J Mol Evol (1995) 41:563–572

JOURNAL OF MOLECULAR EVOLUTION

Fungal Origin by Horizontal Transfer of a Plant Mitochondrial Group I Intron in the Chimeric *CoxI* Gene of *Peperomia*

Jack C. Vaughn,[1] Matthew T. Mason,[1] Ginger L. Sper-Whitis,[1] Peter Kuhlman,[2] Jeffrey D. Palmer[2]

[1] Department of Zoology, Miami University, Oxford, OH 45056, USA
[2] Department of Biology, Indiana University, Bloomington, IN 47405, USA

Received: 2 September 1994 / Accepted: 28 April 1995

Abstract. We present phylogenetic evidence that a group I intron in an angiosperm mitochondrial gene arose recently by horizontal transfer from a fungal donor species. A 1,716-bp fragment of the mitochondrial *coxI* gene from the angiosperm *Peperomia polybotrya* was amplified via the polymerase chain reaction and sequenced. Comparison to other *coxI* genes revealed a 966-bp group I intron, which, based on homology with the related yeast *coxI* intron aI4, potentially encodes a 279-amino-acid site-specific DNA endonuclease. This intron, which is believed to function as a ribozyme during its own splicing, is not present in any of 19 *coxI* genes examined from other diverse vascular plant species. Phylogenetic analysis of intron origin was carried out using three different tree-generating algorithms, and on a variety of nucleotide and amino acid data sets from the intron and its flanking exon sequences. These analyses show that the *Peperomia coxI* gene intron and exon sequences are of fundamentally different evolutionary origin. The *Peperomia* intron is more closely related to several fungal mitochondrial introns, two of which are located at identical positions in *coxI,* than to identically located *coxI* introns from the land plant *Marchantia* and the green alga *Prototheca.* Conversely, the exon sequence of this gene is, as expected, most closely related to other angiosperm *coxI* genes. These results, together with evidence suggestive of co-conversion of exonic markers immediately flanking the intron insertion site, lead us to conclude that the *Peperomia coxI* intron probably arose by horizontal transfer from a fungal donor, using the double-strand-break repair pathway. The donor species may have been one of the symbiotic mycorrhizal fungi that live in close obligate association with most plants.

Key words: Plant mitochondria — *CoxI* gene — Group I intron — Intronic ORF — Horizontal gene transfer — Evolutionary origin — VA mycorrhiza

Correspondence to: J.C. Vaughn

Introduction

Group I and group II introns (reviewed by Lambowitz and Belfort 1993) are ribozymes which function in the catalysis of their own excision/splicing. These two intron classes differ in secondary structure, in mechanism of splicing, and in the proteins they encode. Group I introns are the most phylogenetically widespread class of introns; they are found in nuclear rRNA genes of many diverse eukaryotes, in the genomes of a few eubacteria and their phage, and, most abundantly, in certain organellar genomes. Most chloroplast group I introns have been found in the green algal genus *Chlamydomonas* (e.g., Turmel et al. 1993a,b), although one notably ancient group I intron is broadly present in tRNA-Leu genes among plant and algal chloroplasts and their cyanobacterial ancestors (Kuhsel et al. 1990; Xu et al. 1990). Among mitochondria, group I introns are most numerous in fungi (Clark-Walker 1992) and in the non-vascular plant *Marchantia* (Oda et al. 1992); a few are reported in certain green algae (Wolff et al. 1993; Turmel et al. 1993a).

Cytochrome c oxidase catalyzes the transfer of electrons from cytochrome c to oxygen during the final step of the respiratory chain (reviewed by Capaldi et al. 1983). In most eukaryotes, subunits I–III of this complex are encoded by the corresponding mitochondrial genes *coxI–coxIII,* while the remaining subunits are nuclear-encoded (reviewed by Gray 1992). The *coxI* gene is by far the most intron-rich organellar gene known, containing a grand total of 30 different introns (mostly group I) among all organisms examined thus far. As a consequence, a highly disproportionate share of the total set of introns present in many mitochondrial genomes resides in *coxI,* including nearly half in the fungi *Podospora anserina* (16 of 33 introns; Cummings et al. 1989, 1990) and *Allomyces macrogynus* (12 of 26 introns; Paquin and Lang 1993), over a quarter in *Marchantia polymorpha* (9 of 32 introns; Oda et al. 1992; Ohta et al. 1993; Ohyama et al. 1993), and two-thirds in *Schizosaccharomyces pombe* (4 of 6 introns; Lang 1993). It has been argued that at least five (four group I and one group II) of the nine introns in *Marchantia coxI* were inherited by vertical transmission from a common ancestor shared by "plants" (i.e., land plants and green algae) and fungi, based in part on the observation that these introns are located at identical positions in certain plant and fungal *coxI* genes, and in part on analysis of their sequence relationships (Ohta et al. 1993; Ohyama et al. 1993; Wolff et al. 1993).

All of the 23 introns reported thus far in mitochondria of vascular plants belong to group II (reviewed by Schuster and Brennicke 1994), and no introns of any type have been reported in *coxI* genes of any vascular plant. During a survey of *coxI* genes in vascular plants (Sper-Whitis and Vaughn, unpublished), we were therefore surprised to find an intron within *coxI* of the angiosperm *Peperomia polybotrya.* Here we show that this is the first example of a group I intron found in any vascular plant mitochondrial genome. Two independent lines of evidence are presented, based on intron distributional and also phylogenetic analysis, to support the conclusion that this intron arose via horizontal transfer, most likely from a fungal source.

Materials and Methods

Preparation of Total Genomic DNA. Total cell DNA was isolated from fresh leaves of *Peperomia polybotrya* using a cetyltrimethylammonium bromide procedure (Doyle and Doyle 1987), which included 30-min digestion with 50 μg/ml pancreatic RNase A. The DNA was further purified on CsCl step-gradients made in 10 mM Tris-HCl, pH 8.0 (1.72 g CsCl/ml—bottom, 1.50 g/ml—middle), onto which DNA in 1.0 M CsCl was layered, followed by ultracentrifugation in a Beckman 75-Ti rotor at 36,000 rpm for 12 h at 25°C. DNA, which banded at or slightly below the 1.72 g/ml interface, was both localized and quantitated following tube fractionation by Hoechst 33258 fluorescence in a microspectrofluorometer (Hoefer Scientific).

CoxI Gene Isolation. An alignment was prepared (not shown) of all plant mitochondrial *coxI* gene sequences available from the GenBank/EMBL data base and utilized for design of oligonucleotide primers. Both amino acid and nucleotide sequence conservation were considered, and primers located as close to the ends of the *coxI* gene as possible were sought. The upstream primer had the (noncoding) sequence 5′-(TTATTATCACTTCCGGTACT)-3′, while the downstream primer had the (coding) sequence 5′-(AGCATCTGGATAATCTGG)-3′. This primer combination would be expected to result in amplification of about 45% of the *coxI* gene. DNA fragment amplification was done via the polymerase chain reaction (PCR) with Taq DNA polymerase using the following parameters: 30 cycles of 1-min denaturation at 94°C, 1-min annealing at 47°C, and 2-min extension at 72°C. The *Peperomia* PCR fragment was cloned into the pCRII plasmid vector (Invitrogen) and transformed into competent INVαF′ *E. coli* cells. Recombinant plasmids were isolated from single colonies using standard techniques (Sambrook et al. 1989).

