The Lichenologist **36**(6): 365–380 (2004)
DOI: 10.1017/S0024282904014501 *Printed in the United Kingdom*

Group I intron-like insertions in SSU rDNA of *Cladonia gracilis* and *C. rangiferina*

Michele D. PIERCEY-NORMORE, Georg HAUSNER and Ewan A. GIBB

Abstract: During a study on population genetics of three species of *Cladonia* using the nuclear ribosomal 18S DNA, two species contained group I intron-like sequences located in positions 788 and 940 with reference to the *E. coli* 16S rDNA gene. The intron in position 940 was not typical of group I introns and had previously been described only from the *Parmeliaceae* and *Lecanoraceae*, but here we report the occurrence of this intron in the *Cladoniaceae*. The intron in position 788 had characteristics of group I introns and had previously been reported from the *Physciaceae*. In this paper we provide for the first time a putative RNA fold for the nS788 intron, compare the secondary RNA structure of the *Cladonia* group I introns (nS788 and nS940) to that of other taxa, and infer a phylogenetic history among 15 members of the *Lecanorales* based on the evolutionary history of the nS788 group I intron. The RNA fold for nS788 contained two optional hairpins, P2.1 and a potentially newly described hairpin referred to as P5.1. The phylogenetic hypothesis supports the monophyly of the genus *Cladonia*. It also supports the separation of two large groups in the *Physciaceae*; the *Buellia*-group and the *Physcia*-group.

Key words: *Cladonia*, group I intron 788 and 940, phylogeny, SSU rDNA

Introduction

Nucleotide insertions such as spliceosomal introns, degenerate group I introns, and group I introns have been reported in nuclear ribosomal DNA (rDNA) of fungi including basidiomycetes (Hibbett 1996), a mycorrhizal deuteromycete (Shinohara *et al.* 1996), anamorphs of fungi parasitic on insects (Ito & Hirano 1999), plant pathogenic ascomycetes (Suga *et al.* 2000; Fouly & Wilkinson 2000; Gibb & Hausner 2003), and lichenized ascomycetes (Gargas *et al.* 1995; Grube *et al.* 1996; 1999; Ivanova *et al.* 1998; Stenroos & DePriest 1998; Thell 1999; Myllys *et al.* 1999, 2002; Bhattacharya *et al.* 2002). The first report of group I intron-like insertions in lichen fungi was from the *Cladonia chlorophaea* (Flörke ex Sommerf.) Spreng. complex (DePriest & Been 1992). Since then other reports of insertions from the genus *Cladonia* P. Browne have been from *C. arbuscula* (Wallr.) Flot., *C. mitis* Sandst. (Myllys *et al.* 2002), *C. grayi* G. Merr. ex Sandst., *C. merochlorophaea* Asah. (DePriest & Been 1992), *C. subtenuis* (Abbayes) Mattick (Beard & DePriest 1996), *C. gracilis* (L.) Willd., *C. multiformis* G. Merr., and *C. rangiferina* (L.) Nyl. (Piercey-Normore 2004). The genus *Cladonia* contains about 400 species and is morphologically diverse and widely distributed.

The majority of ribosomal insertions are located near conserved residues thought to be functionally important in the tertiary structure of the mature ribosome (Jackson *et al.* 2002). Many of the insertions in the small subunit (SSU) of ribosomal DNA (rDNA) have been recognized as group I introns because of the secondary RNA structures and conserved core sequences. Group I introns at the RNA level are thought to be ribozymes, enzymes that are usually

M. D. Piercey-Normore: Department of Botany, University of Manitoba, Winnipeg, Manitoba, R3T 2N2 Canada
G. Hausner and E. A. Gibb: Department of Microbiology, University of Manitoba, Winnipeg, Manitoba, R3T 2N2 Canada

cis-acting and catalyse excision of the intron from the rRNA precursor.

Nuclear group I introns are often devoid of recognizable open reading frames (ORFs) but like their counterparts in the mitochondrial genomes they retain characteristic secondary and tertiary conserved structures. Group I introns have diagnostic features such as the conserved catalytic core sequences P, Q, R, S, a U preceding the 5′ splice site and a G to which it pairs, a G-C base pair in P3, and the G preceding the 3′ splice site (Cech 1988). Recently some nuclear group I introns belonging to the I C1 and the I E subgroups have been found to encode putative ORFs resembling homing endonucleases (HEGs) of the HIS-CYS family (Belfort *et al.* 2002). Examples of such HEG encoding introns have been reported in *Nectria galligena* (Johansen & Haugen 1999), *Protomyces pachydermus* (Nishida *et al.* 1998), *Beauveria bassiana* (Yokoyama *et al.* 2002), *Coemansia mojavensis* (Tanabe *et al.* 2002), *Capronia pilosella* and *Pleopsidium chlorophanum* (Haugen *et al.* 2004). When ORFs are present within a group I intron they are usually inserted in any one of several loops that emanate from within the conserved secondary structure (reviewed in Haugen *et al.* 2004).

Although group I introns encoding HEGs are considered to be mobile via a double strand break repair mechanism (Dujon 1989; Belfort *et al.* 2002), they are thought to undergo loss and gain events through evolution (Haugen *et al.* 2004). Group I introns might also be able to transpose into new sites within rRNA genes via RNA intermediates through reverse splicing (Roman & Woodson 1995). Ribosomal RNAs are present in relatively large amounts and thus offer targets for insertion of an intron RNA by reverse splicing. The resulting recombinant rRNA molecule would then have to be reverse transcribed into DNA and inserted into the nuclear DNA by recombination. This model of transposition would explain how group I introns that lack ORFs could avoid being lost or dispersed into new positions, or be transferred horizontally between different species.

Evidence has been presented to show that, within some fungal taxa, group I introns were initially acquired horizontally at a specific site, but after divergence they were transmitted vertically (Hibbett 1996; Bhattacharya *et al.* 1994, 1996, 2002). Vertical inheritance of some group I intron-like insertions has been suggested in lichen fungi by comparison of phylogenies produced from spacer regions in nuclear ribosomal DNA with those produced from group I intron-like insertion sequences (Thell 1999; Myllys *et al.* 1999; Thell & Miao 1999). If these insertions are indeed transmitted vertically then they might be useful in phylogenetic studies. In addition, although these elements appear to be lost sporadically in some lineages, their presence or absence might be a useful molecular marker in designating strains or populations (Coates *et al.* 2002).

In this study we noted and examined two group I intron-like elements (nS788 and nS940) in the SSU rDNA of six strains representing two species of the genus *Cladonia*. Using an evolutionary comparative approach DNA sequences were aligned from related taxa and putative RNA secondary structures were generated for the nuclear group I intron-like insertions. In addition the group I intron sequence for nS788 was subjected to phylogenetic analysis in order to examine the history of this intron among 14 taxa within the *Lecanorales*.

Materials and Methods

The material

During a larger study on the population genetics of three species of *Cladonia,* that included a total of 115 populations (Piercey-Normore 2004), four specimens among six sequenced samples of *Cladonia gracilis* contained an nS788 intron. Intron names follow the convention of Johansen and Haugen (2001). From the four sequenced samples of *C. rangiferina* in Piercey-Normore (2004), two contained the nS788 intron. These two samples were collected from the same lichen mat. All specimens were collected from sites within Manitoba (Table 1). The specimens were air dried and deposited in the University of Manitoba Herbarium (WIN).

