Multiple Acquisitions via Horizontal Transfer of a Group I Intron in the Mitochondrial *cox1* Gene During Evolution of the Araceae Family

Yangrae Cho and Jeffrey D. Palmer

Department of Biology, Indiana University at Bloomington

A group I intron has recently been shown to have invaded mitochondrial *cox1* genes by horizontal transfer many times during the broad course of angiosperm evolution. To investigate the frequency of acquisition of this intron within a more closely related group of plants, we determined its distribution and inferred its evolutionary history among 14 genera of the monocot family Araceae. Southern blot hybridizations showed that 6 of the 14 genera contain this intron in their *cox1* genes. Nucleotide sequencing showed that these six introns are highly similar in sequence (97.7%–99.4% identity) and identical in length (966 nt). Phylogenetic evidence from parsimony reconstructions of intron distribution and phylogenetic analyses of intron sequences is consistent with a largely vertical history of intron transmission in the family; the simplest scenarios posit but one intron gain and two losses. Despite this, however, striking differences in lengths of exonic co-conversion tracts, coupled with the absence of co-conversion in intron-lacking taxa, indicate that the six intron-containing Araceae probably acquired their introns by at least three and quite possibly five separate horizontal transfers. The highly similar nature of these independently acquired introns implies a closely related set of donor organisms.

Introduction

Group I introns are regarded as mobile genetic elements that have been transmitted laterally as well as vertically (reviewed by Dujon 1989; Perlman and Butow 1989; Lambowitz and Belfort 1993). Two types of group I intron mobility are recognized. Intron ''homing'' refers to an intron's transfer from intron-containing to intronless alleles of the same gene at the same site within these alleles. Intron ''transposition'' refers to movement of an intron to a different location, usually a different gene.

Intron homing is well-documented in genetic crosses, in which group I introns duplicatively spread (home) from intron-containing to intron-lacking alleles by homology-dependent gene conversion. Homing has been observed in various mitochondrial (mt) and chloroplast (cp) genes (e.g., Jacquier and Dujon 1985; Lemieux and Lee 1987; Delahodde et al. 1989; Wenzlau et al. 1989), nuclear rRNA genes of *Physarum* (Muscarella and Vogt 1989), and protein-coding genes of T-even phages (Quirk, Bell-Pedersen, and Belfort 1989). It has been pointed out in each case (Jaquier and Dujon 1985; Zinn and Butow 1985; reviewed in Lambowitz and Belfort 1993; Belfort and Perlman 1995) that the homing mechanism is analogous to mating-type switching in yeast (Szostak et al. 1983). Homing is initiated by an intron-encoded protein which searches for its recognition site in the recipient DNA and generates a staggered double-strand break within such a long recognition site that there are few target sites per genome (reviewed in Lambowitz and Belfort 1993). The uncleaved intron-containing gene is thought to serve as a template for double-strand break repair, and two resulting Holliday junctions are thought to be resolved. As a consequence of this process, co-conversion of recipient exonic sequences flanking the intron insertion site by donor sequences is always observed where assayed; co-conversion tract lengths often exceed a few hundred nucleotides (nt) (Jacquier and Dujon 1985; Muscarella and Vogt 1989; Mueller, Smith, and Belfort 1996).

When they occur among mating taxa, these homing events lead to an essentially vertical history of intron transmission (which we term ''vertical homing''). A number of cases of cross-species homing of group I introns via horizontal transfer (''horizontal homing'') have also been inferred (Lang 1984; Biniszkiewicz, Cesnaviciene, and Shub 1994; Turmel et al. 1995; Vaughn et al. 1995; Bhattacharya, Friedl, and Damberger 1996; Hibbett 1996; Cho et al. 1998*b*; Nishida, Tajiri, and Sugiyama 1998). Group I intron transpositions (which can occur by either cross-species or intragenomic horizontal transfer) have also been inferred in a number of cases (reviewed in Michel and Dujon 1986; Lambowitz and Belfort 1993; Turmel, Mercier, and Cote 1993; Bhattacharya, Friedl, and Damberger 1996).

One family of group I introns illustrates all three of these mobility phenomena—vertical homing, horizontal homing, and transposition—especially well. In most cases, including all examined angiosperms, the introns from this family contain an open reading frame (ORF) whose product is thought to have both site-specific endonuclease and RNA maturase (splicing) activities (Delahodde et al. 1989; Wenzlau et al. 1989; Henke, Butow, and Perlman 1995). Members of this intron family, which belongs to the IB1 subclass of group I introns, are found at four different locations across two mt genes (*cox1* and *cob*) and occur in a wide diversity of eukaryotes, including several diverse fungi, the sea anemone *Metridium senile*, the slime mold *Dictyostelium discoideum*, two green algae, the liverwort *Marchantia polymorpha*, and many diverse flowering plants (Cho et al. 1998*b*; Watanabe et al. 1998 and references therein).

A broad-level survey of angiosperms (with a sampling of usually one species each from nearly 200 families) suggested over 30 separate events of horizontal

Abbreviations: cp, chloroplast; mt, mitochondrial; nt, nucleotides; PCR, polymerase chain reaction.

Key words: group I intron, horizontal transfer, mitochondrial DNA, exonic co-conversion, Araceae, phylogeny, *rbcL*, *cox1*.

Address for correspondence and reprints: Jeffrey D. Palmer, Department of Biology, Indiana University, Bloomington, Indiana 47405. E-mail: jpalmer@bio.indiana.edu.

Mol. Biol. Evol. 16(9):1155–1165. 1999

homing of this intron, from an uncertain set of donors and always into the same site within the *cox1* gene (Cho et al. 1998*b*; the intron is located after nucleotide 726 in the *cox1* coding region, at the same site as the related ai4 intron from *Saccharomyces cerevisiae* [Bonitz et al. 1980]). This raised the question of how often this intron is acquired on a finer-scale phylogenetic level, e.g., within a family of angiosperms. To address this question, we surveyed 14 genera in the large (over 100 genera and 2,500 species) and florally distinctive (spathe and spadix) monocot family Araceae (e.g., taro, elephants-ear, jack-in-the-pulpit). We found that the six genera which possess the intron all contain it in this same *cox1* location and that they probably acquired it by at least three and most likely five separate horizontal transfers.

