Curr Genet (2000) 37: 12–23

ORIGINAL PAPER

Anja Besendahl · Yin-Long Qiu · Jungho Lee
Jeffrey D. Palmer · Debashish Bhattacharya

The cyanobacterial origin and vertical transmission of the plastid tRNA^{Leu} group-I intron

Received: 1 August / 22 September 1999

Abstract We have surveyed the distribution and reconstructed the phylogeny of the group-I intron that is positioned in the anticodon loop of the tRNA^{Leu} gene in cyanobacteria and several plastid genomes. Southern-blot and PCR analyses showed that the tRNA^{Leu} intron is found in all 330 land plants that were examined. The intron was also found, and sequenced, in all but one of nine charophycean algae examined. Conversely, PCR analyses showed that the tRNA^{Leu} group-I intron is absent from the red, cryptophyte and haptophyte algae, although it is present in three members of the heterokont lineage. Phylogenetic analyses of the intron indicate that it was present in the cyanobacterial ancestor of the three primary plastid lineages, the Rhodophyta, Chlorophyta, and Glaucocystophyta. Its present-day distribution in plastids is consistent with a history of strictly vertical transmission, with no losses in land plants, several losses among green algae, and nearly pervasive loss in the Rhodophyta and its secondary derivatives.

Key words Group-I intron · Phylogeny · Plastids · tRNA^{Leu}

Communicated by C.W. Birky, Jr.

A. Besendahl · D. Bhattacharya (✉)
Department of Biological Sciences,
University of Iowa, 239 Biology Building,
Iowa City, IA 52242–1324, USA
Tel.: +1-319-335 1977; Fax: +1-319-335 1069
e-mail: dbhattac@blue.weeg.uiowa.edu

Y.-L. Qiu · J. Lee
Institute of Systematic Botany,
University of Zürich, Zollikerstrasse 107,
8008 Zürich, Switzerland

J.D. Palmer
Department of Biology,
Indiana University, Bloomington, IN 47405, USA

Introduction

Group-I introns have a distinct and conserved RNA secondary structure essential for splicing and often undergo auto-catalytic excision from coding regions (Kruger et al. 1982). The existence of "self-splicing" group-I introns and other RNA-enzymes (ribozymes, e.g., group-II introns, RNase P) is interpreted as evidence for a primordial "RNA world" (Gilbert 1986) in which RNA functioned both in catalysis and in information storage (Pace and Marsh 1985). Some group-I introns may therefore trace their ancestry directly to the RNA world (Gilbert 1986; Moore et al. 1993) although it is also entirely possible that this class of introns arose only after the putative transition from an RNA world to a largely DNA world. Group-I introns provide an excellent system to gain insights into the origin and distribution of a widespread class of ribozymes. These analyses include phylogenetic reconstruction because all group-I introns contain a common, conserved catalytic core [the P, Q, R, and S domains (Cech 1988)]. These core sequences are used to align RNA secondary structure elements of group-I introns from different insertion sites in different genes for phylogenetic analyses (Bhattacharya et al. 1994, 1996; De Jonckheere and Brown 1994; Gielly and Taberlet 1994; Hibbet 1996; Friedl et al. in press).

Many different lineages of group-I introns have been identified in organellar and nuclear genomes (e.g., Turmel et al. 1993; Bhattacharya et al. 1994; Damberger and Gutell 1994; Hibbet 1996), in phages and in eubacteria (Lambowitz and Belfort 1993; Biniszkiewicz et al. 1994), and in eukaryotic viruses (Yamada et al. 1994; Nishida et al. 1998). Of these, the group-I intron that interrupts the anticodon loop of the tRNA$^{Leu\text{-}UAA}$ gene in cyanobacteria is remarkable because evolutionary analyses suggest that it originated in the common ancestor of this group (2.7–3.5 billion years ago: Schopf 1993; Brocks et al. 1999; Summons et al. 1999) and is therefore the most ancient intron yet found (Paquin

et al. 1997). The $tRNA^{Leu}$ group-I intron has been vertically transmitted in some cyanobacterial lineages, and lost in others (Kuhsel et al. 1990; Paquin et al. 1997). Recent evidence also suggests that this intron may have been laterally transferred in some cyanobacteria (e.g., *Nostoc* spp.; see Rudi and Jakobsen 1999). In addition, the $tRNA^{Leu}$ gene in some cyanobacteria (e.g., *Microcystis* spp.) contains a divergent intron (a member of cluster-II in Rudi and Jakobsen 1999) that is unrelated to the widespread $tRNA^{Leu}$ intron (cluster I). Primary and secondary structural analyses show that the cluster-II introns share a common ancestry with introns in the $tRNA^{Arg\text{-}CCU}$ gene in α-purple bacteria (Paquin et al. 1999) and have apparently arisen from a $tRNA^{Arg}$-like intron by lateral transfer (Rudi and Jakobsen 1997; Paquin et al. 1999). These results underline the complex evolutionary history of the $tRNA^{Leu}$ introns in cyanobacteria (Rudi and Jakobsen 1999). The intron that is found at the same position in the $tRNA^{Leu}$ gene in the plastid genomes of several diverse algae and land plants is closely related to the widespread (cluster I) intron in cyanobacteria (there are no cluster-II-type introns found in plastid $tRNA^{Leu}$ genes) and is believed to have entered eukaryotes through cyanobacterial endosymbiosis (Kuhsel et al. 1990; Paquin et al. 1997).

There has been no comprehensive analysis of the presence/absence or phylogeny of the $tRNA^{Leu}$ group-I intron in plastids. The presumed cyanobacterial origin of the intron in plastids has also not been rigorously tested. Paquin et al. (1997), for example, included only three plastid $tRNA^{Leu}$ group-I intron sequences in their study of this intron in the cyanobacteria. To address these issues, we have used PCR and Southern-blot analyses to determine the presence/absence of the $tRNA^{Leu}$ group-I intron in a diverse group of algae in the Rhodophyta, Chlorophyta, Haptophyta, Heterokonta, and Cryptophyta, as well as in 330 diverse land plants. Twelve new intron sequences were determined and phylogenetic methods were used to elucidate the evolutionary history of these and the published $tRNA^{Leu}$ group-I intron sequences.

Materials and methods

Strains used in the analyses

The algae and plants used in the PCR analyses to determine the presence/absence of the $tRNA^{Leu}$ group-I intron are listed in Table 1. Genomic DNAs from 330 taxonomically diverse vascular plants were also tested, by Southern-blot analysis, for the presence

Table 1 Algae and plants used in the PCR analyses to determine the presence/absence of the $tRNA^{Leu}$ group-I intron. The GenBank accession number are given for taxa that contain the intron

Species	Strain	Class/division	Accession no.
Equisetum scirpoides	GUBG[a]	Sphenopsida	AF182539
Psilotum triquetrum	GUBG[a]	Psilopsida	AF182358
Mnium hornum	GUBG[a]	Bryopsida	AF182360
Coleochaete orbicularis	UTEX LB 2651[e]	Charophyceae	AF182368
Cosmarium botrytis	SAG 136.80[b]	Charophyceae	AF182362
Closterium cornu	SAG 132.80[b]	Charophyceae	AF182366
Closterium ehrenbergii	SAG 134.80[b]	Charophyceae	AF182367
Closterium littorale	SAG 611-7[b]	Charophyceae	AF182363
Genicularia spirotaenia	SAG B 54.86[b]	Charophyceae	AF182365
Penium margaritaceum	SAG B 22.82[b]	Charophyceae	AF182364
Cylindrocystis brebissonii	SAG 615-1[b]	Charophyceae	–
Roya obtusa	SAG B 168.80[b]	Charophyceae	AF182361
Bangia fuscopurpurea	SAG B59.81[b]	Rhodophyta	–
Compsopogon coeruleus	SAG 36.94[b]	Rhodophyta	–
Flintiella sanguinaria	SAG B40.94[b]	Rhodophyta	–
Galdieria sulphuraria	SAG 107.79[b]	Rhodophyta	–
Porphyra leucosticta	SAG B55.88[b]	Rhodophyta	–
Porphyridium aerugineum	SAG B43.94[b]	Rhodophyta	–
Rhodosorus marinus	SAG 116.79[b]	Rhodophyta	–
Stylonema alsidii	SAG B2.94[b]	Rhodophyta	–
Nemalionopsis tortuosa	SAG 42.9[b]	Rhodophyta	–
Thorea violacea	SAG 51.94[b]	Rhodophyta	–
Chrysochromulina acantha	CCMP 408[c]	Haptophyta	–
Emiliania huxleyi	CCMP 318[c]	Haptophyta	–
Pavlova lutheri	CCMP 375[c]	Haptophyta	–
Phaeocystis globosa	CCMP 374[c]	Haptophyta	–
Ochromonas danica	CBS[d]	Heterokonta	–
Dictyota dichotoma	Collected in Helgoland, Germany	Heterokonta	AF182369
Chroomonas sp.	SAG B980-1[b]	Cryptophyta	–
Pyrenomonas helgolandii	SAG B 28.87[b]	Cryptophyta	–

[a] Göttingen University Botanical Garden
[b] Sammlung von Algenkulturen Göttingen
[c] Provasoli-Guillard National Center for Culture of Marine Phytoplankton
[d] Carolina Biological Supply
[e] University of Texas Algal Collection

of the $tRNA^{Leu}$ group-I intron. The taxa included 20 ferns, five Lycopsida, one Psilopsida (*Psilotum nudum*), two Equisetopsida (*Equisetum arvense*, *Equisetum hyemale*), six cycads, *Gingko biloba*, nine conifers, four Gnetopsida, and 284 monocots and dicots (complete list available from J.D.P; see also Qiu et al. 1998).