TABLE 1. *Taxa in the* Lecanorales *containing the nS788 intron used in this study showing the familial classification, source of the DNA sequence, GenBank accession number, length (bp), and the GC content (%).*

Taxon	Classification in *Lecanorales*	Source of DNA sequence	GenBank Accession no.*	Intron length (bp)	GC content (%)
Cladonia gracilis (L.) Willd. MN319	*Cladoniaceae*	This study: *Normore* 1091, Nesosap Lake, MB.	AY575022	240	57·1
Cladonia gracilis MN356	*Cladoniaceae*	This study: *Normore* 1130, Kisseynew Lake, MB.	AY575023	225	55·5
Cladonia gracilis MN441	*Cladoniaceae*	This study: *Normore* 898a, Long Point, MB.	AY575025	242	57·1
Cladonia gracilis MN750	*Cladoniaceae*	This study: *Normore* 1319, Payuk Lake, MB.	AY575024	242	55·4
Cladonia rangiferina (L.) Nyl. MN945	*Cladoniaceae*	This study: *Normore* 1641a, Sandilands Prov. Forest, MB.	AY575027	240	55·0
Cladonia rangiferina MN946	*Cladoniaceae*	This study: *Normore* 1641a, Sandilands Prov. Forest, MB.	AY575026	241	54·9
Gymnoderma coccocarpum Nyl.	*Cladoniaceae*	Zhou, Q. and Wei, J.: direct GenBank submission.	AF523362	253	49·5
Lecanora dispersa (Pers.) Sommerf.	*Lecanoraceae*	Gargas *et al.* 1995	L37734	183	60·7
Diploicia canescens (Dicks.) A. Massal.	*Physciaceae*	Bhattacharya *et al.* 2002	AJ421684	215	53·0
Diplotomma epipolium (Ach.) Arnold	*Physciaceae*	Helms, G.W.F., Rambold, G. and Friedl, T.: direct GenBank submission.	AJ506969	210	54·8
Physcia aipolia (Ehrh. ex Humb.) Fürnr.	*Physciaceae*	Bhattacharya *et al.* 2002	AJ421687	212	57.6
Physcia stellaris (L.) Nyl.	*Physciaceae*	Bhattacharya *et al.* 2002	AJ421688	212	57·5
Physconia perisidiosa (Erichsen) Moberg	*Physciaceae*	Bhattacharya *et al.* 2002	AJ421689	209	69·4
Acarospora complanata H. Magn.	*Acarosporaceae*	Lutzoni *et al.* 2001	AF356653	249	61·4
Acarospora dissipata H. Magn.	*Acarosporaceae*	Lutzoni *et al.* 2001	AF356655	243	63·0

*GenBank accession numbers for the nS940 intron are listed in the Materials and Methods.

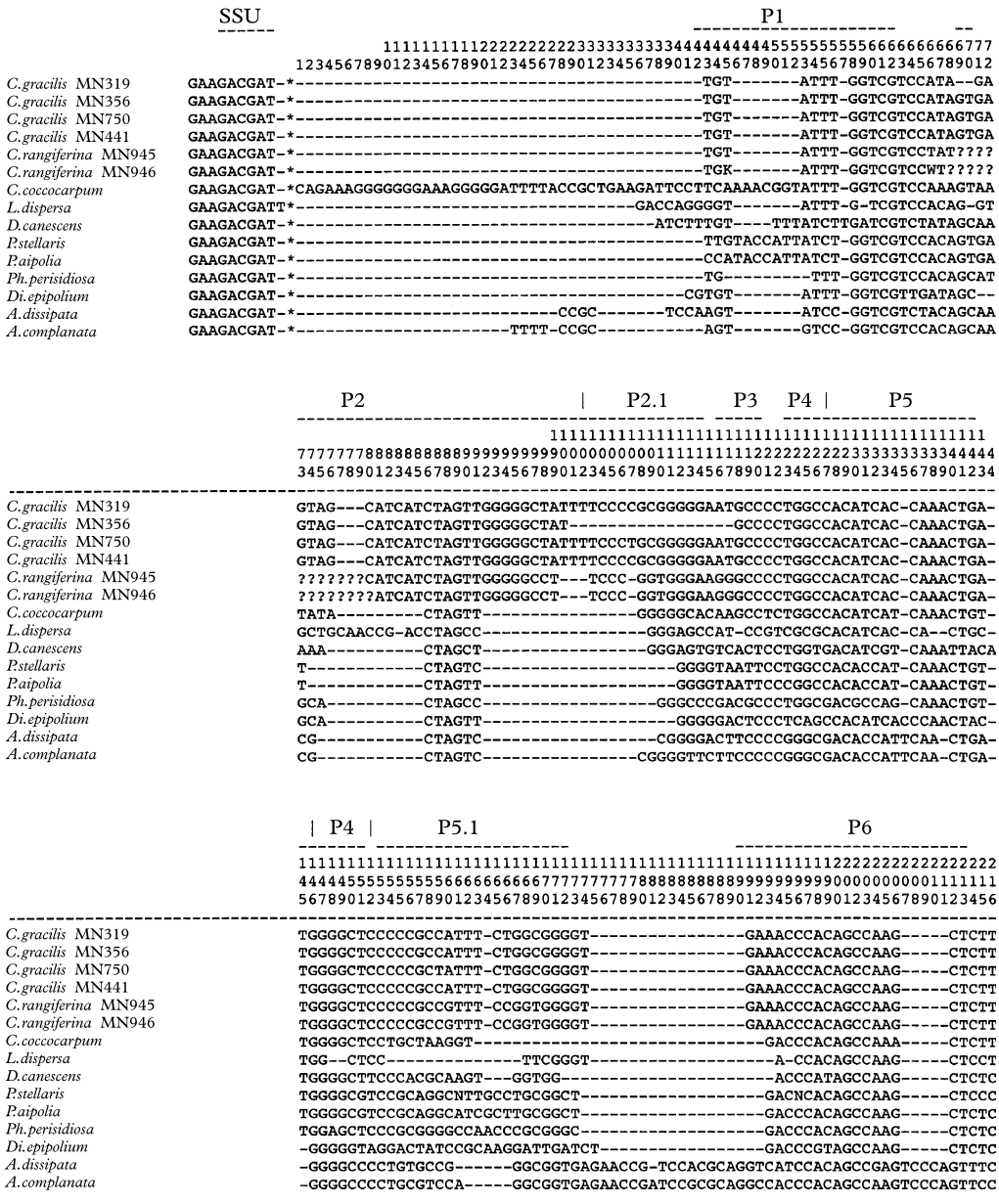


FIG. 1. (continued on p. 369).

DNA isolation and PCR amplification

Total DNA was extracted from a single podetium using a 10% CTAB (cetyltrimethylammonium bromide) extraction (Grube *et al.* 1995). DNA was resuspended in sterile distilled H_2O and amplified by the polymerase chain reaction (PCR). Amplification of a portion of the nuclear SSU of rDNA of the fungal partner was from fungal specific primers nu-SSU-0819-5′ and nu-SSU-1750-3′ (Gargas & DePriest 1996). Amplifications were performed in 20 µl reaction volumes for screening of intron size classes and in 300 µl volumes for sequencing of each sample. Amplification buffer contained 200 mM Tris-HCl (pH 8·4), 500 mM KCl with 2·5 units Taq DNA polymerase (GibcoBRL, Life Technologies) 200 µmol/l of each

https://doi.org/10.1017/S0024282904014501 Published online by Cambridge University Press

