Curr Genet (1996) 29: 582–586

ORIGINAL PAPER

Martin Grube · Andrea Gargas
Paula T. DePriest

A small insertion in the SSU rDNA of the lichen fungus *Arthonia lapidicola* is a degenerate group-I intron

Received: 30 Deember 1994/1 November 1995

Abstract Insertions of less than 100 nt occurring in highly conserved regions of the small subunit ribosomal DNA (SSU rDNA) may represent degenerate forms of the group-I introns observed at the same positions in other organisms. A 63-nt insertion at SSU rDNA position 1512 (relative to the *Escherichia coli* SSU rDNA) of the lichen-forming fungus *Arthonia lapidicola* can be folded into a secondary structure with two stem loops and a pairing of the insertion and flanking sequences. The two stem loops may correspond to the P1 and P2, and the insertion-flanking pairing to the P10, of a group-I intron. Considering these small insertions as degenerate introns provides important clues to the evolution and catalytic function of group-I introns.

Keywords Ribosomal DNA · Small subunit · 18*s* · Degenerate introns · Ascomycetes

Introduction

Insertions are reported from numerous positions in the nuclear small subunit ribosomal DNA (SSU rDNA) (Garrett et al. 1994; Gargas et al. 1995 b). Many of these insertions have been identified as group-I introns. However, a few insertions cannot be classified as this type of intron because they are too small and lack the necessary secondary structure and sequence elements. For example, insertions of 67 and 78 nucleotides have been reported from the SSU rDNA of the ascomycetous fungi *Porpidia crustulata* (Gargas et al. 1995 b) and *Phialophora americana* (Rogers et al. 1993), respectively. Our investigations have identified a third insertion of a similar size from the SSU rDNA of the lichen-forming ascomycete *Arthonia lapidicola*. Since each of the three small insertions is located at a position where group-I introns are known from the SSU rDNA of other organisms (Gargas et al. 1995 b), it is likely that they are "proto-" or "degenerate" introns, the latter proposed by Gargas et al. (1995 b). In the present study we have compared the primary sequence and putative secondary structures of the small insertion from *A. lapidicola* to those from *P. crustulata* and *P. americana*. Since these small insertions occur in phylogenetically related organisms (Gargas et al. 1995 a), with two at the same SSU rDNA position, it is reasonable that they are related either as descendants of a common ancestral sequence or as products of similar molecular mechanisms.

Materials and methods

Collection of Material. Five to ten ascomata (fungal reproductive structures) were selectively removed from a lichen thallus of *A. lapidicola*.

DNA isolation and PCR amplification. DNA was isolated from the ascomata following the protocol of Grube et al. (1995) using a guanidine thiocyanate extraction buffer, as in Armaleo and Clerc (1991), and DNA precipitation onto glassmilk (Geneclean II, Bio 101). The DNA was resuspended in dH_2O and 0.1–10 ng of genomic DNA used for PCR amplification, as reported in Grube et al. (1995), between primers 5′-GCTCTTTCTTGATTTTGT-3′ (NSR 1241-5′, DePriest 1993) in the small subunit (SSU) and 5′-TCCTCCGCTTATTGATATGC-3′ (ITS4, White et al. 1990) in the large subunit (LSU). The PCR product sizes were determined by comparison to polynucleotide size markers with agarose-gel electrophoresis.

DNA sequencing. PCR products, prepared as above, were sequenced with the PRISM Ready Reaction DyeDeoxy Terminator

