Spliced Leader RNA–Mediated trans-Splicing in Phylum Rotifera

Natalia N. Pouchkina-Stantcheva and Alan Tunnacliffe
Institute of Biotechnology, University of Cambridge, Cambridge, United Kingdom

In kinetoplastids, *Euglena*, and four metazoan phyla, trans-splicing has been described as a mechanism for the generation of mature messenger RNAs (mRNAs): 5′-ends of precursor mRNAs are replaced by a short spliced leader (SL) exon from a small SL RNA. Although the full phylogenetic range is unknown, trans-splicing has not been found in vertebrates, insects, plants, or yeast. In animal groups where it does occur, i.e., nematodes, cnidarians, platyhelminths, and primitive chordates, SL RNAs do not show sequence relatedness across phyla. The apparently sporadic phylogenetic distribution and the lack of SL RNA homology have led to opposing hypotheses on its evolution, involving either an ancient origin followed by loss in multiple lineages or independent acquisition in several taxa. Here we present evidence for the occurrence of trans-splicing in bdelloid rotifers (Bdelloidea, Rotifera). A common 23-nt sequence, representing the SL exon—diagnostic of SL RNA–mediated trans-splicing—was found at the 5′-end of at least 50%–65% of mRNAs from *Adineta ricciae* and *Philodina* sp. The trans-splicing pattern in bdelloid rotifers can be unusually complex, as observed in transcripts from a heat shock protein gene, *hsp82-1*, where the SL exon was spliced to three alternative positions. Bdelloid rotifer SL RNAs were found to be 105 or 106 nt long and comprised the SL sequence, a conserved splice donor site and an intron containing a putative spliceosome–binding motif. Intriguingly, some similarity of rotifer SL RNA sequence and predicted secondary structure was seen to that of the predominant SL1 RNA of nematodes, although it is unlikely that this demonstrates homology. In addition, sequence corresponding to the rotifer SL exon was found at the 5′-end of a number of full-length complementary DNA (cDNA) clones in a rice (*Oryza sativa*) database. None of these cDNAs gave a close match with homologous plant genes, suggesting that a small but significant portion of the rice expressed sequence tag database represents sequences derived from rotifers. In summary, the description of SL-mediated trans-splicing in Rotifera extends its representation to at least five metazoan phyla, making it increasingly probable that this is a phylogenetically widespread and therefore ancient phenomenon.

Introduction

Trans-splicing is an mRNA processing event which accurately joins sequences derived from separately transcribed RNAs. In one form of trans-splicing, a leader sequence is spliced from the 5′-end of a small RNA (spliced leader RNA [SL RNA]) to pre-mRNA and constitutes the 5′ terminal exon of the mature mRNA. It was first discovered in trypanosomans (Murphy, Watkins, and Agabian 1986; Sutton and Boothroyd 1986) and later reported in flagellated protozoa (*Euglena*) (Tessier et al. 1991), nematodes (Krause and Hirsh 1987), flatworms (Rajkovic et al. 1990; Davis et al. 1994), hydra (Cnidaria) (Stover and Steele 2001), and primitive chordates, *Ciona intestinalis* (Vandenbergh et al. 2001) and *Oikopleura dioica* (Ganot et al. 2004). However, this mRNA-processing pathway is apparent only in vertebrates, insects, yeast, and plants. In the nematode *Caenorhabditis elegans*, for example, over 80% of pre-mRNAs are trans-spliced to the SL, SL1 (Krause and Hirsh 1987); the same, almost invariant sequence is also used by many other nematode species (Blaxter and Liu 1996). Approximately 25% of genes in *C. elegans* are arranged in clusters and are transcribed as polycistronic pre-mRNAs (Zorio et al. 1994): a second major trans-SL, SL2, is used during processing of these precursors into mature mRNAs. Several other minor SLs have also been described in *C. elegans* (Blaxter and Liu 1996). The addition of the SL sequence to mRNA contributes a m′G cap in kinetoplastids (Campbell, Sturm, and Yu 2000) or a trimethylguanosine cap in cnidarians and nematodes (Blumenthal 1995; Stover and Steele 2001) and influences mRNA stability and translational efficiency of trans-spliced mRNA (Maroney et al. 1995; Zeiner et al. 2003).