Materials and Methods

Total DNAs for 12 of the 14 species of Araceae examined in this study were generously provided by Yin-Long Qiu. DNAs of the other two species were extracted according to the method of Doyle and Doyle (1987) and further purified by CsCl banding.

Specimens are vouchered at the Indiana University Herbarium; (voucher numbers are given in parentheses after species names): *Peltandra virginica* (Qiu 96064), *Dieffenbachia* sp. (Qiu 96007), *Anthurium scherzerianum* (Qiu 96004), *Scindapsus aureus* (Qiu 96005), *Spathiphyllum clevelandii* (Qiu 94140), *Orontium aquaticum* (Qiu 97112), *Xanthosoma mafatta* (Qiu 95063), *Philodendron oxycardium* (Qiu 96053), *Zamioculcas zamiifolia* (Qiu 96006), *Arisaema triphyllum* (Qiu 96011), *Amorphophallus rivieri* (Qiu 96054), *Pistia stratiotes* (no voucher, IU green house), *Zantedeschia aethiopica* (Qiu 96009), and *Lemna gibba* (no voucher). While *Lemna* is traditionally placed in a separate family (Lemnaceae), molecular data strongly indicate that this ''family'' actually belongs within the Araceae (French, Chung, and Hur 1995), and it was thus included in this study.

For Southern blots, 2 μg of total genomic DNA was digested with *Hin*dIII and electrophoresed in a 0.7% agarose gel. The DNA in the gel was depurinated for 5 min in 0.25M HCl, denatured for 30 min in 1.5 M NaCl/0.5 M NaOH, and neutralized for 30 min in 3 M NaCl/0.5 M Tris (pH 7.2). The DNA was transferred overnight from the gel to a Nytran membrane (Schleicher and Schuell) and cross-linked for 23 s to the membrane using a U.V. crosslinker (Fotodyne). The *cox1* intron probe was a gel-purified 766-nt fragment internal to the *cox1* intron from *Veronica ugrestis* which was made by *Dra*I digestion of a PCR product generated using primers IP53 and IP56 (see below). The *cox1* coding region probe was a gel-purified 341-nt PCR product from *Beta vulgaris*, made using primers IP54 (CTATGGCTTAGGTCCG) and IP53. Probes were ^{32}P-labeled by random priming. Prehybridizations, hybridizations, and washings were carried out at 66°C. The membrane was prehybridized for 3 h in 4 × SSC, 0.5% nonfat dry milk, and 0.5% SDS and hybridized overnight in this solution plus probe added to 2 ng/ml. The membrane was washed in 2 × SSC and 0.1% SDS for 1 h and in 0.5 × SSC and 0.025% SDS for 30 min.

Polymerase chain reaction (PCR) amplifications were performed at 55°C annealing temperature in a 10-μl volume containing 20–50 ng DNA, 1.7 mM $MgCl_2$, 5% acetamide, 1 × reaction buffer (50 mM Tris [pH 8.5], 20 mM KCl, 0.5 mg/ml BSA (Sigma), 0.5 g Ficoll 400 (Sigma), 4 mg Xylene Cyanol FF), and 0.05 U *Taq* polymerase (Sigma) in an Idaho thermocycler for 37 cycles, each with a 1-min extension time. A nearly full length region of the cp *rbcL* gene (1,396 nt of the usually 1,428-nt gene) was amplified from 12 species of Araceae using primers AC1 (TCACCACAAACAGAAACGAAAGC) and AC4 (TTGATTTCCTTCCAGACTTCACAA). *cox1* genes of intron-containing species were amplified using three pairs of primers: cox42F (GGATCTTCTCCACTAACCACAAA) and cox657R (GCGGGATCAGAAAAGGTTGTA); IP53 (GGAGCAGTTGATTTAGC) and IP56 (GAGCAATGTCTAGCCC); and INT1.2KF (AGCATGGCTAGCTTTCCTAGA) and coxI1.6KR (AAGGCTGGAGGGCTTTGTAC). These amplified a ca. 600-nt region of the 5′ exon, a ca. 1,650-nt region containing the entire intron and flanking exonic sequences, and a ca. 950-nt region that includes the very end of the intron and most of the 3′ exon, respectively. For intron-lacking species, primers cox42F/cox657R (ca. 600 nt product) and IP53/coxI1.6KR (ca. 1,000 nt) were used to amplify the same aggregate length of coding region (1,512 of ca. 1,590 nt) as for the intron-containing species.

PCR products were electrophoresed in 1.0% agarose gels and purified using Gene Clean kits (BIO101). Purified PCR products made using primers cox42F/cox657R and AC1/AC4 were used directly as template DNAs for ABI automated sequencing (Applied Biosystems Inc.). These four PCR primers were also used as ABI sequencing primers. Three additional internal primers (rbcL2F, CAATGCATGCAGTTATTGA; rbcL2R, GGGACGACCATACTTGTT; rbcL3R, CATATGCCAAACATGAATAC) were used to sequence the AC1/AC4 PCR products. PCR products made using primers IP53/IP56 and INT1.2K/cox1.6KR were ligated to the TA cloning vector (Invitrogen) and sequenced using LiCor automated sequencing methods using two vector primers (LiCor and Epicentre Technologies). For IP53/IP56 products, two dye-labeled internal primers (LiCor and Epicentre Technologies; INT800R, ACGATGTAGTTGAAGTAGTCGTGCTGAAT; INT600F, ATGGGACTTGAAGATCTACCACTACTACGA) were used in addition to the vector primers. All nucleotide sequences were determined on both strands for over 90% of the sequence reported. The cp restriction site data set was kindly provided by J. French and corresponds to that used in French, Chung, and Hur (1995), except that all sites from the *rbcL* region were excluded from the combined phylogenetic analyses to avoid counting the same characters twice.

