Evolutionary Transfer of ORF-containing Group I Introns between Different Subcellular Compartments (Chloroplast and Mitochondrion)

Monique Turmel,* Vincent Côté,* Christian Otis,* Jean-Patrick Mercier,* Michael W. Gray,† Kim M. Lonergan,†¹ and Claude Lemieux*

Program in Evolutionary Biology, Canadian Institute for Advanced Research; *Département de biochimie, Université Laval; and †Department of Biochemistry, Dalhousie University

We describe here a case of homologous introns containing homologous open reading frames (ORFs) that are inserted at the same site in the large subunit (LSU) rRNA gene of different organelles in distantly related organisms. We show that the chloroplast LSU rRNA gene of the green alga Chlamydomonas pallidostigmatica contains a group I intron (CpLSU-2) encoding a site-specific endonuclease (I-Cpal). This intron is inserted at the identical site (corresponding to positions 1931–1932 of the Escherichia coli 23S rRNA sequence) as a group I intron (AcLSU-1.m1) in the mitochondrial LSU rRNA gene of the amoeboid protozoan Acanthamoeba castellanii. The CpLSU-2 intron displays a remarkable degree of nucleotide similarity in both primary sequence and secondary structure to the AcLSU-1.m1 intron; moreover, the Acanthamoeba intron contains an ORF in the same location within its secondary structure as the CpLSU-2 ORF and shares with it a strikingly high level of amino acid similarity (65%; 42% identity). A comprehensive survey of intron distribution at site 1931 of the chloroplast LSU rRNA gene reveals a rather restricted occurrence within the polyphyletic genus Chlamydomonas, with no evidence of this intron among a number of non-Chlamydomonad green algae surveyed, nor in land plants. A parallel survey of homologues of a previously described and similar intron/ORF pair (C. reinhardtii chloroplast CrLSU/A. castellanii mitochondrial AcLSU-1.m3) also shows a restricted occurrence of this intron (site 2593) among chloroplasts, although the intron distribution is somewhat broader than that observed at site 1931, with site-2593 introns appearing in several green algal branches outside of the Chlamydomonas lineage. The available data, while not definitive, are most consistent with a relatively recent horizontal transfer of both site-1931 and site-2593 introns (and their contained ORFs) between the chloroplast of a Chlamydomonas-type organism and the mitochondrion of an Acanthamoeba-like organism, probably in the direction chloroplast to mitochondrion. The data also suggest that both introns could have been acquired in a single event.

Introduction

Group I introns have widespread but irregular distributions among very diverse organisms and genetic systems, being found in bacteria and bacteriophages as well as in the nuclei, chloroplasts, and mitochondria of eukaryotes (see Cech 1988; Michel and Westhof 1990; Reinhold-Hurek and Shub 1992; Saldanha et al. 1993). To date, more than 200 group I introns, inserted in a variety of genes, have been identified (Damberger and Gutell 1994). These have been classified into 11 subgroups (IA1–IA3, IB1–IB4, IC1–IC3, and ID) on the basis of comparative sequence analysis (Michel and Westhof 1990).

Although group I introns occur in a wide diversity of organisms and genomes, the great majority of these genetic elements appear to be of recent vintage (Cavalier-Smith 1991; Palmer and Logsdon 1991). Persuasive evidence for an ancient origin has been reported only for the intron in the tRNA Leu genes (UAA anticodon) of chloroplasts and cyanobacteria, which has been conserved in structure and position over at least a billion years (Kuhsel et al. 1990; Xu et al. 1990). The view that most group I introns originated subsequent to the proliferation of a few ancestral introns is supported not only by their restricted and idiosyncratic distribution but also by the proven mobility of this class of elements. All introns that have been found to be mobile in genetic crosses between intron-plus and intron-minus strains encode DNA endonucleases that cleave the intronless alleles at their cognate site and thereby initiate the invasion of all alleles by the same endonuclease-encoding introns (Dujon 1989; Lambowitz and Belfort 1993). As loss of group
I introns is relatively frequent through integration of reverse-transcribed cDNA copies into the genome (Dujon 1989), this type of intron mobility, called homing, is thought to ensure the maintenance of a given intron insertion site within a particular lineage. Another type of intron mobility, allowing introns to move to their cognate site or a new gene location, is initiated by reversal of the self-splicing reaction, followed by reverse transcription and homologous recombination (Woodson and Cech 1989).

Transfer of a group I intron to divergent species or a different cell compartment by the reverse-splicing mechanism or by a mechanism analogous to intron homing is also anticipated if an intron encounters a new genetic environment during promiscuous events that do not involve sexual exchange. Given that migration of DNA from one cell compartment to another and incorporation of foreign DNA sequences into genomes have both been documented (Palmer 1985, 1991; Gray 1992), such intron movements via lateral transfer are predicted to occur at a detectable frequency. However, only one of the several group I introns reported to be genetically mobile and/or to encode site-specific endonucleases conferring this invasive potential has been found to cohabit divergent genomes (Lonergan and Gray 1994). This intron is the chloroplast intron encoding the I-CreI endonuclease in the large subunit (LSU) rRNA gene of the green alga \textit{Chlamydomonas reinhardtii} (Dürrengerber and Rochaix 1991). It is structurally very similar to the mitochondrial IA3 intron AcLSU·m3 occupying the same insertion site in the LSU rRNA gene of the amoeboid protozoan \textit{Acanthamoeba castellanii}, a resemblance that has been attributed to transferred of a group I intron between chloroplasts and mitochondria during evolution (Lonergan and Gray 1994).