M. Grube
Institut für Botanik, Karl Franzens Universität, Holteigasse 6, A-8010 Graz, Austria

A. Gargas · P.T. DePriest (✉)
Department of Botany, NHB-166, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, USA

Communicated by H. Bertrand

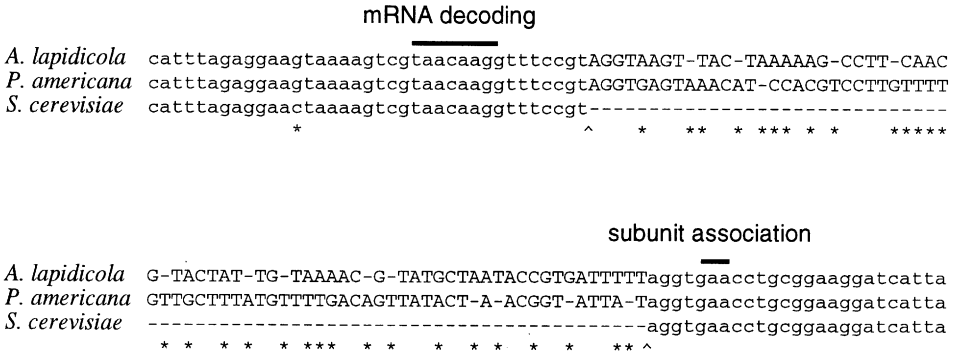

Fig. 1 Alignment of 3′ SSU rDNA sequences from the ascomycetous fungi *Arthonia lapidicola*, *Phialophora americana* and *Saccharomyces cerevisiae*. *A. lapidicola* and *P. americana* contain insertions of 63 and 67 nt, respectively, relative to the sequence of *S. cerevisiae*. The insertions from *A. lapidicola* and *P. americana* are 57% similar in sequence. Gaps inserted for sequence alignment are indicated by *dashes* (−). Differences in nucleotide sequence are indicated by *asterisks* (*). The proposed functions for included regions, mRNA decoding and subunit association, are indicated above the aligned sequences

Cycle Sequencing Kit (Applied Biosystems) from a number of oligonucleotide primers including 5′-GGAAGTAAAAG-TCGTAACAAGG-3′ (ITS 5, White et al. 1990), ITS 4 (as above) and 5′-AATGATCCTTCCGCAGGT-3′ (NSR 1797-3′, DePriest 1993) following the manufacturers instructions. The cycle sequencing reactions were run on a polyacrylamide gel in a 373A Automatic Sequencing Apparatus (Applied Biosystems). The resulting sequences were aligned to reference sequences using PileUp software (GCG, Intelligenetics, Madison, Wis.) with gap penalty = 0.0 and gap extension = 0.0.

Results and discussion

In the present study we have identified a 63-nt insertion from the lichen-forming *A. lapidicola*. By comparison to SSU rDNA sequences, the insertion is located between nucleotides 1771 and 1772 of *Saccharomyces cerevisiae* SSU rDNA (Fig. 1; insertion position 1512 in the nomenclature of Gargas et al. 1995 b). Splicing removal of the insertion sequence during rRNA processing would restore a sequence homologous to that of *S. cerevisiae* SSU rRNA. The putative junction has the U at the 5′ splice site characteristic of group-I introns (Cech 1990; Michel and Westhof 1990; Gargas et al. 1995 b). However, the insertion sequence is too small to form the nine stem-loop regions, and lacks the G at the 3′ splice site and conserved core elements typical of group-I introns (Davies et al. 1982; Michel et al. 1982; Cech 1990). The sequence can be folded into a putative secondary structure with two stem-loop regions and pairing of a flanking and an internal region (Fig. 2 A).

The nucleotide sequence of this insertion was compared to the nucleotide sequences of insertions from the SSU rDNA of two ascomycetous fungi; a 67-nt insertion from the deuteromycete *P. americana* (Rogers et al. 1993) and a 78-nt insertion from the lichen-forming *P. crustulata* (Gargas et al. 1995 b). The *P. americana* insertion is at position 1512 where the small insertion occurs in *A. lapidicola* and group-I introns are reported from green algae (Wilcox et al. 1991); the *P. crustulata* insertion is at position 516 where group-I introns are reported from green algae and amoebae (Embley et al. 1992; Huss et al. 1993 a; Gast et al. 1994). Although similar in size, these small insertions differ in primary sequence (up to 56% sequence divergence) and are not unambiguously homologous. Even the two occurring at the same SSU rDNA position were only 57% similar in sequence (Fig. 1). However, both insertions could be folded into a putative secondary structure strikingly similar to that of *A. lapidicola* (Fig. 2 B, C).