SL RNAs are less than 150 nt in length and, in addition to the SL exon, consist of a conserved 5′ splice donor site and an intron, which includes a putative spliceosome–binding motif. The lack of sequence conservation and the apparently sporadic occurrence of trans-splicing among different phyla has prompted debate about its evolutionary origin. The two opposing hypotheses are that either trans-splicing arose in an ancient eukaryote and that this facility has subsequently been lost in multiple lineages or that trans-splicing has evolved independently in several taxa. Current information on the phylogenetic distribution of trans-splicing does not allow one hypothesis to be conclusively favored over the other, but it has been argued that an ancient origin becomes increasingly more likely (Nilsen 2001; Vandenbergh et al. 2001). Similarity between primary sequences of SL RNAs from different phyla would also favor an ancestral origin, but such similarity has not been observed to date. Here we present evidence for trans-splicing in the phylum Rotifera, obtained during studies on the molecular basis of anhydrobiosis (extreme desiccation tolerance) in two species of bdelloid rotifers, *A. ricciae* and *Philodina* sp.
Methods

*Adineta ricciae* (a gift of Claudia Ricci) was maintained largely as described for *Adineta vaga* (Lapinski and Tunnacliffe 2003). *Philodina* sp. was obtained from Carolina Biological Supply Company (Burlington, N. C., www.carolina.com). Upon receipt, cultures were transferred to tissue culture flasks and collected during the week following receipt. Cultures were washed in water extensively prior to harvest. Culture medium was carefully removed and replaced with either carbonated spring water (pH 7.8) or medium with elevated NaCl concentration (50–100 mM) to cause rotifers to detach from the flask. Animals were collected by centrifugation (3 min at 4,000 rpm; Heraeus Multifuge 3 S-R) in 50-ml Falcon tubes, washed, and resuspended by gentle pipetting in 500 μl distilled water (dH₂O). Rotifers were left for a few hours to settle to minimize stress to the animals and then collected on a Nitex nylon filter (20 μm pore size), covered with a second Nitex filter and placed between two wet Whatman 3MM filter papers. Samples were air-dried in a petri dish at 26°C for 6 or 8 h for *A. ricciae* and *Philodina* sp., respectively. After drying, the Nitex membrane was transferred to a test tube, frozen in liquid nitrogen and stored at −80°C until required.

Total rotifer RNA was isolated using TRI Reagent (Sigma-Aldrich, Poole, UK) according to manufacturer’s instructions. Between 20 and 40 ng of total RNA was used per first-strand complementary DNA (cDNA) synthesis reaction performed in the presence of SMART oligonucleotide (Super SMART PCR cDNA Synthesis Kit; BD Biosciences, Cowley, UK). As a result, cDNA populations were enriched in full-length cDNAs. Resulting libraries were subcloned into pCRII-TOPO vector (Invitrogen, Paisley, UK); cDNAs were sequenced by MWG Biotech (Edersberg, Germany). SL-primed cDNA libraries were constructed by conventional first-strand cDNA synthesis followed by PCR amplification with SL-based (GGCTTTGTCATACC). Products were subcloned into pCRII-TOPO vector (Invitrogen, Crawley, UK) and cloned into pCRII-TOPO. Inserts were sequenced at the Department of Genetics, University of Cambridge. Lasergene sequence analysis software (DNASTAR, www.dnastar.com) and the Mfold server (www.bioinfo.rpi.edu/applications/mfold/old/ma; Zuker 2003) were used for sequence analysis. Pairwise alignments were performed using ClustalW (Thompson, Higgins and Gibson 1994) on the San Diego Supercomputer Center Workbench Web site (http://workbench.sdsc.edu) with gap-opening penalty set to 15 and gap-extension penalty set to 6.66. The nucleotide sequence data reported in this paper have been submitted to the EMBL/GenBank Data Libraries with accession numbers AY823992, AY823993, AY823994, and AY942832 through AY942842.