In the present study, we have identified a group I intron encoding a novel site-specific endonuclease (I-CpAl) in the chloroplast LSU rRNA gene of the green alga \textit{Chlamydomonas pallidostigmatica}. This intron shares remarkable similarity in both primary sequence and secondary structure with the \textit{A. castellanii} mito-
chondrial IB4 AcLSU·m1 intron (Lonergan and Gray 1994) inserted at the same position.

Material and Methods

DNA Sequencing and Analysis

A 2163-bp segment of the \textit{C. pallidostigmatica} (SAG 9.83) chloroplast LSU rRNA gene that contains part of exon 1, intron 1, exon 2, intron 2, and part of exon 3 (Turmel et al. 1993a) was PCR-amplified under previously described conditions (Turmel et al. 1993b) from a chloroplast DNA (cpDNA)-enriched preparation (10 ng) (Turmel et al. 1993a) using the primers 5'-ACAGGTCTCCGCAAAGTCTGTA-3' (#53) and 5'-AT-

TACGCCCTTCTGACAGGCC-3' (#56). The PCR product was partially sequenced by the dideoxy chain termination method using the dSNA cycle sequencing system of Bethesda Research Laboratories (Gaithersburg, Md.). Sequencing reactions were initiated with two oligonucleotides complementary to exon 3 (including primer #56), three primers complementary to intron 2 and one primer complementary to intron 1. Sequence analysis was carried out with the Genetics Computer Group (1991) software package, with the exception of the sequence alignments shown in figure 1, which were generated using CLUSTAL V (Higgins and Sharp 1988).

Endonuclease Assays

I-CpAl and I-CeuI were synthesized in vitro and assayed as described by Côté et al. (1993) with the following modifications. Regions encompassing the CpsLSU·2 and CeLSU·5 intron open reading frames (ORFs) were PCR-amplified from \textit{C. pallidostigmatica} and \textit{Chlamydomonas eugametos} (UTE9) cpDNA-enriched preparations (10 ng), respectively, using Taq DNA polymerase (BIO/CAN Scientific, Mississauga, Canada) and the following primers: for the CpsLSU·2 ORF, 5'-TAATACGACTCTATAGGAGATAGAGAGAATTTGAAAAATTG-3' (#618) and 5'-GGCTGATCCTCATTTAAATC-3' (#619); for the CeLSU·5 ORF, 5'-TAATACGACTCTATAGGAGAACGTAGTGACACAGCTAA-3' (#621) and 5'-ATTAGCGCTTTCGTAGAGGCC-3' (#56). Aliquots (1 µg) of in vitro–synthesized RNA were translated in the rabbit reticulocyte lysate system of Promega (Madison, Wisc.) (50 µl) at 30°C for 1.5 hr in the presence of 40 µCi of [32P]methionine (1,200 Ci/mmole). Following this reaction, endonuclease assays were carried out at 37°C for 1 hr in 50 µl of 20 mM Tris-HCl (pH 8.0), 10 mM MgCl2, and 1 mM dithiothreitol using as DNA substrates 32P-labeled, PCR-amplified fragments that were generated from total cellular DNA preparations (10 ng) of \textit{Chlamydomonas gigantea} (UTE9 LB 848), \textit{Chlamydomonas callosa} (SAG 9.72), \textit{Chlamydomonas moewusii} (UTE9 97), and \textit{Chlamydomonas mutabilis} (SAG 34–72) with the following primers: for \textit{C. gigantea} and \textit{C. moewusii}, 5'-CAGGTCTCCGCAAAGTCTGTA-3' (#53) and 5'-AGTCGCCATCTTCACGGGACA-3' (#517); for \textit{C. callosa}, oligonucleotide #53 and 5'-AGGTGTTAAAAATTTGAGTG-3' (#628); and for \textit{C. mutabilis}, 5'-GTTTTACAAGATGCTCAGATT-3' (#630) and oligonucleotide #517. For each of these PCR reactions, 50 ng of each of the primers were end-labeled using [γ-32P]ATP (6,000 Ci/mmole) and T4 polynucleotide kinase and these labeled primers were then added to the reaction mixture along with 450 ng of each of the unlabeled primers. Cleavage reactions were initiated by the addition of 5–10 µl of the translation mixture con-
Fig. 1.—Comparison of the chloroplast CpsU-2 and mitochondrial AcLSU- m1 introns. A, Comparison of secondary structure models. The CpsU-2 intron, whose sequence is reported in the GenBank accession number L36830, is represented according to the structural conventions proposed by Burke et al. (1987). Outside the ORF, residues that are conserved between the CpsU-2 and AcLSU- m1 introns are indicated with white letters on a black background. Splice sites are denoted by arrows between the exon (lower case) and intron (upper case) nucleotides, while the start and stop codons of the ORF are denoted by asterisks. The sequence (EMBL accession number U03732) and secondary structure of AcLSU- m1 are taken from Lonergan and Gray (1994). B, Sequence alignment of the ORFs. Identical residues are indicated with white letters on a black background. Dashes represent gaps to optimize the alignment. C, Sequence alignment of the predicted proteins encoded by the ORFs. Conserved amino acids are marked with asterisks below the alignment. The conserved motif LAGLI-DADG is indicated. Other details are as above.

taining I-CpaI or I-CeuI and were stopped by the addition of EDTA to 10 mM. All DNA samples were electrophoresed in 5% polyacrylamide-9 M urea gels.