Initial alignments of nucleotide sequences were generated with PILEUP (Genetic Computer Group 1991) and then refined by hand using McClade (Maddison and Maddison 1992) and the editor of PAUP* v

4.0 (Swofford 1998). The *rbcL* alignment (1,347 nt = 1,396 nt minus 49 nt of primer sequence) included 2 published sequences (M91630, M96963) and 12 sequences determined in this study (AJ005623–AJ005632, AJ007543, AJ007544). The *cox1* coding sequence alignment (1,370 nt; this excluded primer sequences, the 20-nt co-conversion region, and terminal regions of somewhat ambiguous sequence) consisted of 13 sequences from the Araceae (AJ007545–AJ007555), all but two of which were determined in this study. The *cox1* intron alignment (947 nt) included four Araceae sequences determined in this study (AJ007545–AJ007548), 29 sequences from Cho et al. (1998*b*; AJ223411–AJ223439), and one sequence from Vaughn et al. (1995; X94594). Gaps were present only in the intron alignment and were excluded from all phylogenetic analyses. The 3′ exonic co-conversion region in *cox1*was also excluded from analyses. The unsequenced *cox1* gene from *Lemna* was treated as missing data.

All phylogenetic analyses were carried out using PAUP* v 4.0 d63 (Swofford 1998) on a Power Macintosh (model 9600/300). Analyses of a combined data set consisting of *rbcL* and *cox1* sequences (again, excluding the 20-nt-long *cox1* co-conversion region) and cp restriction sites were performed with parsimony only. The shortest tree was found using a branch-and-bound search via furthest search addition. Bootstrapping was performed using 100 bootstrap replicates via 10 random-addition replicates and tree bisection-reconnection (TBR) branch swapping. For intron phylogeny construction, parsimony analyses were performed with stepwise addition, 10 replicates of random addition, and TBR branch-swapping options; Neighbor-Joining analyses were performed using Jukes-Cantor corrected distances and TBR branch-swapping options; and maximum-likelihood analyses used an empirical nucleotide frequency, a transition/transversion ratio of 0.5, and the HKY model. For bootstrap analyses of the intron data set, parsimony analysis was performed under as-is addition and subtree pruning-regrafting branch-swapping conditions, while Neighbor-Joining analysis used the same conditions as above.

Results and Discussion

Distribution of the *cox1* Intron in Araceae

To determine the distribution of this *cox1* intron in the Araceae, a Southern blot survey was carried out using total DNAs from a single species from each of 14 genera in the family. A *cox1* coding region probe from the dicot *Beta* hybridized to a single band in every species, with some variation in signal strength, presumably owing to variable DNA loadings and variable proportions of mtDNA in the total DNA preparations used (fig. 1*A*). In contrast, an intron probe from the dicot *Veronica* hybridized well to only 5 of the 11 species shown in figure 1*B* and 6 of 14 overall, with no signals detectable in eight of the DNAs. The intron probe usually hybridized to two bands, one in common with the exon probe (cf. fig. 1*A* and *B*) and one unique. Subsequent sequencing of the *cox1* intron (see below) revealed that the two

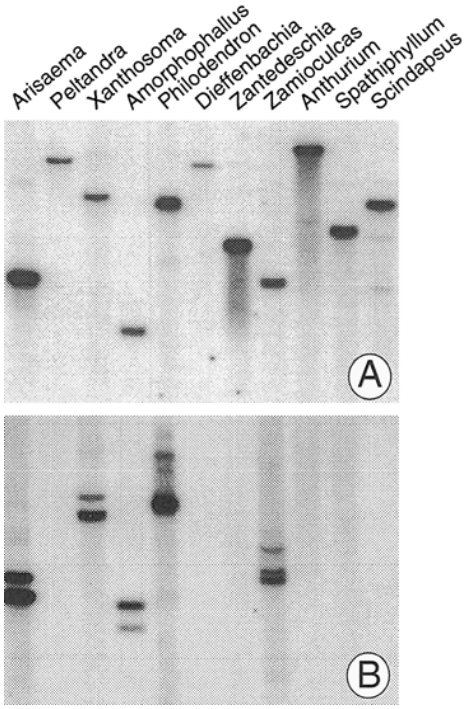

FIG. 1.—Southern hybridizations showing presence or absence of the *cox1* group I intron in the Araceae. Total DNAs from the 11 indicated species were digested with *Hin*dIII and analyzed by Southern blot hybridizations, using 341-nt coding region (*A*) and 766-nt intron (*B*) probes from the *cox1* gene.

*Hin*dIII bands in the intron lanes are the result of a *Hin*dIII site within the intron.

Mapping Intron Distribution on Araceae Phylogenies

We sequenced 1,347 nt of the cp *rbcL* gene and either 1,390 or 2,356 nt of the mt *cox1* gene (depending on whether or not it contained the 966-nt intron; see next section) from each of these 14 Araceae taxa (except for two *rbcL* sequences that were already available and *cox1* from *Lemna*, which is as yet unsequenced). Phylogenetic analyses of the coding regions from these two genes were performed in order to generate an ''organismal'' phylogenetic framework in which to interpret the observed distribution of the intron in terms of alternative models of intron gain and loss. Unfortunately, trees generated using either *rbcL* alone, *cox1* alone, or both genes combined were poorly resolved (data not shown). Therefore, we combined these two sequence data sets with previously published cpDNA restriction sites for these same taxa (French, Chung, and Hur 1995) and carried out a parsimony analysis. This produced a single shortest tree (fig. 2*A*) of 939 steps which is moderately well resolved in the sense that several nodes have high bootstrap support, while several others are poorly supported. As expected, this tree is highly congruent with published trees based on two of its constituent data sets—cpDNA restriction sites (all 14 of our Araceae taxa were examined in French, Chung, and Hur [1995]) and *rbcL* sequences (only three of these taxa were included in Chase et al. [1993] and Duvall et al. [1993]); there is no prior analysis of *cox1* sequences for the family. Of