Results

To further investigate the phylogenetic distribution of *trans*-splicing, we examined cDNAs from two species of bdelloid rotifer (*Bdelloidea, Rotifera*). Interest in these organisms has been stimulated by their ability to reproduce asexually and to undergo anhydrobiosis. Analysis of full-length cDNA libraries from bdelloid rotifers *A. ricciae* and *Philodina* sp. revealed a number of cDNAs with a common 23-nt sequence (fig. 1). This finding suggested that rotifer mRNAs undergo SL addition at their 5′-ends. The putative SL sequences were found on cDNAs encoding a variety of proteins, including structural proteins, metabolic enzymes, and regulatory proteins. However, not all cDNAs seemed to contain the SL sequence; of 14 nonredundant, nuclear-encoded 5′ cDNA ends from *A. ricciae*, 7 lacked the leader sequence. Because some of the latter group of cDNAs might represent incomplete mRNA copies, we performed reverse transcriptase–polymerase chain reaction (RT-PCR) using a primer based on SL sequence in combination with a gene-specific primer for each case. Two of the

Rotifer SL RNA was cloned using a modified protocol for RNA 3′-end ligation–mediated PCR (LM PCR) (Elbashir, Lendeckel, and Tuschl 2001; Hitchcock et al. 2004): 5 μM of chimeric RNA-DNA adapter 5′-(5′-P)rCrArGdCdTdCdCdAdGdTdAdAdCdCdTdAdCdC dideoxycytidine (ddC; Dharmacon, Dallas, Tex.) was ligated to 50–100 ng of total rotifer RNA using 20 U of T4 RNA ligase (New England BioLabs, Hitchin, UK) in a final volume of 20 μl. The 5′-end of the hybrid oligonucleotide was phosphorylated, and 3′ modification (ddC) was used to block undesirable 3′-end ligations. The ligation reaction was incubated at 15°C for 1 h. The ligated RNA was phenol-chloroform purified, ethanol precipitated, and resuspended in 8 μl of diethylpyrocarbonate H₂O. The whole ligation was used in a first-strand synthesis reaction driven by Invitrogen SuperScript III reverse transcriptase with 2.5 μg reverse adapter primer (GGC TAG GTT ACA TGG AG). The reaction was performed at 50°C for 30 min; 2 μl of RT mix served as template during second-strand synthesis, where 10 μM each of reverse adapter primer and SL-specific 21-bp nested forward primer (GGC TTA TTA CAA CTT ACC AAG) were added. PCR products were resolved on 2% agarose gels. Products in the region of 80–100 bp were isolated with Qiaquick Gel Extraction Kit (Qiagen, Crawley, UK) and cloned into pCRII-TOPO. Inserts were sequenced at the Department of Genetics, University of Cambridge. Lasergene sequence analysis software (DNASTAR, www.dnastar.com) and the Mfold server (www.bioinfo.rpi.edu/applications/mfold/old/ma; Zuker 2003) were used for sequence analysis. Pairwise alignments were performed using ClustalW (Thompson, Higgins and Gibson 1994) on the San Diego Supercomputer Center Workbench Web site (http://workbench.sdsc.edu) with gap-opening penalty set to 15 and gap-extension penalty set to 6.66. The nucleotide sequence data reported in this paper have been submitted to the EMBL/GenBank Data Libraries with accession numbers AY823992, AY823993, AY823994, and AY942832 through AY942842.

Methods

*Adineta ricciae* (a gift of Claudia Ricci) was maintained largely as described for *Adineta vaga* (Lapinski and Tunnacliffe 2003). *Philodina* sp. was obtained from Carolina Biological Supply Company (Burlington, N. C., www.carolina.com). Upon receipt, cultures were transferred to tissue culture flasks and collected during the week following receipt. Cultures were washed in water extensively prior to harvest. Culture medium was carefully removed and replaced with either carbonated spring water (pH 7.8) or medium with elevated NaCl concentration (50–100 mM) to cause rotifers to detach from the flask. Animals were collected by centrifugation (3 min at 4,000 rpm; Heraeus Multifuge 3 S-R) in 50-ml Falcon tubes, washed, and resuspended by gentle pipetting in 500 μl distilled water (dH₂O). Rotifers were left for a few hours to settle to minimize stress to the animals and then collected on a Nitex nylon filter (20 μm pore size), covered with a second Nitex filter and placed between two wet Whatman 3MM filter papers. Samples were air-dried in a petri dish at 26°C for 6 or 8 h for *A. ricciae* and *Philodina* sp., respectively. After drying, the Nitex membrane was transferred to a test tube, frozen in liquid nitrogen and stored at −80°C until required.