Determination of the Cleavage Pattern of I-CpaI

The C. gigantea cpDNA region containing the target site of I-CpaI was PCR-amplified in the presence of
the unlabeled primers #53 and #517. The purified PCR product was used as a template in sequencing reactions performed with the dsDNA cycle sequencing system of Bethesda Research Laboratories (Gaithersburg, Md.). These sequencing reactions were initiated with the 5' end-labeled primers #53 and #517 (see above). Sequencing ladders were resolved in a 5% polyacrylamide-9 M urea gel alongside the C. gigantea DNA substrate cleaved with 1-Cpal.

Results
The Chlamydomonas Chloroplast CpLSU·2 Intron Is Structurally Very Similar to the Acanthamoeba Mitochondrial AcLSU·m1 Intron

The chloroplast LSU rRNA gene of C. pallidostigma\textit{tica} contains five group I introns (CpLSU·1–CpLSU·5) (Turmel et al. 1993a), whose insertion positions relative to the Escherichia coli 23S rRNA sequence are indicated in table 1. In the present study, we determined the sequence of the CpLSU·2 intron inserted at a position corresponding to residues 1931–1932 in domain IV of the E. coli 23S rRNA. We found that its primary sequence and potential secondary structure are remarkably similar to those of the AcLSU·m1 intron that has recently been discovered at the same insertion site in the mitochondrial LSU rRNA gene of A. castellanii (Lonergan and Gray 1994) (fig. 1A). These chloroplast and mitochondrial introns are more similar to one another than either of these introns is to any other known IB4 intron (data not shown). Aside from the size of some peripheral loops and base-paired regions, the secondary structure of CpLSU·2 differs from that of AcLSU·m1 only in lacking the helical regions P2 and P9.1 and also the interactions P9.0 and P10, which are widespread but not universal among group I introns. These two introns (873 and 755 bp, respectively) share a strikingly similar ORF (50% identity at the nucleotide level) in the loop extending P8 (figs. 1A and 1B). The C. pallidostigma\textit{tica} chloroplast CpLSU·2 ORF encodes a potential protein of 152 amino acids with a single copy of the motif LAGLI-DADG (the presence of this motif defines one of the four major structural families of group I-encoded proteins [see Shub et al. 1994]). This protein sequence can be aligned over its entire length, except for the last 12 amino acid residues, with the corresponding A. castellanii mitochondrial AcLSU·m1 intron-encoded protein comprising 142 amino acids (fig. 1C). The two intron-encoded proteins display 42% identity and 65% similarity at the amino acid level; significantly, each of them is related to the site-specific endonuclease I-Ceu1 encoded by the IB4 CeLSU·5 intron in the C. eugametos chloroplast LSU rRNA gene (Gauthier et al. 1991; Turmel et al. 1991). This intron is inserted at the same site as CpLSU·1, only eight nucleotides upstream of the CpLSU·2 and AcLSU·m1 intron position (see table 1). Clearly, the AcLSU·m1- and CpLSU·2-encoded proteins are more similar to each other than either of these proteins is to the I-Ceu1 endonuclease, as evidenced by pairwise comparisons indicating 27% and 23% identity and 50% similarity between the derived amino acid sequences of I-Ceu1 and those of the AcLSU·m1- and CpLSU·2-encoded proteins, respectively. Moreover, the CpLSU·2 and AcLSU·m1 intron-encoded proteins display more sequence similarity to one another than either of these proteins does to any of its homologues containing the LAGLI-DADG motif (data not shown).

The CpLSU·2 Intron ORF Encodes a Site-Specific Endonuclease

To determine whether the CpLSU·2 intron-encoded protein is endowed with a double-strand DNA endonuclease activity that is specific for the exon 2-exon 3 junction in the C. pallidostigma\textit{tica} LSU rRNA gene,

