNAs in the Dicyostelium discoideum transcriptome. Nucleic Acids Res. 28:329–334.

11. Deutscher, M. P. 2006. Degradation of RNA in bacteria: comparison of mRNA and stable RNA. Nucleic Acids Res. 34:659–666.

12. Eichinger, L., J. A. Pachecat, G. Glöckner, M. A. Rajandream, R. Suengue, M. Berriman, J. Song, R. Olsen, K. Szatrowski, Q. Xu, B. Tunggal, S. Koonin, and W. Bork. 2001. The genomes of two social amoebae. Nature 413:93–97.

13. Gingeras, T. R., J. J. Drenkow, A. Piccolboni, S. Bekiranov, G. Helt, H. Tammana, and T. R. Gingeras. 1995. Detection and analysis of non-coding RNAs from mammalian pre-mRNA sequences. Nucleic Acids Res. 23:1332–1349.

14. Guthrie, C., and B. Patterson. 1988. Spliceosomal snRNAs. Annu. Rev. Biochem. 57:417–441.

15. Hall, T. A. 1988. Posttranscriptional modifications in the U1 small nuclear RNA genes. p. 201–227. In H. Grosjean and R. Beve (ed.), Modification and editing of RNA. ASM Press, Washington, D.C.

16. Myerski, E., and C. Brantl. 1991. A phylogenetic study of U4 snRNA sequences from metazoans and the existence of a conserved secondary structure corresponding to ‘free’ U4 snRNA. Biochimie 73:17–28.

17. Newman, A., and C. Norman. 1991. Mutations in yeast U5 snRNA alter the specificity of 5’ splice-site cleavage. Cell 65:115–123.

18. Newman, A., and J. L. Manley. 1992. U5 snRNA interacts with exon sequences at 5’ and 3’ splice sites. Cell 68:743–754.

19. Notredame, C., D. G. Higgins, and J. Heringa. 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302:205–217.

20. Notrott, S., K. Hartmann, F. Fabrizio, H. Urlaub, I. Vidovic, R. Ficner, and R. Lührmann. 1999. Functional interaction of a novel 15.35kD [U4/U5/U6] tri-snRNP protein with the 5’-stem-loop of U4 snRNA. EMBO J. 18:6119–6129.

21. Orum, H., H. Nielsen, and J. Engberg. 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. 227:134–141.

22. Patton, R. T., and R. Pederson. 1988. The Mr 70,000 protein of the U1 small nuclear ribonucleoprotein particle binds to the 5’ stem-loop of U1 RNA and interacts with Sm domain proteins. Proc. Natl. Acad. Sci. USA 85:747–751.

23. Read, R. L., R. G. Martinho, S. W. Wang, A. M. Cormack, and C. J. Norbury. 2002. Cytoplastic poly(A) polymerases mediate cellular responses to stress. Proc. Natl. Acad. Sci. USA 99:12079–12084.

24. Saitoh, S., A. Chabes, W. H. McDonald, L. Thelander, J. R. Yates, and P. Russell. 2002. Cid13 is a cytoplasmic poly(A) polymerase that regulates ribonucleic reductase mRNA. Cell 109:563–573.

25. Scherly, D., W. Boelens, W. J. van Vennrooij, N. A. Dathan, J. Hamm, and I. V. Dijkman. 1985. Isolation and characterization of rabbit anti-m 3’-end formation and transcription initiation can occur independent of the cap and poly(A) tail in Drosophila ovariens extracts. Development 126:499–4996.

26. Schedl, P., A. T. Bankier, R. E. L. Chisholm, R. Gibbs, W. F. Loomis, M. Madera, B. A. Konfortov, F. Rivero, A. Tivey, S. Sugano, B. White, D. Walker, D. H. White, C. Churcher, J. Cooper, S. Haydock, N. van Hoof, A. Llanos-Cuentas, F. Chappuis, J. Arevalo, and J. C. Dujardin. 2005. Yeast Trf5p is a nuclear poly(A) activity but no DNA polymerase activity. Mol. Cell. Biol. 25:10183–10189.