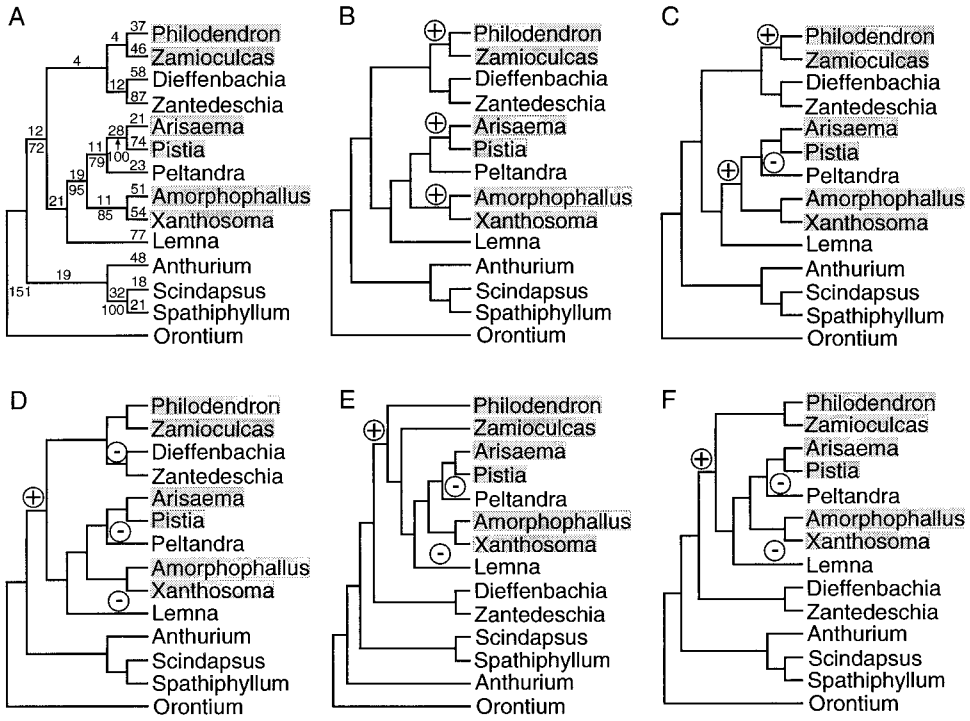

FIG. 2.—*A,* ''Organismal'' phylogeny of the 14 Araceae examined in this study. Shown is the single shortest tree (of 939 steps) from a parsimony analysis of a combined data set of cp *rbcL* sequences, mt *cox1* sequences (excluding the co-conversion region; see fig. 5), and cpDNA restriction sites. Above the branches are numbers of parsimony steps; below are bootstrap scores above 50%. Names of intron-containing taxa are shaded in this tree and in *B–E.* The tree is rooted on *Orontium,* which was basalmost among these Araceae in analyses of both cp restriction sites (French, Chung, and Hur 1995) and *rbcL* sequences (unpublished data) in which other monocot families were used as unequivocal outgroups. *B,* One of two shortest intron gain (plus sign)/loss (minus sign) scenarios mapped onto the shortest organismal tree shown in *A.* This scenario posits three separate intron gains. *C,* The other shortest intron gain/loss scenario (two gains and one loss) mapped onto the shortest organismal tree. *D,* A one-step-longer scenario (only one gain and three losses) mapped on the same tree. *E* and *F,* The two next-shortest organismal trees (of 940 steps), onto which are mapped one-gain/two-loss scenarios of intron evolution.

particular note, this tree strongly supports French, Chung, and Hur's (1995) conclusion that, contrary to Stockey, Hoffman, and Rothwell (1997), the aquatic *Pistia* does not belong to the aquatic Araceae-segregate family Lemnaceae (here represented solely by *Lemna*).

We used this combined evidence tree (fig. 2*A*) as a current best estimate of organismal phylogeny for the Araceae. A well-supported clade (95% bootstrap support) composed of the five taxa from *Arisaema* through *Xanthosoma* contains four of the six intron-containing species and the intronless *Peltandra.* There are two most parsimonious scenarios to account for the intron's presence in four of five members of this clade (all intron optimization scenarios were generated manually, with gain and loss of the intron weighted equally). A two-gain scenario posits separate gains in the common ancestor of *Arisaema*/*Pistia* and in that of *Amorphophallus*/*Xanthosoma,* while a gain/loss scenario posits gain of the intron in the common ancestor of all five taxa, followed by intron loss in *Peltandra.*

Looking across the entire Araceae phylogeny, there are two most-parsimonious, three-step scenarios, each of which simply extends the above two subtree models by postulating a separate gain of the intron in the common ancestor of the remaining two intron-containing taxa, *Philodendron* and *Zamioculcas.* Thus, figure 2*B* shows a three-gain scenario, and figure 2*C* shows a two-gain/one-loss scenario. Loss of a character is often thought to be easier and more common than its gain (but see below for the perhaps exceptional behavior of this intron); therefore, it is worth pointing out that an alternative intron gain/loss scenario only one step longer than these two shortest ones (fig. 2*B* and *C*) would postulate but a single gain of the *cox1* intron among all examined Araceae, followed by three separate losses (fig. 2*D*).

We also examined the set of two parsimony trees one step longer (940 steps) than the best tree (fig. 2*A–D*) (there are no two-step-longer trees, and there are 13 three-step-longer trees) to see whether any of them led to more parsimonious scenarios of intron gain and loss than those discussed above. While both trees lead to equally parsimonious intron-evolution models to those in figure 2*B* and *C,* they differ in that each postulates only a single intron gain followed by two losses (fig. 2*E* and *F*). Overall, then, mapping intron presence/absence on the shortest and next-shortest parsimony trees of the Araceae leads to fairly simple scenarios, with few (one to three) intron gains invoked, few (one to three) losses, and varying but potentially substantial (fig. 2*D–F*) degrees of vertical transmission of the intron.

