mRNA-type introns in U6 small nuclear RNA genes: implications for the catalysis in pre-mRNA splicing

Tokio Tani[1] and Yasumi Ohshima

Department of Biology, Faculty of Science, Kyushu University, Hakozaki, Higashi-ku, Fukuoka 812 Japan

U6 small nuclear RNA is one of the spliceosomal RNAs involved in pre-mRNA splicing. In the fission yeast *Schizosaccharomyces pombe*, the U6 RNA gene was found to have an intron similar to a nuclear pre-mRNA intron, and it was proposed that the U6 intron might be inserted erroneously during pre-mRNA splicing. Using the polymerase chain reaction, we analyzed the U6 RNA genes of 52 organisms. In addition to the five species of *Schizosaccharomyces*, we found that the yeast species *Rhodotorula hasegawae* and *Rhodosporidium dacryoidum* also have mRNA-type introns in their U6 genes; however, in all the other organisms tested, we found no intron within the region of the U6 gene examined. Four introns and one intron are present in the *R. hasegawae* and *R. dacryoidum* U6 genes, respectively; and these introns are located at sites differing from the location of the *Schizosaccharomyces* U6 intron. Most of the U6 introns locate within the conserved domain, which is strikingly similar in structure to the catalytic center of the negative strand of the satellite RNA of tobacco ring spot virus. The introns of the *S. pombe* and *R. dacryoidum* U6 genes are located immediately adjacent to the nucleotides that were shown to be essential for the second step of the splicing reaction. These results support the notion that U6 RNA has a catalytic role in pre-mRNA splicing and that U6 introns originated from insertion of an excised intron during pre-mRNA splicing.

[*Key Words:* U6 small nuclear RNA; pre-mRNA splicing; mRNA-type intron; (−)sTRSV; catalytic RNA]

Received January 3, 1991; revised version accepted March 1, 1991.

The nuclei of eukaryotic cells contain a class of metabolically stable small RNAs called small nuclear RNAs (snRNAs; Busch et al. 1982). In mammals, there are six major snRNAs, U1–U6. These RNAs are complexed with proteins, forming small nuclear ribonucleoprotein particles, or snRNPs. Of these, U4 and U6 RNAs are base-paired and are present in the same snRNP (Rinke et al. 1985). The U1, U2, U5, and U4/U6 snRNPs are involved in pre-mRNA splicing as part of the spliceosome (for review, see Maniatis and Reed 1987). The U1 snRNP interacts with the 5′ splice site, and the U2 snRNP interacts with the branch site; the U5 snRNP may bind to the 3′ splice site. In contrast, there is no evidence that U4/U6 snRNP interacts directly with the pre-mRNA. It has been suggested that either U4 RNA is released from the spliceosome or that the U4/U6 interaction is destabilized prior to, or concomitant with, the 5′ splice site cleavage (Lamond et al. 1988; Blencowe et al. 1989). However, the exact function of the U4/U6 snRNP in pre-mRNA splicing has remained unknown.

U6 RNA has several unique properties. First, the U6 RNA is the most highly conserved spliceosomal snRNA (Brow and Guthrie 1988). Second, although the majority of the snRNAs have a trimethylguanosine cap at the 5′ end, the 5′-end modification of U6 RNA is a γ-monomethyl phosphate ester (Singh and Reddy 1989). Third, U6 RNA is the only spliceosomal snRNA that lacks an Sm-binding site (Reddy and Busch 1988). The assembly of U6 RNA into an snRNP depends on its association with U4 RNA, which contains the Sm-binding site. Fourth, U6 RNA is transcribed by RNA polymerase III (Dahlberg and Lund 1988; Kleinschmidt et al. 1990), whereas other snRNAs are transcribed by RNA polymerase II (Dahlberg and Lund 1988).

The apparent constraints on both the size and the sequence of U6 RNA among species argue for a central role of U6 RNA in the splicing process. We found that the U6 RNA gene of the fission yeast *Schizosaccharomyces pombe* has an intron of 50 bp in the conserved U4/U6 interaction domain (Tani and Ohshima 1989). Although U6 RNA encodes no protein, the intron has the distinctive structure for the nuclear pre-mRNA intron; that is, it contains the consensus sequences identified at the 5′ and 3′ splice sites and the branch site in nuclear pre-mRNA. This was the first example of an mRNA-type intron in an snRNA gene or a gene transcribed by RNA polymerase III. The discovery of an mRNA-type intron in a highly conserved region of the *S. pombe* U6 gene led to the hypothesis that U6 RNA functions as a catalytic element during pre-mRNA splicing and that the site of

[1]Corresponding author.

intron insertion reflects the catalytic center of U6 RNA (Brow and Guthrie 1989). In this scenario, the intron is erroneously inserted into U6 RNA in the spliceosome during pre-mRNA splicing and its cDNA replaces the chromosomal U6 gene in *S. pombe*. To test this hypothesis, we have systematically examined whether organisms other than *S. pombe* have an intron in the U6 RNA gene at or near the site of the *S. pombe* intron.

Results

Analysis of U6 genes in various organisms

To examine whether U6 genes in organisms other than *S. pombe* have an intron, DNAs from phylogenetically diverse organisms were subjected to the polymerase chain reaction (PCR) (Saiki et al. 1988). We used as PCR primers 20-mer oligonucleotides corresponding to the highly conserved regions flanking the site of the *S. pombe* U6 intron. In the case of intronless U6 genes, the expected size of the amplified products is ~50 bp. If there is an intron between the primers, the amplified products would be much larger than 50 bp. We first analyzed 24 kinds of organisms from animals, plants, and fungi. As shown in Figure 1A, organisms other than *S. pombe* produced DNAs of ~50 bp, indicating that there are no introns in the U6 genes of these organisms. We observed longer amplified products from *Drosophila*, horseshoe crab, and *Tetrahymena* DNAs in addition to the 50-bp products. However, sequence analysis of the longer products revealed that they were generated because of the tandem repetition of the U6 genes in these organisms (data not shown). The weak larger band in salmon DNA appears to be an unrelated amplification product, because it was not detected in the reaction using the primers corresponding to other regions of U6 RNA or the buffer containing 1.5 mM $MgCl_2$.

We then analyzed U6 genes of various kinds of yeasts. DNAs from 18 yeast genera were subjected to the PCR reaction. As shown in Figure 1B, DNAs from most of the yeast genera, except for *Cryptococcus laurentii*, *Rhodosporidium dacryoidum*, and *Rhodotorula hasegawae*, produced only 50-bp products, indicating the absence of an intron in the U6 genes. In the cases of *R. dacryoidum* and *R. hasegawae*, however, we observed single amplified products of ~130 and 110 bp, respectively. The weak larger product in *C. laurentii* DNA was not detected in the reaction using the buffer containing 1.5 mM $MgCl_2$, suggesting that it is an unrelated amplification product (data not shown). Therefore, we did not analyze this band further.