DNA extraction, PCR and Southern methods, and sequencing

Total cellular DNA was isolated from certain taxa (all examined algae, plus *Equisetum scirpoides*, *Psilotum triquetrum*, and *Mnium hornum*) using the Plant DNeasy Mini Kit according to the manufacturer's instructions (Qiagen) or by using the CTAB method (Doyle and Doyle 1990). Tissue was disrupted with a mortar and pestle in the presence of liquid-N_2 prior to the DNA isolation. The $tRNA^{Leu}$ group-I intron was isolated from members of the Streptophyta with the following 5′ and 3′ primers, respectively: 5′-t^{Leu}, $^{5'}$TGGYGAAATYGGTAGACGCWRCGG AC$^{3'}$; 3′-t^{Leu}, $^{5'}$TGG GGATAGAGGGACTTGAACCCTCACGATTTTTA$^{3'}$ (Kuhsel et al. 1990). We made new PCR primers that were specific for the Rhodophyta and Cryptophyta on the basis of the published sequences of the $tRNA^{Leu}$ genes of *Cyanidioschyzon merolae* (Genbank, D63675), *Cyanidium caldarium* (D63676), *Porphyra purpurea* (U38804, Rhodophyta), and *Guillardia theta* (AF041468, Cryptophyta). None of these taxa contain the $tRNA^{Leu}$ group-I intron. The sequences of these primers are: 5′- Red, $^{5'}$TGTGGTAGAC ACAACAGACT$^{3'}$; 3′-Red, $^{5'}$GGACTTGAACCC TCACGATT$^{3'}$. The reaction conditions for all the PCRs were as follows: 94 °C for 10 min as pre-treatment, followed by 35 cycles at 94 °C for 1 min (denaturation), 55 °C, 60 °C or 65 °C for 2 min (primer-annealing), and 72 °C for 4 min (primer-extension). We included a final step at 72 °C for 10 min to complete the primer extension. The PCR products encoding the $tRNA^{Leu}$ group-I introns were gel-purified in 1.0% agarose gels and then cloned using the TA Cloning Kit according to the manufacturer's instructions (Invitrogen). Sequences were determined with the Prism Ready Reaction Dyedeoxy Terminator kit (Perkin Elmer) using an ABI 373 A sequencing machine. DNA synthesis was initiated from the T7 and M13 reverse primer sites encoded in the pCR2.1 plasmid. We verified the presence of PCR-amplifiable plastid DNA in the above taxa with PCR reactions using primers complementary to the 5′- and 3′-termini of the plastid-encoded small subunit ribosomal RNA (rRNA) and the PCR reaction conditions described above. The rRNA primers were as follows: SG1, $^{5'}$GTGCTGCAGAGAGTTY-GATCCTGGCTCAGG$^{3'}$; SG2, $^{5'}$CACGGATCCAAGGAGG TG ATCCANCCNCACC$^{3'}$ (Huss and Giovannoni 1989).

Genomic DNA from 330 vascular plants examined by Southern blots were isolated using the CTAB method (Doyle and Doyle 1990) and further purified with CsCl/ethidium bromide ultracentrifugation. Between 3 and 10 μg of DNA was digested from each species with the restriction endonuclease *Bam*HI and size-fractionated on agarose gels. The amount of plastid DNA loaded per lane was adjusted according to the results of initial hybridizations using a chloroplast 16 S rRNA gene probe from *Nicotiana tabacum* in order to obtain a relatively even intensity of hybridization signal. All Southern transfers were done with Immobilon nylon membranes (Millipore). The blots were probed with a random-prime labeled (^{32}P-dATP) $tRNA^{Leu}$ intron probe from *N. tabacum*. Hybridizations were done at 60 °C for 18 h in 5 × SSC, 50 mM Tris (pH 8.0), 0.1% SDS, 10 mM EDTA, and 2 × Denhardt's solution. Filters were washed twice for 30 min at 60 °C in 2 × SSC, 0.5% SDS prior to autoradiography.

Splicing assays

Linearized plasmid fragments containing cloned $tRNA^{Leu}$ introns from *P. triquetrum* and *C. botrytis* were used to generate RNA transcripts for *in vitro* splicing assays. The intron constructs contained 20 nt of 5′- and 41 nt of 3′-flanking tRNA sequence, including the regions required for the P1 and P10 interactions (Cech et al. 1994, see Fig. 3). Transcription was by T7 RNA polymerase according to the manufacturer's instructions (Stratagene). If splicing did not go to completion in the transcription buffer [40 mM Tris-HCl (pH 8.0), 8 mM $MgCl_2$, 50 mM NaCl, 2 mM spermidine, 30 mM DTT, 400 μM rNTPs] then the $MgCl_2$ concentration was either increased to 158 mM, or the spermidine concentration was increased to 7 mM, and the RNAs incubated at 37 °C for an additional 30 min. These conditions favor intron splicing (see Oliveira and Ragan 1994 for details). RNA splicing products were separated on polyacrylamide (4–6%)/7 M urea gels and stained with ethidium bromide prior to photography. These gels were blotted onto nylon membranes (Nytran, Schleicher and Schuell) and probed with sequences which included both intron and flanking exon regions. Probe preparation and detection of hybridizing fragments was done using a non-radioactive method according to the manufacturer's instructions (Gene Images, Amersham).

Phylogenetic analyses

A set of 75 $tRNA^{Leu}$ intron sequences was manually aligned (alignment available from GenBank) using the conserved P, Q, R, and S elements to identify homologous regions (Michel and Westhof 1990; Bhattacharya et al. 1994). A total of 191 aligned sequence positions (152 parsimony informative characters) were submitted to LogDet transformation and distance bootstrap phylogenetic analyses [done with the computer program, PAUP* V4.0b2 (Swofford 1999)]. In the LogDet transformation, gaps were excluded from the analysis and the neighbor-joining (Saitou and Nei 1987) method was used to build the phylogeny. A bootstrap analysis (2000 replications) was carried out using this method. For the distance bootstrap analysis (2000 replications), distance matrices were calculated according to the Maximum Likelihood model in PAUP*. Gaps were excluded in the distance calculations, the transition/transversion ratio was set at 2.0, and site-to-site rate heterogeneity was modeled as a gamma distribution (Yang 1994) with four discrete rate categories ($\alpha = 0.5$). Neighbor-joining trees were built from the distance matrices. The bootstrap values resulting from the LogDet transformation and distance analysis were included at the branches of the LogDet tree (for shared monophyletic groups).

The stability of key nodes in the LogDet tree was assessed with the Kishino–Hasegawa test [Kishino and Hasegawa 1989 (within PAUP*)]. In this analysis, we changed the topology of the LogDet tree using the MacClade computer program (version 3.07, Maddison and Maddison 1997) and compared these rearranged trees (in terms of the number of total parsimony steps) to the "best" tree to test different hypotheses about $tRNA^{Leu}$ intron evolution.

A total of 25 published cyanobacterial and all published plastid $tRNA^{Leu}$ group-I introns were included in our phylogenetic analyses. The divergent cluster-II introns were not analyzed because of their independent origins (Rudi and Jakobsen 1997; Paquin et al. 1999). We also excluded the highly degenerate, AT-rich intron sequences in the plastid genome of apicomplexans (e.g., Wilson et al. 1996). We did not analyze all the available plastid $tRNA^{Leu}$ group-I intron sequences from the angiosperms. They are too numerous to warrant inclusion here and most of these sequences have been used for systematic studies at the level of species/genera (e.g., Gielly and Taberlet 1994, 1996; Kita et al. 1995; Kajita et al. 1998). There are also many unpublished $tRNA^{Leu}$ group-I intron sequences available from the Bryopsida (e.g., Cox and Hedderson, see GenBank) and the Lycopsida (e.g., *Huperzia* spp., *Lycopodium* spp.; Wikstrom Kenrick, and Chase, see GenBank) of which we have included only a fraction in our phylogenies. The EMBL/GenBank/DDBJ accession numbers of the $tRNA^{Leu}$ sequences determined in this study are shown in Table 1. We found nine substitutions that differentiate the $tRNA^{Leu}$ group-I intron sequence of *M. hornum* that already exists in GenBank from that determined in this study. We refer to these sequences as *M. hornum* 1 and 2, respectively. The remaining introns included in our phylogenetic analyses were already available from the sequence databases.