Phylogeny of the *cox1* Intron

We next examined phylogenies of the *cox1* intron itself to evaluate their congruence with organismal es-

	Xanthosoma	Philodendron	Zamioculcas	Amorphophallus	Arisaema	Pistia
Xanthosoma	-	0.4	1.1	0.8	0.8	2.2
Philodendron	0.7	-	1.0	0.5	0.7	2.1
Zamioculcas	0.7	0.6	-	0.7	1.2	2.3
Amorphophallus	0.3	0.4	0.4	-	0.7	2.5
Arisaema	0.4	0.6	0.4	0.2	-	1.8
Pistia	1.3	1.5	1.3	1.1	1.0	-
Peltandra	0.5	0.6	0.6	0.2	0.4	1.3
Zantedeschia	1.3	1.3	1.2	1.1	1.0	1.8
Dieffenbachia	0.7	0.4	0.4	0.4	0.4	1.0
Anthurium	1.0	1.0	0.9	0.8	0.7	1.6
Scindapsus	0.8	0.8	0.8	0.6	0.7	1.7
Spathiphyllum	0.7	0.6	0.6	0.4	0.4	1.3
Orontium	1.8	1.9	1.7	1.6	1.7	2.6

FIG. 3.—Divergence of the *cox1* intron (above the diagonal) and coding sequence (below the diagonal). Shown are Jukes-Cantor corrected distances ×100. The 3′ co-conversion region was excluded from the coding sequence calculations.

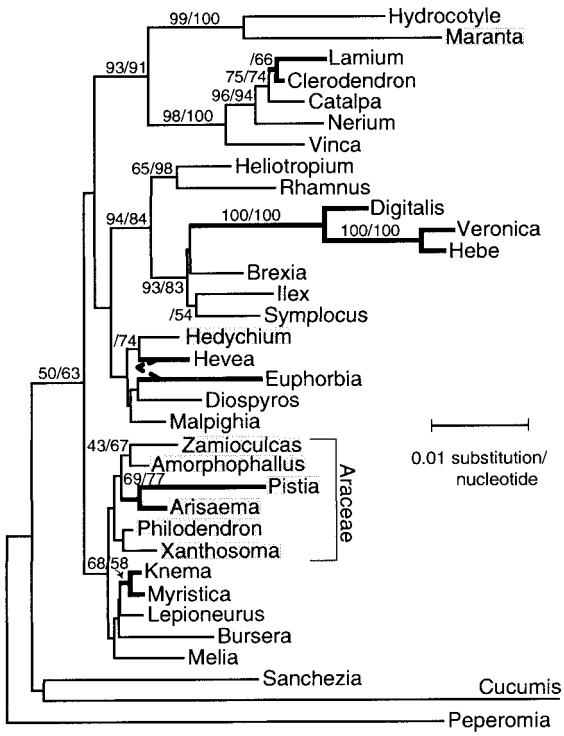

FIG. 4.—Phylogenetic relationships of the *cox1* introns from Araceae and other angiosperms. The phylogeny shown was created by a neighbor-joining analysis of a Jukes-Cantor-corrected distance matrix of a 947-character alignment of the entire *cox1* intron. This tree is arbitrarily rooted. Bootstrap scores above 50% are shown from parsimony (before the slash) and neighbor joining (after the slash) analyses. Five small intron clades of inferred vertical transmission are indicated by heavy branches. Monocots are shaded.

timates of phylogeny as an independent source of evidence on the relative balance between vertical and horizontal transmission of the intron during the evolution of the Araceae. Sequencing of nearly the full length of the *cox1* gene from the six taxa identified by Southern blots (fig. 1) as likely containing a *cox1* intron shows that they each contain a single intron, whose location within the *cox1* coding sequence is identical among all Araceae, as well as all other angiosperms known to contain a *cox1* intron (Vaughn et al. 1995; Adams, Clements and Vaughn 1998; Cho et al. 1998*b*). The six Araceae *cox1* introns are identical in length (966 nt) and highly similar in sequence (97.7%–99.6% identity; fig. 3).

Figure 4 shows an arbitrarily rooted phylogeny of the six Araceae introns, together with 28 *cox1* introns from other families of angiosperms (Vaughn et al. 1995; Cho et al. 1998*b*). As reported elsewhere (Cho et al. 1998*b*), there is considerable, often robust, incongruency between the overall topologies of the intron phylogeny (fig. 4) and those of the organisms in which they reside. To illustrate one example, the introns from monocots (shaded names in fig. 4), which, as organisms, are unquestionably monophyletic, fall into three separate groups, each affiliated with a different group of dicots. The intron from the monocot *Maranta* groups with that of *Hydrocotyle,* a member of the dicot Asteridae II clade, with 99%–100% bootstrap support, while the intron from *Ilex,* the only other Asteridae II included, groups with high bootstrap support with a different lineage of dicots.

Given the extreme incongruence between intron and organismal phylogenies, such that only a few terminal groups of vertically transmitted introns are recognized (bold branches in fig. 4; Cho et al. 1998*b*), the clustering of all six Araceae *cox1* introns was surprising. Except for the clustering of the *Arisaema* and *Pistia* introns, the Araceae intron branching order (fig. 4) is incongruent with that of the host organisms (fig. 2*A*). However, bootstrap levels within the Araceae part of the intron phylogeny in figure 4 are generally low, and results from the Kishino-Hasegawa (1989) test indicate that this intron topology is not significantly better than those of maximum-parsimony and maximum-likelihood trees of the six Araceae introns constrained to match the organismal tree (fig. 2*A*) for these six taxa. Thus, one cannot rule out the possibility that the introns have a history of entirely vertical transmission within the family, as is also suggested by mapping the intron's distribution on the two next-shortest trees of the Araceae (fig. 2*E* and *F*).