Structure of the R. dacryoidum *U6 gene*

Sequence analysis of the amplified product from *R. dacryoidum* DNA revealed a 76-bp insertion in the U6

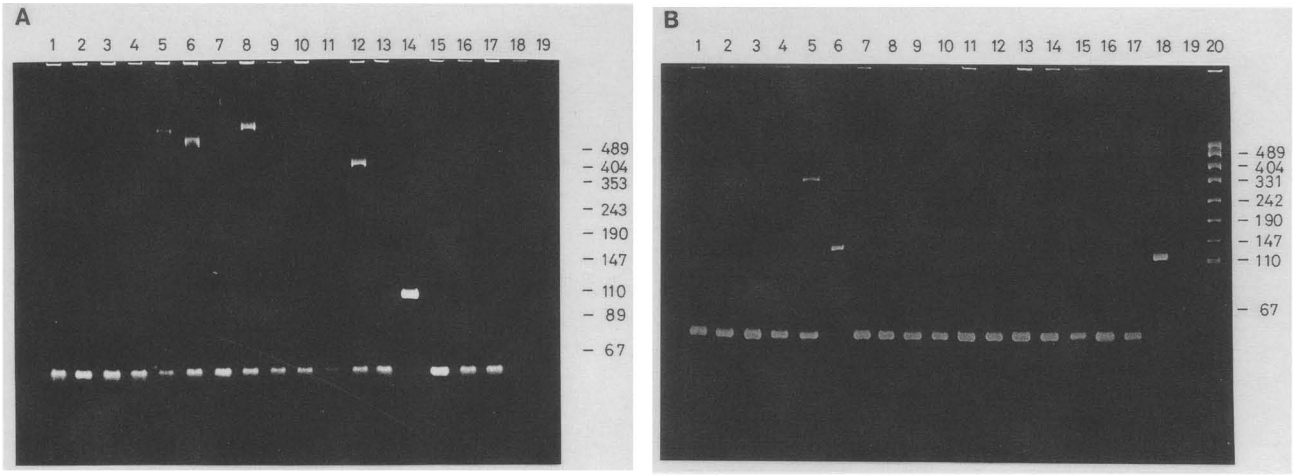

Figure 1. PCR analysis of U6 RNA genes in various organisms. (*A*) Analysis of U6 genes of phylogenetically diverse organisms. U6 sequences were amplified with an upstream primer (corresponding to positions +21–40 of *S. pombe* U6 RNA) and a downstream primer (positions +52–71). Genomic DNAs examined were as follows: Human (lane *1*), rat (lane *2*), pig (lane *3*), calf (lane *4*), salmon (lane *5*), *Drosophila melanogaster* (lane *6*), *Caenorhabditis elegans* (lane *7*), horseshoe crab (lane *8*), squid (lane *9*), sea anemone (lane *10*), sponge (lane *11*), *Tetrahymena thermophila* (lane *12*), *Physarum polycepharum* (lane *13*), *S. pombe* (lane *14*), *Dictyostelium discoideum* (lane *15*), carrot (lane *16*), rice (lane *17*), *Escherichia coli* (lane *18*), mock reaction without DNA (lane *19*). Planaria, *Neurospora crassa*, *Aspergillus oryzae*, seaweed, *Ascaris suum*, and *Lentinus* DNAs produced only 50-bp bands (data not shown). Markers are *Hpa*II fragments of pUC18. (*B*) U6 genes of various kinds of yeasts. DNAs examined were as follows: *Hanseniaspora guilliermondii* (lane *1*), *Saccharomycopsis lipolytica* (lane *2*), *Hansenula anomala* (lane *3*), *Nadsonia commutata* (lane *4*), *Cryptococcus laurentii* (lane *5*), *Rhodosporidium dacryoidum* (lane *6*), *Selenozyma peltata* (lane *7*), *Schizoblastosporion kobayasii* (lane *8*), *Filobasidium capsuligenum* (lane *9*), *Ambrosiozyma philentoma* (lane *10*), *Stephanoascus ciferrii* (lane *11*), *Lipomyces tetrasporus* (lane *12*), *Kluyveromyces lactis* (lane *13*) *Leucosporidium scottii* (lane *14*), *Arthroascus javanensis* (lane *15*), *Oosporidium margaritiferum* (lane *16*), *Leucosporidium antarcticum* (lane *17*), *Rhodotorula hasegawae* (lane *18*), mock reaction without DNA (lane *19*), pUC19/*Hpa*II (lane *20*).

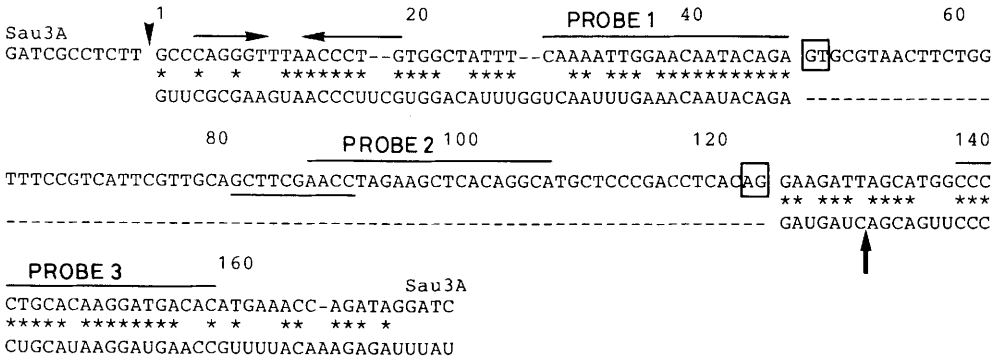

Figure 2. Nucleotide sequence of the U6 gene of *R. dacryoidum*. The *upper line* shows the noncoding strand sequence of the *Rhodosporidium* U6 gene. The *lower line* shows the sequence of *S. cerevisiae* U6 RNA. Stars indicate identical nucleotides; dashes denote gaps introduced to optimize sequence homology. The 5′ end of the coding region was determined by RNA sequencing and indicated by arrowhead. An inverted repeat, which can form a 5′ hairpin structure, is shown by the horizontal arrows. The GT and AG sequence found at the donor and acceptor sites are boxed. Overlines represent the sequences complementary to the probes used for Northern blot analysis. The vertical arrow indicates the position of the intron present in the *S. pombe* U6 RNA gene. The B-box-like sequence present in the intron is underlined.

gene. Because this insertion was present next to the 3′ end of the upstream primer, we obtained the flanking sequences of the amplified region using the inverse PCR method (Fig. 2; Ochman et al. 1989). The insertion contains GT and AG at the 5′ and 3′ ends, respectively; hence, it may be an mRNA-type intron. Interestingly, it is present 7 bp upstream of the site of the *S. pombe* intron.