DNA Sequencing and Sequence Analysis. The cloned 1,716-bp *coxI* gene fragment was largely sequenced on both strands via the dideoxy chain termination method (Sanger et al. 1977), using a number of different primers. Sequence manipulations and searches of the GenBank/EMBL data base were carried out using the GCG software package from Genetics Computer Group, University of Wisconsin (1991), version 7.

Sources of Gene Sequences. GenBank accession numbers for the following plant *coxI* gene sequences included in our analyses are: *Beta vulgaris,* X57693; *Marchantia polymorpha,* M68929; *Glycine max,* M16884; *Oenothera berteriana,* X05465; *Oryza sativa,* X15990; *Peperomia polybotrya,* X87336; *Pisum sativum,* X14409; *Prototheca wickerhamii,* X68721; *Raphanus sativus,* X57692; *Sorghum bicolor,* M14453; *Thuja plicata,* X64833; *Triticum aestivum,* Y00417; *Zea mays,* X02660. Fungal *coxI* accession numbers are: *Aspergillus nidulans,* X00790; *Kluyveromyces lactis,* X57546; *Neurospora crassa,* X14669; *Podospora anserina,* X55026; *Saccharomyces cerevisiae,* J01481; *Schizosaccharomyces pombe,* X00886. Accession number for the *S. cerevisiae cob* gene is V00686. In addition, unpublished *coxI* gene sequences (Sper-Whitis and Vaughn, manuscript in preparation) from the following plants were utilized in this study: *Arabidopsis thaliana; Cabomba* sp.; *Ephedra equisetifolia; Ginkgo biloba; Isoetes flaccida; Lycopodium digitatum; Nephrolepis exaltata; Psilotum nudum; Zamia* sp. Finally, the *coxI* gene sequence from the mycorrhizal fungus *Gigaspora rosea* is from Adams and Vaughn (unpublished).

Phylogenetic Analyses of CoxI Gene Sequences. Sequences were aligned by computer and refined by eye with reference to RNA secondary structure elements, or to chemical characteristics of amino acids when relevant, and were used to construct six data sets consisting of (1) exon nucleotide sequences; (2) exon amino acid sequences; (3) entire-intron nucleotide sequences; (4) conserved group I intron core nucleotide sequences; (5) intronic open reading frame (ORF) nucleotide sequences; and (6) intronic ORF amino acid sequences. For the exon data sets, only the 702-bp region covered by the *Peperomia* PCR fragment described earlier, exclusive of terminal primers, was used for analysis, and only edited versions of the plant sequences were utilized (Hiesel et al. 1994).

Parsimony analyses were run on Macintosh Quadra 800 and 950 computers using PAUP v.3.1 (Swofford 1993). A branch-and-bound algorithm (Hendy and Penny 1982) was initially used to find the shortest tree for each data set. This was followed by 100 simple heuristic searches on randomized bootstrap data sets (Felsenstein 1985) to determine internal support for the clades found in the most parsimonious trees. Branch lengths were calculated under a model of accelerated transformation (ACCTRAN).

Maximum likelihood analyses were performed using fastDNAml v.1.0.6 (Olsen et al. 1994) running on a Sun SparcServer 1000 with six processors. This computer is approximately two orders of magnitude

faster than a Macintosh Quadra for such analyses. One hundred bootstrap data sets were generated and analyzed using the script fastDNAml_boot with a transition/transversion ratio of $T = 2.0$. In order to improve the chances of attaining the optimal tree for each replicate, sequences were added up to eight times in random order until two matching topologies were reached, and global swapping was allowed over the default number of branches. Consensus trees were generated from the resulting sets of trees with the PHYLIP tool CONSENSE.

Neighbor-Joining analyses were performed using version 3.5c of the software package PHYLIP (Felsenstein 1993), recompiled for use on PowerPC microprocessors by Don Gilbert at Indiana University, and run on a Macintosh Quadra 800 with a PowerMacintosh Upgrade Card from Apple Computer Corp. One hundred randomized bootstrap data sets were generated by invoking the PHYLIP tool SEQBOOT. Distance matrices were calculated by DNADIST (with the Kimura two-parameter model) for nucleotide sequence sets or PROTDIST (using PAM matrices) for amino acid sequence sets, and NEIGHBOR was used for tree construction. The trees output from NEIGHBOR was condensed into one consensus tree with CONSENSE.

Results

The Peperomia CoxI Gene Contains a Group I Intron and Intron-ORF

During a survey of plant mitochondrial *coxI* gene evolution and RNA editing, we (Sper-Whitis and Vaughn, unpublished) used primers near the ends of the gene to PCR amplify about half of the *coxI* gene from a variety of vascular plants. All plants yielded products of the size expected for an uninterrupted, intron-lacking gene except for the angiosperm *Peperomia polybotrya,* whose approximately 1,750-bp product was much larger than expected. Complete sequencing of this product revealed the presence of a single large intron of 966-bp located at precisely the same position as well-characterized group I introns in *coxI* genes of *Saccharomyces cerevisiae* (intron "aI4"; Bonitz et al. 1980), the fungus *Podospora anserina* (intron "aI9"; Cummings et al. 1989), the liverwort *Marchantia polymorpha* (intron "aI7"; Ohta et al. 1993), and the green alga *Prototheca wickerhamii* (intron "aI3"; Wolff et al. 1993). A search of the GenBank/EMBL data base revealed that the most closely related sequences to the *Peperomia* intron are two related group I introns from *Saccharomyces,* introns aI4 and bI4. (The latter is in the *cob* mitochondrial gene; Nobrega and Tzagoloff 1980.)

Figure 1 shows that the *Peperomia* and aI4 introns are readily alignable over essentially their entire lengths; excluding gaps they are 58% identical in sequence (553 nucleotide identities over 951 positions). Furthermore, secondary structure analysis reveals that the *Peperomia* intron can be folded into a canonical group I intron secondary structure, one which is virtually superimposable on the aI4 structure (Fig. 2). We therefore conclude that the *Peperomia coxI* interruption is indeed a group I intron and that the yeast and *Peperomia* introns belong to the same subgroup IB1 (Michel and Westhof 1990). The validity of the structure derived for the *Peperomia* intron is based on the powerful comparative approach (Fox and Woese 1975), in which the existence and extent of suspected helices are proven by the demonstration of numerous examples of compensatory base changes in the stems (Noller et al. 1981). We also conclude that the *Peperomia coxI* PCR product is derived from mitochondrial DNA, as the exon portions contain C at ten positions at which the mitochondrial *coxI* transcripts undergo C → U RNA editing (Sper-Whitis and Vaughn, unpublished), whereas nuclear genes derived from plant mitochondria have edited T's at such positions (Nugent and Palmer 1991; Covello and Gray 1992; Grohmann et al. 1992).