Coordinate Variation in Mutation Rates in Araceae cp and mt Genomes

As a bit of a digression from the central issue of this paper—the evolutionary history of this group I intron—some interesting observations about mutation rate variation arise from the intron phylogeny and other phylogenies. The terminal branch lengths leading to the two clearly vertically transmitted introns (see next section), from *Pistia* and *Arisaema,* are very unequal (fig. 4). The same relationship (*Pistia* faster than *Arisaema*) was also seen in analyses of *cox1* coding sequences (e.g., in parsimony analyses, the *Pistia* and *Arisaema* terminal branches were 13 steps long and 1 step long across all sites, respectively, and 7 and 0 steps long at silent sites) and is also evident in the divergence matrices in figure 3. According to the maximum-likelihood relative-rate

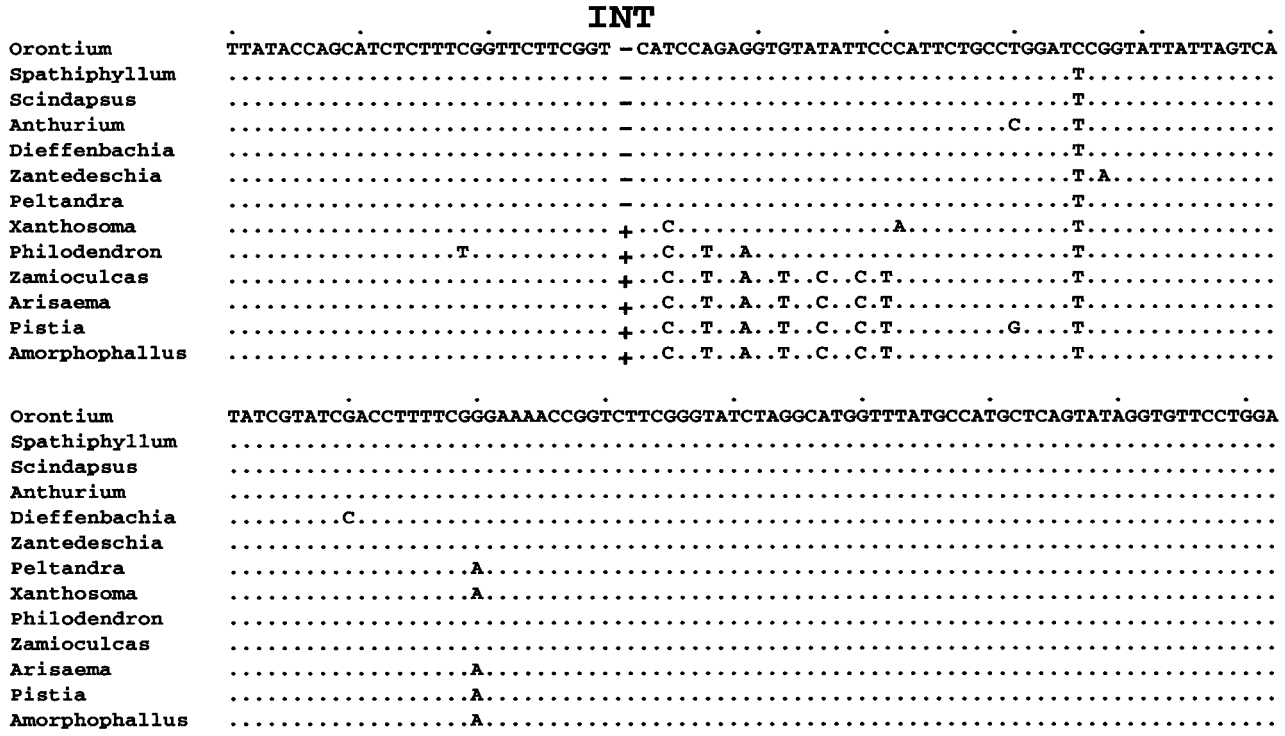

FIG. 5.—Co-conversion of exonic sequences immediately 3′ of the intron insertion site (marked ''INT''). Periods in the alignment indicate identity to the *Orontium* reference sequence; differences only are shown. ± signs at the ''INT'' column indicate intron presence/absence.

test of Muse and Gaut (1994), the mt substitution rate is significantly higher (95% confidence) across the entire *cox1* gene in *Pistia* than in *Arisaema.* Only recently have a few cases of variation in mtDNA mutation rate been reported within angiosperms, and these all involve much more divergent taxa than these two Araceae but no greater magnitude of rate variation than observed here (Eyre-Walker and Gaut 1997; Laroche et al. 1997).

Interestingly, the *Pistia* terminal branch length is several times longer than that of *Arisaema* in all chloroplast analyses too: 61 versus 20 steps in a maximum-parsimony analysis of the same *rbcL* and restriction site data analyzed in figure 2*A*, 32 versus 6 steps in a maximum-parsimony analysis of *rbcL* silent sites, and 8.0-fold higher in a maximum-likelihood analysis of all *rbcL* sites. Thus, it appears that both organellar genomes have several-fold higher mutation rates in *Pistia* than in its sister *Arisaema.* If the same effect is also observed for nuclear protein genes in these taxa, then some population genetic or life history factor, such as generation time differences, is probably driving a coordinate change in mutation rate across all three genomes in one plant lineage or the other (Eyre-Walker and Gaut 1997).

Evidence for Multiple Intron Horizontal Transfers from Co-conversion Tracts

We have presented three lines of evidence that are consistent with a history of few gains (as few as one) and largely vertical transmission of the *cox1* intron during evolution of the Araceae. These are as follows: (1) The six Araceae introns are extremely similar to one another—they are identically located within the *cox1* gene, are identical in length, and are 97.7%–99.6% identical in sequence (fig. 3). (2) A phylogeny of these intron sequences (fig. 4), while superficially incongruent with an organismal tree for these taxa (fig. 2*A*), is not significantly better than intron phylogenies which match the organismal tree. (3) The intron's phylogenetic distribution suggests a history of largely or entirely vertical transmission of the intron in the Araceae (figs. 2*B–F*). Indeed, both next-shortest trees reconstruct the intron's history as entirely vertical—a single ancestral gain followed by two subsequent losses. Despite all this, however, key evidence that leads us to favor an alternative model—of largely horizontal transmission of the *cox1* intron—comes from co-conversion data.