To confirm that the intron-like insertion in the *R. dacryoidum* U6 gene is actually an intron, two experiments were performed. First, we used Northern blot analysis and searched for the U6 precursor containing the intron (Fig. 3A). Total RNA isolated from *R. dacryoidum* was subjected to Northern blot analysis using oligonucleotide probes complementary to the presumed exon 1, intron, or exon 2 regions. Cells were exposed to heat for 30 min prior to RNA extraction so as to accumulate the precursor (Potashkin and Frendewey 1989; Tani and Ohshima 1989; Frendewey et al. 1990). Using exon probes 1 and 3, we observed mature U6 RNA of ~115 nucleotides and a presumed precursor of ~190 nucleotides, which appeared to contain the intron. In contrast, we detected only the presumed precursor band with the intron probe 2. Second, we determined the partial nucleotide sequence of the mature U6 RNA using the dideoxy method with a reverse transcriptase (Fig. 3B). The intron-like sequence, which extends from GT to AG, was accurately removed. On the basis of these results, we concluded that the U6 gene of *R. dacryoidum* has a 76-bp mRNA-type intron, 7 bp upstream of the site of the *S. pombe* intron. The *R. dacryoidum* intron shows no significant sequence homology with that of *S. pombe*.

The R. hasegawae *U6 gene has four mRNA-type introns*

Sequence analysis of the amplified product from *R. hasegawae* DNA revealed a 63-bp intron-like insertion in the U6 gene. We then determined the complete nucleotide sequence of the *R. hasegawae* U6 gene, using the inverse PCR method (Fig. 4). Surprisingly, a comparison of the gene sequence with the *S. pombe* U6 RNA sequence suggested the presence of four mRNA-type introns, which have GT and AG at the 5′ and 3′ ends, in the *R. hasegawae* U6 gene.

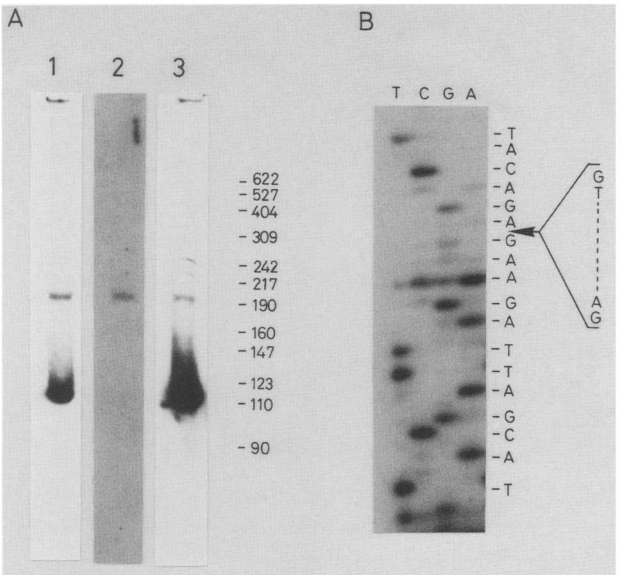

Figure 3. (*A*) Northern blot analysis. *R. dacryoidum* total RNA (4 μg) was electrophoresed on a 5% polyacrylamide/7 M urea gel, transferred to a nylon membrane (Biodyne), and hybridized separately with probe 1 (lane *1*), probe 2 (lane *2*), or probe 3 (lane *3*). Positions and sizes in bases of markers (pBR322/*Hpa*II) are shown at *right*. (*B*) Sequence near the splice junction of the *Rhodosporidium* U6 RNA. Kinase-labeled probe 3 was used as a sequencing primer. The RNA sequence predicted by the cDNA sequence is written at *right*. The arrow shows the position of the intron-like sequence.

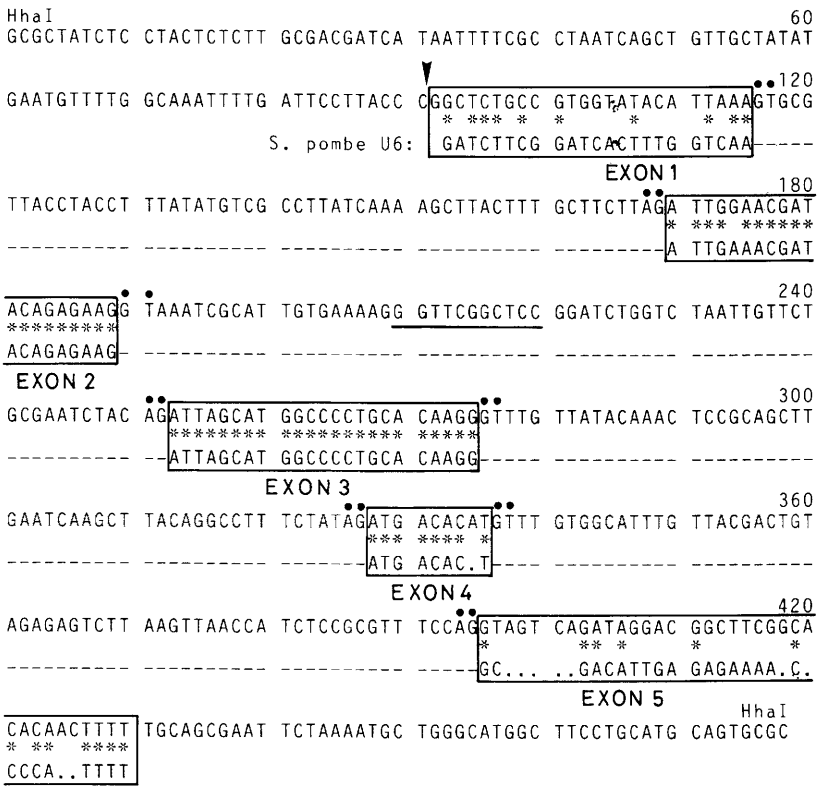

Figure 4. The nucleotide sequence of the *R. hasegawae* U6 gene. The *upper line* shows the noncoding strand sequence of the *R. hasegawae* U6 gene. The *lower line* shows the sequence of *S. pombe* U6 RNA. The identical nucleotides are denoted by asterisks (*). Gaps are introduced to optimize sequence homology and are represented by dots. The presumed exon regions are boxed. Small solid circles over nucleotides indicate the GT and AG sequences present at the ends of possible introns. The 5′ end of the coding region, which was determined by RNA sequencing, is denoted by an arrowhead. The B-box-like sequence present in the presumed second intron is underlined. The 3′ end is tentatively assigned to fall within a (T) residue tract characteristic of transcription termination sites of RNA polymerase III.