The *Peperomia* intron contains one long ORF of 840 nucleotides, beginning with an ATG codon located 81 nucleotides from the intron's 5′-terminus. Unlike the homologous ORF located in the yeast aI4 intron (Bonitz et al. 1980), the *Peperomia* intronic ORF is not in-frame with its 5′-flanking exon. Nevertheless, the two ORFs are 59% identical in nucleotide sequence over the 834 positions in the region of overlap beginning with codon ATG in *Peperomia* (Fig. 1). Excluding gaps, the two inferred proteins share 52% identity across the 229-residue region of overlap between *Peperomia* codons #48 and 276 (Fig. 3). The yeast aI4 intronic ORF has been shown to encode a site-specific DNA endonuclease, named I-*Sce*II (Delahodde et al. 1989). The recognition sequence for the putative *Peperomia* endonuclease is predicted to be identical to that of yeast I-*Sce*II, because the *coxI* gene nucleotide sequence surrounding the intron insertion site in the two species is virtually identical across some 20 nucleotides. We have named this potential new DNA endonuclease I-*Pepo*I to distinguish it from the enzyme I-*Ppo*I encoded by a nuclear rRNA gene group I intron in *Physarum polycephalum* (Muscarella et al. 1990).

The Peperomia Group I Intron Originated by Horizontal Transfer

Two lines of evidence support the hypothesis that the *Peperomia coxI* intron arose by horizontal transfer, rather than by vertical transmission through the plant lineage. First, parsimony considerations of intron presence/absence in vascular plants favor the notion that the intron was recently introduced into the *Peperomia* lineage. The 19 other vascular plants for which *coxI* sequences have been determined all lack an intron at this, or any other, position (Fig. 4). A hypothesis of strictly vertical evolution would require loss of the intron from all vascular plant lines except for the one leading to *Peperomia.* Based on current sampling and phylogenetic concepts (Donoghue 1994), this would require a minimum of at least six and more likely ten independent losses, one each from the ten groups of vascular plants shown in Fig. 4 other than the Piperales itself. A hypothesis of horizontal evolution, however, would postulate only a single event.

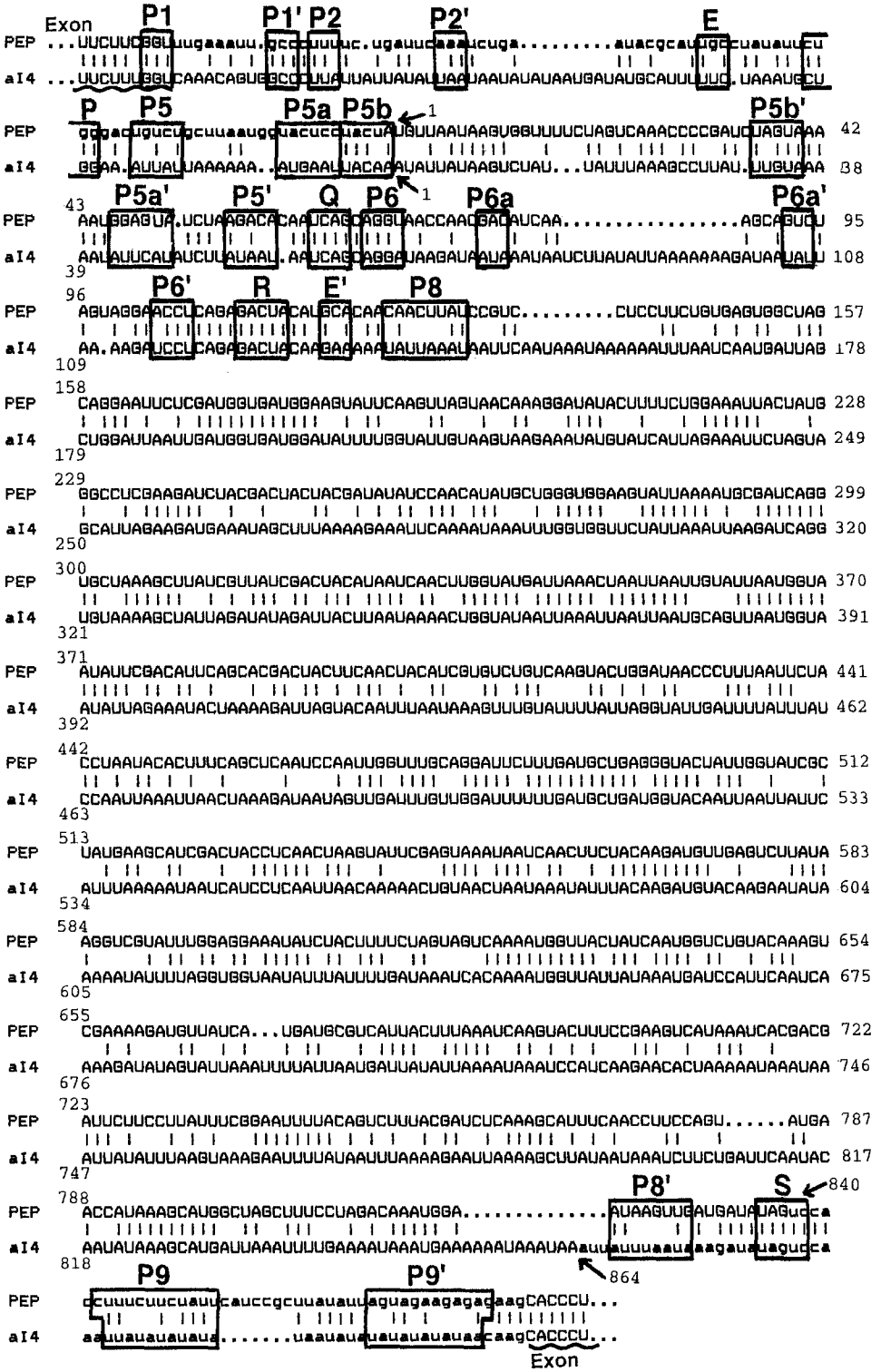

Fig. 1. Alignment of introns located at identical positions in *coxI* genes of *Peperomia* (PEP) and *S. cerevisiae* (aI4). Nucleotides within intron ORFs are in *capital letters*, while non-ORF nucleotides are in *lowercase letters*. Flanking exon sequences are indicated by *wavy lines*. Sequence elements that form intramolecularly paired stem regions (P1/P1′, P2/P2′, etc.) of the intron secondary structure (Fig. 2) are *boxed*. *Numbering* of nucleotides begins with the start codon AUG in *Peperomia* and the corresponding nucleotide in intron aI4. Gaps inserted to improve alignment are shown by *dots*. The *Peperomia* ORF contains 840 nucleotides, while the corresponding aI4 sequence contains 864 nucleotides. The *Peperomia* sequence has been assigned GenBank accession number X87336.