Results and discussion

Distribution of the tRNA^{Leu} group-I intron in the Streptophyta

PCR-amplification using primers that recognize conserved regions flanking the anticodon loop of the tRNA^{Leu} gene gave products of a size expected for an intron-containing gene (300–500 bp; see Fig. 1A) for most members of the Streptophyta. These include all examined members of the Zygnematales (i.e., *Cosmarium botrytis*, *Closterium cornu*, *Closterium ehrenbergii*, *Closterium littorale*, *Genicularia spirotaenia*, *Penium margaritaceum* and *Roya obtusa*) except for *Cylindrocystis brebissonii*. In this alga, a fragment of only 81-bp was amplified; this is the size expected for a tRNA^{Leu} gene lacking any intron (Paquin et al. 1997). To verify the result with *C. brebissonii*, the 81-bp fragment was cloned and sequenced. The sequence, 5′-uggcgaaaucguuagacgcugcggacUU***AA**AAUCCGUUG GUUUGUGuaaaaaucgugaggguucaag ucccucuaucccc-3′ (with primer sequences shown in lower case), contains the anticodon (bold-face) and the base-paired regions of the anticodon loop (underlined, the intron insertion site is marked with the asterisk). We postulate that the tRNA^{Leu} group-I intron has been lost in *C. brebissonii*. PCR products expected to contain the tRNA^{Leu} group-I intron were also found in the charophyte *Coleochaete orbicularis*, in the bryophyte *M. hornum* (and in other mosses, see Fig. 3), and in *P. triquetrum* and *E. scirpoides*. All putatively intron-containing products were sequenced and shown to contain a group-I intron positioned at the identical site between the U and the A of the UAA anticodon loop.

Southern analyses of land-plant total cellular DNAs identified tRNA^{Leu} intron-hybridizing fragments (see Fig. 1B) in all 330 diverse vascular plants examined. These results underscore the remarkable stability of the tRNA^{Leu} group-I intron in land plants (Gielly and Taberlet 1994, 1996; Kita et al. 1995; Daros and Flores 1996; Kajita et al. 1998). Recent sequencing studies also show that the tRNA^{Leu} intron is found in several dozen species of Bryopsida (Cox and Hedderson, see GenBank) and in the Lycopsida (Wikstrom Kenrick and Chase, see GenBank).

The tRNA^{Leu} group-I intron in rhodophytes, cryptophytes, haptophytes, and heterokonts

We used PCR methods to search for group-I introns in the tRNA^{Leu} genes of two cryptophytes, two heterokonts, four haptophytes, and ten red algae. The plastids of the cryptophytes, heterokonts, and haptophytes are most likely derived from independent secondary endosymbioses of red algae by three disparate eukaryotic hosts (Douglas et al. 1991; Medlin et al. 1995, 1997; Daugbjerg and Andersen 1997 a,b; Fraunholz et al. 1997; Douglas and Penny 1999; Oliveira and Bhattacharya, in press). The red-algal endosymbionts appear to be either members of the Bangiophycidae or taxa ancestral to this group. The subclass Bangiophycidae is a paraphyletic assemblage of unicellular, filamentous, or simple blade-forming algae that diverge at the base of the Rhodophyta and from which the monophyletic subclass Florideophycidae is derived (Freshwater et al. 1994; Ragan et al. 1994). It is possible therefore that a tRNA^{Leu} group-I intron that was present in the red-algal plastid could have been transferred, via secondary endosymbiosis, into plastids derived from these taxa. Our analyses show, however, that a group-I intron does not interrupt the plastid tRNA^{Leu} gene of the eight bangiophytes [which represent the three orders of the subclass (Garbary and Gabrielson 1990)] and the two florideophytes that we examined. One of these putatively intron-lacking PCR products of 81 nt, was cloned and sequenced from *Compsopogon coeruleus*, and

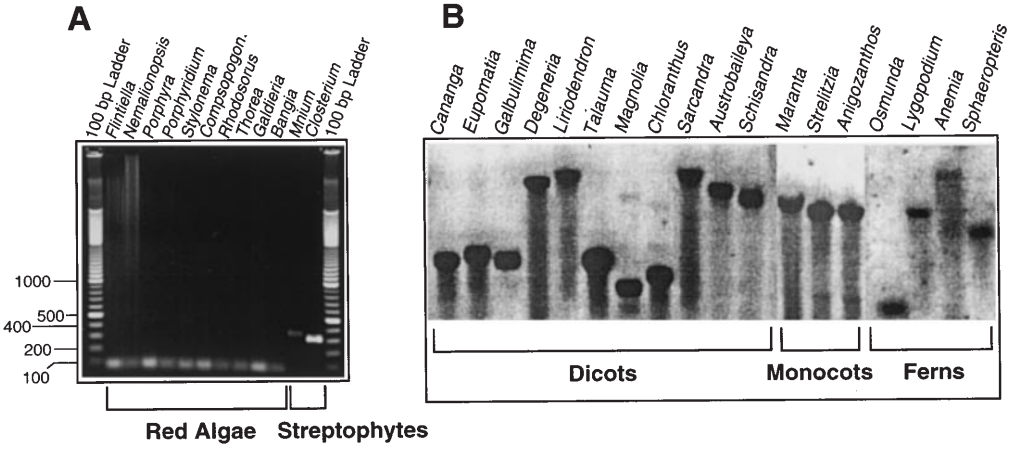


Fig. 1A,B Presence or absence of the tRNA^{Leu} group-I intron in diverse algae and land plants. **A** agarose gel (1.0%) showing tRNA^{Leu} gene PCR products either containing an intron [both streptophytes (*M. hornum*, *C. littorale*) examined here] or lacking an intron (all ten red algae). The 100-bp ladder is from New England Biolabs. **B** Southern-blot analysis showing hybridization of a tRNA^{Leu} intron probe from *N. tabacum* to each of 18 representative angiosperms and fern DNAs

found to contain an intronless tRNA^{Leu} gene (data not shown). These results are consistent with sequence data from the bangiophytes *Cyanidioschyzon merolae*, *Cyanidium caldarium* (Ohta 1997) and *P. purpurea* (Reith and Munholland 1995) that show intron-lacking tRNA^{Leu} genes. Kuhsel et al. (1990) reported that *Smithora naiadum* (Compsopogonales) also lacks a plastid tRNA^{Leu} intron. The tRNA^{Leu} gene of *Gracilaria lemaneiformis* (Floridiophycidae) was originally thought to contain the intron (Kuhsel et al. 1990) but this PCR product was subsequently found to be artefactual (J.D. Palmer, unpublished data). Together, these data suggest that most, perhaps all, extant Rhodophyta lack this group-I intron in their tRNA^{Leu} genes.

The PCR analyses of the cryptophyte and haptophyte algae also failed to reveal the presence of the tRNA^{Leu} intron, a result that is supported by sequence data from the cryptophyte *Guillardia theta* (Douglas and Penny 1999). Within the heterokonts, we found taxa with and without the tRNA^{Leu} intron. Previous work has shown that *Ochromonas danica*, *Dictyota dichotoma*, *Costaria costata*, and *Vaucheria bursata* all contain the intron (Kuhsel et al. 1990). Our phylogenetic analyses with these and other tRNA^{Leu} introns positioned the sequences from *O. danica* and *D. dichotoma* with strong bootstrap support *within* land plants (see below and Fig. 3A). This surprising result caused us to question the authenticity of these two ''algal'' sequences and we therefore did PCR analyses with newly extracted genomic DNA from both of these taxa and from *V. bursata*. These analyses showed that the newly tested *O. danica* DNA does not contain the tRNA^{Leu} intron whereas the *D. dichotoma* DNA yielded an intron that falls within a clade including the other heterokont intron sequences. The *V. bursata* intron that we sequenced was identical to that reported by Kuhsel et al. (1990). Conversely, PCR analysis with the genomic DNA of *O. danica* originally used by Kuhsel et al. (1990) resulted in the amplification of the original land plant-like intron (verified by sequence analysis). These results suggest that the intron sequences reported by Kuhsel et al. (1990) to be from *O. danica* and *D. dichotoma* were instead derived from contaminating DNA of different, but related, seed plants (see Fig. 3A). According to this interpretation, 3 of the 5 examined heterokonts contain this intron [two members of the Phaeophyceae (*C. costaria*, *D. dichotoma*) and the xanthophyte *V. bursata*]. The chrysophyte *O. danica* and the diatom *Odontella sinensis* (Kowallik et al. 1995) both lack the intron. Interestingly, phaeophytes and xanthophytes form a monophyletic group in small-subunit ribosomal DNA and *rbc*L phylogenies (e.g., Daugbjerg and Andersen 1997b; Potter et al. 1997).