Total rotifer RNA was isolated using TRI Reagent (Sigma-Aldrich, Poole, UK) according to manufacturer’s instructions. Between 20 and 40 ng of total RNA was used per first-strand complementary DNA (cDNA) synthesis reaction performed in the presence of SMART oligonucleotide (Super SMART PCR cDNA Synthesis Kit; BD Biosciences, Cowley, UK). As a result, cDNA populations were enriched in full-length cDNAs. Resulting libraries were subcloned into pCRII-TOPO vector (Invitrogen, Paisley, UK); cDNAs were sequenced by MWG Biotech (Edersberg, Germany). SL-primed cDNA libraries were constructed by conventional first-strand cDNA synthesis followed by PCR amplification with SL-based (GGCTTTGTCATACC). Products were subcloned into pCRII-TOPO vector (Invitrogen, Crawley, UK) and cloned into pCRII-TOPO. Inserts were sequenced at the Department of Genetics, University of Cambridge. Lasergene sequence analysis software (DNASTAR, www.dnastar.com) and the Mfold server (www.bioinfo.rpi.edu/applications/mfold/old/ma; Zuker 2003) were used for sequence analysis. Pairwise alignments were performed using ClustalW (Thompson, Higgins and Gibson 1994) on the San Diego Supercomputer Center Workbench Web site (http://workbench.sdsc.edu) with gap-opening penalty set to 15 and gap-extension penalty set to 6.66. The nucleotide sequence data reported in this paper have been submitted to the EMBL/GenBank Data Libraries with accession numbers AY823992, AY823993, AY823994, and AY942832 through AY942842.
Fig. 1.—A common sequence at the 5′-ends of multiple bdelloid rotifer mRNAs. The common 23–nt sequence at the 5′ termini of bdelloid rotifer (Adineta ricciae; A_r{1–7}) and Philodina sp. (P_1–4) mRNAs is shown in upper case; divergent downstream sequences are in lower case. The translation start codon is indicated in bold upper case. Single nucleotide transitions within the SL sequence are highlighted. Closest matches of SL-containing transcripts were determined by BlastX; the nucleotide sequence data presented in the figure have been submitted to the EMBL/GenBank Data Libraries with following accession numbers: Ar1 (AY942832), similar to RAD23 homolog B (Danio rerio); Ar2 (AY942833), similar to unnamed protein product CAF98644 (Tetraodon nigroviridis); Ar3 (AY942834), similar to AAS91553/AmphiHMGI12 (Branchiostoma belcheri tiginaeae); Ar4 (AY942835), similar to ribosomal protein L18a (Canis familiaris); Ar5 (AY942836), similar to 60S ribosomal protein L37 (Xenopus tropicalis); Ar6 (AY942837), similar to ribosomal protein S7 (Petromyzon marinus); Ar7 (AY942838), similar to a DnaD-like protein (AAH75905) from D. rerio; P1 (AY942839), similar to S14e ribosomal protein (Dascillus cervinus); P2 (AY942840), similar to fba1 (Schizosaccharomyces pombe); P3 (AY942841), similar to guanine nucleotide-binding protein beta polypeptide 2-like 1 (D. rerio); and P4 (AY942842), similar to putative 42-9-9 protein (Mus musculus).

7 sequences were successfully amplified, showing that at least 9 out of 14 A. ricciae mRNAs contain the SL sequence. Similar estimates were obtained for Philodina sp. cDNAs, where four out of eight nonredundant clones apparently represented trans-spliced mRNAs. Therefore, at least 50%–65% of bdelloid rotifer mRNAs were found to contain a SL. In three cases, single nucleotide transitions were detected within the SL sequence (fig. 1). This might represent some variation within the SL exon in multiple copies of the SL RNA gene but will require further analysis for confirmation. To date, however, we have not observed a major SL variant, like SL2 of C. elegans, in rotifers.