Second, formal phylogenetic analysis, carried out using three different tree-building methods (parsimony, maximum likelihood, and neighbor-joining), shows that the intron and exon sequences of *Peperomia coxI* have fundamentally different phylogenetic histories. Phylogenetic analyses of 702 nucleotides of *coxI* exonic sequences produced a tree (Fig. 5B) that agrees extraordinarily well (i.e., in all topological respects and with

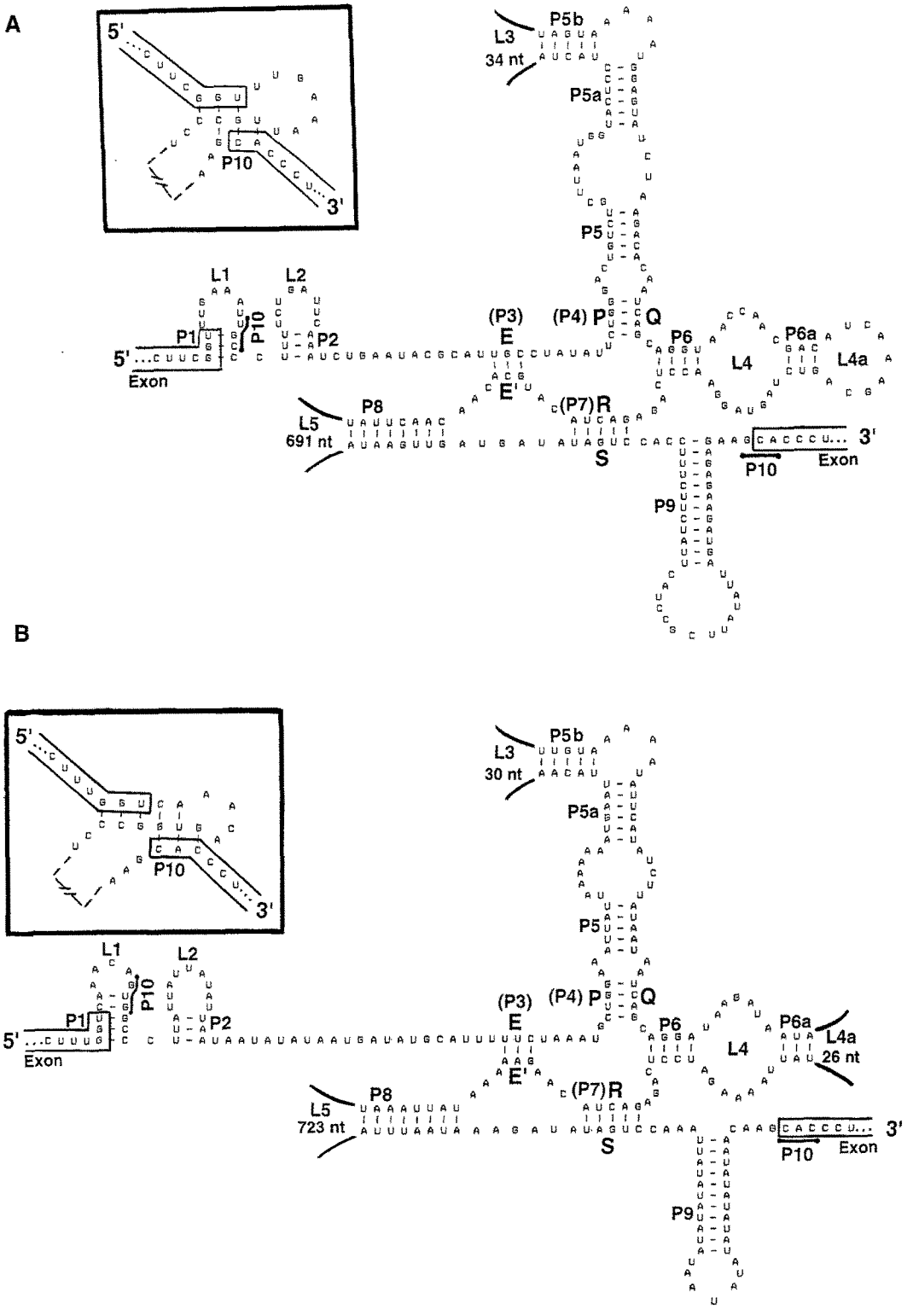

Fig. 2. Predicted secondary structures of *coxI* introns located at identical positions in *Peperomia* **(A)** and *S. cerevisiae* **(B).** Stems (P) and loops (L) are *numbered* according to the standard scheme for group I introns (Burke et al. 1987). Exonic boundary sequences are *boxed. Insets* show P1 and P10 interactions believed to help align the flanking exons for splicing.

virtually all bootstrap values quite high for all three methods) with current views of relationships among the six fungi (Bruns et al. 1992; Wilmotte et al. 1993) and nine plants (Donoghue 1994) examined. Of paramount importance, the *Peperomia coxI* coding region is indisputably of plant origin, clustering with other land plant sequences with 100% bootstrap support and specifically with other seed plant (*Raphanus* through *Thuja*) se-

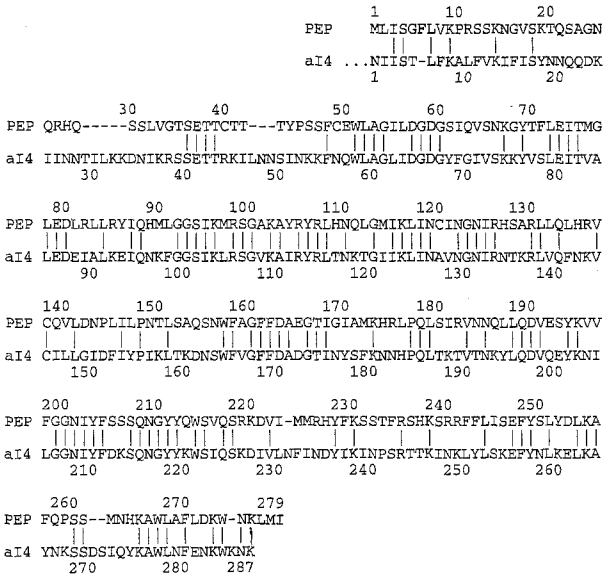

Fig. 3. Predicted amino acid sequences encoded by *coxI* gene intronic ORFs in *Peperomia* (PEP) and *S. cerevisiae* (aI4). The *numbering* of amino acid residues begins with the first methionine in the *Peperomia* sequence and the corresponding residue in the yeast sequence. The latter residue is 15 amino acids downstream of the putative initiator methionine of the yeast polypeptide. Gaps inserted to improve the alignment are shown by *dashes*. Assignment of codons is based on the universal code, except for the yeast mitochondrial CTN and TGA codons, which respectively specify threonine instead of leucine (Li and Tzagoloff 1979) and tryptophan instead of termination (reviewed by Clark-Walker 1992).