Sequence and secondary structure conservation of the tRNA^{Leu} intron in the Streptophyta

All tRNA^{Leu} group-I intron sequences determined in this study were included in an alignment with previously sequenced tRNA^{Leu} introns from cyanobacteria and other plastids. The P, Q, R, and S regions (i.e., P3, P4, P7, J8/7, respectively) that form the catalytic core of group-I introns (Cech 1988; Michel and Westhof 1990) were used to guide this alignment. These regions are conserved in all tRNA^{Leu} group-I introns, as are more peripheral regions such as P2 and P5. Among the Streptophyta introns in our study, the alignable regions varied from 65% and 100% sequence identity (between *C. littorale-Lycopodiella lateralis*, and *Avena sativa-Hordeum vulgare-Triticum aestivum*, respectively).

The secondary structure of the *G. spirotaenia* tRNA^{Leu} intron was used to map the regions of identity within the Zygnematales intron lineage (Fig. 2). The only significant differences among these homologous introns was in the sequence composition and/or length of peripheral loops and helices (e.g., P6, P8 in *G. spirotaenia*, see Fig. 2). An extension of P8 was found in all the Zygnematales and in many angiosperm tRNA^{Leu} group-I introns. Conserved regions outside of the catalytic core included the GAAA tetraloop in L2 and a T-shaped P9 domain. Within the Zygnematales, the tRNA^{Leu} group-I introns varied from 74% to 92% sequence identity. In comparison, the highly conserved small subunit rRNA of these taxa varied from 85% to 99% sequence identity (D. Bhattacharya, unpublished data).

The phylogeny and splicing of the tRNA^{Leu} group-I intron in the Streptophyta

The 191 aligned sequence positions of tRNA^{Leu} introns were submitted to LogDet transformation and distance bootstrap analyses. The resulting trees are summarized in Fig. 3A. Within the Streptophyta and the Chlorophyta, the tRNA^{Leu} group-I intron phylogeny generally agrees with the phylogeny of the green algae and the land plants based on nuclear, mitochondrial and plastid-encoded molecular markers (Manhart 1994; Graham 1996; Kranz and Huss 1996; Lewis et al. 1997; Bhattacharya and Medlin 1998; Duff and Nickrent 1999). The most important difference between the intron and ''host'' cell trees is the unexpected polyphyly of the Chlorophyta (for a review of the phylogeny of the Chlorophyta see Friedl 1997). The intron sequences of *Bryopsis plumosa* and *Derbesia marina* (Ulvophyceae) appear as a highly divergent clade that is positioned outside of the Chlorophyta (here represented by two members of the Trebouxiophyceae, *Chlorella* spp.). The two ulvophyte introns group with the heterokont sequences in both LogDet and distance trees, with 78% bootstrap support in the latter analysis. Although at face value this placement suggests an independent origin of the tRNA^{Leu} intron in the ulvophytes, we think it probably reflects an artefactual misplacement (i.e., a long-branch ''attraction'') of the ulvophyte intron lineage in our phylogenetic analyses due to the high divergence of these sequences. Indeed, forced monophyly of the ulvophyte and chlor-

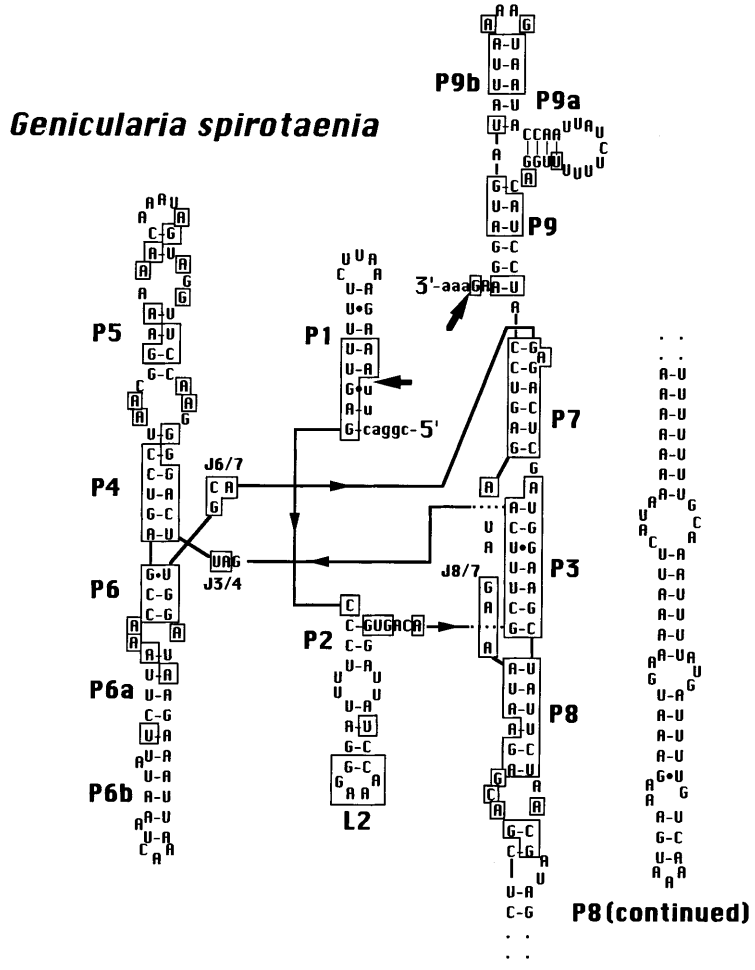


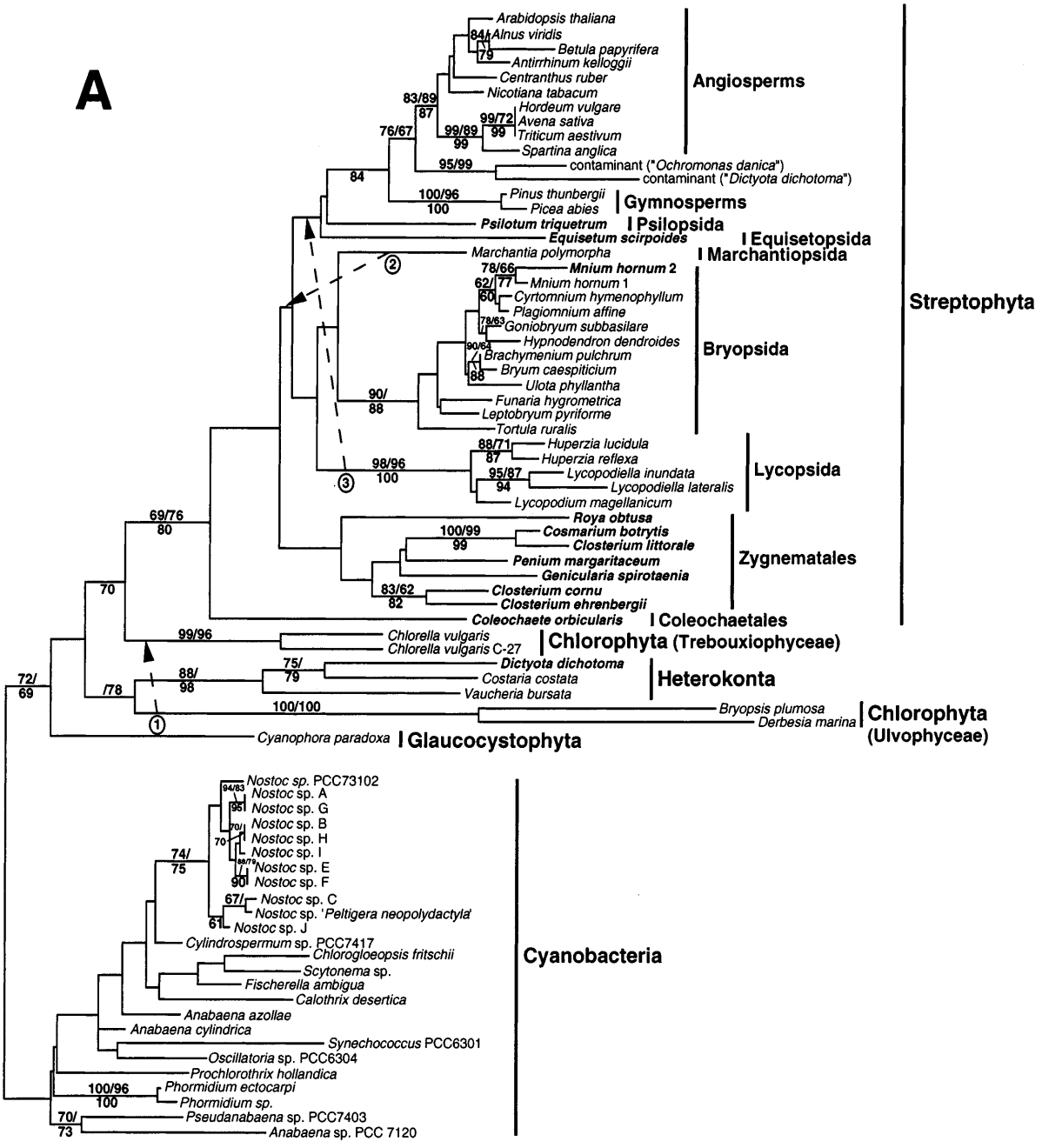
Fig. 2 Putative secondary structure of the group-I intron within the $tRNA^{Leu}$ precursor of *G. spirotaenia* (Zygnematales) drawn according to the convention of Cech et al. (1994). The 5′ and 3′ splice junctions are marked with *arrows* as are the locations of the pairing segments P1–P9. Positions that are identical in all Zygnematales $tRNA^{Leu}$ group-I introns are *boxed. Solid lines* connect secondary structure elements that are believed to interact in close proximity