27. Sharp, M. 1984. A protein and definition of its binding site on U1 snRNA. EMBO J. 3:6119–6125.

28. Sussman, M. 1998. Identification of transcribed sequences in Drosophila melanogaster. Genome Res. 8:1055–1068.

29. Sussman, M. 2005. The genome of the social amoeba Dicyostelium discoideum. PLoS Biol. 3:e243–e263.

30. Stelter, E., C. Guthrie. 1988. Two conserved domains of yeast U2 snRNA are separated by 945 nonessential nucleotides. Cell 55:41–48.

31. Staley, J. P., and C. Guthrie. 1998. Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell 92:315–326.

32. Stolc, V., Z. Gauhar, C. Mason, G. Halasz, F. M. van Batenburg, S. A. Rikf, S. Hua, T. Herreman, W. Tongprasit, P. E. Barbano, H. J. Bussemaker, and K. P. White. 2004. A gene expression map for the euchromatic genome of Drosophila melanogaster. Genome Res. 14:595–609.

33. Stolc, V., P. M. Samanta, W. Tongprasit, H. Sethi, S. Liang, D. C. Nelson, A. Hegeman, C. Nelson, D. Rancour, S. Bednarek, E. Lr, Uriq, Zhao, R. L. Wrobel, C. Newmann, B. G. Fox, G. N. Phillips, Jr., J. L. Markley, and M. R. Sussman. 2005. Identification of transcribed sequences in Arabidopsis thaliana by using high-resolution genome tiling arrays. Proc. Natl. Acad. Sci. USA 102:4453–4458.

34. von Hoof, A., P. Lennertz, and R. Parker. 2000. Yeast exosome core activity.
accumulate 3’-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. Mol. Cell. Biol. 20:441–452.

55. Waibel, F., and W. Filipowicz. 1990. RNA-polymerase specificity of transcription of Arabidopsis U snRNA genes determined by promoter element spacing. Nature 346:199–202.

56. Wang, L., C. R. Eckmann, L. C. Kadyk, M. Wickens, and J. Kimble. 2002. A regulatory cytoplasmic poly(A) polymerase in Caenorhabditis elegans. Nature 419:312–316.

57. Wickens, M., P. Anderson, and R. J. Jackson. 1997. Life and death in the cytoplasm: messages from the 3’ end. Curr. Opin. Genet. Dev. 7:220–232.

58. Will, C. L., and R. Lührmann. 2001. Spliceosomal U5snRNP biogenesis, structure and function. Curr. Opin. Cell Biol. 13:290–301.

59. Win, T. Z., S. Draper, R. L. Read, J. Pearce, C. J. Norbury, and S. W. Wang. 2006. Requirement of fission yeast Cid14 in polyadenylation of rRNAs. Mol. Cell. Biol. 26:1710–1721.

60. Wise, J. A., and A. M. Weiner. 1981. The small nuclear RNAs of the cellular slime mold Dictyostelium discoideum. Isolation and characterization. J. Biol. Chem. 256:956–963.

61. Wyatt, J. R., E. J. Sontheimer, and J. A. Steitz. 1992. Site-specific cross-linking of mammalian U5 snRNP to the 5′ splice site before the first step of pre-mRNA splicing. Genes Dev. 6:2542–2553.

62. Wyers, F., M. Rougemaille, G. Badis, J. C. Rousselle, M. E. Dufour, J. Boulay, B. Regnault, F. Devaux, A. Namane, B. Seraphin, D. Libri, and A. Jacquier. 2005. Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell 121:725–737.

63. Yu, Y.-T., E. C. Scharl, C. M. Smith, and J. A. Steitz. 1999. The growing world of small nuclear ribonucleoproteins, p. 487–524. In R. F. Gesteland, T. R. Cech, and J. F. Atkins (ed.), The RNA world, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

64. Yuo, C.-Y., M. Ares, Jr., and A. M. Weiner. 1985. Sequences required for 3′ end formation of human U2 small nuclear RNA. Cell 42:193–202.