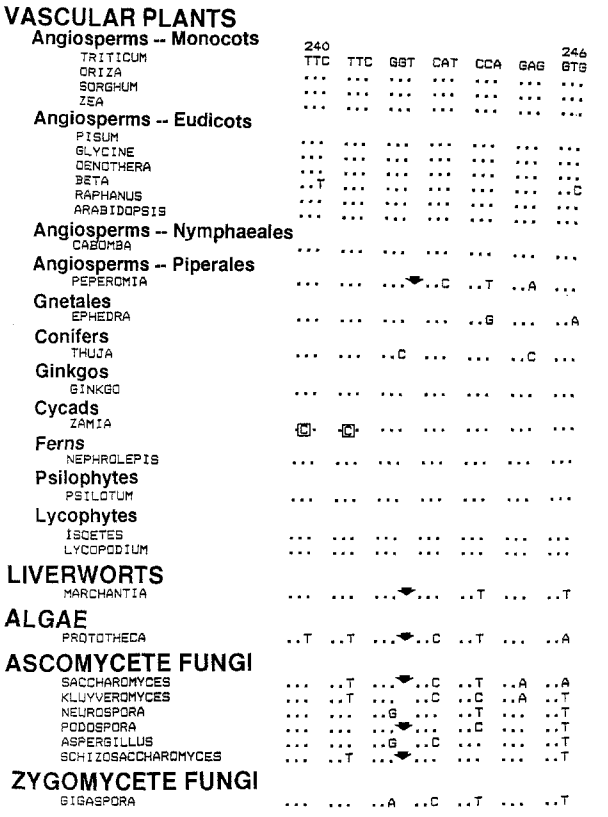

Fig. 4. Alignment of *coxI* exonic sequences from plants and fungi that flank the *Peperomia* group I intron insertion site. Codons are *numbered* relative to the wheat sequence (Bonen et al. 1987). Only nucleotides differing from those in wheat are shown. *Boxed* nucleotides in *Zamia* undergo C-to-U RNA editing (Sper-Whitis and Vaughn, unpublished). Sequences appear in approximate order of organismal relatedness (Bruns et al. 1992; Wilmotte et al. 1993; Donoghue 1994). *Arrows* mark the position of the group I intron in those *coxI* genes which contain it. Sources of sequences are given in Materials and Methods.

quences, again with 100% bootstrap support. That the *Peperomia* coding region is of the expected phylogenetic affinity dispels the possibility that the fungal-like intron in this species is merely an artifact of fungal DNA contamination of the *Peperomia* DNA used for PCR amplification. The poor resolution among the *Peperomia* and other angiosperm *coxI* sequences is not unexpected, both because relationships among the major lineages of angiosperms are still poorly understood (Donoghue 1994) and because of the limited number of changes among the angiosperm sequences.

The intron phylogenetic analyses included those introns that showed the highest similarity with the related *coxI* introns from *Peperomia* and *Marchantia* in Blast searches and which also contain a (related) endonuclease ORF, regardless of the intron's genic position. *CoxI* intron aI2a from *Schizosaccharomyces pombe* (Trinkl and Wolf 1986) is inserted at the same position as the *Peperomia* intron, but was excluded because it lacks an ORF. Phylogenetic analysis of the entire intron nucleotide sequence resolved the nine plant and fungal group I introns into three major groups, each with moderate-to-high bootstrap support with all three methods (Fig. 5A). One clade consists of two introns that occupy a second location (Fig. 5A) in *coxI* genes of *Aspergillus* and *Schizosaccharomyces* and which themselves are related by a relatively recent lateral transfer event (Lang 1984). The second group contains *coxI* introns, located at the same position as in *Peperomia,* from two "plants": the nonvascular plant *Marchantia* and the green alga *Prototheca.* The third clade contains the third plant intron—the *Peperomia coxI* intron—and four fungal mitochondrial introns.

The clear-cut conflict between the position of *Peperomia* in the *coxI* exon and intron trees (Fig. 5B) is strong evidence for a chimeric origin of this gene. The *Peperomia coxI* exons are clearly of land plant origin, whereas the intron is, almost equally clearly, not. This latter point is evident in three ways. First, the exclusion of the *Peperomia* intron from a plant-specific clade (i.e., *Marchantia* plus *Prototheca*) is strongly supported by bootstrap analysis (82–94%). Second, a parsimony analysis in which the *Peperomia* and *Marchantia* introns were constrained to be monophyletic (i.e., as expected for vertical transmission of the intron) produced a shortest tree that was fully 39 steps longer than the shortest tree in the unconstrained analysis of Fig. 5A. Third, the *Marchantia* and *Prototheca* introns were strongly united, to the exclusion of the *Peperomia* intron, in all phylogenetic analyses (carried out using all three methods) performed on

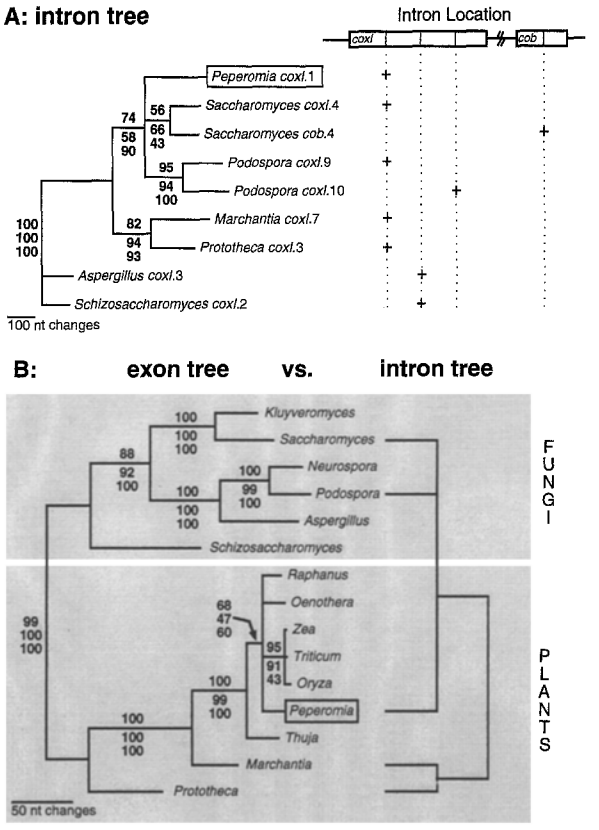

Fig. 5. Phylogenetic trees of mitochondrial *coxI* intron **(A)** and exon **(B)** sequences. The relative order of the related introns included in these analyses is shown at *upper right* (the three *coxI* introns are at nucleotide positions 720, 731, and 807 according to the *Saccharomyces* coordinate system; the *cob* intron is at position 760). Tree **A** shows the single shortest tree (of 2,313 steps) constructed from a parsimony analysis of the 956 variable positions in the entire-intron nucleotide data set, while **B** shows the most parsimonious tree (of 889 steps) for the 535 variable positions in the exon nucleotide data set. In both trees, bootstrap values from parsimony analysis are shown above branches, with maximum likelihood bootstrap values in the middle and neighbor-joining values at bottom. Branches with less than 50% bootstrap support in parsimony analysis are shown collapsed. The intron tree is rooted on the *Aspergillus* and *Schizosaccharomyces* introns, while the exon tree is rooted according to the traditional plant/fungal split. In **B,** a skeletal version of the intron tree is contrasted with the exon tree to emphasize the extent to which the two trees are congruent. Note that while the *Peperomia coxI* intron is most closely related to fungal introns, its exon sequence is most closely related to other angiosperm sequences.