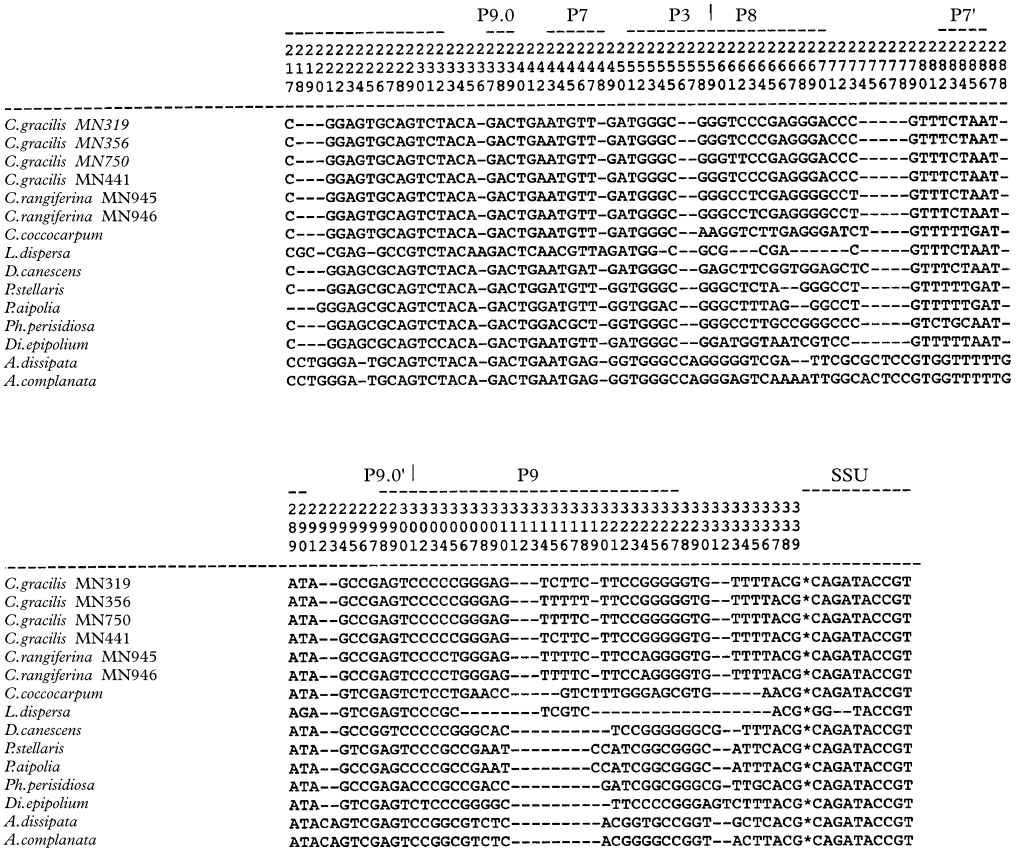


FIG. 1. *(continued)*.

FIG. 1. Nucleotide sequence alignment for nS788 showing the hairpin stems above the alignment. The intron sequence is separated from the flanking SSU regions by an asterisk. Base positions in the intron are numbers located above the alignment, indels are indicated by dashes, and missing data indicated by question marks.

dNTP, 0·5 µmol/l of primer, and between 10 and 50 ng of DNA. Amplification conditions in Fisher Scientific Techne Genius and Biometra T-gradient were: initial template denaturing at 94°C for 5 minutes, then denaturing at 94°C for 1 minute, annealing at 54°C for 1 minute, and extension at 72°C for 2 minutes, for 33 cycles. PCR products were agarose gel purified by freeze squeezing blocks of agarose and subsequent precipitation with 0·2 volumes 5 M NaCl and 2·5 volumes 100% ethanol. The size and quantity of the PCR products were estimated by resolving the DNA fragments with a 1 Kb Plus ladder (Invitrogen) on a 1% agarose gel stained with ethidium bromide.

DNA sequencing

Double stranded products were sequenced using BigDye Terminators Version 3.0 on a 377 and 377XL ABI DNA Sequencing Instrument (University Core DNA and Protein Services, University of Calgary, Calgary, Alberta). Sequencing primers were those described for PCR in addition to nu-SSU-1203-5′ and nu-SSU-1465-3′ previously described in Gargas & DePriest (1996). Sequences were assembled into full-length sequences using Sequencher 3.1.1 (Gene Codes Corporation, Ann Arbor, MI, USA). DNA was sequenced in both directions. The samples of fungal SSU were aligned manually in Se-Al v 1.0 and final adjustments were made in PAUP* 4.0b10 (Swofford, 2003). All sequences for nS788 have been deposited in GenBank under the accession numbers listed in Table 1. Additional sequences similar to the nS788 insertion were recovered from GenBank (see Table 1) with the Basic Local Alignment Search Tool (BLAST, Altschul *et al.* 1997) using default conditions. Accession numbers for nS940 sequences are AY575028 (MN750), AY575029 (MN441), AY575030 (MN319), AY575031 (MN356), AY575032 (MN946), AY575033 (MN945), and AY575034 (MN550).

Phylogenetic analysis

Aligned sequences from fungal SSU intron-like insertion nS788 were subjected to two methods of phylogenetic analysis, neighbour joining and maximum parsimony (MP), using PAUP* 4.0b10. Three separate MP analyses were performed depending on the nucleotide positions excluded from the data set: (1) all nucleotides were included; (2) only the stem regions were included (positions 43–45, 56–64; 71–160, 192–215, 243–249, 254–272, 284–290, 301–310, 318–328 in Fig. 1); and (3) only the loops were included in the analyses (positions 46–55, 65–70, 161–191, 216–242, 250–253, 273–283, 291–300, 311–317, 329–339 in Fig. 1). The data was partitioned into stem and loop regions based on PQRS comparisons with other introns (Michel and Westhof 1990) and the original sequence alignment. Alignment gaps were treated as missing data. Neighbour joining was performed using uncorrected 'P' distance options. Maximum parsimony was performed using the options TBR (Tree bisection and reconnection) branch swapping, collapse zero length branches, and acctran character-state optimization. Heuristic searches were conducted using 100 random addition replicates and bootstrap searches of 100 resamplings (Felsenstein 1985).

Incongruence tests were performed for tree topologies using the Kishino-Hasagawa and Templeton tests performed in PAUP* using parsimony treescore options (Kishino-Hasagawa and nonparametric tests, respectively).

The genus *Acarospora* was assigned as the outgroup because it was the most dissimilar in generic comparisons in this study (Table 2) and was basal to *Cladonia, Lecanora*, and *Physcia* in Stenroos & DePriest (1998). Other taxa with the nS788 intron were present in BLAST search outputs but the introns did not align without multiple insertions and deletions. They were excluded from analyses in this study.

Secondary structure

The secondary structure of the intron and the naming of the major helices were predicted by following the conventions for group I introns as described by Burke *et al.* (1987) and the final RNA folds were presented according to a format suggested by Cech *et al.* (1994). The sequence alignment allowed initial detection of the conserved core sequences. The stem loop structures were then folded using the mfold web server [http://www.bioinfo.rpi.edu/applications/mfold/old/rna/form1.cgi (Zuker 2003)]. If more than one fold was produced the final fold was based on comparisons with previously published group I intron folds, maximization of the hydrogen bonding forming solid stems, and the largest negative delta g value (free energy).

Results and Discussion

The presence of insertions in the SSU rDNA of *C. rangiferina* and *C. gracilis* was discovered during a population study on the SSU rDNA sequence variation in fungal and algal partners in the genus *Cladonia* (Piercey-Normore 2004). The length of the SSU coding region between primer sites 819 and 1750 bp in the fungal partner was constant within and between the species of *C. rangiferina* and *C. gracilis* at 894 bp (Piercey-Normore 2004) excluding short ambiguous regions at each end of the sequence. The length of the PCR amplified product for the same region in *C. rangiferina* and *C. gracilis* was longer, ranging from 1100 to 2200 bp. Sequence analysis revealed one to five insertions ranging in size from 218 bp to 244 bp. Four sequenced specimens of *C. gracilis* and two sequenced specimens of *C. rangiferina* contained an intron at position 788 and another intron at position 940, with reference to the *E. coli* 16S rDNA gene (Gutell 1993).

It was not unexpected to find introns in *Cladonia* since it is a large genus with a broad distribution and already had previous reports of introns present. However, this is the first report of introns in positions 788 and 940 for the genus *Cladonia*. Using BLAST (Altschul *et al.* 1997) we found five significant matches with an *E* value less than 10^{-3} for regions homologous to the sequence in intron nS788. Although a BLAST search produced no significant matches for nS940, the intron appeared to be homologous to those in Grube *et al.* (1999) from the same location.