To demonstrate the presence of four introns in the *R. hasegawae* U6 gene, we determined the nucleotide sequence of the mature U6 RNA and performed Northern blot analysis to detect the U6 precursor, as described above. For the RNA sequencing, we used four distinct primers complementary to the presumed exons (Fig. 5A). We determined 87 bases from the 5′ end of the mature U6 RNA. All of four intron-like insertions were precisely removed in the mature U6 RNA. The nucleotide sequence of the mature U6 RNA was identical to that expected from the presumed exons shown in Figure 4.

Figure 5B shows the results of Northern blot analysis. Oligonucleotides complementary to presumed exons and introns were synthesized and used as probes. Total RNA used in the experiment, except for lane 1b, was isolated from *R. hasegawae* that had been exposed to heat for 30 min prior to RNA extraction to inhibit splicing. Using the exon probes, we detected major bands of ~115 nucleotides, which appeared to correspond to the mature U6 RNA, and several slower and minor bands (odd-numbered lanes). In contrast, only the minor bands were detected with the intron probes (even-numbered lanes), thus suggesting that the minor bands correspond to the U6 RNA precursor and intermediates of the splicing reaction. The length of the slowest minor RNA (~340 bases), which was hybridized with both the intron and exon probes, is consistent with that of the presumed U6 RNA precursor (342 bases). As the presumed exon 4 is only 9 bp in length, the exon 4 probe (probe 7) is extended by several bases of exons 3 and 5; therefore, we did not observe the possible precursor band using the exon 4 probe (lane 7). The minor bands were not observed using the total RNA from nonheated *R. hasegawae* (lane 1b), a finding that also suggests that the minor bands correspond to the U6 RNA precursor and the intermediates. On the basis of the results described above, we concluded that four mRNA-type introns are present in the *R. hasegawae* U6 gene.

U6 introns are not present in other species of Rhodosporidium *and* Rhodotorula

To investigate whether or not U6 genes of close relatives of *S. pombe*, *R. dacryoidum*, and *R. hasegawae* have an intron, DNAs from four other species of *Schizosaccharomyces*, five other species of *Rhodosporidium*, and another species of *Rhodotorula* were subjected to PCR analysis as described above. As reported recently (Frendewey et al. 1990; Reich and Wise 1990), U6 genes in other *Schizosaccharomyces* species (*S. japonicus*, *S. malidevoruns*, *S. japonicus versatilis*, and *S. octosporus*) were found to contain mRNA-type introns of well-conserved sequences at the same site as the *S. pombe* U6 intron (data not shown). In contrast, no introns were found in the U6 genes of the *Rhodosporidium* species tested, except for *R. dacryoidum* (Fig. 6). Similarly, we detected a 50-bp amplified product using DNA from *Rhodotorula pailida*, a close relative of *R. hasegawae*, indicating that there is no intron in the region between two primers used (data not shown). These results suggest that the *R. dacryoidum* and *R. hasegawae* U6 genes have

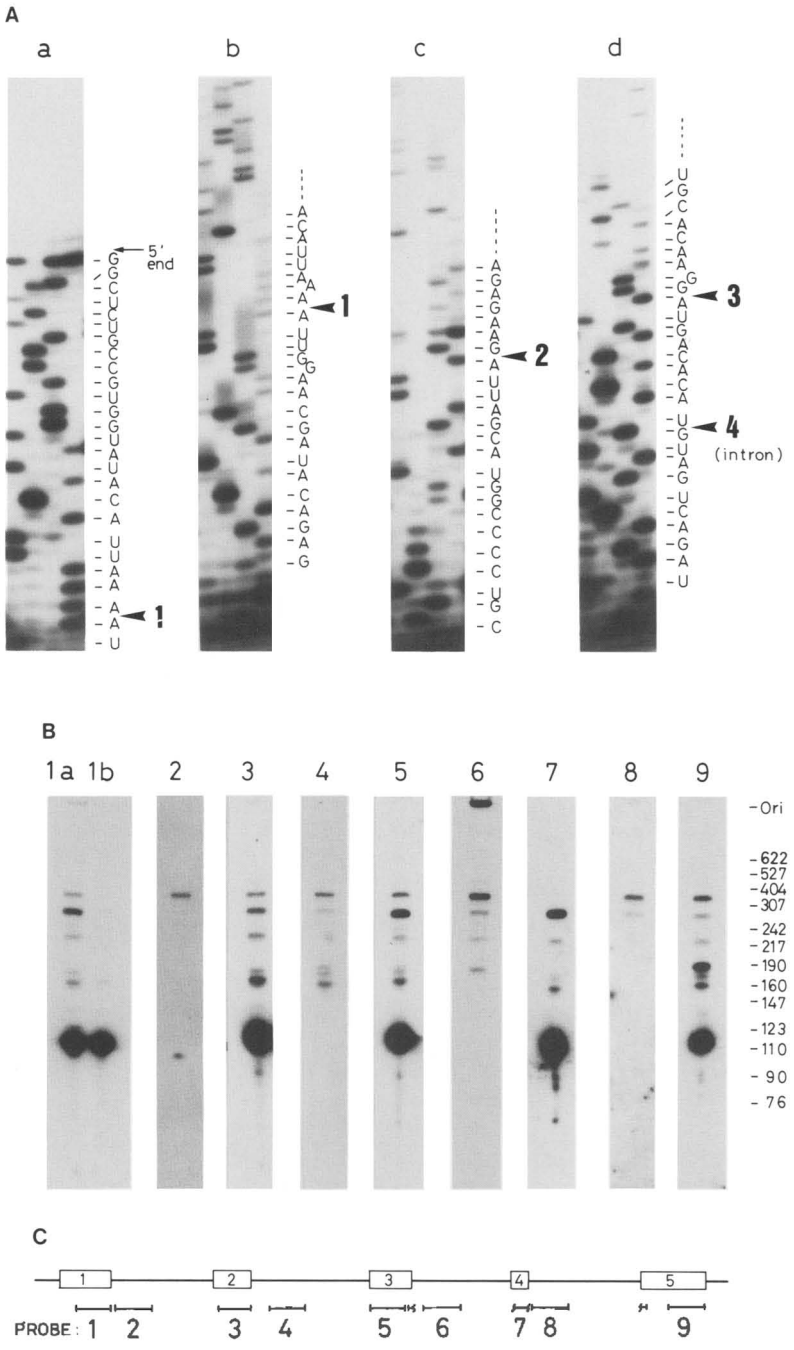

Figure 5. (*A*) Nucleotide sequence of the mature U6 RNA of *R. hasegawae*. Ten micrograms of total RNA from nonheated *R. hasegawae* was used for RNA sequencing. The kinase-labeled primer complementary to the presumed exon 2 (*a*), exon 3 (*b*), exon (3)-4-(5) (*c*), or exon 5 (*d*) was used for the reaction. The RNA sequences predicted by the cDNA sequence are at *right*. Arrowheads show the positions of the intron sequences. (*B*) Northern blot analysis of *R. hasegawae* total RNA. Total RNA (4 μg) was electrophoresed on a 5% polyacrylamide/7 M urea gel, transferred to a nylon membrane (Biodyne), and hybridized separately with the oligonucleotides probes complementary to the presumed exons or introns. The regions complementary to each probe are represented schematically in *C*. The lane numbers correspond to the probe numbers used for hybridization. The bands at the top of lanes *1a* and *6* correspond to the pockets of the gel. Markers were kinase-labeled *Hpa*II fragments of pBR322 DNA. (*C*) Schematic representation of the exon/intron organization of the *R. hasegawae* U6 gene.

gained the mRNA-type introns rather recently during evolution.