The two stem-loops in the putative secondary structures can be interpreted as P1 and P2 pairings and the pairing of flanking regions to internal sequences as the P10 pairing characteristic of a group-I intron (Davies et al. 1982; Michel et al. 1982; Cech et al. 1994). The P1 pairing of group-I introns is a helix composed of flanking exons and intron sequences with a specific pairing of the U immediately preceding the intron/exon junction to a G in the introns internal guide sequence (IGS) (Waring and Davis 1984). A U•G pairing is found in the putative secondary structures of each of the three insertions (Fig. 2) and therefore the helix is designated as a P1 comparable to that of group-I introns. The remaining stem-loop lacks features specific to the P3, P4, P5 or P7 helices of group-I introns (Cech 1990), but cannot be differentiated from P2, P6, P8 or P9. Because of the stem-loop's proximity to the putative P1, we have designated it as a P2. P2 is not essential to the catalytic function of group-I introns since it has been deleted from introns without affecting their self-splicing *in vitro* (Szostak 1986). The exon sequences flanking the putative P2 pair with the IGS of the P1 stem (Fig. 2, insets), as if there has been a ligation of the P2 stem loop to the 3′ flanking exon. The 3′ exon-IGS pairing is characteristic of the P10 of group-I introns proposed to align the splice sites for cleavage and exon ligation (Waring and Davis 1984; Burke et al. 1990; Michel et al. 1990; Suh and Waring 1990).