### Table 1

<table>
<thead>
<tr>
<th>Designation</th>
<th>Subgroup</th>
<th>Source</th>
<th>Location</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>CpLSU·1</td>
<td>IB4</td>
<td>C. pallidostigma\textit{tica}</td>
<td>Chloroplast</td>
<td>1923</td>
</tr>
<tr>
<td>CeLSU·5</td>
<td>IB4</td>
<td>C. eugametos</td>
<td>Chloroplast</td>
<td>1923</td>
</tr>
<tr>
<td>CpLSU·2</td>
<td>IB4</td>
<td>C. pallidostigma\textit{tica}</td>
<td>Chloroplast</td>
<td>1931</td>
</tr>
<tr>
<td>AcLSU·m1</td>
<td>IB4</td>
<td>A. castellanii</td>
<td>Mitochondrion</td>
<td>1931</td>
</tr>
<tr>
<td>CpLSU·3</td>
<td>IB4</td>
<td>C. pallidostigma\textit{tica}</td>
<td>Chloroplast</td>
<td>2262</td>
</tr>
<tr>
<td>CpLSU·4</td>
<td>IA1</td>
<td>C. pallidostigma\textit{tica}</td>
<td>Chloroplast</td>
<td>2449</td>
</tr>
<tr>
<td>CrLSU</td>
<td>IA3</td>
<td>C. reinhardii</td>
<td>Chloroplast</td>
<td>2593</td>
</tr>
<tr>
<td>AcLSU·m3</td>
<td>IA3</td>
<td>A. castellanii</td>
<td>Mitochondrion</td>
<td>2593</td>
</tr>
<tr>
<td>CpLSU·5</td>
<td>IA3</td>
<td>C. pallidostigma\textit{tica}</td>
<td>Chloroplast</td>
<td>2596</td>
</tr>
<tr>
<td>CeLSU·6</td>
<td>IA3</td>
<td>C. eugametos</td>
<td>Chloroplast</td>
<td>2596</td>
</tr>
</tbody>
</table>

*Position of the corresponding residue in the E. coli 23S rRNA sequence that immediately precedes the intron.
we undertook the synthesis of this protein \textit{in vitro} and tested its endonuclease activity in cleavage assays. A protein of the appropriate size (18 kDa) was produced \textit{in vitro} by T7 RNA polymerase-mediated transcription of a PCR-amplified fragment carrying the CpLSU · 2 intron ORF and subsequent translation of the resulting RNA in a rabbit reticulocyte lysate. The \textit{in vitro} translation product was tested in cleavage assays with a substrate for the putative \textit{C. pallidostigma}tica intron-encoded endonuclease. Because the chloroplast LSU rRNA gene of \textit{C. gigantea} has no intron and shows perfect sequence identity with its \textit{C. pallidostigma}tica homologue in the region of the CpLSU · 2 insertion site (Turmel et al. 1993a; M.T., C.O., and C.L., unpublished data), we used as the cleavage substrate a 260-bp end-labeled PCR product derived from this region of the \textit{C. gigantea} cpDNA. Upon incubation at 37°C in 20 mM Tris-HCl (pH 8.0), 10 mM MgCl$_2$, and 1 mM dithiothreitol, an aliquot of the translation product proved very efficient in cleaving specifically this intronless substrate (fig. 2). Thus, in contrast to the I-Chil endonuclease but like I-CeI, both of which have been tested in similar assays (Côté et al. 1993), the CpLSU · 2-encoded endonuclease is stable under the conditions of the translation reaction. As expected, it failed to cleave a 255-bp chloroplast LSU rRNA gene fragment from \textit{C. callosa}, which differs from the \textit{C. gigantea} DNA substrate by the presence of an intron sequence at the same position as CpLSU · 2 (fig. 2). Following the nomenclature proposed by Dujon et al. (1989), the \textit{C. pallidostigma}tica endonuclease is named I-Cpal.

I-Cpal and I-CeI Recognize Distinct DNA Sequences and Cleave at Distinct Sites

The cleavage site generated by I-Cpal was mapped by coelectrophoresing sequencing ladders of each strand of the DNA substrate alongside the cleavage products (fig. 3). These sequencing reactions were initiated with the same $^{32}$P-end-labeled oligonucleotides that were used to produce the substrate, thus allowing accurate sizing of the cleavage products. I-Cpal cleaves its DNA substrate only 9 bp downstream of the cleavage site reported for I-CeI (Marshall and Lemieux 1991). Each of these endonucleases generates 4-nucleotide 3' extensions (C(T/ G)AA) that are 5–6 bp downstream of the site at which the intron is inserted in the intron-plus allele (fig. 3).

The finding that a very short distance separates the double-strand breaks produced by I-CeI and I-Cpal raises the possibility that these related endonucleases recognize identical sequences. As the recognition sequence of I-CeI has been characterized (Marshall and Lemieux 1992), this hypothesis can be readily tested by assaying I-Cpal on a DNA substrate in which the I-CeI recognition sequence is interrupted by an intron at the same position as CeLSU · 5, that is, a substrate resistant to cleavage by I-CeI. We performed such an assay with a DNA substrate derived from the chloroplast LSU rRNA gene of \textit{C. mutabilis} and found that it is cleaved by I-Cpal, but not by I-CeI (fig. 4). As expected, each of the endonucleases cleaved the intronless \textit{C. moewusii} substrate, which was used as the positive control (fig. 4).
FIG. 3.—Cleavage pattern of I-C4pal. The 260-bp *C. gigantea* cpDNA region containing the cleavage site was PCR-amplified in the presence of the unlabeled primers #53 and #517, and the resulting product was used as a template in sequencing reactions initiated with these primers labeled with ³²P at their 5' termini. Sequencing products (A, C, G, and T lanes) were resolved in a 5% polyacrylamide/urea gel alongside the 260-bp *C. gigantea* DNA fragment that had been cleaved by I-C4pal. Arrows indicate the positions of the 3' termini of the cleavage products. On the sequence below the autoradiograms, the insertion sites of the introns CeLSU·5, CpLSU·1, CpLSU·2, and AcLSU·m1 are denoted by arrowheads, whereas the cleavage sites of I-CelI and I-C4pal are denoted by staggered lines.