Intriguingly, Blast searches revealed that a number of full-length cDNAs derived from rice (Oryza sativa) cDNA libraries (Kikuchi et al. 2003) contained a sequence identical to that found at the 5′-end of rotifer cDNAs (table 1). Ten such clones were identified; they encode a range of proteins with excellent matches to sequences from various species, the best of which is listed in the column “Organism” of table 1. However, none of these sequences matched closely with homologous plant genes. Because a trans-splicing pathway has not been reported in plants, it is likely that the SL-containing cDNAs represent foreign—probably rotifer—material in rice libraries. In support of this, we compared one of the sequences of table 1 (no. 4), encoding a 60S ribosomal protein L57, with A. ricciae cDNA for the same protein (Ar5; fig. 1). ClustalW analysis showed a percent identity of 85.5% at the nucleotide level, indicating that the expressed sequence tag in question is likely to originate from a rotifer, although probably a different species. If correct, this would suggest the occurrence of trans-splicing in a third (or more) rotifer species. Interestingly, one of the “rice” cDNAs (no. 10, table 1) contains a SL variant identical to that seen in A. ricciae clone Ar1 (fig. 1).

As first realized for nematodes, a conserved 5′-sequence can be used to construct full-length cDNA libraries (Bektesh and Hirsh 1998). To demonstrate this principle in rotifers, RT-PCR was performed on poly(A)-containing mRNA using both SL and oligo(dT) primers. A selection of complete cDNAs was generated in this way, as determined by the presence of a major open reading frame just downstream of the SL sequence and a poly(A) tract at the 3′-end, and included sequences encoding a variety of proteins (data not shown).

To provide additional evidence for trans-splicing in bdelloid rotifers, we analyzed transcript copies from a hsp82 (82 kDa heat shock protein) gene in A. ricciae. A single partial hsp82 genomic sequence from A. ricciae is recorded in sequence databases (accession number AY394701); the corresponding gene is termed here hsp82-1. To map the 5′-ends of its transcripts, we performed gene-specific RT-PCR using the SL sequence as 5′ primer; gene-specific 3′ primers were designed from

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Nearest BlastX Matches (“Hit ID”) with Nonredundant Sequences from a Rice Expressed Sequence Tag Database Containing Bdelloid Rotifer SL-like 5′ Additions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clone ID</td>
<td>Leader Sequence</td>
</tr>
<tr>
<td>1</td>
<td>AK100559 -GGCTTATTACAACCTACCAAGAG</td>
</tr>
<tr>
<td>2</td>
<td>AK071809 -GGCTTATTACAACCTACCAAGAG</td>
</tr>
<tr>
<td>3</td>
<td>AK069174 -GGCTTATTACAACCTACCAAGAG</td>
</tr>
<tr>
<td>4</td>
<td>AK063177 -GGCTTATTACAACCTACCAAGAG</td>
</tr>
<tr>
<td>5</td>
<td>AK063233 -GGCTTATTACAACCTACCAAGAG</td>
</tr>
<tr>
<td>6</td>
<td>AK063098 -GGCTTATTACAACCTACCAAGAG</td>
</tr>
<tr>
<td>7</td>
<td>AK062639 -GGCTTATTACAACCTACCAAGAG</td>
</tr>
<tr>
<td>8</td>
<td>AK062512 -GGCTTATTACAACCTACCAAGAG</td>
</tr>
<tr>
<td>9</td>
<td>AK062346 -GGCTTATTACAACCTACCAAGAG</td>
</tr>
<tr>
<td>10</td>
<td>AK100502 -GGCTTATTACAACCTACCAAGAG</td>
</tr>
</tbody>
</table>
III-L for the longest (842 bp) clone (fig. 2A). All three clones included the previously unreported 5′-end of the hsp82-1 coding region (48 bp), together with different extents of the upstream UTR (5′ UTR); genomic sequence corresponding to most of III-L was also obtained after PCR amplification from A. ricciae genomic DNA.