A group I intron at position 940 in *C. gracilis* and *C. rangiferina*

The insertion in position 940 was unusual for a group I intron because the 5′ splice site followed a G rather than a U. This difference was also reported for nS940 in the *Physciaceae* (Grube *et al.* 1999). However, other diagnostic characters for group I introns (Cech 1988) were present in nS940. The conserved catalytic core sequences of nS940 (P, Q, R, S) were typical of the consensus sequences reported in Cech (1988) with little variation among different strains (Fig. 2). The PQRS core sequences

TABLE 2. *Similarity (%) of the nS788 intron DNA sequence based on alignment in Fig. 1 for taxa used in this study. Ranges of similarities represent those genera that have more than one species or individual in this study (see Table 1). Values in bold are similarities among individuals within a single genus.*

	Cladonia	*Gymnoderma*	*Lecanora*	*Diploicia*	*Diplotomma*	*Physcia*	*Physconia*	*Acarospora*
Cladonia	**90·8–100**							
Gymnoderma	74·3–75·9							
Lecanora	78·1–81	73·4						
Diploicia	75–77·2	64·5	68·5					
Diplotomma	62·2–67·3	62·2	65·6	64·2				
Physcia	70·4–74·1	69·6–71·1	68·9–71·7	70·9–72·1	64·1–64·8	**93·8**		
Physconia	69·8–72·1	67·4	74·5	66·4	62·4	74·4–76·7		
Acarospora	63·3–67·8	64·5–65·7	65·3–68·8	61·9–62·9	55·1–55·4	63·8–67·9	65–65·3	**90·3**

TABLE 3. *Phylogenetic information of the nS788 derived trees and datasets showing total number of characters, number of uninformative and informative characters, number of MP trees retrieved, tree lengths (steps), the consistency and retention indices (CI and RI, respectively), and the measure of phylogenetic signal (g1 statistic).*

Dataset	No. characters	No. trees	Tree length (steps)	Variable uninformative characters	Informative characters	CI	RI	g1 statistic
Combined data	339	4	370	64	109	0·7378	0·6689	− 0·96*
Stems	180	23	257	38	77	0·7276	0·6618	− 0·75*
Loops	117	7	100	22	31	0·7800	0·7412	− 0·67*

*Indicates statistical significance by comparison with a standard value in Hillis & Huelsenbeck (1992).

A. nS940 insertion

Taxon	'P'	'Q'	'R'	'S'
C.gracilis MN319	AACUGCGUGGAC	UACCCGCAG	CUUCGGUCCACAGAUCAAG-UGGUGG	ACUUAUGAUAUGAUCGAC
C.gracilis MN356	.....	.....	.....G.....	.....
C.gracilis MN441	.....	.....	.....-.....	.....
C.gracilis MN750	.....	.....	.....-.....	.....
C.rangiferina MN550	.....	.....	.....-.....	.....
C.rangiferina MN946	.....	.....	.....-.....	.....
C.rangiferina MN945	.....	.....	.....-.....	.....

B. nS788 insertion

Taxon	'P'	'Q'	'R'	'S'
C.gracilis MN319	GGCCACAUCAC	CUGAU-GGGGC	CUACA-GACUGAAUGUU-G	UAAU-AUA--GCCG
C.gracilis MN356	.....	.....	.....	.....
C.gracilis MN750	.....	.....	.....	.....
C.gracilis MN441	.....	.....	.....	.....
C.rangiferina MN945	.....	.....	.....	.....
C.rangiferina MN946	.....	.....	.....	.....
C.coccocarpum	.....U	...U.....	.....	.G.....U..
L.dispersa	---.....	...C....--.	.....A....C..C...A.	......G....U..
D.canescens	..UG.....GU	U.ACAU.....	.....A...	.....
P.stellaris	A.....	UACGG....UA	.C.....	.....U..
P.aipolia	.....C..U	...U.....	.....G.....	.G.....
Ph.perisidiosa	.....C..U	...U.....	.....G.....	.G.....U..
Di.epipolium	...G..GC..G	...U....A..	.....G.C.C...	C.....
A.dissipata	...G...C..U	....-.....	.....AG..	.UU.G...CA.U..
A.complanata	...G...C..U	....-.....	.....AG..	.UU.G...CA.U..

FIG. 2. Putative structural P, Q, R, and S core sequences for taxa in this study. A, the nS940 intron; B, the nS788 intron.

in nS940 were more similar to those in Cech (1988) than were the nS788 core sequences. Consequently, nS940 was easier to fold than nS788.

The sequence alignment for nS940 (available through EMBL) covered 246 positions with similarity between *C. gracilis* and *C. rangiferina* ranging from 88·2% to 91·4%. Similarities within *C. gracilis* (94·6–98·6%) and *C. rangiferina* (98·7–100%) were higher reflecting genetic diversity within populations. The similarity in the nS940 intron sequences between the *Cladoniaceae* in this study and the *Physciaceae* in Grube *et al.* (1999) was 62–73%, comparable to the similarity scores among nS788 sequences across the same families (Table 2).

The putative secondary structure of nS940 is very similar to that proposed by Grube *et al.* (1999) (Fig. 3). They used the nS940 intron sequence to infer a phylogeny of the *Parmeliaceae* and *Lecanoraceae*. Therefore, further characterization of nS940 was not carried out in this study. However, a maximum parsimony analysis using nS940 sequences in this study clearly separated *C. gracilis* from *C. rangiferina* (data not shown) as expected from

other data (Stenroos *et al.* 1997; DePriest *et al.* 2000; Stenroos *et al.* 2002) suggesting that this intron is vertically transmitted.

A group I intron at position 788 in *C. gracilis* and *C. rangiferina*

The insertion in position 788 contained characteristics of group I introns (Fig.1). These include the conserved catalytic core sequences P, Q, R, S, the U preceding the 5′ splice site and the G to which it pairs, a G-C base pair in P3, and the G preceding the 3′ splice site (Cech 1988). The core sequences for nS788 were difficult to locate since base composition varied from published consensus sequences (Cech 1988). However, the core sequences were revealed when the intron sequences were aligned with those of more distantly related taxa (Fig. 1) identifying four conserved regions corresponding to the putative P, Q, R, and S elements. The high GC content in all the taxa containing intron nS788 (Table 1) may explain the deviation of these core sequences from those previously published for other introns. The GC content of the same intron in distantly related species was similar (53–69%) except in the longest sequence from *Gymnoderma coccocarpum* (49·5%).

Similarity of the nS788 intron sequences among the *Lecanorales* in this study ranged from 62 to 74%. Although the core sequence elements were conserved among taxa in this study, they were different from the consensus sequences reported in Cech (1988). A number of group I intron-like insertions located at position 788 have been reported from SSU rDNA of lichen fungi. Since many insertions have been reported at nS788 (Jackson *et al.* 2002), and only one insertion was reported at nS789 in *Lecanora dispersa* (Gargas *et al.* 1995), the homology of nS788 to Ldi.nS789 was questioned. Although Ldi.nS789 (218 bp) was considerably shorter than nS788 (Table 1), the similarity between Ldi.nS789 and nS788 was comparable to that of nS788 and other genera in this study (Table 2).