Taken together, these results indicate that the recognition sequences of I-CelI and I-C4pal are distinct, although we cannot eliminate the possibility that they overlap.

Restricted Distribution of LSU rDNA Introns at Site 1931 among the Chloroplasts of Green Algae

The striking similarity between the chloroplast CpLSU·2 and mitochondrial AcLSU·m1 introns provides strong evidence for the transfer of a group I intron between chloroplasts and mitochondria during evolution. Similar evidence for the movement of an LSU rDNA intron inserted at another site (corresponding to residues 2593–2594 of the *E. coli* 23S rRNA) has been obtained by Lonergan and Gray (1994), who found that the *A. castellanii* mitochondrial IA3 AcLSU·m3 intron resembles the *C. reinhardtii* chloroplast IA3 intron CrLSU encoding the I-CreI endonuclease. Such an interorganellar transfer event might have occurred either intracellularly within a common, chloroplast-containing ancestor of *Acanthamoeba* and *Chlamydomonas* or through horizontal transfer between separate cells (Lonergan and Gray 1994). To distinguish between these possibilities with regard to the transfer involving introns inserted at the same site as the *C. pallidostigmatica*
chloroplast CpLSU \( \cdot \) 2 intron (this site is hereafter designated site 1931), we have examined the distribution of group I introns at this site among green algae encompassing a broad phylogenetic range. An intracellular event would predict the presence of introns similar to CpLSU \( \cdot \) 2 and AcLSU \( \cdot \) m1 in all major lineages of green algae, as any common photosynthetic ancestor of *Acanthamoeba* and *Chlamydomonas* would probably have appeared before the divergence of land plants and green algae (Lonergan and Gray 1994). In contrast, a more restricted distribution would be consistent with a horizontal transfer event between the organelles of cells of different genera.

Our intron survey was carried out in the course of a more global study of the chloroplast LSU rRNA gene, which aims at elucidating the phylogenetic relationships among green algae by comparative analysis of the rRNA-encoding sequences, and also at probing the evolution of group I introns in this gene. The complete sequences of the chloroplast LSU rRNA genes of 48 green algae are available at present. All these sequences except one (that of *Chlorella ellipsoida*) were determined in our laboratories (M.T. and C.L.). Sequences of the rRNA-encoding regions and of some introns from 17 *Chlamydomonas* species have been published (Turmel et al. 1991, 1993a; Côté et al. 1993); the rRNA-encoding sequences of the 31 remaining green algae as well as the other intron sequences will be published in due course together with detailed phylogenetic analyses.

Of the 48 green algae examined, six *Chlamydomonas* species (including *C. pallidostigmatica*) and three green algae from other genera were found to harbor a group I intron at site 1931 of the chloroplast LSU rRNA gene (fig. 5). All of the nine chloroplast introns belong to the IB4 subgroup (M.T., C.O., J.-P.M., and C.L., unpublished data), as does *A. castellanii* mitochondrial AcLSU \( \cdot \) m1. Except for that from *Haematococcus lacustris*, the chloroplast introns inserted at site 1931 all show more sequence similarity to each other (64\%–94\% identity) than to the mitochondrial AcLSU \( \cdot \) m1 intron (57\%–63\% identity) and form a clade that is supported

![Image of a diagram](image)