I-S, II-I, and III-L all share the same sequence (33 to +1) upstream of the initiation codon (ATG) and the −148 to −34 region is identical in both II-I and III-L (fig. 2B). Upstream of position −33, the gene sequence contains a splice acceptor site (AG) at position −35, a splice intermediate branch point (BP1: TACAAC) at position −89, and two pyrimidine tracts between them, consistent with trans-splicing of the SL exon to abut position −33, as seen in cDNA clone I-S. Similarly, the SL sequence of cDNA clone II-I adjoins sequence at position −148 immediately following a splice acceptor site, which is preceded by a pyrimidine tract and near consensus branch point (BP2: CTCGAT), all located within a 48-bp intron of the hsp82-T gene. This intron, which has a rare (GC); position −196 splice donor sequence (Burset, Seledtsov, and Solovyev 2001), has been removed in cDNA clone III-L, demonstrating that the splice acceptor at −150 can be used for either trans- or cis-splicing. As expected, if trans-splicing is occurring, the SLs within cDNAs I-S and II-I were not present in the hsp82-1 genomic sequence. The three forms of hsp82-1 cDNAs could be due to either alternative trans-splicing of the SL exon to different splice acceptors in the same long precursor mRNA or to the use of alternative promoters generating different species of pre-mRNA, or both. Several putative promoter elements were identified within the 5′ UTR of hsp82-1 genomic DNA. These include two heat shock element (HSE) motifs (hse-1 and hse-2 in fig. 2), a CCAAT box, a GAGA box/CT-element, and a number of TATA boxes. Mapping of transcription initiation sites will be required to distinguish between these possibilities.

Individual cDNAs corresponding to SL RNAs of both A. ricciae and Philodina sp. were generated by 3′-end RNA LM PCR. Sequence analysis of independent cloned LM PCR products showed that rotifer SL RNA can be either 105 or 106 nt long and includes the SL exon itself (23 nt), a splice donor dinucleotide (GT), and a putative Sm-binding motif (fig. 3A). Five variants of the Sm-like motif were identified with a consensus AAYUYUGA (where Y is a pyrimidine), similar to the consensus, AGCUUUGG, of the primitive chordate C. intestinalis (Vandenberghe et al. 2001). A number of species-specific differences are also discernible within the introns of A. ricciae and Philodina sp. SL RNAs. For instance, in Philodina sequences there were several nucleotide substitutions within the putative Sm-binding site and near the 3′-end of the intron, which were never present at the same positions in Adinetna cDNAs. Similarly, a major variant seen in several A. ricciae SL RNAs was not found in Philodina sp. This variant showed an interesting feature, i.e., complementary nucleotide transitions at two independent sites: ATC → GCT and GAT → AGC in the 75–92 nt SL RNA region, together with G → A and C → T at positions 81 and 86. These changes have implications for SL RNA secondary structure (see below).
Secondary structure prediction of rotifer SL RNA was performed using Version 3.1 of Mfold without folding constraints (Zuker 2003). Rotifer SL RNA was predicted to fold into three stem loops, two of which were located towards the 5'-end of the molecule (fig. 3B). Part of the SL exon and the splice donor dinucleotide (GU) were located in the double-stranded region of hairpin I, similar to that predicted for all other SL RNAs, while the putative Sm–binding site was found in the loop region of hairpin II. Lowering the temperature from the default 37°C to 26°C, at which rotifers are grown, had no effect on predicted folding. In most SL RNAs described to date, secondary structure predictions locate the possible Sm-binding motif to a single-stranded region, but this feature is imposed in Mfold by the investigator. We chose not to apply such a constraint to rotifer SL RNA because the proposed Sm-binding site is located in a more 5'-position compared to analogous motifs of all other SL RNAs with the exception of that of C. intestinalis. Forcing the Sm motif region to adopt a single-stranded structure led to the abolishment of hairpin II and part of hairpin III and resulted in unlikely base pairing between the 5'- and 3'-ends of rotifer SL RNA. The complementary transitions in the variant A. ricciae sequence mentioned above are located within hairpin III and are predicted to base pair with each other. The maintenance of base pairing in hairpin III of both major SL RNA variants in A. ricciae is consistent with the predicted structural role for these residues.