The secondary structure of nS788 intron had some helices that were variable in size

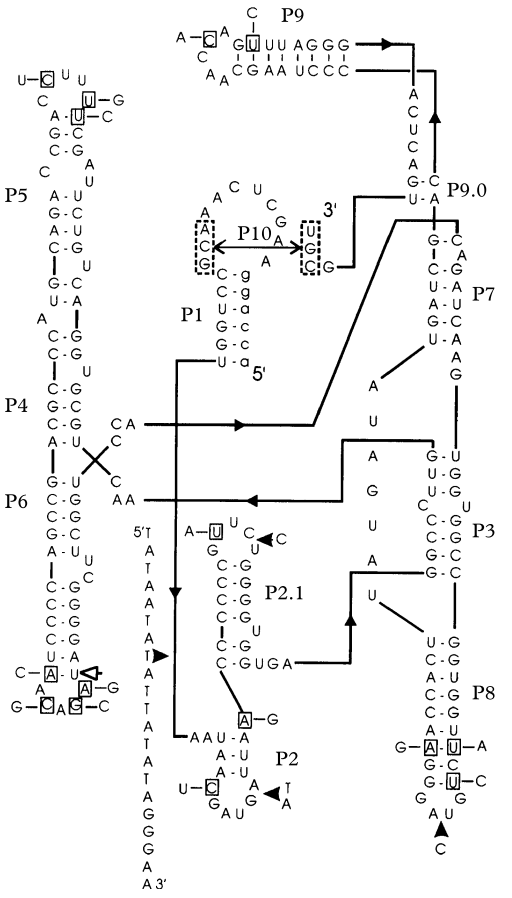
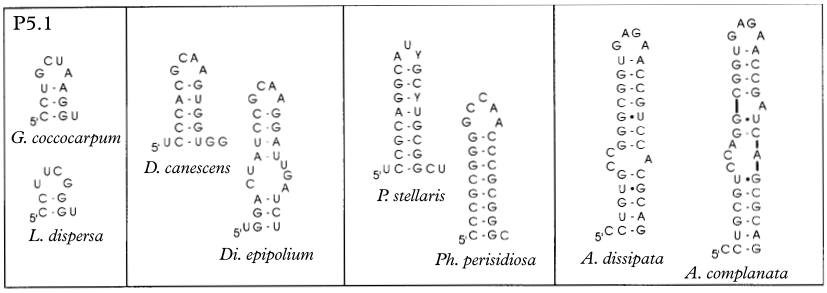
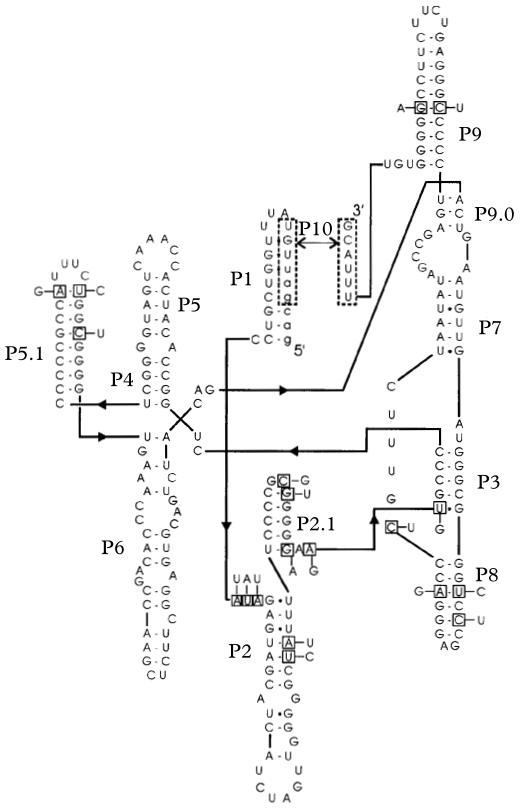


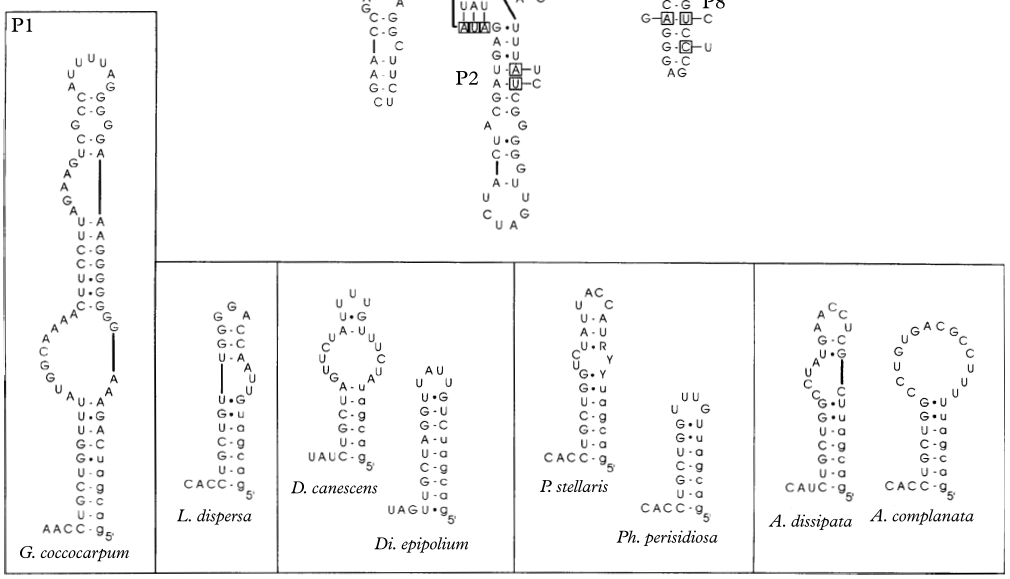
FIG. 3. Putative secondary structure model of intron nS940 from *Cladonia gracilis* (MN319) following the model proposed by Grube *et al.* (1999). Intron sequences are in upper case letters and exon SSU rDNA sequences are in lower case letters. Compensatory substitutions between *C. gracilis* and *C. rangiferina* are indicated by squares.

and in their presence or absence. This is typical of group 1A introns (Michel & Westhof 1990). A helix designated as P5.1 was present in both *C. gracilis* and *C. rangiferina*, reduced in size in *G. coccocarpum* and *L. dispersa*, and well developed in both species of *Acarospora* (Fig. 4). The presence of the larger 5.1 hairpin in the outgroup, *Acarospora*, suggests that the larger hairpin may be the ancestral condition. However, the lack of the P5.1 in other group I introns and the variation in size observed in this study suggest that P5.1 is an optional rather

https://doi.org/10.1017/S0024282904014501 Published online by Cambridge University Press







https://doi.org/10.1017/S0024282904014501 Published online by Cambridge University Press

than a required hairpin. Intermediate sizes of P5.1 were present in the other species (Fig. 4). The length of the entire nS788 intron sequence did not reflect the size of the P5.1 hairpin since *L. dispersa* had the shortest and *G. coccocarpum* the longest intron sequence and both had the shortest hairpins in the P5.1 position. The presence of the required P1 helix was conserved across all taxa with the base of the helix always having the same base pair matches. The P1 helix was large in *G. coccocarpum* (a member of the *Cladoniaceae*) but the size and base pairings varied among the other taxa (Fig. 4). In addition, a P2.1 helix was present in most members studied but it was absent in *C. gracilis* (MN356) suggesting that it was an optional helix.

The majority of the base substitutions in the secondary structure of nS788 in *Cladonia* are compensatory changes based on the putative secondary structure model (Fig. 4). For example base number 105 in P5.1 changes from C in *C. gracilis* to U in *C. rangiferina*. Both nucleotides can pair with the G on the opposite side of the stem. In the entire nS788 intron sequence there were three compensatory changes in the stems, one change in the loops, and two insertions for individuals within *C. gracilis* (see Fig. 1). Similarly, many of the substitutions between *C. gracilis* and *C. rangiferina* were also compensatory. Between the two species (*C. gracilis* and *C. rangiferina*) there were nine compensatory changes in the stems, four changes in the stems were not compensatory, and five changes were in the loops between the two species (see Fig. 4). So overall, the compensatory nature of most nucleotide substitutions can be viewed as support for the proposed RNA fold.