**Fig. 4.**—Cleavage specificity of I-CeuI and I-CpalI. An unprogrammed rabbit reticulocyte translation mix (- Enzyme, + Lysate) and a translation mix containing in vitro-synthesized RNAs from the CeLSU \( \cdot \) 5 and CpLSU \( \cdot \) 2 intron ORFs (+ Enzyme, + Lysate) were each incubated with an intronless DNA substrate (260 bp) derived from the *C. moewusii* (moe) chloroplast LSU rRNA gene and with an intron-containing substrate (282 bp) derived from the *C. mutabilis* (mut) gene. The group I intron in the latter substrate is inserted at the same position as CpLSU \( \cdot \) 1 and CeLSU \( \cdot \) 5. DNA samples were electrophoresed in a 5% polyacrylamide/urea gel alongside the corresponding DNA substrates (- Enzyme, - Lysate). The *C. mutabilis* and *C. moewusii* chloroplast LSU rRNA gene sequences containing the cleavage sites of I-CeuI and I-CpalI are aligned below the autoradiograms. Identical residues in these sequences are boxed. Residues that are not conserved between the *A. castellanii* mitochondrial and *C. moewusii* chloroplast LSU rRNA gene sequences are denoted by asterisks. The intron (lower case) in the *C. mutabilis* gene disrupts the minimal recognition sequence of I-CeuI (Marshall and Lemieux 1992), thereby preventing cleavage by this endonuclease.
Fig. 5.—Distribution of chloroplast LSU rDNA introns at sites 1931 and 2593 in a variety of photosynthetic organisms. Taxa featuring introns at sites 1931 and 2593 are denoted by circles and squares, respectively. Filled symbols denote introns with an internal ORF, whereas open symbols denote introns lacking an ORF. The phylogeny was inferred from chloroplast LSU rRNA gene sequences. It was computed with fastDNAml (version 1.0.6) (Olsen et al. 1993) from a data set of 2334 semiconserved positions, using the Anacystis nidulans LSU rRNA gene sequence as an outgroup and 100 bootstrap replicates, with 10 randomizations of order and termination of rounds finding the same tree twice. A consensus of the 100 best trees was generated using the CONSENSE program of the PHYLIP package (version 3.5c) (Felsenstein 1993) and branch lengths were estimated using fastDNAml with the user tree option. The percentage of bootstrap resamplings that support the topological elements is shown to the left of each grouping. The scale bar indicates 10 changes per 100 nucleotides. Origins of the green algal strains and of the chlorarachniophyte-like species Pedinomonas minitissima (the taxa spanning the thick vertical line) are indicated in parentheses: CCMP, Center for the Culture of Marine Phytoplankton at the Bigelow Laboratory for Ocean Sciences; NIES, National Institute for Environmental Studies at Tsukuba; SAG, Sammlung von Algenkulturen at Göttingen; UTEX, Culture Collection of Algae at the University of Texas. The LSU rRNA gene sequences of 17 Chlamydomonas species (see accession numbers in Turmel et al. 1993a) and of Chlorella ellipsoidea (M36158) have been published. Accession numbers for published nongreen algal sequences are as follows: O. sativa, X15901; N. tabacum, J01446 and Z00044; M. polymorpha, X01647 and X04465; A. longa, X14386; E. gracilis, X13310; P. littoralis, X61179 and S36159; and P. palmata, Z18289.

by a bootstrap value of 97% in a neighbor-joining analysis (fig. 6). In contrast, the H. lacustris intron is more similar to the A. castellanii mitochondrial AcLSU·m1 intron (63% identity) than to most of its chloroplast homologues (53%–63% identity) and clusters with the chloroplast IB4 CeLSU·5 intron (site 1923) used as an outgroup in the aforementioned neighbor-joining analysis of site-1931 introns (fig. 6). Like AcLSU·m1, six of the chloroplast introns (including CpLSU·2) feature, in the loop extending P8, an ORF that encodes a poten-
I introns were identified at this site among the 15 green algae encompassing lineages basal to those featuring *Chlamydomonas* species. This restricted intron distribution suggests that a horizontal transfer event accounts for the interorganellar movement of LSU rDNA intron sequences at site 1931.

Broad Distribution of LSU rDNA Introns at Site 2593 among the Chloroplasts of Green Algae

At the time Lonergan and Gray (1994) published their study, the presence of chloroplast introns at site 2593 had been described for only four *Chlamydomonas* species (Turmel et al. 1993a). Among the 48 green algae examined here, we found chloroplast LSU rDNA introns at this site to be more prevalent than those at site 1931, occurring in four green algae occupying branches outside the *Chlamydomonas* lineages in addition to nine taxa clustering within these lineages (fig. 5). Like their counterparts at site 1931, all chloroplast introns at site 2593 appear to be homologues of each other, as revealed by comparative analysis of primary sequences and secondary structures (M.T., C.O., J.-P.M., and C.L., unpublished data). Several of them show a higher sequence identity with the mitochondrial AcLSU·m3 than with some of their chloroplast homologues (fig. 6); however, none has a sequence that is especially similar to that of this mitochondrial intron. The mitochondrial and chloroplast LSU rDNA introns at site 2593 are more similar to one another than either of these introns is to any other IA3 intron (data not shown). Our neighbor-joining analysis indicates that the relationships among most site-2593 introns are uncertain (fig. 6). Indeed, only three pairs of introns form clades that are supported by bootstrap values greater than 85%. Like AcLSU·m3 (Lonergan and Gray 1994), nine of the chloroplast introns (including CrLSU; Rochaix et al. 1985) have, in the loop extending P6, an ORF encoding a potential protein with one copy of the LAGLI-DAGD motif; the remaining four chloroplast introns lack any ORF (see fig. 5). The ORFs of AcLSU·m3 and of the chloroplast introns are all homologous in sequence (M.T., C.O., J.-P.M., and C.L., unpublished data).

Discussion

The data presented here strongly suggest that all known chloroplast and mitochondrial group I introns inserted at sites 1931 and 2593 of the LSU rRNA gene are homologues of one another. To date, no nuclear introns have been found at these sites. As proposed by Lonergan and Gray (1994), three alternative scenarios could account for the inferred homology between the chloroplast and mitochondrial group I introns: (1) vertical transmission of an ancient intron that was present...
in a eubacterium ancestral to Cyanobacteria and Proteobacteria (the progenitors of chloroplasts and mitochondria, respectively) and that was also present in both proto-mitochondrial and proto-chloroplast genomes; (2) lateral transfer of an intron between chloroplast and mitochondrial genomes in the same cell of a photosynthetic common ancestor of *Chlamydomonas* and *Acanthamoeba* (or simultaneous acquisition of a foreign intron by the two organelle genomes in such a common ancestor); and (3) intron transfer via intercellular exchange of genetic information between Acanthamoeboid and Chlamydomonad organisms. On the basis of a number of studies providing evidence for DNA movement between organelles in the same cell and considering the available phylogenetic information on group I intron distribution, Lonergan and Gray (1994) favored (but could not distinguish between) scenarios 2 and 3.