Discussion

Evolutionary origin of the U6 introns

We examined the U6 genes of 52 organisms using PCR analysis and found that U6 genes in most of the organisms tested had no introns. Therefore, it is likely that mRNA-type introns found in some of the U6 genes were not present in the ancestral U6 gene and arose independently during the course of evolution. PCR analysis using DNAs from several species of genera *Rhodosporidium* and *Rhodotorula* showed that only the *R. dacryoidum* and *R. hasegawae* U6 genes have mRNA-type introns, suggesting that those introns have arisen rather recently during evolution. In contrast, in the case of U6 genes in *Schizosaccharomyces*, the intron insertion appears to have occurred earlier, before separation of the species within the genus *Schizosaccharomyces*. However, we cannot exclude the possibility that at least six introns were present in the ancestral U6 gene and they were lost in most organisms except for *Schizosaccharo-*

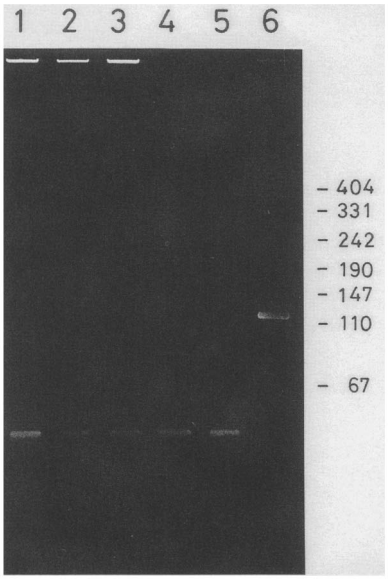

Figure 6. Analysis of U6 genes of *Rhodosporidium* species. DNAs from six species of *Rhodosporidium* were subjected to PCR analysis. Species examined were as follows: *R. diobovatum* (lane *1*), *R. malvinellum* (lane *2*), *R. toruloides* (lane *3*), *R. capitatum* (lane *4*), *R. bisporidiis* (lane *5*), and *R. dacryoidum* (lane *6*). Markers are *Hpa*II fragments of pUC19.

myces, *R. dacryoidum*, and *R. hasegawae*. If this is the case, it is necessary to suppose that two of six introns resided only 3 bp apart, which is an unprecedented occurrence.

Structural similarity between U6 RNA and the catalytic domain of the negative strand of the tobacco ring spot virus satellite RNA

We found four mRNA-type introns in the *R. hasegawae* U6 gene and one intron in the *Schizosaccharomyces* and *R. dacryoidum* U6 genes, respectively. Most of the mRNA-type introns are present in a restricted region in the presumed secondary structure of U6 RNA (Fig. 7A). Two hairpin structures can be formed in common in U6 RNA from human to *Saccharomyces cerevisiae*. One possible hairpin structure at the 5′ end of U6 RNA was shown to be required for the capping of U6 RNA with a γ-methyltriphosphate cap (Singh et al. 1990). The other possible hairpin structure proposed here lies in the highly conserved region downstream of the invariant sequence ACAGAGA. As shown in Figure 7A, most U6 introns are within the conserved region, including the second hairpin structure.

Interestingly, we found that the conserved region of U6 RNA has a remarkable structural similarity with the catalytic domain of the negative strand of the satellite RNA of tobacco ring spot virus [(−)sTRSV]. The (−)sTRSV is 359 bases in length and undergoes self-catalyzed cleavage during replication (Hampel and Tritz 1989). Ligation of cleaved segments of (−)sTRSV has also been observed in vitro (Buzayan et al. 1986). The cleavage reaction in vitro can be carried out under very mild conditions of temperature, pH, and salt similar to physiological conditions (Hampel et al. 1990). The catalytic domain of (−)sTRSV has a novel structural motif different from the well-known "hammerhead" model proposed for other virus-related RNA or viroids (Forster and Symons 1987). The minimal catalytic complex of (−)sTRSV consists of a 50-base catalytic RNA, which contains a hairpin structure, and a 14-base substrate RNA (Fig. 7B). The 50-base catalytic RNA has a sequence ACAGAGAAG close to the cleavage site. The identical ACAGAGAAG sequence and a similar possible hairpin structure are also found in the U6 RNA. Interestingly, deletion or substitution mutations that destroy the possible second hairpin structure of U6 RNA abolished