two subsets of the entire-intron data set—namely, the group I intron core sequences and the intronic ORF sequences (the latter analyzed using both amino acid and nucleotide sequences; data not shown).

While we can firmly conclude that the *Peperomia coxI* intron is not of vertical, land plant origin, the evidence for a specifically fungal origin of this intron is somewhat less strong. This greater uncertainty is largely a reflection of the overall moderate level of bootstrap support (varying from 58% to 90%, depending on the method used) for the grouping of this intron with four fungal introns (Fig. 5A). However, additional evidence for a fungal origin comes from consideration of amino

Table 1. Percent amino acid identity among related intronic ORFs[a]

	Pp aI1	*Mp* aI7	*Pw* aI3	*Sc* aI4	*Sc* bI4	*Pa* aI9	*Pa* aI10
P. polybotrya (*Pp*)							
intron aI1	—	19%	26%	52%	47%	41%	38%
M. Polymorpha (*Mp*)							
intron aI7		—	35%	24%	26%	26%	24%
P. wickerhamii (*Pw*)							
intron aI3			—	30%	33%	36%	30%
S. cerevisiae (*Sc*)							
intron aI4				—	64%	46%	40%
S. cerevisiae (*Sc*)							
intron bI4					—	45%	40%
P. anserina (*Pa*)							
intron aI9						—	51%

[a] Number of amino acids included in pairwise comparisons ranges from 217 to 229

acid sequence relationships of the relevant intron ORFs (Table 1). The *Peperomia* ORF is consistently higher in amino acid identity (38–51%) to the four fungal intron ORFs than it is to either the *Prototheca* (26%) or *Marchantia* (19%) ORFs, or the latter two are to the fungal sequences (24–36%).

Discussion

We have discovered the first group I intron in a vascular plant mitochondrial genome and the first intron of any type in a vascular plant *coxI* gene. Two separate lines of evidence, intron presence/absence and phylogenetic analysis of intron sequences, both indicate a horizontal origin of the *Peperomia coxI* intron. The phylogenetic analyses also point to a specifically fungal origin of this intron, while the intron distribution data are neutral on this question. Conversely, the distribution data suggest a fairly recent origin of this intron within a *Peperomia*-specific lineage of flowering plants, at most 150 million years ago and possibly much more recently, while the phylogenetic analyses are essentially silent on the timing of intron transfer.

The *Peperomia* transfer adds to the growing body of reports of horizontal transfer of genetic information between distantly related species (reviewed in Smith et al. 1992; Kidwell 1993; Lambowitz and Belfort 1993). Such lateral transfers involve a variety of sequences, including introns, transposons, and housekeeping genes and occur on both evolutionary and developmental time scales (e.g., *Agrobacterium* plasmid transformation of the plant nucleus during bacterial infections, leading to crown-gall or hairy-root tumors; Zambryski et al. 1989).

Of greatest relevance to our study are the various claims made for lateral transfer of introns between disparate species. One of the most clear-cut cases involves the two identically positioned *coxI* group I introns that

form the outgroup clade in Fig. 5A. Horizontal transfer of this intron between recent ancestors of *Aspergillus* and *Schizosaccharomyces* has been invoked because the inferred amino acid sequences encoded by the two intronic ORFs are much more similar in sequence than are the proteins encoded by the two host *coxI* genes themselves (Lang 1984; Waring et al. 1984). A similar comparison of levels of exon and intron conservation has led to the proposal of recent horizontal transfer of a group II intron in *coxI* genes of *Saccharomyces* and *Kluyveromyces* (Hardy and Clark-Walker 1991). A unique case, of interorganellar transfer, has been proposed for a group I intron present at the same position in the 26S-like rRNA genes of the *Chlamydomonas reinhardtii* chloroplast and the *Acanthamoeba castellanii* mitochondrion (Lonergan and Gray 1994). A group I intron may also have been exchanged between fungal mitochondria and phage T4 (Michel and Dujon 1986), while phage group I introns in general show evidence of being mobile genetic elements (Belfort 1990; Lambowitz and Belfort 1993).

The *Peperomia* case stands out from these intron transfers in two respects. First, this is the only strong evidence for the horizontal acquisition of an intron by a vascular plant organellar genome; the numerous other introns in these genomes are notably uniform in distribution and vertical in descent (Palmer 1991; Nugent and Palmer 1993; Wolfe and Palmer, unpublished), and thus contrast vividly with the mobile group I and II introns of fungal mitochondrial and phage genomes. Second, we regard this as one of the strongest cases made for long-distance lateral transfer of an organellar intron. This is in part because the evidence that the *Peperomia coxI* intron is not of land plant origin seems so clear, and in part because ours is, perhaps, the first case where formal phylogenetic analysis has been applied to exon and intron sequences of the same gene, here showing convincingly that these sequences have traveled fundamentally different evolutionary pathways. The phylogenetic analysis also highlights two previously inferred (Bonitz et al. 1980; Cummings et al. 1990) intragenomic, intraspecific lateral transfers of this same *coxI* intron (Fig. 5A), i.e., its duplicative transpositions from a *Peperomia*-like position to either a secondary *coxI* location (in *Podospora*) or to a different gene (*cob* in *Saccharomyces*). Combining these three interkingdom or intragenomic transfers with the above-mentioned *Aspergillus/Schizosaccharomyces* transfer (Lang 1984; Waring et al. 1984) and the ancient intragenomic duplication implied at the base of Fig. 5A, it is clear that this *coxI* intron has indeed been very promiscuous during its evolution.

The *Peperomia* intron transfer warrants at least three different sorts of further study. First, by determining the intron's distribution within the *Peperomia* lineage (Piperales), it should be possible to estimate when this intron was acquired. This study will be greatly facilitated by the extraordinarily low mutation rate in plant mitochondrial genomes (Wolfe et al. 1987), which should make tracking the intron by Southern hybridization or PCR analysis extremely easy and reliable, and by the low incidence of intron loss in these genomes (Nugent and Palmer 1993), which should reduce the chance that subsequent loss of the intron will obscure the timing of its origin.