We have shown here that among green algae examined to date, the occurrence of homologous chloroplast LSU rDNA introns at site 1931 is limited to those belonging to *Chlamydomonas* lineages (see fig. 5). This observation argues against vertical inheritance of an AcLSU·m1-type intron from a common eubacterial ancestor of chloroplasts and mitochondria (scenario 1) or from a common photosynthetic, eukaryotic ancestor of *Acanthamoeba*, chlorophytes and metaphytes (scenario 2). Such schemes would require multiple independent losses to account for the observed distribution. Consequently, the view that the homologous chloroplast and mitochondrial LSU rDNA introns at site 1931 originated via movement of genetic information between different organelles of different organisms (scenario 3) is the one best supported by our results.

In contrast, at site 2593, we observed a phylogenetically broader intron distribution, with introns being found at this site in four green algae occupying earlier-diverging branches relative to *Chlamydomonas* lineages (see fig. 5). This difference in intron distribution at sites 1931 and 2593 may not be significant because our sampling of green algae was heavily biased in favor of members of *Chlamydomonas* lineages and also because all classes of green algae were not represented. In this context, it is noteworthy that, during recent sequencing of the chloroplast LSU rRNA gene from the prasinophyte *Pterosperma cristanum*, a green alga probably representing one of the most early-diverging green algal branches (O'Kelly, 1992), we found at site 1931 a group I intron structurally homologous to other chloroplast LSU rDNA introns residing at this site (M.T., C.O., and C.L., unpublished data). Note that a relatively broad intron distribution at both sites 1931 and 2593 would not necessarily support scenario 2 because this scheme also requires a broad representation of homologous LSU rDNA introns at each site in mitochondrial DNAs (mtDNAs) spanning the phylogenetic range between *Acanthamoeba* and *Chlamydomonas*. The data currently available for the mitochondrial LSU rRNA genes of a few land plant species as well as for those of the green algae *C. reinhardtii*, *C. eugametos*, and *Prototheca wickamii* indicate that there are no introns at these sites in these mtDNAs (see Lonergan and Gray 1994), thus arguing against scenario 2. To unequivocally resolve the issue of intracellular versus intercellular transfer of LSU rDNA introns between organelles at sites 1931 and 2593, additional phylogenetic information on both chloroplast and mitochondrial LSU rDNA intron distributions will be needed.

Prior to our study, there were no reports documenting the presence of homologous group I introns sharing common insertion sites in the chloroplast genome. At first glance, our finding of such a homology for the introns at sites 1931 and 2593 of the chloroplast LSU rRNA gene from several green algae might suggest that all introns at each site originated by vertical inheritance of a mobile, endonuclease-encoding intron from a common green algal ancestor, with frequent losses accounting for the absence of introns from members of clades featuring these genetic elements (see fig. 5). However, given that introns at both sites can integrate into intronless alleles by homing or by reverse self-splicing followed by reverse transcription and homologous recombination, some of the observed introns may well have originated from rare events of horizontal transfer. Our phylogenetic analysis of chloroplast intron sequences at sites 1931 and 2593 of LSU rRNA gene coding regions is consistent with the vertical transmission of introns in some lineages, although we observed some noncongruence between the intron trees and LSU rRNA gene trees (figs. 5 and 6). For example, the green algae *H. lacustris*, *Stephanosphaera pluvialis*, and *Dunaliella parva* cluster in a clade supported by a bootstrap value of 100% in the LSU rRNA gene tree, while their introns at site 1931 are found in three distinct, well-supported clades in the intron tree. Such inconsistencies may reflect the horizontal transfer of intron sequences from one lineage to another or from another source and/or radical divergence of vertically inherited intron sequences within some lineages. Recently, Bhattacharya et al. (1994) used phylogenetic analysis of intron and small subunit (SSU) rRNA gene sequences to demonstrate strict vertical inheritance of group I introns sharing a common insertion site (site 1506) and lacking an endonuclease-encoding gene in the SSU rRNA gene from two lineages of green algae belonging to the order Zygoughmatales (Charophyceae). On the other hand, the results of these authors did not provide any insights on the evolutionary origin.
of the nuclear SSU rDNA group I introns found at the same site in two other green algae, two red algae, and two fungi. Thus, both horizontal transfer and common ancestry may have led to the complex distribution of nuclear SSU rDNA group I introns at site 1506. For nuclear rDNA group I introns occupying different insertion sites, evidence has been reported that they have arisen independently (Van Oppen et al. 1993; Bhattacharya et al. 1994).