Phylogenetic treatment of the nS788 intron

Analysis of the nS788 intron data supported the monophyly of the genus *Cladonia* relative to the other taxa examined in the *Lecanorales* with 85% level of confidence based on bootstrap analysis (Fig. 5A). The genus was also clearly separated from the other taxa included in our analyses irrespective of whether the intron sequences were subdivided into subsets that examined the phylogenetic history of the stem regions separately from the evolution of the intron loop (non-pairing) regions (Figs 5B and 5C respectively). Based on the analysis of the entire nS788 alignment each of the two species within *Cladonia* (*C. gracilis* and *C. rangiferina*) formed monophyletic groups with bootstrap support of 93% and 99%, respectively (Fig. 5A). Both species are relatively distantly related within the genus (DePriest *et al.* 2000; Stenroos *et al.* 1997, 2002). *Cladonia gracilis* is in sect. *Cladonia* of the genus *Cladonia* (*sensu* Ahti) and *C. rangiferina* is in sect. *Tenues* (*sensu* Ahti). *Cladonia rangiferina* was for a while considered to be in a separate genus, *Cladina*, but was recently resynonymized with *Cladonia* (Ahti & DePriest 2001). Although sequence variation was present between and within species of *Cladonia* (Table 2), two samples of *C. rangiferina* from a single mat produced identical intron sequences suggesting that *C. rangiferina* grew clonally. Clonal growth of reindeer lichens was also supported by other studies of introns in Piercey-Normore (2004) and Beard & DePriest (1996).

The phylogeny based on the entire sequence of the nS788 intron is not in disagreement with the phylogenies proposed by Scheidegger *et al.* (2001), Nordin &

FIG. 4. Putative secondary structure model of intron nS788 from *Cladonia gracilis* (MN441). Intron sequences are in upper case letters and exon SSU rDNA sequences in lower case letters. The P5.1 and P1 hairpin loops are boxed above and below the folding. Compensatory substitutions between *C. gracilis* and *C. rangiferina* are indicated by squares. Insertions and deletions are indicated by solid and open arrows, respectively. Substitutions between *Physcia stellaris* and *P. aipolia* were indicated by ambiguity codes Y (C or T) and R (A or G) for the P1 and P5.1 helices.

Mattson (2001), and Grube & Arup (2001) for the *Physciaceae* and that proposed by Stenroos & DePriest (1998) for the placement of *Acarospora* as outgroup. Representatives of the family *Cladoniaceae* in this study formed a sister group to the *Buellia* group of the *Physciaceae* (Fig. 5A; Bhattacharya *et al.* 2002) even though the two families are distantly related (Stenroos & DePriest 1998). As a member of the *Cladoniaceae*, the genus *Gymnoderma* should form a sister lineage to *Cladonia* as was reported in Stenroos & DePriest (1998). However, in Fig. 5A *Lecanora dispersa* is placed between *G. coccocarpum* and *Cladonia* making the *Cladoniaceae* paraphyletic. The remainder of the tree, consisting of the *Physciaceae*, was similar to the phylogeny proposed by Grube & Arup (2001). Although there was low bootstrap support, the combined phylogeny (Fig. 5A) separated the *Physcia* group from the *Buellia* group (*sensu* Rambold *et al.* 1994) by placing the *Physcia* group in a basal position. In Grube & Arup (2001) the *Physcia* and *Buellia* groups form sister clades. The lack of bootstrap support (Fig. 5A) and the smaller number of samples in this study compared to Grube & Arup (2001) may account for this difference in topology between these two studies.

The stems and loops within the secondary structure model of group I introns are constrained in different ways and may provide different types of phylogenetic information. Analyses were run that included stem regions only (eliminating loops), loop regions only (eliminating stems), and the combined regions, to determine the degree of phylogenetic signal in nS788. All three data sets contained significant phylogenetic signal (Table 3). Phylogenetic trees produced from each of the stem (Fig. 5B) and loop (Fig. 5C) data sets were similar in topology to those produced by the combined data set (entire intron sequence) (Fig. 5A) but there was much less resolution in the stem tree. The lack of congruence between the stem trees and the combined trees (P=0·0001–0·0126), was probably due to the lack of resolution in the stem trees affecting the incongruence test. Even though there were fewer informative characters and more homoplasy in the loop trees than in the stem trees (Table 3) the loop tree (Fig. 5C) was more similar to the combined tree (Fig. 5A) than the stem tree (Fig. 5B) was to the combined tree (Fig. 5A). The incongruence tests indicated no significant difference between the loop trees and the combined trees (P=0·3701–0.8530). Dixon and Hillis (1993) reported that the highest level of support was achieved when the entire data set (stems and loops) was used in the analysis.

Stem regions are conserved across taxa to retain secondary structure in the introns of the SSU rDNA. The majority of changes in the stem regions were compensatory changes (Fig. 4). However, because of the constraint in these regions allowing for only certain types of changes there may be problems associated with convergence of characters and hence taxa. The loops are thought to be less constrained and would allow a larger number of neutral substitutions. Neutral substitutions may provide homoplastic signal but not convergence and therefore provide a true phylogenetic signal. However, saturation may occur in the loops if the taxa are too distantly related. Saturation may provide low phylogenetic signal and lead to variable topologies. Another problem with the loops is that the many changes due to lack of sequence constraint make it difficult to decide on the homologous regions. The loops may be homologous at the structural level but the nucleotide sequences, especially in distantly related taxa, may not be homologous. A comparative phylogenetic study of more distant taxa using secondary structure of the introns as a character, such as was done in Toor *et al.* (2001), may provide more reliable phylogenetic signal than the DNA sequence alone.

The presence of an intron at position 789 of the nuclear SSU gene has been described in *Lecanora dispersa* (Gargas *et al.* 1995). This intron, Ldi.nS789, is inserted 1bp away from the insertion of intron nS788 examined in this study. The close physical location of nS788 to nS789 does not presuppose that

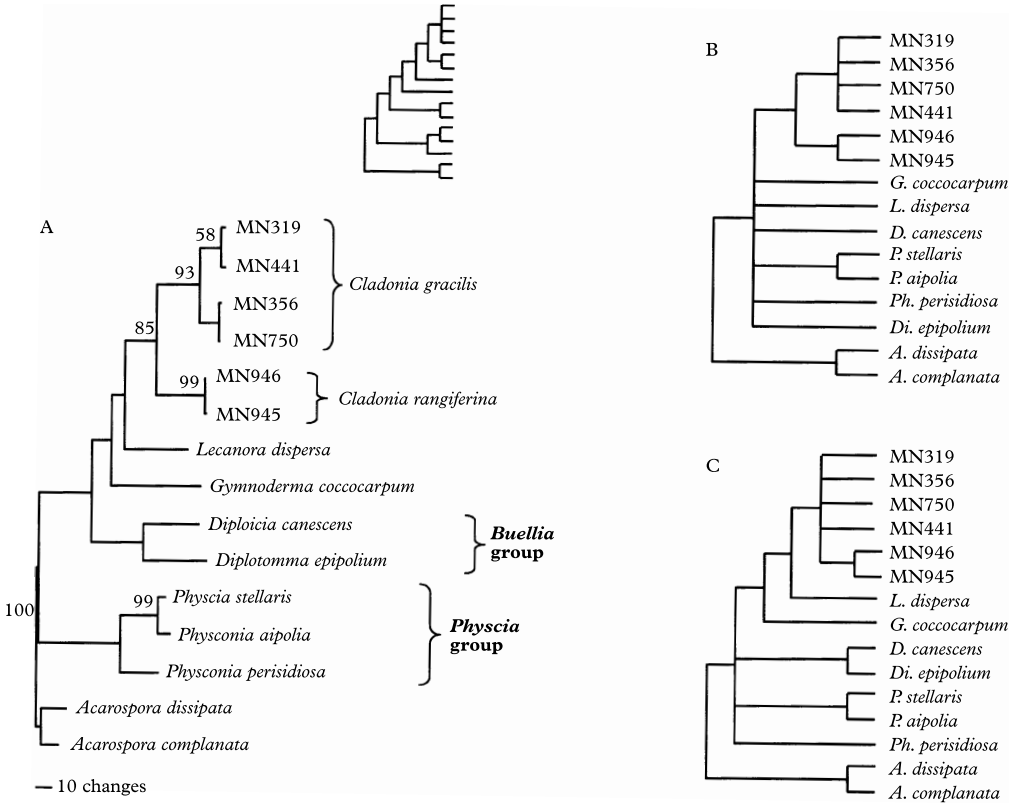


FIG. 5. Phylogenetic relationships inferred from the nS788 intron sequence alignment of the *Lecanorales*. Since the neighbour joining trees supported the MP trees, only the MP trees are presented. A, 50% majority rule consensus of 4 MP trees inferred from the entire nS788 intron sequence alignment. (Insert shows topology of the strict consensus tree with taxa in the same order); B, strict consensus of 23 MP trees inferred from the stem regions of the nS788 intron sequence; C, strict consensus of 7 MP trees inferred from the loop regions of the nS788 intron sequence. See Methods for positions corresponding to stems and loops and Table 3 for tree descriptive statistics. Two species of *Acarospora* have been used as the outgroup taxa. Bootstrap support greater than 50% is indicated above the branches.