ophyte introns (rearrangement 1 in Fig. 3A) results in a topology that is not significantly different ($P = 0.57$, Fig. 3B) from the LogDet tree.

Other differences between the intron tree and current views on the phylogeny of the Streptophyta are the placement of the Lycopsida and of the liverwort, *Marchantia polymorpha*. Fossil and molecular sequence data suggest that liverworts may be the first land plants (rearrangement 2 in Fig. 3A) and that the lycophytes are the sister group of the Euphyllophytina (rearrangement 3 in Fig. 3A, see Mishler et al. 1994; Kenrick and Crane 1997; Qiu et al. 1998; Duff and Nickrent 1999). The Kishino–Hasegawa test shows that placement of the *M. polymorpha* intron at the base of the land plants results in a tree that is significantly different ($P = 0.03$) from the LogDet tree shown in Fig. 3A, whereas rearrangement 3 results in a topology that is not significantly different ($P = 0.16$) from the LogDet tree. It is unclear how best to interpret these results because: (1) the branching order of the different streptophyte clades in Fig. 3A does not have significant bootstrap support, and (2) we do not have any hornwort intron sequences in our analyses and only a single liverwort sequence. It is, therefore, possible that the instability of the branch points in our $tRNA^{Leu}$ intron tree reflects the limited resolving power of the short intron sequences and inadequate taxon sampling. Removal of the divergent ulvophyte introns and the land plant contaminants (i.e., *O. danica* and *D. dichotoma*) from a bootstrap analysis with the LogDet transformation does not alter the topology of the tree shown in Fig. 3A. Analysis of the reduced data set shows, however, greater bootstrap support for the monophyly of the Chlorophyta/Streptophyta (70%, Fig. 3A).

Fig. 3A,B Phylogeny of $tRNA^{Leu}$ group-I introns. **A** tree built with the LogDet transformation and the neighbor-joining method based on comparisons of 191 aligned sequence positions using the PAUP* (V4.0b2, Swofford 1999) computer program. The tree should be regarded as unrooted, i.e., it is arbitrarily shown rooted on the branch between plastids and cyanobacteria. The results of bootstrap analyses are shown on the tree. The *values* shown above the branches and *to the left* of the slash-marks are bootstrap values using the LogDet method; values *to the right* of the slash-marks were made with the neighbor-joining method (maximum-likelihood model), and values shown *below* the branches were made using the LogDet method with the introns from the ulvophytes and the two land plant "contaminants" (see text) excluded from the analysis. Only bootstrap values ≥60% are shown. The *broken lines* indicate the tree rearrangements that were tested with the Kishino–Hasegawa test. Taxa that were sequenced in this study are in *boldface*. **B** results of the Kishino–Hasegawa test with rearranged trees 1–3. The LogDet tree shown in **A** was used as the "best" tree in these comparisons

In summary, the phylogenetic and presence/absence data support the hypothesis that, with the possible exception of the liverwort *M. polymorpha*, the $tRNA^{Leu}$ group-I intron has been vertically inherited (with some losses) since the origin of the Streptophyta. Possible explanations for the long-term vertical inheritance of the $tRNA^{Leu}$ intron in the Streptophyta include a low intron-loss rate in this lineage (e.g., Qiu et al. 1998) or the



B

Rearrangement	Length	Difference	s.d.(Diff.)	P*
LogDet	1091	(best)		
1	1093	2	3.47	0.57
2	1099	8	3.71	0.03*
3	1093	2	1.41	0.16

evolution of an intron function that selects for its retention. There is, however, no direct evidence to support either of these explanations. One implication of long-term intron vertical inheritence that can be tested biochemically is the loss of self-splicing ability. It is possible, for example, that over time the $tRNA^{Leu}$ group-I intron has become dependent on, and co-evolved with, a regulatory factor provided by the "host" cell to facilitate splicing. Consistent with this idea, in vitro splicing analyses show that the $tRNA^{Leu}$ introns from *P. triquetrum* and *C. botrytis* are unable to catalyze their self-excision from a precursor $tRNA^{Leu}$ transcript under the experimental conditions employed (Fig. 4, see also Cech 1988). These introns did not self-splice in any of the buffers used in *in vitro* splicing assays [i.e., only transcription buffer (lane 1), transcription buffer with $MgCl_2 = 158$ mM (lane 2), transcription buffer with spermidine = 7 mM (lane 3)]. First-step splicing products (i.e., intron + 3′-exon) of predicted sizes 489 nt and 540 nt in *P. triquetrum* and *C. botrytis*, respectively, are absent, as are linear introns (*P. triquetrum*, 450 nt; *C. botrytis*, 501 nt) and ligated exons (199 nt; 81 nt of exon sequence + 118 nt of pCR2.1 plasmid vector). The ligated exons would be slightly larger than the exon fragment of size 118 nt that is clearly resolved in the lanes containing the negative control (pCR2.1). Experiments are underway to express introns from different members of the Streptophyta in *Escherichia coli* to test for the effects of bacterial proteins/RNAs in facilitating $tRNA^{Leu}$ intron splicing. Although we hypothesize that an external factor may facilitate splicing of the $tRNA^{Leu}$ intron in vivo, it is also possible that our in vitro conditions did not allow for proper folding of the intron

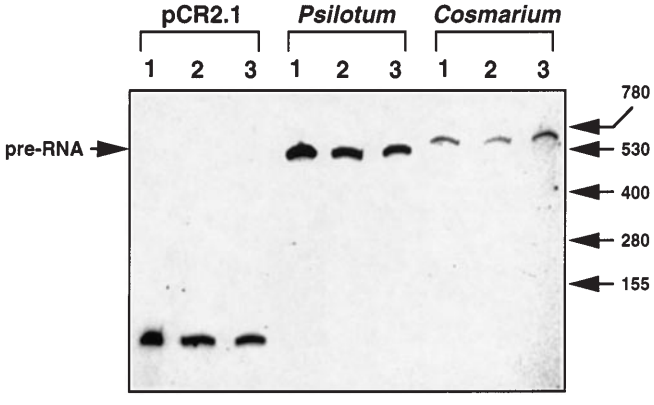


Fig. 4 Results of *in vitro* splicing assays with $tRNA^{Leu}$ group-I introns from two members of the Streptophyta. Linear DNA fragments containing the cloned $tRNA^{Leu}$ introns of *P. triquetrum* and *C. botrytis* were created with *Kpn*I/*Pvu*I double-digests. The negative control contained only the *Kpn*I/*Pvu*I fragment of the pCR2.1 plasmid (118 nt). To facilitate splicing, transcription products were either further incubated at 37 °C for 30 min in the transcription buffer (*lane 1*), or $MgCl_2$ was added to a concentration of 158 mM (*lane 2*), or spermidine was added to a concentration of 7 mM (*lane 3*). The pre-RNAs containing the *P. triquetrum* and *C. botrytis* introns are of sizes 649 nt and 700 nt, respectively. The RNA ladder is from BRL

RNA, thereby hindering self-splicing (Uhlenbeck 1995; Pan et al. 1997). In support of our results, however, previous analyses of plastid $tRNA^{Leu}$ group-I introns showed that the introns of *Cyanophora paradoxa* and *M. polymorpha* also cannot self-splice (Xu et al. 1990).