All six of the intron-containing Araceae show differences in a 20-nt exonic tract immediately downstream of the *cox1* intron that we attribute to changes incurred by co-conversion of the donor *cox1* gene during intron homing (fig. 5). These differences extend in a 3′ gradient away from the intron insertion site and occur identically at a given site relative to the clearly ancestral condition present in the seven sequenced intron-lacking Araceae. Not only do these seven Araceae all have the same sequence in this region, but this sequence is identical to that found in all of the many diverse intron-lacking angiosperms, both monocots and dicots (except for a single nucleotide difference in *Beta*; Vaughn et al. 1995; Cho et al. 1998*b*). Considering the very high overall identities of the *cox1* coding sequence among these 13 taxa (97.4%–99.8%; fig. 3; also see fig. 5), it is extremely unlikely that these differences could have arisen through so many individual point mutations in

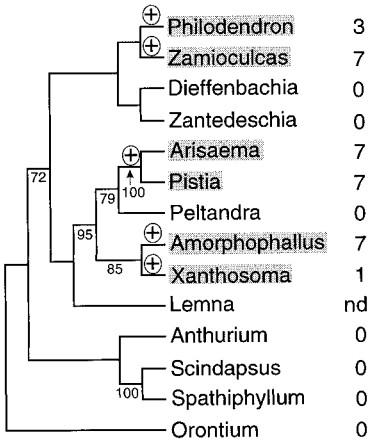

FIG. 6.—Our preferred scenario of *cox1* intron evolution. Tree topology is the same as in figure 2*A–D,* with bootstrap scores above 50% shown below branches. Ovaled plus signs indicate five inferred acquisitions of the *cox1* intron among the six Araceae taxa (shaded names) which contain it. Numbers at the right indicate the number of exonic nucleotides immediately downstream of the intron insertion site that are changed by co-conversion (see fig. 5).

such a short region. Instead, we attribute them to co-conversion.

The differences in the lengths of these co-conversion tracts, together with the absence of such tracts in all intron-lacking taxa, provide key evidence in favor of as many as five separate acquisitions of the intron within the family (fig. 6). We argue that, regardless of how closely related they are, any two taxa whose co-conversion tracts differ in length probably acquired their introns separately. For example, *Amorphophallus* and *Xanthosoma* are sister taxa with 85% bootstrap support, and thus are inferred to have received their introns by vertical transmission according to all parsimony models of intron distribution (fig. 2*B–E*). However, because their co-conversion tracts differ in length, and substantially so (fig. 5), we conclude that they most likely acquired their introns by two separate, and recent, horizontal transfers. By the same logic, we conclude that *Philodendron* and *Zamioculcas,* which cluster weakly in the shortest organismal tree (fig. 2*A*), also acquired their introns separately (fig. 6).

The absence of any evidence for co-conversion in all intron-lacking taxa argues that they never had the intron, for if they had, they should have retained the co-conversion signatures, as has been observed in an unequivocal case of loss of this intron in a lineage of dicots (Cho et al. 1998*a*). This inference is most meaningful for *Peltandra,* which is nested within a strongly supported clade of five taxa, the other four of which contain the intron, and which had been inferred to have lost the intron in all but one of the previously considered models for intron evolution (fig. 2*C–F*).

The *Arisaema*/*Pistia* pair is the only clade of inferred vertical transmission of the intron among the six intron-containing Araceae. *Arisaema* and *Pistia* form a strongly supported pair (100% bootstrap support) in the organismal tree (fig. 2*A*), their introns form a moderately well supported clade (69% and 77% bootstrap support, respectively; fig. 4), and they have identical co-conversion tract lengths (fig. 5). Overall, then, we postulate as many as five separate intron gains via horizontal transfer to account for the intron's presence in only 6 of the 14 examined Araceae taxa (fig. 6).

These co-conversion tract length differences can be made to fit more strictly vertical models of intron transmission if one postulates either high rates of back mutation within co-conversion tracts or persistent heteroplasmy of *cox1* alleles that differ in either presence/absence of the intron and/or the length of its co-conversion tract. However, we regard both phenomena as unlikely and therefore favor the five-intron-gain model of figure 6. Consider back mutation in the context of the three-gain model of figure 2*B,* as this postulates the fewest back mutations, four in *Philodendron* and six in *Xanthosoma,* each from an ancestral co-conversion tract with seven changes (as in *Zamioculcas* and *Amorphophallus,* respectively). (The optimizations in fig. 2*C–F* postulate an additional seven back mutations in each intronloss lineage). Outside of the 20-nt co-conversion region, there are only eight differences between *Philodendron* and *Zamioculcas* in the remaining 1,370 nt of *cox1* coding region sequenced, and only four differences between *Xanthosoma* and *Amorphophallus* (figs. 3 and 5). In other words, there are roughly equal numbers of differences (4 vs. 8 and 6 vs. 4) in these two regions despite their 68-fold disparity in length.

Such a pronounced and unidirectional (i.e., reverse) mutational bias could occur if there were intense selection for the ancestral sequence in the co-conversion region. However, such selection can be rejected because six of the seven co-conversion differences (all but the 3′-most) are silent and because the silent substitution rate in plant mtDNA is a close approximation of the neutral mutation rate. (There is only modest codon usage bias in plant mtDNA evolution [Lonsdale 1989; Macfarlane, Wahleithner, and Wolstenholme 1990], with this bias favoring, if anything, retention of the co-conversion sequence [about two thirds of third codon positions are T and A, with T > A > G > C]. Also, the overall point mutation rate across the entire mt genome [most of which is noncoding] is similar to the silent rate in protein genes, which implies that the silent rate cannot be constrained by other gene-specific pressures, e.g., by selection to maintain pairing between mRNAs and guide RNAs thought to be involved in the RNA editing that is common in plant mitochondria [Palmer and Herbon 1988; Wolfe 1996]). Furthermore, if back-mutation were to occur, it should abolish the 3′ co-conversion gradient shown in figure 5 (such abolition is observed in the only lineage of plant *cox1* introns in which back-mutation is evident [Cho et al. 1998*a*]; also note that this involves a special case of a plant mitochondrial lineage with an unusually high mutation rate and that despite this, there is much less back-mutation than must be postulated for the Araceae). Again, then, we reject the back-mutation model (of vertical intron transmission) to account for these co-conversion tract length differences.