Second, sequencing of the exonic boundaries of the intron from taxa that phylogenetically circumscribe the time of its acquisition may shed light on the mechanism of intron transfer. Current thinking is that group I introns transpose using the double-strand-break repair pathway (Szostak et al. 1983; Clyman and Belfort 1992; Lambowitz and Belfort 1993), in which the site-specific DNA endonuclease encoded by an intronic ORF of a donor genome makes a staggered, double-strand break within an intronless target site. The break is then enlarged exonucleolytically, followed by recombination of donor and host sequences within the exonic regions flanking the intron, and finally gap-repair using the donor exonic sequences as templates. As a consequence, co-conversion of markers surrounding the site of endonuclease cleavage is observed—i.e., the exonic sequences immediately flanking the newly inserted intron are derived from the donor sequence and can differ from those in the host (Bell-Pedersen et al. 1989).

Inspection of plant and fungal *coxI* exon sequences that flank the *Peperomia* intron insertion site hints at the possibility of a co-conversion event in the specific history of this *Peperomia* gene. That is, the *Peperomia coxI* sequence is notably divergent, at three of 21 positions flanking the intron site, compared to the absolute identity in this region of 15 of the 21 known land plant sequences (Fig. 4). Furthermore, the *Peperomia* sequence matches the *Saccharomyces* sequence at these three divergent positions (Fig. 4), consistent with a fungal origin of both the *Peperomia* intron (Fig. 5A, Table 1) and, by co-conversion, its flanking exonic sequences. However, these three changes may be fortuitous, independent changes. The five other variable plant sequences also have multiple changes; other fungi differ from *Saccharomyces* at these three positions (all of which are silent); and the alga *Prototheca* shares two of the three changes with *Peperomia.* To properly test the co-conversion hypothesis, one would need to sequence exonic boundary sequences from throughout the *Peperomia* lineage that has the intron and from related taxa lacking the intron. A 1:1 correspondence between intron presence and C/T/A at these three positions would be strong evidence for co-conversion.

Finally, since the *Peperomia* transfer may well be quite recent, there is a good chance that one can identify the specific, putatively fungal donor lineage by sequencing *coxI* genes and introns from a broad range of additional fungi. One possible source of the *Peperomia coxI* intron is within the symbiotic, obligate inter- and intracellular vesicular mycorrhizal (VAM) fungi known to

grow in close association with root cells in nearly all plant species (reviewed by Bonfante-Fasolo 1984). We are currently exploring this interesting hypothesis.

Acknowledgments. We thank C.A. Makaroff and K.G. Wilson for valuable discussions and P.A. Schultz for sharing her VAM spore DNA isolation procedure prior to its publication. We also thank K.L. Adams for sharing his *Ephedra* and *Gigaspora coxI* gene sequences prior to their publication. J.F. Keegan provided *P. polybotrya* specimens from the Miami University greenhouses. This work was supported by funds provided to J.C.V. from a Howard Hughes Biomedical Research Grant to Miami University, by Howard Hughes Summer Research Internship awards to M.T.M. and G.L.S.-W., by a Sigma Xi grant to G.L.S.-W., by an NSF postdoctoral fellowship to P.K., and by NIH grant GM-35087 to J.D.P.

References

Belfort M (1990) Phage T4 introns: self-splicing and mobility. Annu Rev Genet 24:363–385

Bell-Pedersen D, Quirk SM, Aubrey M, Belfort M (1989) A site-specific endonuclease and co-conversion of flanking exons associated with the mobile *td* intron of phage T4. Gene 82:119–126

Bonen L, Boer PH, McIntosh JE, Gray MW (1987) Nucleotide sequence of the wheat mitochondrial gene for subunit I of cytochrome oxidase. Nucleic Acids Res 15:6734

Bonfante-Fasolo P (1984) Anatomy and morphology of VA mycorrhizae. In: Powell CL, Bagyaraj DJ (eds) VA Mycorrhiza. CRC Press, Boca Raton, FL pp 5–33

Bonitz SG, Coruzzi G, Thalenfeld BE, Tzagoloff A, Macino G (1980) Assembly of the mitochondrial membrane system. Structure and nucleotide sequence of the gene coding for subunit 1 of yeast cytochrome oxidase. J Biol Chem 255:11927–11941

Bruns TD, Vilgalys R, Barns SM, Gonzalex D, Hibbett DS, Lane DJ, Simon L, Stickel S, Szaro TM, Weisburg WG, Sogin ML (1992) Evolutionary relationships within the fungi: analyses of nuclear small subunit rRNA sequences. Mol Phylogen Evol 1:231–241

Burke JM, Belfort M, Cech TR, Davies RW, Schweyen RJ, Shub DA, Szostak JW, Tabak HF (1987) Structural conventions for group I introns. Nucleic Acids Res 15:7217–7221

Capaldi RA, Malatesta F, Darley-Usmar VM (1983) Structure of cytochrome c oxidase. Biochim Biophys Acta 726:135–148

Clark-Walker GD (1992) Evolution of mitochondrial genomes in fungi. Int Rev Cytol 141:89–127

Clyman J, Belfort M (1992) *Trans* and *cis* requirements for intron mobility in a prokaryotic system. Genes Dev 6:1269–1279

Covello PS, Gray MW (1992) Silent mitochondrial and active nuclear genes for subunit 2 of cytochrome c oxidase (*cox2*) in soybean: evidence for RNA-mediated gene transfer. EMBO J 11:3815–3820

Cummings DJ, Michel F, McNally KL (1989) DNA sequence analysis of the 24.5 kilobase pair cytochrome oxidase subunit I mitochondrial gene from *Podospora anserina:* a gene with sixteen introns. Curr Genet 16:381–406

Cummings DJ, McNally KL, Domenico JM, Matsuura ET (1990) The complete DNA sequence of the mitochondrial genome of *Podospora anserina.* Curr Genet 17:375–402

Delahodde A, Goguel V, Becam AM, Creusot F, Perea J, Banroques J, Jacq C (1989) Site-specific DNA endonuclease and RNA maturase activities of two homologous intron-encoded proteins from yeast mitochondria. Cell 56:431–441

Donoghue MJ (1994) Progress and prospects in reconstructing plant phylogeny. Ann Missouri Bot Gard 81:405–418

Doyle JJ, Doyle JL (1987) A rapid DNA isolation for small quantities of fresh tissue. Phytochem Bull 19:11–15

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

Felsenstein J (1993) PHYLIP (phylogeny inference package), version 3.5c. Distributed by the author, Dept of Genetics, University of Washington, Seattle