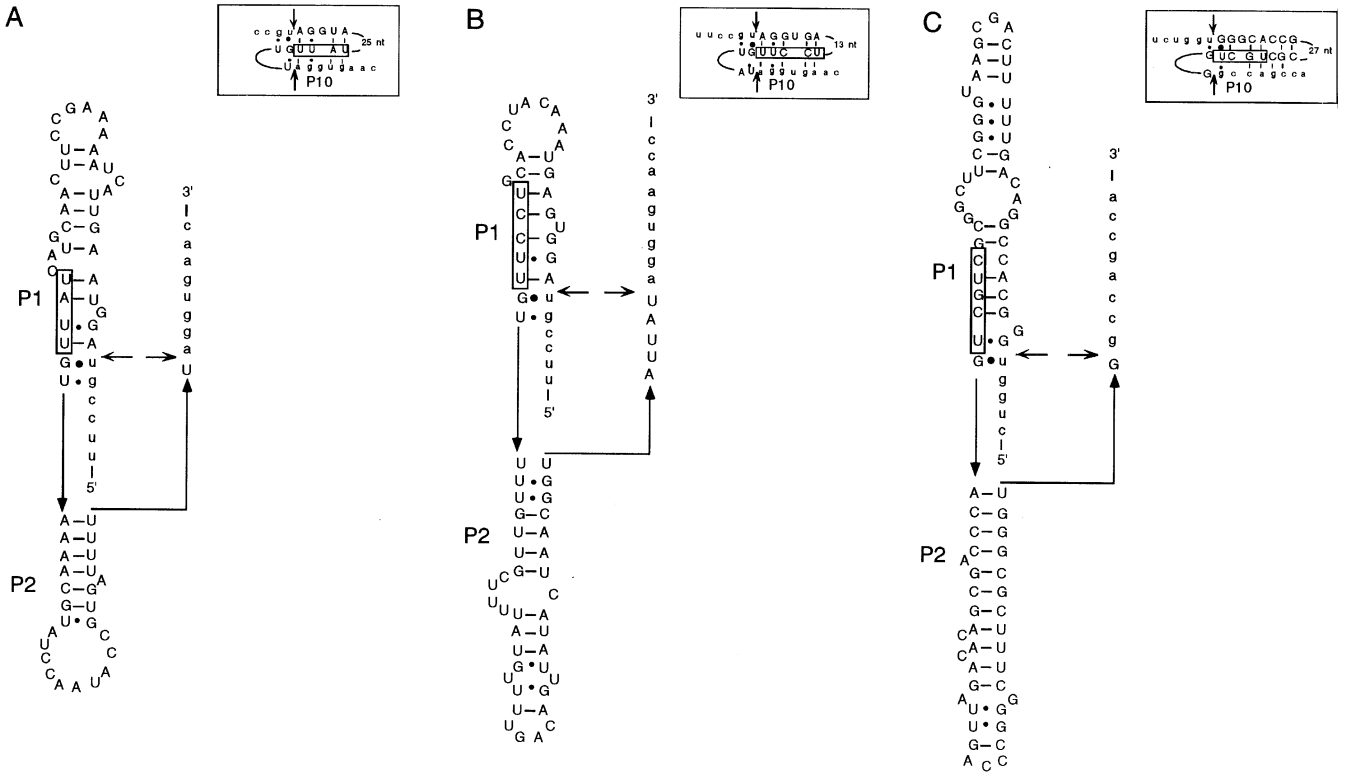

Fig. 2A–C Putative secondary structures of degenerate introns from the SSU rDNA. **A** a degenerate intron of 63 nt at position 1512 in the SSU rDNA of the lichen-forming fungus *A. lapidicola*. **B** a degenerate intron of 67 nt at position 1512 of the SSU rDNA of the deuteromycete *P. americana*. **C** a degenerate intron of 78 nt at position 516 of the SSU rDNA of the lichen-forming fungus *P. crustulata*. *Upper case* letters indicate the primary sequence of the intron, and *lower case* letters the SSU rDNA flanking regions. *Arrows* (→) indicate the 5′ and 3′ putative splice sites that restore conserved SSU rDNA sequences. P1 and P2 indicate the putative stem-loop regions of group-I introns. The putative internal guide sequence (IGS) is enclosed in a *box*. Canonical base pairing is indicated by *dashes* (–), and the non-canonical pairing of G and U by *dots* (•). A *large dot* (●) indicates the pairing of the flanking U and internal G at the 5′ splice site. The arrangement of stem loops is according to Cech et al. (1994). An *inset* shows the putative P10 pairing of the internal guide sequence (IGS) with the nucleotides of the P1 and flanking SSU rDNA sequences

The putative secondary structures of these small insertions suggest that group-I intron-like 5′ stems occur in the absence of a central, conserved core. However, the missing core regions are known experimentally to catalyze the splicing of group-I introns during processing of the RNA transcript (Cech 1990). Rogers et al. (1993) reported that the mature rRNA of *P. americana* lacked the 67-nt insertion, consistent with its splicing removal during rRNA processing. We predict that the 63-nt insertion is also lacking from the mature rRNA of *A. lapidicola*. However, since cultures are not available for this fungus, the mature ribosomal RNA subunits were not examined directly for the presence of these small insertion sequences. It is improbable that these small insertions are capable of self-splicing either *in vitro* or *in vivo*; future research should investigate mechanisms by which small insertions can be spliced from RNA transcripts in the absence of their catalytic cores. The splicing removal of small insertions proposed by Rogers et al. (1993) suggests that P1, P2 and P10 are sufficient for defining splicing junctions. Even without catalytic regions, it is reasonable that degenerate introns could be spliced by trans-acting factors such as ribozymes. Szostak (1986) demonstrated that an engineered ribozyme could catalyze splicing of a 63-nt construct containing the 5′ exon, P1 and P2. The small insertions of SSU rDNA may be naturally occurring substrates for such ribozyme-catalyzed reactions.

These small insertions may represent either incipient- or vestigial-forms of group-I introns present at the same position in other organisms. If they are incipient, then the stems and loops are pieces of introns that assemble gradually until an intron functions catalytically – an idea quite untenable given the intricate tertiary interactions required (e.g. Michel and Westhof 1990). If these small insertions represent the vestiges of an initially complete and recognizable intron, then what mechanisms delete portions of group-I introns? The observation of three small insertions, similar in size and secondary structure, suggest that such deletion mechanisms are not random. The occurrence of small insertions in different positions and organisms predicts that these deletion mechanisms are universal

and recurrent. It is possible that the degenerate introns are the products of miss-splicing of a larger group-I intron to a cryptic 5′ splice site immediately following P2, with subsequent reverse transcription of the processed rRNA transcript and homologous recombination with the rDNA repeat (Levra-Juillet et al. 1989). This miss-splicing event could account for both the deletion of P3 through P9 of the group-I intron and the absence of the characteristic G at the 3′ terminus of the insertion (Fig. 3). The maintenance of a P10 pairing between the flanking regions and the P1 stem suggests that either the deletions occurred recently and the sequences have not degenerated or else that selective pressures constrain the evolution of these sequences. It is unlikely that either of these types of small insertions are limited to pseudogenes of the rDNA since the insertions retain higher-order pairings such as that of P10.