nS788 is homologous to nS789 and was horizontally transferred from 1bp downstream. The size of the helices, P1 and P5.1 (Fig. 4), and the nucleotide sequence of the catalytic core regions (Fig. 2) in Ldi.nS789 are very different from those in nS788 in *Cladoniaceae* and *Physciaceae*. In addition, the SSU sequences flanking the intron splice site in *L. dispersa* are different from the splice sites in the species of *Cladonia* in this study. The 5′ splice site in *L. dispersa* contains an extra 'T' and the four bases immediately following the intron at the 3′ splice site in *L. dispersa* are different from those in the *Cladonia* sequences (Fig. 1). *Lecanora dispersa* was analysed using nS789 sequence aligned with nS788 data set, and the phylogenetic analysis produced a paraphyletic *Cladoniaceae* (Fig. 5A), offering further argument that Ldi.nS789 is not homologous with nS788. Grube *et al.* (1999) reported a similar occurrence with nS940 and nS943 and suggested that they were not homologous. It has been noted in other reports that in many instances nuclear group I introns in phylogenetic analysis group according to their insertion sites (Bhattacharaya *et al.* 1996) suggesting that these ribozymes adapt to specific target sites. Therefore, even minor 'shifts/sliding' events might be sugges-

https://doi.org/10.1017/S0024282904014501 Published online by Cambridge University Press

tive of analogous rather then homologous insertions.

In conclusion, we provide a putative secondary structure model of nS788 found in *C. rangiferina*, *C. gracilis* and the *Physciaceae*. This is the first report proposing a secondary RNA structure for the nS788 intron. The intron Ldi.nS789 is not homologous to nS788 in *Cladonia* producing a paraphyletic family *Cladoniaceae* when included in the data set. The intron nS788 is vertically transmitted since the genus *Cladonia* was monophyletic and the tree is in agreement with previously published trees. In addition, the comparison of stem and loop regions showed that the different types of constraint may lead to variable phylogenetic hypotheses. This warrants further study of the two regions, stems and loops, with larger taxonomic comparisons and sample sizes. The comparison of nS788 nucleotide sequence between the *Cladoniaceae* and the *Physciaceae* provides similarity values comparable to those of nS940 between the *Cladoniaceae*, the *Parmeliaceae*, and the *Lecanoraceae*.

Finally, group I introns are of interest not only as taxonomic markers or phylogenetic signals, but also because they are also viewed as catalytic RNAs. Catalytic RNAs are model systems for understanding ribozyme-mediated cleavage reactions (Landweber *et al.* 1998; Doudna & Cech 2002). Group I ribozymes also have practical applications because artificially modified group I introns can serve as trans-cleaving ribozymes that can inactivate specific nuclear or viral gene products by cleaving mRNAs (Johansen *et al.* 1997).

Funding for this study was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC). Specimens collected in provincial parks were collected under Manitoba Conservation permit numbers 12605 and 12627, kindly provided by H. Hernandez.

References

Ahti, T. & DePriest, P. T. (2001) New combinations of *Cladina* epithets in *Cladonia* (Ascomycotina: *Cladoniaceae*). *Mycotaxon* **78:** 499–502.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J. Zhang, Z., Miller, W. & Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25:** 3389–3402.

Beard, K. H. & DePriest, P. T. (1996) Genetic variation within and among mats of the reindeer lichen, *Cladina subtenuis*. *Lichenologist* **28:** 171–182.

Belfort, M., Derbyshire, V., Parker, M.M., Cousineau, B. & Lambowitz, A.M. (2002) Mobile introns: pathways and proteins. In *Mobile DNA II* (N. L. Craig, R. Craigie, M. Gellert & A. M. Lambowitz, eds): 761–783. Washington DC: ASM Press.

Bhattacharya, D., Surek, B., Rüsing, M., Damberger, S. & Melkonian, M. (1994) Group I introns are inherited through common ancestry in the nuclear-encoded rRNA of *Zygnematales* (*Charophyceae*). *Proceedings of the National Academy of Sciences (USA)* **91:** 9916–9920.

Bhattacharya, D., Friedl, T. & Damberger, S. (1996) Nuclear-encoded rDNA group I introns: origin and phylogenetic relationships of insertion site lineages in the green algae. *Molecular Biology and Evolution* **13:** 978–989.

Bhattacharya, D. Friedl, T. & Helms, G. (2002) Vertical evolution and intragenic spread of lichen-fungal group I introns. *Journal of Molecular Evolution* **55:** 74–84.

Burke, J. M., Belfort, M., Cech, T. R., Davies, R. W., Schweyen, R. J., Shub, D. A., Szostak, J. W. & Tabak, H. F. (1987) Structural conventions for group-I introns. *Nucleic Acids Research* **15:** 7217–7222.

Cech, T. R. (1988) Conserved sequences and structures of group I introns: building an active site for RNA catalysis – a review. *Gene* **73:** 259–271.

Cech, T. R., Damberger, S. H. & Gutell, E. R. (1994) Representation of the secondary and tertiary structure of group I introns. *Nature Structural Biology* **1:** 273–280.

Coates, B. S., Hellmich, R. L. & Lewis, L. C. (2002) Nuclear small subunit rRNA group I intron variation among *Beauveria* spp. provide tools for strain identification and evidence of horizontal transfer. *Current Genetics* **41:** 414–424.

DePriest, P. T. & Been, M. D. (1992) Numerous group I introns with variable distributions in the ribosomal DNA of a lichen fungus. *Journal of Molecular Biology* **228:** 315–321.

DePriest, P. T., Piercey-Normore, M., Sikaroodi, M., Kärkkäinen, K., and Oksanen, I, Yahr, R. & Ahti, T. (2000) Phylogenetic analyses of *Cladonia* and *Cladina* (lichen-forming Ascomycota). In *Conference Abstracts of PEET III: The Monographic Process*. Washington, DC March 2000: 13.

Dixon, M. T. & Hillis, D. M. (1993) Ribosomal RNA secondary structure: compensatory mutations and implications for phylogenetic analysis. *Molecular Biology and Evolution* **10:** 256–267.

Doudna, J. A. & Cech, T. R. (2002) The chemical repertoire of natural ribozymes. *Nature* **418:** 222–228.

Dujon, B. (1989) Group I introns as mobile genetic elements: facts and mechanistic speculations – a review. *Gene* **82:** 91–114.

Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39:** 783–791.

Fouly, H. M. & Wilkinson, H. T. (2000) A group I intron in the nuclear small subunit ribosomal DNA of *Gaeumannomyces graminis*. *Current Microbiology* **40:** 291–296.

Gargas, A. & DePriest, P. T. (1996) A nomenclature for fungal PCR primers with examples from intron-containing SSU rDNA. *Mycologia* **88:** 745–748.