The inability of plastid $tRNA^{Leu}$ introns to self-splice is in contrast to their cyanobacterial homologs, which readily undergo auto-excision (Xu et al. 1990). This is an interesting observation when seen from the perspective of the widely different sequence divergence rates of the intron in plastids and cyanobacteria. Within cyanobacteria, the pairwise distances (uncorrected "p", PAUP*) between all introns vary from 1.6% between the two *Phormidium* species to 31.0% between *Anabaena* sp. PCC7120 and *Synechococcus* PCC6301 (the *Nostoc* J intron is used to represent all members of this genus). The average distance is 16.3% among all cyanobacterial introns. In comparison, there are greater distances between the plastid introns. To determine this, we used *Arabidopsis thaliana* as the standard and compared all other introns to it. In this comparison the distances varied from 3.2% betwen *A. thaliana* and *Alnus viridis* to 50.4% between *A. thaliana* and *Derbesia marina*. The average distance between *A. thaliana* and all other plastid introns was 22.6%. These results are readily apparent in the differences in branch lengths within the cyanobacterial and plastid intron groups in Fig. 3A. In addition, detailed analyses of the cyanobacterial $tRNA^{Leu}$ introns by Paquin et al. (1997) have shown high primary sequence and secondary structure conservation. As an example, 83% of sites were invariant in at least 10 of 13 cyanobacterial introns in their study (Paquin et al. 1997). The higher divergence within plastids is especially notable given that plastids are a younger group than cyanobacteria, in effect representing but a single, derived lineage of cyanobacteria.

The slower evolution of cyanobacterial introns may reflect stronger selective constraints because of the primary and secondary structural requirements of auto-excision. The plastid introns may have become dependent on a splicing factor early in plastid evolution (as suggested by the lack of self-splicing ability in all studied plastid introns), and this may have led to a relaxation in the constraints on intron sequence divergence. This scenario would provide an explanation for the two characteristics that most clearly differentiate plastid and cyanobacterial $tRNA^{Leu}$ introns, self-splicing ability and sequence divergence rates. On the other hand, these characteristics may be unrelated, as it appears that chloroplast molecules (such as small subunit rRNAs) tend, in general, to evolve faster than cyanobacterial homologues (e.g., see trees in Bhattacharya and Medlin 1995; Palmer and Delwiche 1998).

$tRNA^{Leu}$ group-I intron evolution and plastid phylogeny

The tree shown in Fig. 3A is consistent with the phylogeny of plastids inferred from plastid-encoded small-

subunit rRNA or other plastid protein-coding regions (Bhattacharya and Medlin 1995; Helmchen et al. 1995; Delwiche and Palmer 1997; Martin et al. 1998). The three primary plastid lineages identified by these studies are the green algal/land plant plastids (chloroplasts), the cyanelles of the glaucocystophytes, and the red-algal plastids (rhodoplasts), all of which are likely to have arisen from a single common cyanobacterial endosymbiosis (Bhattacharya and Medlin 1995; Delwiche and Palmer 1997; Palmer and Delwiche 1998). Consistent with this idea, the $tRNA^{Leu}$ group-I intron sequences, which represent all three lineages of primary plastids, form a monophyletic group, with moderate bootstrap support for this result (see also Paquin et al. 1997). The intron data, like the coding-region analyses (see above references), do not allow us to distinguish between a single primary endosymbiosis, or multiple primary endosymbioses involving closely related cyanobacteria, as the source of the primary plastid lineages.

An additional observation, based on our PCR analyses of ten red algae and previous data from four others, is the surprising absence of the $tRNA^{Leu}$ group-I intron in all examined rhodoplasts. Given that a $tRNA^{Leu}$ group-I intron was almost certainly present in the common ancestor of the red algae, this finding may be interpreted in the following manner. The $tRNA^{Leu}$ group-I intron was lost early in red-algal evolution *prior* to the origin of the cryptophyte and haptophyte plastids from (intron-lacking) red algae. The presence of the $tRNA^{Leu}$ group-I intron in phaeophyte (*Costaria costata*, *Dictyota dichotoma*) and xanthophyte (*V. bursata*) members of the Heterokonta suggests that its plastid originated even earlier, from a basal unsampled (or extinct) red-algal lineage that still contained the intron. Such a scenario would require at least two independent losses of the $tRNA^{Leu}$ intron within heterokonts, in chrysophytes (*O. danica*) and diatoms (*O. sinensis*, see Fig. 5). Alternatively, the endosymbioses that gave rise to the cryptophyte, haptophyte, and heterokont plastids may have occurred nearly simultaneously but involved different red algae that either still contained the intron (heterokonts), or had already lost the sequence (cryptophytes and haptophytes). In support of this view, molecular-clock calculations suggest that both the

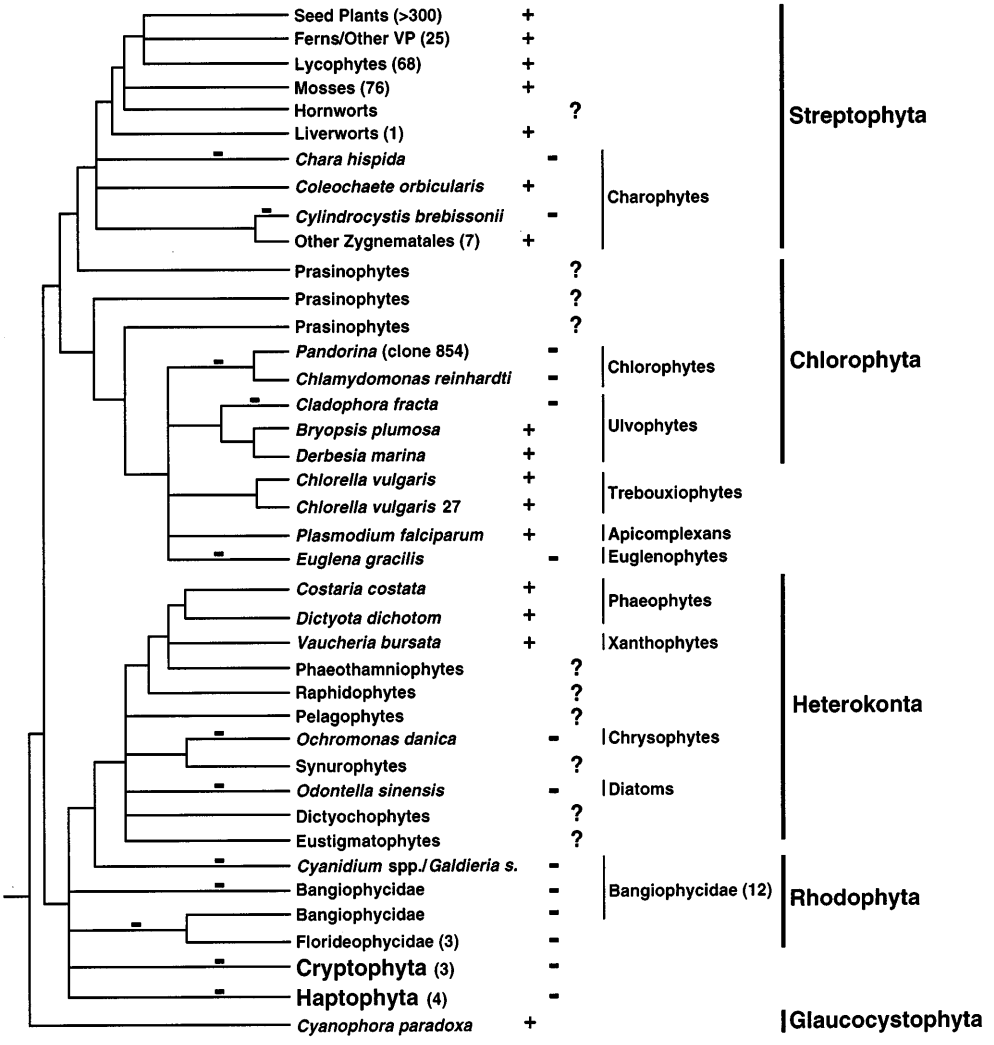


Fig. 5 Phylogic distribution of the $tRNA^{Leu}$ group-I intron among plastids. The three primary plastid lineages are the Chlorophyta/Streptophyta, Rhodophyta, and Glaucocystophyta. The plastids of the Cryptophyta, Haptophyta, and Heterokonta are derived from red-algal secondary endosymbioses, whereas the plastids of the Euglenophyta and the apicomplexans are thought to be derived from green-algal secondary endosymbioses (see text for references on which the tree topology is based). Intron-containing and -lacking lineages are shown with (+) and (−) signs, respectively. The number of taxa for which intron presence/absence data exist is shown after the species/lineage names. Lineages for which no intron data are available are marked with a question mark. (−) signs on the tree indicate inferred intron losses

haptophyte and heterokont plastids originated close in time, shortly before the Permian-Triassic boundary (Medlin et al. 1997). We can also imagine another scenario in which the red algae never contained the $tRNA^{Leu}$ group-I intron and that its existence in some heterokonts is explained by an independent lateral transfer of the intron into the common ancestor of the Phaeophyceae/Xanthophyceae.