Comparative analysis of aberrant introns is essential for understanding the evolutionary potential of group-I introns. The three putative degenerate introns that we describe here will be joined by other insertions similar in size and structure as more SSU rDNA sequences are reported. The production of degenerate group-I introns apparently parallels that of group-III introns, known to be reduced group-II introns that retain their intron-flanking and stem-loop regions (Christopher and Hallick 1989). These observations suggest that introns in general evolve, or degenerate, by the loss of specific regions or structures, perhaps through RNA-mediated processes (Levra-Juillet et al. 1989). Independent reductions of distinct types of introns indicates that this is a fundamental pattern in the evolution of genes and genomes. The primary evolution of genes and genomes may not occur by clock-like accumulation of nucleotide substitutions, but rather by profound changes such as those evidenced by the degeneration of these intron sequence elements.

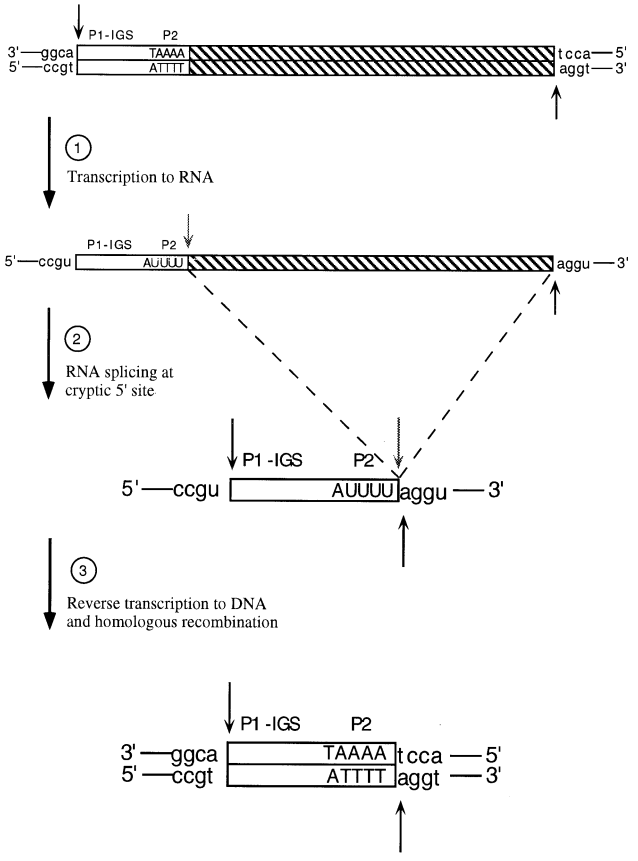

Fig. 3 Production of small insertions from group-I introns by miss-splicing to a cryptic 5′ site. An *open box* indicates the 63-nt insertion of *A. lapidicola*. The relative positions of P1, including the IGS, and P2 are labeled above this box. *Diagonal shading* indicates the putative group-I intron sequence, including the catalytic core and 3′ terminus, deleted from the small insertion. *Upper case* letters indicate the nucleotide sequences in the 3′ terminus of the insertion and *lower case* letters indicate the nucleotide sequences flanking the insertion. *Black arrows* (↑) indicate the positions of the 5′ and 3′ splice sites, *gray arrows* indicate the position of a putative cryptic 5′ splice site. DNA sequences are represented by double-stranded models, RNA sequences are represented by single-stranded models

Acknowledgments M.G. thanks A. Tehler and J. Hafellner for discussions of fungal systematics, J. Poelt for his generous gift of fungal material and G. Högenauer for research support. This work was supported in part by FWF P9588BIO to J.H., a Smithsonian Institution Visiting Scientist Fellowship to M.G., a Smithsonian Institution Postdoctoral Fellowship to A.G., and a Smithsonian Institution Scholarly Studies Grant and a National Museum of Natural History Research Initiatives Award to P.T.D.