A common ancestry for most of the chloroplast LSU rDNA introns inserted at sites 1931 and 2593 would imply the occurrence of numerous events of intron loss during the evolution of green algae. Intron losses have been predicted to be much less frequent for endonuclease-encoding introns than for introns not encoding such proteins because, in the former case, all newly arising intronless alleles should, in principle, be replaced with intron-containing alleles by the homing process if the integrity of the intron ORF has been preserved. For this reason, group I intron-encoded endonucleases have been thought to contribute to the maintenance of corresponding mobile introns at their insertion sites. The sporadic intron distributions we observed at both sites 1931 and 2593 does not support this notion. At both sites, introns lacking an endonuclease gene have been identified in green algae from different lineages, suggesting that independent losses of endonuclease genes (by an unknown mechanism) can take place relatively frequently, leading to the rapid disappearance of mobile introns among descendants. Obviously, the ORFs of introns at both sites 1931 and 2593 are not essential for splicing of these introns or for other functions; otherwise they would have been retained in all cases.

The moderate sequence identities (57%-73%) between the green algal chloroplast intron sequences and their *A. castellanii* mitochondrial counterparts at sites 1931 and 2593 suggest that the postulated, interorganellar DNA movements responsible for the homology between these chloroplast and mitochondrial introns must have occurred a long time ago. On the other hand, the observed level of similarity would be compatible with very recent transfer events if the newly introduced introns rapidly diverged in sequence following their insertion. Intron content and distribution in *A. castellanii* mtDNA could be rationalized in terms of a relatively recent origin of AcLSU·m1 and AcLSU·m3. In this highly compact mitochondrial genome, the LSU rRNA gene is the only one interrupted with group I introns (Burger et al. 1995). If AcLSU·m1 and AcLSU·m3 were of relatively ancient origin, group I introns might be expected to be present in additional genes as a result of intragenomic intron proliferation. For example, in the cpDNAs of *Chlamydomonas* species, five genes (*psaB, psbA, psbC, SSU, and LSU rRNA genes*) have been found to contain group I introns, some of which are closely related (Turmel et al. 1993b; see also Côté and Turmel 1994), thus suggesting that some group I introns were present in the common ancestor of all *Chlamydomonas* species but have since migrated to other locations. Because members of the genus *Acanthamoeba* are phagocytic organisms that live in soil and water, also the habitat of numerous kinds of green algae, group I introns appear more likely to have been transferred from the chloroplast of a green alga to the mitochondrion of an Acanthamoeboid protozoan than in the reverse direction. To account for a high level of sequence identity between the single group I intron in the nuclear SSU rRNA gene of *Acanthamoeba griffini* and one of the two nuclear SSU rDNA introns of *D. parva*, Gast et al. (1994) have recently proposed that an Acanthamoeboid organism acquired a group I intron by horizontal transfer from phagocytized green algae. Of the 29 other isolates of *Acanthamoeba* examined by these authors, none was found to harbor a nuclear SSU rDNA intron at the same site as that of *A. griffini*, results indicative of a recent transfer event. Our finding of endonuclease-encoding LSU rDNA introns at both sites 1931 and 2593 in *H. lacustris* and *S. pluvialis* (see fig. 5) is consistent with the possibility that green algal chloroplast introns occupying these sites were transferred simultaneously to the mtDNA of an Acanthamoeboid organism.

Although the *Chlamydomonas* chloroplast CpLSU·2 and CrLSU introns each encode a double-strand DNA endonuclease specific for the intronless alleles at the cognate site (Dürrenberger and Rochaix 1991; present study), there is no evidence of coconversion of intron-flanking sequences, a process that has been shown to accompany intron homing (see Dujon 1989; Lambowitz and Belfort 1993). Indeed, there is no indication that AcLSU·m1 and AcLSU·m3 are flanked by chloroplast-specific rRNA sequences or that CpLSU·2 and CrLSU are flanked by mitochondrion-specific rRNA sequence. We thus favor the view that the interorganellar exchange of rDNA introns occurred by the reverse-splicing mechanism. Recent reports suggest that movements of group I or group II introns by this mechanism may be relatively frequent *in vivo* (Mueller et al. 1993; Sellem et al. 1993; Thompson and Herrin 1994). On the other hand, homing could have promoted the rapid spread of newly integrated introns in one or a few organelle DNA molecules to all other identical DNA molecules in the organelles. Given that the genetic code of *Acanthamoeba* mtDNA differs from the universal code (TGA specifies Trp) (Burger et al. 1995) and the fact that the AcLSU·m1 and AcLSU·m3 ORFs contain in-
ternal TGA codons (Loneragan and Gray 1994), only the homologous chloroplast ORFs are expected to be expressed into active homing endonucleases in both chloroplasts and mitochondria.

Acknowledgments

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada (to M.T. and C.L.), Le Fonds pour la Formation de Chercheurs et l’Aide à la Recherche (to M.T. and C.L.), and the Medical Research Council of Canada (to M.W.G.). M.T. and C.L. are Scholars, and M.W.G. a Fellow, in the Evolutionary Biology Program of the Canadian Institute for Advanced Research, which we thank for salary support.

LITERATURE CITED


Interorganellar Transfer of Group I Introns

545


Mitchell L. Sogin, reviewing editor

Received October 26, 1994

Accepted February 13, 1995