A noteworthy feature of rotifer SL RNAs is their resemblance to those of nematodes. The SL exon itself, at 23 nt is almost identical in size to the 22 nt of all major nematode SLs, whereas SL sequences from other organisms vary in size from 16 to 51 nt; a three stem-loop structure is also predicted for both rotifer and nematode SL RNAs. Interestingly, their sequences also appear to be related: an alignment of the A. ricciae SL RNA sequence with SL1 RNA from C. elegans (fig. 4A) showed similarity in the SL exon itself, around the exon-intron junction and the SL exon.
also throughout the intron, and gave a ClustalW score of 32 (indicating percent sequence identity adjusted for gap penalties; see Methods). Alignments of rotifer SL RNA with SL1 RNA from other nematodes gave similar results, e.g., with SL1 A RNA of Oscheius (formerly Dolichorhabditis) sp. CEW1 (Evans et al. 1997; fig. 4B), which gives a score of 26. The similarity between rotifer SL and nematode SL1 RNA sequences is comparable to that between C. elegans SL1 and SL2 RNAs, which gives a score of 21 in the alignment of figure 4C and between C. elegans and Oscheius sp. SL1 RNAs, with a score of 29 (alignment not shown). Although intriguing, these findings do not demonstrate homology between SL RNA genes from bdelloid rotifers and nematodes; it remains possible that the similarity observed is due to chance. If so, we might expect alignments of rotifer SL RNA sequences with those of some nematodes to give a considerably lower percentage identity. Indeed, the Wuchereria bancrofti SLG1 RNA (Dassanayake, Chandrasekharan, and Karunanayake 2001) gives a score of only 9 after alignment with A. ricciae SL RNA. Comparisons of rotifer SL RNA sequences with those of nonnematode species, e.g., cestode or larvaccean, showed no similarity (data not shown).

Discussion

Bdelloid rotifers, one of three classes of the phylum Rotifera, have several features of interest to geneticists and biochemists, including oblate asexual reproduction (Mark Welch and Meselson 2000) and anhydrobiosis, a reversible state of metabolic arrest induced by desiccation (reviewed in Tunn acliffe and Lapinski 2003). During studies on the molecular basis of anhydrobiosis in bdelloid rotifers, we found a number of full-length cDNAs to contain a common sequence at the 5′ termini, which is diagnostic of SL RNA-mediated trans-splicing. The 23-nt SL sequence was attached to a range of different cDNAs from two bdelloid species grown in the laboratory and probably one or more other species, which contaminate rice plants. Because at least 50%–65% of cDNA sequences analyzed contained the SL exon, trans-splicing is an important mRNA-processing pathway in bdelloid rotifers.