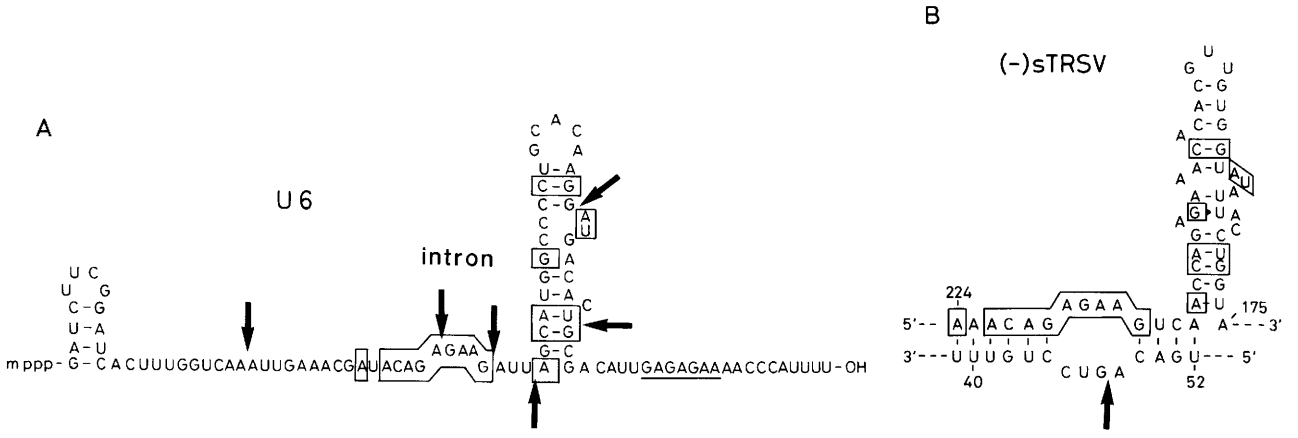

Figure 7. (*A*) Positions of mRNA-type introns in U6 RNA. The possible secondary structure of *S. pombe* U6 RNA is shown. Arrows indicate the positions of introns found in the *S. pombe*, *R. dacryoidum*, and *R. hasegawae* U6 genes. The homologous sequences between U6 RNA and the catalytic domain of (−)sTRSV are boxed. The sequence involved in the possible base-pairing with the 5′ end of U2 RNA is underlined. (*B*) The structure of the catalytic domain of (−)sTRSV. The arrow denotes the cleavage site. The numbering of nucleotides is according to Hampel and Tritz (1989). The boxes delineate nucleotide sequences identical to the *S. pombe* U6 RNA sequence.

splicing activity in a splicing complementation assay (Fabrizio et al. 1989; Fabrizio and Abelson 1990; Vankan et al. 1990). In addition, single nucleotide substitutions within the invariant ACAGA sequence or the sequence around the insertion site of the *S. pombe* U6 intron, which is located at the base of the hairpin structure, block splicing either at the first or second step in vitro (Fabrizio and Abelson 1990). In vivo analysis using *S. cerevisiae* also shows most of the single lethal nucleotide substitutions cluster in the same two domains (Madhani et al. 1990). Furthermore, the catalytic RNA of (−)sTRSV cleaves the 5′ side of the GU sequence, which is shared by the 5′ end of an mRNA-type intron. Preliminary experiments have shown that a synthetic RNA that is 49 bases in length and has 78% sequence identity with the conserved region of *S. pombe* U6 RNA can cleave a 15-base RNA substrate that contains GU sequence in vitro (T. Tani and Y. Ohshima, unpubl.). Taken together, we propose that the conserved region that includes the second hairpin structure of U6 RNA functions as a catalytic domain during pre-mRNA splicing.

A model for the catalysis of pre-mRNA splicing by U6 RNA

Catalytic RNA of (−)sTRSV is thought to abstract a proton from the 2′-hydroxyl group adjacent to the susceptible phosphodiester bond and cleaves the bond to give a 2′-,3′-cyclic phosphate and 5′-OH termini. This is not the case in pre-mRNA splicing. We propose that in the first step of the pre-mRNA splicing, U6 RNA abstracts a proton from the 2′-hydroxyl group of adenosine in the branch site, leading to cleavage of the 5′ splice site and formation of 3′-OH and a 2′-,5′-phosphodiester bond. Using the psoralen cross-linking method, U6 RNA was shown to base-pair with U2 RNA, which also base-pairs with the branch site sequence of a pre-mRNA during pre-mRNA splicing (Hausner et al. 1990). The region of U6 RNA involved in the base-pairing with the 5′ end of U2 RNA is located downstream from the possible "catalytic domain" of U6 RNA. This leads to the close positioning of the catalytic domain of U6 RNA with the branch site in the splicing reaction. In the second step of the splicing reaction, U6 RNA may possibly abstract a proton from the 3′-hydroxyl group of the 5′ exon, leading to excision of a lariat intron and ligation of the exons. Because U6 RNA has no sequence complementarity with the consensus sequences of the 5′ and 3′ splice sites, close interaction between U6 RNA and the splice sites would require some splicing factor(s), such as U1 snRNP.

Reverse splicing of an excised intron into U6 RNA

In our proposal, the catalytic domain of U6 RNA is positioned in close proximity to the 3′ splice site at the moment of intron excision and exon ligation. Thus, reintegration of an excised intron into the U6 RNA can occur via reverse splicing. In that case, the position of mRNA-type introns marks the catalytic domain of U6 RNA as proposed by Brow and Guthrie (1989). It is noteworthy that the introns of *S. pombe* and *R. dacryoidum* are present immediately adjacent to the nucleotides that are essential for the second step of the splicing reaction (Fabrizio and Abelson 1990). This is consistent with the idea that the U6 introns arose via an aberrant reversal of the second catalytic step of the splicing reaction. Reverse splicing of an excised intron into RNA was demonstrated with group I and group II self-splicing introns (Woodson and Cech 1989; Augustin et al. 1990; Mörl and Schmelzer 1990).

Does U4 RNA function as an antisense repressor?

The region that forms the possible hairpin structure in U6 RNA participates in a base-pairing interaction with U4 RNA (Brow and Guthrie 1988). Therefore, U6 RNA cannot form the hairpin structure in U4/U6 snRNP. After dissociation of U4 RNA, however, U6 RNA could turn into an "active form." The U4/U6 interaction was shown to be destabilized prior to the 5′ splice site cleavage (Lamond et al. 1988; Blencowe et al. 1989). Taken together, as indicated by Brow and Guthrie (1989), we suggest that U4 RNA functions as an antisense RNA to suppress the catalytic activity of U6 RNA by preventing the formation of the hairpin structure until pertinent spliceosome assembly is established.

B-box-like sequences in the U6 introns

Why have we found the U6 gene with an intron only in yeasts so far? Multiple copies of U6 genes are present in many organisms (Dahlberg and Lund 1988). In these organisms, gene conversion would prevent maintenance of the U6 intron even if it has been newly acquired in one locus. In contrast, only a single copy of the U6 gene has been found in yeasts (Brow and Guthrie 1988; Tani and Ohshima 1989). Therefore, if a presumed reverse-transcribed product of U6 RNA with an intron replaced the chromosomal U6 gene, the intron-containing gene would survive. Frendewey et al. (1990) reported that *Schizosaccharomyces* U6 introns contain a sequence that matches the B-box consensus for RNA polymerase III. Interestingly, similar sequence is present in the introns of the *R. dacryoidum* and *R. hasegawae* U6 genes (GCTTCGAACC in the *R. dacryoidum* U6 intron and GGTTCGGCTCC in the *R. hasegawae* U6 second intron, where the underlined bases match the eukaryotic B-box consensus). Although further work is needed to determine whether these B-box-like sequences are functional, the possibility that this element improves U6 expression offers an alternative explanation why the U6 introns become fixed in the genome, as suggested by Frendewey et al. (1990).

It is surprising that the gene for U6 RNA, with only 100 bases, has four introns in *R. hasegawae*. We do not know whether other genes in *R. hasegawae* have many introns; however, a mechanism that generates and maintains many introns is likely to be present in *R. hase-*

gawae. Northern blot analysis of *R. hasegawae* total RNA suggested the presence of several splicing intermediates. In the analysis using the intron 1 or 4 probe, most of these intermediates were not detected. Introns 1 and 4 therefore appear to be removed efficiently during the splicing process. Characterization of the putative intermediates is under way.