Fox GE, Woese CR (1975) 5S RNA secondary structure. Nature 256: 505–507

Gray MW (1992) The endosymbiont hypothesis revisited. Int Rev Cytol 141:233–357

Grohmann L, Brennicke A, Schuster W (1992) The mitochondrial gene encoding ribosomal protein S12 has been translocated to the nuclear genome in *Oenothera.* Nucleic Acids Res 20:5641–5646

Hardy CM, Clark-Walker GD (1991) Nucleotide sequence of the *coxI* gene in Kluyveromyces lactis mitochondrial DNA: evidence for recent horizontal transfer of a group II intron. Curr Genet 20:99–114

Hendy MD, Penny D (1982) Branch and bound algorithms to determine minimal evolutionary trees. Math Biosci 59:277–290

Hiesel R, von Haeseler A, Brennicke A (1994) Plant mitochondrial nucleic acid sequences as a tool for phylogenetic analysis. Proc Natl Acad Sci 91:634–638

Kidwell MG (1993) Lateral transfer in natural populations of eukaryotes. Annu Rev Genet 27:235–256

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

Lang BF (1984) The mitochondrial genome of the fission yeast *Schizosaccharomyces pombe:* highly homologous introns are inserted at the same position of the otherwise less conserved *coxI* genes in *Schizosaccharomyces pombe* and *Aspergillus nidulans.* EMBO J 3:2129–2136

Lang BF (1993) The mitochondrial genome of *Schizosaccharomyces pombe.* In: O'Brien SJ (ed) Genetic maps, 6th ed, book 3. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 3.118–3.119

Li M, Tzagoloff A (1979) Assembly of the mitochondrial membrane system. Sequences of yeast mitochondrial valine—and an unusual threonine tRNA gene. Cell 18:47–53

Lonergan KM, Gray MW (1994) The ribosomal RNA gene region in *Acanthamoeba castellanii* mitochondrial DNA. A case of evolutionary transfer of introns between mitochondria and plastids? J Mol Biol 239:476–499

Michel F, Dujon B (1986) Genetic exchanges between bacteriophage T4 and filamentous fungi? Cell 46:323

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

Muscarella DE, Ellison EL, Ruoff BM, Vogt VM (1990) Characterization of I-*Ppo,* an intron-encoded endonuclease that mediates homing of a group I intron in the ribosomal DNA of *Physarum polycephalum.* Mol Cell Biol 10:3386–3396

Nobrega F, Tzagoloff A (1980) Assembly of the mitochondrial membrane system. DNA sequence and organization of the cytochrome b gene in *Saccharomyces cerevisiae* D273-10B. J Biol Chem 255: 9828–9837

Noller HF, Kop JA, Wheaton V, Brosius J, Gutell RR, Kopylov AM, Dohme F, Herr W, Stahl DA, Gupta R, Woese CR (1981) Secondary structure model for 23S ribosomal RNA. Nucleic Acids Res 9:6167–6189

Nugent JM, Palmer JD (1991) RNA-mediated transfer of the gene *coxII* from the mitochondrion to the nucleus during flowering plant evolution. Cell 66:473–481

Nugent JM, Palmer JD (1993) Evolution of gene content and gene organization in flowering plant mitochondrial DNA: a general survey and further studies on *coxII* gene transfer to the nucleus. In:

Brennicke A, Kuck U (eds) Plant mitochondria. VCH Publishers, New York, pp 163–170

Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogura Y, Kohchi T, Ohyama K (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. A primitive form of plant mitochondrial genome. J Mol Biol 223:1–7

Ohta E, Oda K, Yamato K, Nakamura Y, Takemura M, Nozato N, Akashi K, Ohyama K, Michel F (1993) Group I introns in the liverwort mitochondrial genome: the gene coding for subunit 1 of cytochrome oxidase shares five intron positions with its fungal counterparts. Nucleic Acids Res 21:1297–1305

Ohyama K, Oda K, Ohta E, Takemura M (1993) Gene organization and evolution of introns of a liverwort, *Marchantia polymorpha,* mitochondrial genome. In: Brennicke A, Kuck U (eds) Plant mitochondria. VCH Publishers, New York, pp 115–129

Olsen GJ, Matsuda H, Hagstrom R, Overbeek R (1994) FastDNAml: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. Comput Appl Biosci 10:41–48

Palmer JD (1991) Plastid chromosomes: structure and evolution. In: Bogorad L, Vasil IK (eds) Cell culture and somatic cell genetics of plants. vol 7A. The molecular biology of plastids. Academic Press, San Diego, pp 5–53

Paquin B, Lang BF (1993) The mitochondrial genome of *Allomyces macrogynus.* In: O'Brien SJ (ed) Genetic maps, 6th ed. book 3. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 3.130–3.132

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci 74:5463–5467

Schuster W, Brennicke A (1994) The plant mitochondrial genome: physical structure, information content, RNA editing, and gene migration to the nucleus. Annu Rev Plant Physiol Plant Mol Biol 45:61–78

Smith MW, Feng D-F, Doolittle RF (1992) Evolution by acquisition: the case for horizontal gene transfers. Trends Biochem Sci 17:489–493

Swofford DL (1993) PAUP: phylogenetic analysis using parsimony, version 3.1. Computer program distributed by the Illinois Natural History Survey, Champaign

Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW (1983) The double strand-break repair model for recombination (review). Cell 33: 25–35

Trinkl H, Wolf K (1986) The mosaic *coxI* gene in the mitochondrial genome of *Schizosaccharomyces pombe:* minimal structural requirements and evolution of group I introns. Gene 45:289–297

Turmel M, Mercier J-P, Cote M-J (1993a) Group I introns interrupt the chloroplast *psa*B and *psb*C and the mitochondrial *rrn*L gene in *Chlamydomonas.* Nucleic Acids Res 21:5242–5250

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

Waring RB, Brown TA, Ray JA, Scazzocchio C, Davies RW (1984) Three variant introns of the same general class in the mitochondrial gene for cytochrome oxidase subunit I in *Aspergillus nidulans.* EMBO J 3:2121–2128

Wilmotte A, van de Peer Y, Goris A, Chapelle S, de Baere R, Nelissen B, Neefs J-M, Hennebert GL, de Wachter R (1993) Evolutionary relationships among higher fungi inferred from small ribosomal subunit RNA sequence analysis. System Appl Microbiol 16:436–444

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054–9058

Wolff G, Burger G, Lang BF, Kuck U (1993) Mitochondrial genes in the colourless alga *Prototheca wickerhamii* resemble plant genes in their exons but fungal genes in their introns. Nucleic Acids Res 21:719–726

Xu MQ, Kathe SD, Goodrich-Blair H, Nierzwicki-Bauer SA, Shub DA (1990) The prokaryotic origin of a chloroplast intron: a self-splicing group I intron in the gene for tRNA-Leu (UAA) of cyanobacteria. Science 250:1566–1570

Zambryski P, Tempe J, Schell J (1989) Transfer and function of T-DNA genes from *Agrobacterium* Ti and Ri plasmids. Cell 56: 193–201