Gargas, A., DePriest, P. T. & Taylor, J. W. (1995) Positions of multiple insertions in SSU rDNA of lichen-forming fungi. *Molecular Biology and Evolution* **12:** 208–218.

Gibb, E. A. & Hausner, G. (2003) A group I intron-like sequence in the nuclear SSU gene of the ophiostomatoid fungus *Gondwanamyces proteae*. *Mycological Research* **107:** 1442–1450.

Grube, M. & Arup, U. (2001) Molecular and morphological evolution in the *Physciaceae* (*Lecanorales*, lichenized Ascomycotina), with special emphasis on the genus *Rinodina*. *Lichenologist* **33:** 63–72.

Grube, M., DePriest, P. T., Gargas, A. & Hafellner, J. (1995) DNA isolation from lichen ascomata. *Mycological Research* **99:** 1321–1324.

Grube, M., Gargas, A. & DePriest, P. T. (1996) A small insertion in the SSU rDNA of the lichen fungus *Arthonia lapidicola* is a degenerate group I intron. *Current Genetics* **29:** 582–586.

Grube, M., Gutmann, B., Arup, U., de los Rios, A., Mattsson, J.-E. & Wedin, M. (1999) An exceptional group I intron-like insertion in the SSU rDNA of lichen mycobionts. *Current Genetics* **35:** 536–541.

Gutell R. R. (1993) Collection of small subunit (16S- and 16S-like) ribosomal RNA structures. *Nucleic Acids Research* **21:** 3051–3054.

Haugen P., Reeb V., Lutzoni, F. & Bhattacharya D. (2004) The evolution of homing endonuclease genes and group I introns in nuclear rDNA. *Molecular Biology and Evolution* **21:** 129–140.

Hibbett, D. S. (1996) Phylogenetic evidence for horizontal transmission of group I introns in the nuclear ribosomal DNA of mushroom-forming fungi. *Molecular Biology and Evolution* **13:** 903–917.

Hillis, D. M. & Huelsenbeck, J. P. (1992) Signal, noise, and reliability in molecular phylogenetic analyses. *Journal of Heredity* **83:** 189–195.

Ito, Y. & Hirano, T. (1999) A group I intron in the 18S ribosomal DNA from the parasitic fungus *Isaria japonica*. *Journal of Molecular Evolution* **48:** 337–340.

Ivanova, N. V., DePriest, P. T., Bobrova, V. K. & Troitskii, A. V. (1998) Group I introns in 18S rDNAs of lichens of the *Umbilicariaceae* family. *Doklady Biochemistry* **363:** 186–189.

Jackson, S. A., Cannone, J. J., Lee, J. C., Gutell, R. R. & Woodson, S. A. (2002) Distribution of rRNA introns in the three-dimensional structure of the ribosome. *Journal of Molecular Biology* **323:** 35–52.

Johansen, S., Einvik, C., Elde, M., Haugen, P., Vader, A. & Haugli, F. (1997) Group I introns in biotechnology: prospects of application of ribozymes and rare-cutting homing endonucleases. In *Biotechnology Annual Reviews* Vol. 3 (M. R. El-Gewely, ed.): 111–150. Amsterdam: Elsevier Science BV.

Johansen, S. & Haugen, P. (1999) A complex group I intron in *Nectria galligena* rDNA. *Microbiology* **145:** 516–517.

Johansen, S. & Haugen, P. (2001) A new nomenclature of group I introns in ribosomal DNA. *RNA* **7:** 935–936.

Landweber, L. F., Simon, P. J. & Wagner, T. A. (1998) Ribozyme engineering and early evolution. *BioScience* **48:** 94–103.

Lutzoni, F., Pagel, M. & Reeb, V. (2001) Major fungal lineages are derived from lichen symbiotic ancestors. *Nature* **411:** 937–940.

Michel, F. & Westhof, E. (1990) Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *Journal of Molecular Biology* **216:** 585–610.

Myllys, L., Lohtander, K., Källersjö, M. & Tehler, A. (1999) Sequence insertions and ITS data provide congruent information on *Rocella canariensis* and *R. tuberculata* (*Arthoniales*, Euascomycetes) phylogeny. *Molecular Phylogenetics and Evolution* **12:** 295–309.

Myllys, L., Stenroos, S. & Thell, A. (2002) New genes for phylogenetic studies of lichenized fungi: glyceraldehyde-3-phosphate dehydrogenase and beta-tubulin genes. *Lichenologist* **34:** 237–246.

Nishida, H., Tajiri, Y. & Sugiyama, J. (1998) Multiple origins of fungal group I introns in the position of nuclear SSU rRNA gene. *Journal of Molecular Evolution* **46:** 442–448.

Nordin, A. & Mattson, J.-E. (2001) Phylogenetic reconstruction of character development in Physciaceae. *Lichenologist* **33:** 3–23.

Piercey-Normore, M. D. (2004) Selection of algal genotypes by three species of lichen fungi in the genus *Cladonia*. *Canadian Journal of Botany* **82:** 947–961.

Rambold, G., Mayrhofer, H. & Matzer, M. (1994) On the ascus types in the *Physciaceae* (*Lecanorales*). *Plant Systematics and Evolution* **192:** 31–40.

Roman, J. & Woodson, S. A. (1995) Reverse splicing of the tetrahymena IVS: evidence for multiple reaction sites in the 23S rRNA. *RNA* **1:** 478–490.

Scheidegger, C., Mayrhofer, H., Moberg, R. & Tehler, A. (2001) Evolutionary trends in the *Physciaceae*. *Lichenologist* **33:** 25–45.

Shinohara, M. L., LoBuglio, K. F. & Rogers, S. O. (1996) Group I intron family in the nuclear ribosomal RNA small subunit genes of *Cenococcum geophilum* isolates. *Current Genetics* **29:** 377–387.

Stenroos, S. K. & DePriest, P. T. (1998) SSU rDNA phylogeny of cladoniform lichens. *American Journal of Botany* **85:** 1548–1559.

Stenroos, S., Ahti, T. & Hyvönen, J. (1997) Phylogenetic analysis of the genera *Cladonia* and

Cladina (*Cladoniaceae*, lichenized Ascomycota). *Plant Systematics and Evolution* **207:** 43–58.

Stenroos, S., Hyvönen, J., Myllys, L., Thell, A. & Ahti, T. (2002) Phylogeny of the genus *Cladonia* s.lat. (*Cladoniaceae*, Ascomycetes) inferred from molecular, morphological, and chemical data. *Cladistics* **18:** 237–278.

Suga, H., Oyabu, K., Ito, M., Kageyama, K. & Hyakumachi, M. (2000) Detection of intron-like sequences in the small subunit rDNA 3′ region of *Fusarium solani*. *Mycological Research* **104:** 782–787.

Swofford, D. L. (2003) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10*. Sunderland, Massachusetts: Sinauer Associates.

Tanabe, Y., Yokota, A. & Sugiyama, J. (2002) Group I introns from Zygomycota: evolutionary implications for the fungal IC1 intron subgroup. *Journal of Molecular Evolution* **54:** 692–702.

Thell, A. (1999) Group I intron versus ITS sequences in phylogeny of cetrarioid lichens. *Lichenologist* **31:** 441–449.

Thell, A. & Miao, V. (1999) Phylogenetic analysis of ITS and group I intron sequences from European and North American samples of cetrarioid lichens. *Annales Botanici Fennici* **35:** 275–286.

Toor, N., Hausner, G. & Zimmerly, S. (2001) Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. *RNA* **7:** 1142–1152.

Yokoyama, E., Yamagishi, K. & Hara, A. (2002) Group I intron containing a putative homing endonuclease gene in the small subunit ribosomal DNA of *Beauveria bassiana* IFO 31676. *Molecular Biology and Evolution* **19:** 2022–2025.

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

Accepted for publication 25 *August* 2004