U3 and U6 introns

Myslinski et al. (1990) reported that an mRNA-type intron is also present in two genes coding for the *S. cerevisiae* U3 RNA (snR17), which is thought to be involved in rRNA processing. U3 RNA is no more a spliceosomal RNA than is mRNA. Therefore, the hypothesis described in the previous section does not hold for introns present in U3 RNA genes. The finding of a U3 RNA intron may extend the possibility of insertion of an mRNA-type intron to a wide variety of RNA species. This leads to the idea that the introns in the U6 and U3 genes and some introns in the pre-mRNA genes have arisen in a basically similar manner, that is, the reverse splicing reaction of an excised intron. The introns in the U6 RNA genes are unique in that the U6 RNA is possibly involved in the removal of introns of the same type and that the site of insertion may be related to functional sites of the RNA. Alternatively, introns of the U3 genes and the pre-mRNA genes might arise by an unknown mechanism completely different from that of the U6 introns.

Identification of aberrant U6 RNA containing an intron from a pre-mRNA, possibly present only in a few molecules in the nucleus, will be required for further testing of the hypothesis.

Materials and methods

Preparation of total DNA

Total DNAs from various organisms were prepared as described (Maniatis et al. 1982) or provided by other researchers. For preparation of yeast DNAs, yeast cells were vortexed with glass beads for 10 min in 50 mM Tris-HCl (pH 8.0) and 50 mM EDTA, and then incubated with 0.1% SDS and 100 μg/ml of proteinase K at 50°C for 2 hr. DNA was precipitated with two volumes of ethanol and 2.5 M ammonium acetate. Various kinds of yeast strains were purchased from the Institution for Fermentation, Osaka, Japan.

Preparation of total RNA

R. dacryoidum and *R. hasegawae* were grown to mid-log phase at 15°C and 23°C, respectively, and then heated for 30 min at 30°C and 35°C to accumulate the U6 precursor. Total RNAs were isolated using glass beads (Domdey et al. 1984).

PCR analysis

PCR reactions were carried out in a solution (100 μl) containing 50 mM KCl, 10 mM Tris-HCl (pH 8.5), 2.5 mM $MgCl_2$, 0.1% Triton X-100, 200 mM dNTP, 0.1 μM of each primer, 2.5 units of *Taq* polymerase, and 1 μg of total DNA. Each amplification cycle consisted of denaturation for 1 min at 90°C, annealing for 2 min at 50°C, and extension time of 2 min at 70°C. Thirty cycles were carried out. Amplified products were analyzed on a 10% polyacrylamide gel.

Inverse PCR

Total DNAs from *R. dacryoidum* and *R. hasegawae* were digested with *Sau*3A and *Hha*I, respectively. The digested DNAs were circularized with T4 DNA ligase in a dilute DNA concentration and subjected to PCR using the reverse primers corresponding to the intron region. Conditions of amplification were the same as described above. The amplified product was subcloned into the M13 vector and sequenced using a 370A DNA sequencer (Applied Biosystems).

Northern blot analysis and RNA sequencing

For Northern blot analysis, total RNA (4 μg) was electrophoresed on a 5% polyacrylamide gel and transferred to a nylon membrane (Biodyne) by capillary action. Hybridization was done at 42°C in a solution containing 6× SSC, 50 mM sodium phosphate, 5× Denhardt's solution, 0.1% SDS, 2 mM EDTA, and 20 μg/ml of tRNA. The filter was washed at room temper-

Table 1. *Oligonucleotides*

Primers for the PCR analysis
Upstream: 5′-CAAATTGAAACGATACAGAG-3′
Downstream: 5′-TCATCCTTGTGCAGGGGCCA-3′
Primers for the inverse PCR
Reverse primers for the *R. dacryoidum* U6 gene:
5′-TCGAAGCTGCAACGAATGAC-3′,
5′-CTCACAGGCATGCTCCCGAC-3′
Reverse primers for the *R. hasegawae* U6 gene:
5′-GATCCGGAGCCGAACCCTT-3′,
5′-TGGTCTAATTGTTCTGCGAA-3′
Probes for Northern blot analysis
Probes for *R. dacryoidum* total RNA
Exon 1:
5′-TCTGTATGGTTCCAATTTTG-3′
Intron:
5′-TGCCTGTGAGCTTCTAGGTT-3′
Exon 2:
5′-GTGTCATCCTTGTGCAGGGG-3′
Probes for *R. hasegawae* total RNA
Exon 1:
5′-TTTAATGTATACCACGGC-3′
Intron 1:
5′-ATAAAGGTAGGTAACGCAC-3′
Exon 2:
5′-CTTCTCTGTATCGTTCCA-3′
Intron 2:
5′-GGAGCCGAACCCTTTTCAC-3′
Exon 3:
5′-TGTGCAGGGGCCATGCTAAT-3′
Intron 3:
5′-AAGCTGCGGAGTTTGTATAA-3′
Exon (3)-4-(5):
5′-CTACATGTGTCATCCTTG-3′
Intron 4:
5′-TCGTAACAAATGCCACAAAC-3′
Exon 5:
5′-GTTGTGTGCCGAAGCCGTC-3′

ature for 20 min in 6× SSC/0.1% SDS, and then at 42°C and 50°C for 20 min and 10 min, respectively. Autoradiography was carried out at −70°C with an intensifying screen.

RNA sequencing was done by the dideoxy method with a reverse transcriptase. The oligonucleotide primer complementary to the presumed exon sequence was kinase labeled and annealed to 10 μg of total RNA. The sample was split into four aliquots and used to direct cDNA synthesis in the presence of ddATP, ddGTP, ddCTP, or ddTTP. Samples were run on an 8% polyacrylamide/8.3 M urea gel.

Oligonucleotides

Oligonucleotides used in this study are listed in Table 1.

Acknowledgments

We thank Drs. S. Iwanaga, Y. Watanabe, Y. Tanaka, T. Kuroiwa, H. Uchimiya, K. Hasunuma, K. Kitamoto, S. Kusakabe, M. Ohuchida, K. Shishido, and Mr. S. Tashiro for providing the genomic DNAs or organisms, and M. Ohara for comments. This research was supported by grants from the Ministry of Education, Science and Culture of Japan.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

Note added in proof

Sequence data described in this paper have been submitted to the EMBL/Gen Bank Data Libraries.

References

Augustin, S., M.W. Müller, and R.J. Schweyen. 1990. Reverse self-splicing of group II intron RNAs in vitro. *Nature* **343:** 383–386.