To illustrate our current knowledge of the evolutionary history of the plastid $tRNA^{Leu}$ group-I intron, we have mapped its distribution on a tree (Fig. 5) that shows the presently accepted relationships of plastids (see Bhattacharya and Medlin 1998; Martin et al. 1998) and of the host cells (see Daugbjerg and Andersen 1997 b; Friedl 1997; Bailey et al. 1998; Bhattacharya et al. 1998; Palmer and Delwiche 1998; Qiu et al. 1998). The $tRNA^{Leu}$ group-I intron presumably originated early in cyanobacterial evolution (see Paquin et al. 1997 for details) and was laterally transferred into the common ancestor of the three primary algal lineages (Chlorophyta, Rhodophyta, Glaucocystophyta) through primary endosymbiosis. These three lineages are thought to have diverged nearly simultaneously from each other (Delwiche and Palmer 1997), with the cyanelles most likely being the first divergence (Martin et al. 1998).

Within the Streptophyta, the intron was present at least in the common ancestor of the charophytes and appears to have been vertically inherited thereafter. No losses have occurred among 330 examined land plants, whereas two losses are apparent in charophytes [*Chara hispida* (Kuhsel et al. 1990) and *C. brebissonii*]. The lack of resolution among members of the Chlorophyta, some of which have the intron and some of which do not, makes it difficult to infer its evolutionary history in this group. In addition, no data are available from the Prasinophyceae, a paraphyletic group that diverges at the base of the Chlorophyta (e.g., *Mantoniella squamata*, Friedl 1997) and the Streptophyta (i.e., *Mesostigma viride*, Bhattacharya et al. 1998). The plastid of *Euglena gracilis* (Hallick et al. 1993), which very likely originated from a green-algal secondary endosymbiosis, lacks the intron. Conversely, the remnant plastid genome of the nonphotosynthetic apicomplexan parasite, *Plasmodium falciparum* (Wilson et al. 1996), which may also be of green-algal origin (Kohler et al. 1997), contains a degenerate intron in the homologous position of the $tRNA^{Leu}$ gene. The *P. falciparum* $tRNA^{Leu}$ intron is extremely A + T-rich (92%) and is reduced in length (134 nt). This intron sequence cannot be aligned with other $tRNA^{Leu}$ introns due to its biased nucleotide content, and therefore its ancestry has not been examined with phylogenetic methods.

Unique evolutionary history of the $tRNA^{Leu}$ group-I intron

The $tRNA^{Leu}$ intron is unique among the known group-I and -II introns that interrupt organellar coding regions. No other group-I or group-II intron has such a wide distribution and has been vertically inherited for over a billion years, i.e., since the cyanobacterial origin of plastids (Kuhsel et al. 1990; Xu et al. 1990). Complete genome sequences suggest that, except for the $tRNA^{Leu}$ intron, all introns that are present in the plastid and mitochondrion were acquired after the endosymbioses that gave rise to these organelles (for a review see Palmer and Delwiche 1998). The plastid genome of *Cyanophora paradoxa* (one of the primary plastid lineages), for example, contains only the $tRNA^{Leu}$ intron (Stirewalt et al. 1995) whereas most (perhaps all) extant rhodophytes have lost the intron. The only other intron known from rhodophyte plastids is a group-II-like intron in the *rpeB* gene of *Rhodella violacea* (Bernard et al. 1992). This distribution suggests that the ancestor of all plastids contained the single intron interrupting the $tRNA^{Leu}$ gene.

Acknowledgements We thank Tom Vincent for carrying out some of the land-plant PCR analyses, Mariana Oliveira (Sao Paulo) for the genomic DNA from the red algae, Linda Medlin (Bremerhaven) for the haptophyte genomic DNA, Klaus Kowallik (Düsseldorf) for his gift of *Dictyota* and *Vaucheria* DNA, Chuck Delwiche (College Park) for the *Ochromonas* culture, and Thomas Proeschold (Göttingen) for help with the algal cultures. D.B. was supported by the College of Liberal Arts (University of Iowa) and a grant from the Carver Foundation. Y.L.Q. acknowledges NIH postdoctoral fellowship GM-17923, and J.D.P acknowledges NIH grant GM-35087.

References

Bailey JC, Bidigare RR, Christensen SJ, Andersen RA (1998) Phaeothamniophyceae Classis Nova: a new lineage of chromophytes based upon photosynthetic pigments, *rbc*L sequence analysis and ultrastructure. Protist 149:245–263

Bernard C, Thomas JC, Mazel D, Mousseau A, Castets AM, Tandeau de Marsac N, Dubacq JP (1992) Characterization of the genes encoding phycoerythrin in the red alga *Rhodella violacea*: evidence for a splitting of the *rpeB* gene by an intron. Proc Natl Acad Sci USA 89:9564–9568

Bhattacharya D, Medlin L (1995) The phylogeny of plastids: a review based on comparisons of small subunit ribosomal RNA coding regions. J Phycol 31:489–498

Bhattacharya D, Medlin L (1998) Algal phylogeny and the origin of land plants. Plant Physiol 116:9–15

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). Proc Natl Acad Sci 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. Mol Biol Evol 13:978–989

Bhattacharya D, Weber K, An SS, Berning-Koch W (1998) Actin phylogeny identifies *Mesostigma viride* as a flagellate ancestor of the land plants. J Mol Evol 47:544–560

Biniszkiewicz D, Cesnaviciene E, Shub DA (1994) Self-splicing group-I intron in cyanobacterial initiator methionine tRNA: evidence for lateral transfer of introns in bacteria. EMBO J 13:4629–4635

Brocks JJ, Logan GA, Buick R, Summons RE (1999) Archean molecular fossils and the early rise of eukaryotes. Science 285:1033–1036

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

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

Damberger SH, Gutell RR (1994) A comparative database of group-I intron structures. Nucleic Acids Res 22:3508–3510

Daros JA, Flores R (1996) A group-I plant intron accumulates as circular RNA forms with extensive 5′ deletions in vivo. RNA 2:928–936

Daugbjerg N, Andersen RA (1997 a) Phylogenetic analyses of the *rbc*L sequences from haptophytes and heterokont algae suggest their chloroplasts are unrelated. Mol Biol Evol 14:1242–1251

Daugbjerg N, Andersen RA (1997 b) A molecular phylogeny of the heterokont algae based on analyses of chloroplast-encoded *rbc*L sequence data. J Phycol 33:1031–1041

De Jonckheere JF, Brown S (1994) Loss of the ORF in the SSU rDNA group-I intron of one *Naegleria* lineage. Nucleic Acids Res 22: 3925–3927

Delwiche DF, Palmer JD (1997) The origin of plastids and their spread via secondary symbiosis. In: Bhattacharya D (ed) Origins of algae and their plastids. Springer-Verlag, Wien, pp 53–86

Douglas SE, Penny SL (1999) The plastid genome from the cryptomonad alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae. J Mol Evol 48:235–244

Douglas SE, Murphy CA, Spencer DF, Gray MW (1991) Molecular evidence that cryptomonad algae are evolutionary chimaeras of two phylogenetically distinct eukaryotes. Nature 350:148–151

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Duff RJ, Nickrent DL (1999) Phylogenetic relationships of land plants using mitochondrial small-subunit rDNA sequences. Am J Bot 86:372–386

Fraunholz MJ, Wastl J, Zauner S, Rensing SA, Scherzinger MM, Maier U-G (1997) The evolution of cryptophytes. In: Bhattacharya D (ed) Origins of algae and their plastids. Springer-Verlag, Wien, pp 163–174

Freshwater DW, Fredericq S, Butler BS, Hommersand MH, Chase MW (1994) A gene phylogeny of the red algae (Rhodophyta) based on plastid *rbc*L. Proc Natl Acad Sci USA 91:7281–7285

Friedl T (1997) Evolution of green algae. In: Bhattacharya D (ed) Origins of algae and their plastids. Springer-Verlag, Wien, pp 87–102

Friedl T, Besendahl A, Pfeiffer P, Bhattacharya D (in press) The distribution of group-I introns in lichen algae suggests that lichenization facilitates intron lateral transfer. Mol Phylog Evol

Garbary DJ, Gabrielson PW (1990) Taxonomy and evolution. In: Cole KM, Sheath RG (eds) Biology of the red algae. Cambridge University Press, New York

Gielly L, Taberlet P (1994) The use of chloroplast DNA to resolve plant phylogenies: non-coding versus *rbc*L sequences. Mol Biol Evol 11:769–777

Gielly L, Taberlet P (1996) A phylogeny of the European gentians inferred from chloroplast $tRNA^{Leu}$ (UAA) intron sequences. Bot J Linn Soc 120:57–75

Gilbert W (1986) The RNA world. Nature 319:618

Graham LG (1996) Green algae to land plants: an evolutionary transition. J Plant Res 109:241–251