Evidence for the complexity of trans-splicing in bdelloid rotifers was obtained through analysis of A. ricciae hsp82-1 cDNAs. The hsp82 gene has been used in rotifers for studies on phylogenetic relationships and asexual reproduction: in four bdelloid genomes examined, two or more copies of hsp82 have been identified (Mark Welch and Meselson 2004). Cytogenetic analysis of Philodina roseola revealed the presence of four hsp82 genes with each copy on a different chromosome (Mark Welch, Mark Welch, and Meselson 2004). These can be grouped as two gene pairs according to sequence similarity: copies 1 and 2 differ by 3.5% and copies 3 and 4 by 6%, but the former differ from the latter by ~47%. To date, only one hsp82 sequence from A. ricciae has been deposited in databases, but cDNAs from this gene, hsp82-1, revealed features characteristic of SL trans-splicing and showed that the SL exon could be spliced to three different positions in the hsp82-1 5′ UTR. Differential trans-splicing patterns have also been reported in trypanosomes (Mair et al. 2000; Nepomuceno-Silva et al. 2001; Manning-Cela, Gonzalez, and Swindle 2002) and can be explained either by alternative trans-splicing of the same pre-mRNA or by use of alternative promoters. Distinguishing between these possibilities will require accurate promoter mapping experiments, but whichever is correct, the long hsp82-1 (III-L) transcript identified must be a product of both conventional cis-splicing and SL-dependent trans-splicing. The former relies on the recruitment of a rare GC donor site, which shows a near perfect match with the consensus exon-intron sequence MAGGCAA/GT found in mammalian GC-AG type introns (Burset, Seledtsov, and Solovyev 2001). The splice acceptor in this intron can also participate in trans-splicing which gives rise to II-I type transcripts. Such “dual-purpose” spliceacceptors have been observed in the Trypanosoma brucei PAP (poly(A) polymerase) gene (Mair et al. 2000) and a trehalose synthase gene of the anhydrobiotic nematode, Aphe lenchus avenae (Goyal et al. 2005). Interestingly, a second copy of the hsp82 gene in A. ricciae, hsp82-2, which is very similar to hsp82-1 in the 5′ UTR, does not contain a splice donor site in the equivalent position, where its sequence is AC instead of GC; hsp82-2 also lacks the most downstream splice acceptor dinucleotide (deletion of AG at position −35; fig. 2) necessary for formation of a transcript like I-S. This suggests that fewer variant transcripts will result from trans-splicing of hsp82-2 pre-mRNAs and, indeed, only cDNAs equivalent to II-I of hsp82-1 have been obtained to date (unpublished data).

The discovery of trans-splicing in Rotifera extends the known phylogenetic range of this phenomenon to five metazoan phyla. In itself, this increases the likelihood of an ancient origin for trans-splicing, coupled with its loss in several lineages, because it is arguably less stringent to lose a biological function than to gain one, as the opposing hypothesis demands. The apparent similarity between rotifer and nematode SL RNAs is intriguing; one explanation might be that the sequences are homologous and that therefore
trans-splicing is ancestral to the divergence between their respective evolutionary lineages. However, previously, SL RNAs from different phylogenetic groups have shown no relatedness (Nilsen 2001) and, indeed, considerable divergence of SL RNA sequences can exist even within a single phylum, i.e., Platyhelmintes (Davis 1997), assuming this group is monophyletic. Given this degree of interfylgetic and intraphylgetic diversity, it seems unlikely that SL RNA sequence would be conserved between rotifers and nematodes. More likely is that the sequence similarity arose by chance. Furthermore, apart from patches of matching sequence in the 5’-end of the SL exon and around the splice donor site, which might be due to functional constraints, the remaining identities are scattered throughout the intron region in the rotifer-nematode SL RNA alignments and often occur at different positions in the C. elegans and Oscheius sp. CEW1 sequences (figs. 4A,B). Further difficulties become apparent when current interpretations of animal phylogeny are taken into account, because rotifers and nematodes are placed in separate taxonomic groups of Bilateria, i.e., Lophotrochozoa and Ecdysozoa, respectively (Halanych 2004). Rotiferan and Nematoda are widely separated in evolution, therefore, and if their SL RNA genes were homologous, we might expect to see a comparable or greater degree of identity between corresponding sequences of rotifers and a nearer evolutionary neighbour. However, although flatworms (Platyhelminthes) are grouped with rotifers in Platyzoa, Lophotrochozoa, there is no similarity between their respective SL RNA sequences.

At present, therefore, the significance of the apparent similarity between rotifer and nematode SL RNA sequences is difficult to assess and clearly its interpretation would be better informed with more data on the phylogenetic distribution of trans-splicing, as called for by Nilsen (2001). In the context of bdellid rotifers, it would be of immediate interest to look for trans-splicing in the other main group of rotifers, the monogononts, and in the closely related acanthocephalans (Garey et al. 1996; Mark Welch 2001). A systematic approach of this kind should result in more rapid progress of our understanding of trans-splicing evolution.

Acknowledgments

We would like to thank Claudia Ricci (University of Milan) for providing Adineta ricciae. This work was funded by BBSRC grant 8/S19912. A.T. is the Anglian Water Fellow in Biotechnology of Pembroke College, Cambridge.

Literature Cited


Billie Swalla, Associate Editor

Accepted March 16, 2005