Blencowe, B.J., B.S. Sproat, V. Ryder, S. Barabino, and A.I. Lamond. 1989. Antisense probing of the human U4/U6 snRNP with biotinylated 2′-OMe RNA oligonucleotides. *Cell* **59:** 531–539.

Brow, D.A., and C. Guthrie. 1988. Spliceosomal RNA U6 snRNA is remarkably conserved from yeast to mammal. *Nature* **334:** 213–218.

———. 1989. Splicing a spliceosomal RNA. *Nature* **337:** 14–15.

Busch, H., R. Reddy, L. Rothblum, and Y.C. Choi. 1982. SnRNAs, SnRNPs, and RNA processing. *Annu. Rev. Biochem.* **51:** 617–654.

Buzayan, J.M., W.L. Gerlach, and G. Bruening. 1986. Non-enzymatic cleavage and ligation of RNAs complementary to a plant virus satellite RNA. *Nature* **323:** 349–353.

Dahlberg, J.E. and E. Lund. 1988. The genes and transcription of the major small nuclear RNAs. In *Structure and function of major and minor small nuclear ribonucleo protein particles* (ed. M.L. Birnstiel), pp. 38–70. Springer-Verlag, Berlin/Heidelberg/New York/London/Paris/Tokyo.

Domdey, H., B. Apostol, R-J. Lin, A. Newman, E. Brody, and J. Abelson. 1984. Lariat structures are in vivo intermediates in yeast pre-mRNA splicing. *Cell* **39:** 611–621.

Fabrizio, P. and J. Abelson. 1990. Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precursor messenger RNA splicing. *Science* **250:** 404–409.

Fabrizio, P., D.S. McPheeters, and J Abelson. 1989. In vitro assembly of yeast U6 snRNP: A functional assay. *Genes & Dev.* **3:** 2137–2150.

Forster, A.C., and R.H. Symons. 1987. Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites. *Cell* **49:** 211–220.

Frendewey, D., I. Barta, M. Gillespie, and J. Potashkin. 1990. Schizosaccharomyces U6 genes have a sequence within their introns that matches the B box consensus of tRNA internal promoters. *Nucleic Acids Res.* **18:** 2025–2032.

Hampel, A. and R. Tritz. 1989. RNA catalytic properties of the minimum (−)sTRSV sequence. *Biochemistry* **28:** 4929–4933.

Hampel, A., R. Tritz, M. Hicks, and P. Cruz. 1990. "Hairpin" catalytic RNA model: Evidence for helices and sequence requirement for substrate RNA. *Nucleic Acids Res.* **18:** 299–304.

Hausner, T-P., L.M. Giglio, and A.M. Weiner. 1990. Evidence for base-pairing between mammalian U2 and U6 small nuclear ribonucleoprotein particles. *Genes & Dev.* **4:** 2146–2156.

Kleinschmidt, A.M., T. Pederson, T. Tani, and Y. Ohshima. 1990. An intron-containing Schizosaccharomyces pombe U6 RNA gene can be transcribed by human RNA polymerase III. *J. Mol. Biol.* **211:** 7–9.

Lamond, A.I., M.M. Konarska, P.J. Grabowski, and P.A. Sharp. 1988. Spliceosome assembly involves the binding and release of U4 small nuclear ribonucleoprotein. *Proc. Natl. Acad. Sci.* **85:** 411–415.

Madhani, H.D., R. Bordonné, and C. Guthrie. 1990. Multiple roles for U6 snRNA in the splicing pathway. *Genes & Dev.* **4:** 2264–2277.

Maniatis, T. and R. Reed. 1987. The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. *Nature* **325:** 673–678.

Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: A laboratory manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Mörl, M. and C. Schmelzer. 1990. Integration of group II intron bl1 into a foreign RNA by reversal of the self-splicing reaction in vitro. *Cell* **60:** 629–636.

Myslinski, E., V. Ségault, and C. Branlant. 1990. An intron in the genes for U3 small nuclear RNAs of the yeast Saccharomyces cerevisiae. *Science* **247:** 1213–1216.

Ochman, H., J.W. Ajioka, D. Garza, and D.L. Hartl. 1989. Inverse polymerase chain reaction. In *PCR technology* (ed. H.A. Erlich), pp. 105–111, Stockton Press, New York/London/Tokyo/Melbourne/Hong Kong.

Potashkin, J. and D. Frendewey. 1989. Splicing of the U6 RNA precursor is impaired in fission yeast pre-mRNA splicing mutant. *Nucleic Acids Res.* **17:** 7821–7831.

Reddy, R. and H. Busch. 1988. Small nuclear RNAs: RNA sequences, structure, and modifications. In *Structure and function of major and minor small nuclear ribonucleoprotein particles* (ed. M.L. Birnstiel), pp. 1–37, Springer-Verlag, Berlin/Heidelberg/New York/ London/Paris/Tokyo.

Reich, C. and J.A. Wise. 1990. Evolutionary origin of the U6 small nuclear RNA intron. *Mol. Cell. Biol.* **10:** 5548–5552.

Rinke, J., B. Appel, M. Digweed, and R. Lührmann. 1985. Localization of a base paired interaction between small nuclear RNAs U4 and U6 in intact U4/U6 ribonucleoprotein particles by psoralen cross-linking. *J. Mol. Biol.* **185:** 721–731.

Saiki, R.K., D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis, and H.A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239:** 487–491.

Singh, R. and R. Reddy. 1989. γ-Monomethyl phosphate: A cap structure in spliceosomal U6 small nuclear RNA. *Proc. Natl. Acad. Sci.* **86:** 8280–8283.

Singh, R., S. Gupta, and R. Reddy. 1990. Capping of mammalian

U6 small nuclear RNA in vitro is directed by a conserved stem-loop and AUAUAC sequence: Conversion of a non-capped RNA into a capped RNA. *Mol. Cell. Biol.* **10:** 939–946.

Tani, T. and Y. Ohshima. 1989. The gene for the U6 small nuclear RNA in fission yeast has an intron. *Nature* **337:** 87–90.

Vankan, P.,C. McGuigan, and I.W. Mattaj. 1990. Domains of U4 and U6 snRNAs required for snRNP assembly and splicing complementation in Xenopus oocyte. *EMBO J.* **9:** 3397–3404.

Woodson, S.A. and T.R. Cech. 1989. Reverse self-splicing of the *Tetrahymena* group I intron: Implication for the directionality of splicing and for intron transposition. *Cell* **57:** 335–345.

mRNA-type introns in U6 small nuclear RNA genes: implications for the catalysis in pre-mRNA splicing.

T Tani and Y Ohshima

Genes Dev. 1991, **5:**
Access the most recent version at doi:10.1101/gad.5.6.1022