References

Armaleo D, Clerc P (1991) Lichen chimeras: DNA analysis suggests that one fungus forms two morphotypes. Exp Mycol 15: 1–10

Burke JM, Esherick JS, Burfeind WR, King JL (1990) A 3′ splice site-binding sequence in the catalytic core of a group-I intron. Nature 344:80–82

Cech TR (1990) Self-splicing of group-I introns. Annu Rev Biochem 59:543–568

Cech TR, Damberger S, Gutell R (1994) Representation of the secondary and tertiary structure of group-I introns. Nature Struct Biol 1:273–280

Christopher DA, Hallick RB (1989) *Euglena gracilis* chloroplast ribosomal protein operon: a new chloroplast gene for ribosomal protein L5 and description of a novel organelle intron category designated group-III. Nucleic Acids Res 17:7591–7608

Davies RW, Waring RB, Ray JA, Brown TA, Scazzocchio C (1982) Making ends meet: a model for RNA splicing in fungal mitochondria. Nature 300:719–728

DePriest PT (1993) Small subunit rDNA variation in a population of lichen fungi due to optional group-I introns. Gene 134:67–71

Embley TM, Dyal PL, Kilvington S (1992) A group-I intron in the small subunit ribosomal RNA gene from *Naegleria andersoni* ssp. andersoni strain PPMFB. Nucleic Acids Res 20:6411

Gargas A, DePriest PT, Grube M, Tehler A (1995 a) Multiple origins of lichen symbioses in fungi suggested by SSU rDNA phylogeny. Science 268:1492–1495

Gargas A, DePriest PT, Taylor JW (1995 b) Positions of multiple insertions in SSU rDNA of lichen-forming fungi. Mol Biol Evol 12:208–218

Garrett RA, Aagaard C, Andersen M, Dalgaard JZ, Lykke-Andersen J, Phan HTN, Trevisanato S, Ostergaard L, Larsen N, Leffers H (1994) Archaeal rRNA operons, intron splicing and homing endonucleases, RNA polymerase operons and phylogeny. System Appl Microbiol 16:680–691

Gast RJ, Fuerst PA, Byers TJ (1994) Discovery of group-I introns in the nuclear small subunit ribosomal RNA genes of *Acanthoamoeba*. Nucleic Acids Res 22:592–596

Grube M, DePriest PT, Gargas A, Hafellner J (1995) Isolation of DNA from lichen ascomata. Mycol Res 99:1321–1324

Huss VAR, Hirmer M, Kessler E (1993 a) Genbank accession X73997, X73998

Huss VAR, Wenzeler P, Kessler E (1993 b) Genbank accession X73991

Levra-Juillet E, Boulet A, Seraphin B, Simon M, Faye G (1989) Mitochondrial introns aI1 and/or aI2 are needed for the in vivo deletion of intervening sequences. Mol Gen Genet 217:168–171

Michel F, Westhof E (1990) Modelling of the three-dimensional architecture of group-I catalytic introns based on comparative sequence analysis. J Mol Biol 216:585–610

Michel F, Jacquier JA, Dujon B (1982) Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. Biochimie 64:867–881

Michel F, Netter P, Xu M-Q, Shub DA (1990) Mechanism of 3′ splice site selection by the catalytic core of the *sun* Y intron of bacteriophage T4:the role of a novel base-pairing interaction in group-I introns. Genes Dev 4:777–788

Rogers SO, Yan ZH, Shinohara M, LoBuglio KF, Wang CJK (1993) Messenger RNA intron in the nuclear 18s ribosomal RNA gene of deuteromycetes. Curr Genet 23:338–342

Suh ER, Waring RB (1990) Base pairing between the 3′ exon and an internal guide sequence increases 3′ splice site specificity in the *Tetrahymena* self-splicing rRNA intron. Mol Cell Biol 10:2960–2965

Szostak JW (1986) Enzymatic activity of the conserved core of a group-I self-splicing intron. Nature 322:83–86

Waring RB, Davies RW (1984) Assessment of a model for intron RNA secondary structure relevant to RNA self-splicing – a review. Gene 28:277–291

White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In:Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) PCR protocols, a guide to methods and applications. Academic Press, Inc., Harcourt Brace Jovanovich, Publishers, New York, pp 315–322

Wilcox LW, Lewis LA, Fuerst PA, Floyd GL (1991) Group-I introns within the nuclear-encoded small-subunit rRNA gene of three green algae. Mol Biol Evol 9:1103–1118