Hallick RB, Hong L, Drager RG, Favreau MR, Monfort A, Orsat B, Spielmann A, Stutz E (1993) Complete sequence of *Euglena gracilis* chloroplast DNA. Nucleic Acids Res 21:3537–3544

Helmchen T, Bhattacharya D, Melkonian M (1995) Analyses of ribosomal RNA sequences from glaucocystophyte cyanelles provide new insights into the evolutionary relationships of plastids. J Mol Evol 41:203–210

Hibbett DS (1996) Phylogenetic evidence for horizontal transmission of group-I introns in the nuclear ribosomal DNA of mushroom-forming fungi. Mol Biol Evol 13:903–917

Huss VAR, Giovannoni SJ (1989) Primary structures of the chloroplast small subunit ribosomal RNA gene from *Chlorella vulgaris*. Nucleic Acids Res 17:9487

Kajita T, Kamiya K, Nakamura K, Tachida H, Wickneswari R, Tsumura Y, Yoshimaru H, Yamazaki T (1998) Molecular phylogeny of Dipetrocarpaceae in Southeast Asia based on nucleotide sequences of matK, trnL intron, and trnL-trnF intergenic spacer region in chloroplast DNA. Mol Phylog Evol 10:202–209

Kenrick P, Crane PR (1997) The origin and early evolution of plants on land. Nature 89:33–39

Kishino H, Hasegawa M (1989) Evaluation of the maximum-likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. J Mol Evol 29:170–179

Kita Y, Ueda K, Kadota Y (1995) Molecular phylogeny and evaluation of the Asian *Aconitum* subgenus *Aconitum* (Ranunculaceae). J Plant Res 108:429–442

Kohler S, Delwiche CF, Denny PW, Tilney LG, Webster P, Wilson RJ, Palmer JD, Roos DS (1997) A plastid of probable green algal origin in Apicomplexan parasites. Science 275:1485–1489

Kowallik KV, Stoebe B, Schaffran I, Kroth-Pancic P, Freier U (1995) The chloroplast genome of a chlorophyll a + c-containing alga, *Odontella sinensis*. Plant Mol Biol Rep 13: 336–342

Kranz HD, Huss VAR (1996) Molecular evolution of ferns and allies, and their relationship to seed plants: evidence from complete 18S rRNA gene sequences. Plant Syst Evol 202:1–11

Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. Cell 31:147–157

Kuhsel MG, Strickland R, Palmer JD (1990) An ancient group-I intron shared by eubacteria and chloroplasts. Science 250:1570–1573

Lambowitz AM, Belfort M (1993) Introns as mobile genetic elements. Annu Rev Biochem 62:587–622

Lewis LA, Mishler BD, Vilgalys R (1997) Phylogenetic relationships of the liverworts (Hepaticae), a basal embryophyte lineage, inferred from nucleotide sequence data of the chloroplast gene *rbc*L. Mol Phylog Evol 7:377–393

Maddison WP, Maddison DR (1997) MacClade, version 3.07. Sinauer, Sunderland, Massachusetts

Manhart JR (1994) Phylogenetic analysis of green plant *rbc*L sequences. Mol Phylog Evol 3:114–127

Martin W, Stöbe B, Goremykin V, Hansmann S, Hasegawa M, Kowallik KV (1998) Gene transfer to the nucleus and the evolution of chloroplasts. Nature 393:162–165

Medlin LK, Cooper A, Hill C, Wrieden S, Wellbrook U (1995) Phylogenetic position of the Chromista plastids based on small subunit rRNA coding regions. Curr Genet 28:560–565

Medlin LK, Kooistra WHCF, Potter D, Saunders GW, Andersen RA (1997) Phylogenetic relationships of the 'golden algae' (haptophytes, heterokont chromophytes) and their plastids. In: Bhattacharya D (ed) Origins of algae and their plastids. Springer-Verlag, Wien, pp 187–220

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

Mishler BD, Lewis LA, Buchheim MA, Renzaglia KS, Garbary DJ, Delwiche CF, Zechman FW, Kantz TS, Chapman RL (1994) Phylogenetic relationships of the "green algae" and "bryophytes." Ann Missouri Bot Gard

Moore MJ, Query CC, Sharp PA (1993) Splicing of precursors to mRNA by the spliceosome. In: Gesteland RF, Atkins JF (eds) The RNA world. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp 303–357

Nishida K, Suzuki S, Kimura Y, Nomura N, Fujie M, Yamada T (1998) Group-I introns found in *Chlorella* viruses: biological implications. Virology 15:319–326

Ohta N (1997) Analysis of a plastid gene cluster reveals a close relationship between *Cyanidioschyzon* and *Cyanidium*. J Plant Res 110:235–245

Oliveira MC, Bhattacharya D (in press) Phylogeny of the Bangiophycidae (Rhodophyta) and the secondary endosymbiotic origin of plastids. Am. J Bot

Oliveira MC, Ragan MA (1994) Variant forms of a self-splicing group-I intron in nuclear small-subunit rRNA genes of the marine red alga *Porphyra spiralis* var *amplifolia*. Mol Biol Evol 11:195–207

Pace NR, Marsh TK (1985) RNA catalysis and the origin of life. Origins of Life 16:97–116

Palmer JD, Delwiche CF (1998) The origin and evolution of plastids and their genomes. In: Soltis DE, Soltis PS, Doyle JJ (eds) Molecular systematics of plants. II. DNA sequencing. Kluwer, Norwell, Massachuesetts

Pan J, Thirumalai D, Woodson SA (1997) Folding of RNA involves parallel pathways. J Mol Biol 273:7–13

Paquin B, Kathe SD, Nierzwicki-Bauer SA, Shub DA (1997) Origin and evolution of group-I introns in cyanobacterial tRNA genes. J Bacteriol 179:6798–6806

Paquin B, Heinfling A, Shub DA (1999) Sporadic distribution of tRNA(Arg)CCU introns among alpha-purple bacteria: evidence for horizontal transmission and transposition of a group-I intron. J Bacteriol 181:1049–1053

Potter D, Saunders GW, Andersen RA (1997) Phylogenetic relationships of the Raphidophyceae and Xanthophyceae as inferred from nucleotide sequences of the 18S ribosomal RNA gene. Am J Bot 84:966–972

Qiu Y-L, Cho Y, Cox C, Palmer JD (1998) The gain of three mitochondrial introns identifies liverworts as the earliest land plants. Nature 394:671–674

Ragan MA, Bird CJ, Rice EL, Gutell RR, Murphy CA, Singh RK (1994) A molecular phylogeny of the marine red algae (Rhodophyta) based on the nuclear small-subunit rRNA gene. Proc Natl Acad Sci USA 91:2776–7280

Reith ME, Munholland J (1995) Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. Plant Mol Biol Rep 13:333–335

Rudi K, Jakobsen KS (1997) Cyanobacterial $tRNA^{Leu}$(UAA) group-I introns have a polyphyletic origin. FEMS Microbiol Lett 156:293–298

Rudi K, Jakobsen KS (1999) Complex evolutionary patterns of $tRNA^{Leu}$ (UAA) group-I introns in cyanobacterial radiation. J Bacteriol 181:3445–3451

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Schopf JW (1993) Microfossils of the early Archaean Apex chert: new evidence of the antiquity of life. Science 260:640–646

Stirewalt V, Michalowski C, Löffelhardt W, Bohnert H, Bryant D (1995) Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*. Plant Mol Biol Rep 13:327–332

Summons RE, Jahnke LL, Hope JM, Logan GA (1999) 2-Methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis. Nature 400:554–557

Swofford DL (1999) PAUP: Phylogenetic analysis using parsimony, V4.0b2. Smithsonian Institution, Washington D.C.

Turmel M, Gutell RR, Mercier JP, Otis C, Lemieux C (1993) Analysis of the chloroplast large subunit ribosomal RNA gene from 17 *Chlamydomonas* taxa. Three internal transcribed spacers and 12 group-I intron insertion sites. Mol Biol Evol 232:446–467

Uhlenbeck OC (1995) Keeping RNA happy. RNA 1:4–6

Wilson RJM, Denny PW, Preiser PR, Rangachari K, Roberts K, Roy A, Whyte A, Strath M, Moore DJ, Moore PW, Williamson DH (1996) Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. J Mol Biol 261:155–172

Xu M-Q, Kathe SD, Goodrich-Blair H, Nierzwicki-Bauer SA, Shub DA (1990) Bacterial origin of a chloroplast intron: conserved self-splicing group-I introns in cyanobacteria. Science 250:1566–1570

Yamada T, Tamura K, Aimi T, Songsri P (1994) Self-splicing group-I introns in eukaryotic viruses. Nucleic Acids Res 22: 2532–2537

Yang Z (1994) Maximum-likelihood Phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol 39:306–314