Under a persistent heteroplasmy model of vertical intron transmission, one or a few early intron gains (e.g., in the common ancestor of the top 10 taxa in fig. 6 or at the positions marked by plus signs in the three-gain model of fig. 2*B*) would have been followed by retention of the original intronless *cox1* allele in essentially all descendant lineages, i.e., by persistent heteroplasmy of intron-plus and intron-minus alleles. Under the extreme form of this model, taxa such as *Peltandra* with secondary ''loss'' of the intron would have sorted out these two alleles, losing (or nearly losing) the intron-containing allele and fixing (or nearly fixing) the intronless allele, whereas sorting out would have occurred in the reverse direction in those taxa which still retain the intron. Differences in co-conversion tract lengths among these latter taxa would reflect vertical homing events in which an intron with the original tract length of seven differences had homed into an intronless allele in an event that generated a shorter co-conversion tract in the recipient allele, with the latter allele then rising to fixation (e.g., in *Philodendron* and *Xanthosoma*).

We find this scenario unlikely for three reasons. First, it postulates a kind of heteroplasmy that is far more persistent, i.e., throughout the evolution of a substantial portion of the large family Araceae, than any heteroplasmy that has yet been documented in plant mt genomes. Second, if anything, an endonuclease-encoding group I intron should be much less prone to persistent heteroplasmy than the ordinary mt sequence, as the pervasive, unidirectional spread of intron-plus into intron-minus alleles via vertical homing should quickly eliminate intronless alleles from a population. Indeed, in genetic crosses, homing group I introns quickly convert 95%–100% of progeny alleles to intron-containing ones (Lemieux and Lee 1987; Muscarella and Vogt 1989; reviewed in Dujon 1989). Third, if such heteroplasmy were so persistent, then one would expect it to have persisted in some extant plants, in which case we should have detected it. This is especially so in PCR reactions using primers flanking the intron, in which an intronless allele should have a large competitive advantage over an allele containing this nearly 1-kb intron. However, no intronless bands were observed in such reactions using any of the six intron-containing DNAs as templates (data not shown). Also, none of the three intronless taxa from the intron-containing ''clade'' shows any detectable hybridization with the *cox1* intron probe (fig. 1*B*). *Arisaema* and *Philodendron* do show faint hybridization to a second band with the coding region probe (fig. 1*A*); however, in the absence of recovery of an intronless PCR product from these taxa (see above), we interpret these faint signals as likely representing either duplications of very short segments of the *cox1* coding region within the mitochondrial genome (short dispersed repeats are common in plant mtDNAs [Itadani et al. 1994; Unseld et al. 1997; unpublished data]) or a copy of the gene in the nucleus (such transfers are also common [Blanchard and Schmidt 1995]).

For these reasons, we find it very unlikely that the entire 10-taxon clade of intron-containing *Araceae* acquired their introns by a single horizontal transfer. We believe that there were almost certainly at least three independent intron introductions in the common ancestors of each of the three sister pairs of intron-containing taxa (cf. figs. 2*B* and 6). We also prefer, based on differences in co-conversion tract lengths, the five-intron-gain model of figure 6, whereby *Philodendron* and *Zamioculcas* acquired their introns independently, as did *Amorphophallus* and *Xanthosoma*. At the same time, we cannot rule out the possibility that for either pair, the intron was acquired in its common ancestor, followed by a brief period of vertical homing during which all intron-lacking alleles became ''filled'' by the intron, with one of these homings happening to generate the shorter co-conversion tract now extant in the *Philodendron* or *Xanthosoma* lineages. Because homing is so pervasive, this critical secondary homing presumably would have occurred soon after intron acquisition by lateral transfer, leading to the prediction that the short and long co-conversion tract lineages (e.g., the *Philodendron* and *Zamioculcas* lineages, respectively) would trace back to the very base of their combined lineage. Conversely, if their introns were acquired separately, then we would predict that further sampling would lead to the discovery of intron- and co-conversion-lacking taxa that are specific relatives of *Philodendron* to the exclusion of *Zamioculcas* (and vice versa). Overall, then, the evidence leads us to favor at least three, most likely five (fig. 6), separate introductions of this invasive intron into Araceae *cox1* genes.

Why More Intron Gains than Losses?

Among the 14 Araceae examined in this study, we infer at least three, quite likely five, separate gains of the *cox1* intron, but no losses (fig. 6). This striking excess of gains over losses probably reflects at least one analytical and two biological factors. Our sampling of but one species from each of 14 diverse genera among the 100+ genera in the Araceae (French, Chung, and Hur 1995), coupled with the intron's recent and frequent invasion of angiosperm mt genomes (see next section), means that our study is biased toward detecting gains over losses. If we had instead sampled all 15 genera and hundreds of species of Araceae (French, Chung, and Hur 1995) from the *Pistia*/*Arisaema* clade of inferred vertical intron transmission, and only these species, then we might have detected an excess of intron losses compared with the putatively single ancestral gain in the group.

The two biological factors responsible for the gain-loss bias are the DNA endonuclease and RNA maturase activities that are known to be specified by the intron-encoded protein in yeast (Delahodde et al. 1989; Wenzlau et al. 1989; Henke, Butow, and Perlman 1995). The endonuclease activity is responsible for the intron's homing ability (see *Introduction*). As described in the preceding section, homing should cause intron-containing alleles of the *cox1* gene to spread and swamp out intron-lacking alleles. Consider a plant that is homoplasmic for the intron. The initial event that would lead to a descendant plant lacking the intron is loss of the intron in a single copy of the *cox1* gene, presumably via a gene conversion event between the gene and a cDNA

made from a spliced *cox1* mRNA. The fact that there are, on average, hundreds of mt genomes per cell, together with the propensity of plant mitochondria to fuse and share their genomes panmictically and to undergo frequent inter- and intramolecular recombination and gene conversion (Lonsdale 1987; Lonsdale et al. 1988), means that there will be a huge excess of intron-containing genes, each producing the homing endonuclease and each a candidate for donation of its intron back to the intronless gene. Thus, although intronless mutations should arise at the same frequency for *cox1* as for any other single-intron-containing mt gene, they are probably fixed at a lower frequency.

The fact that the intron ORF-encoded RNA maturase is necessary for self-splicing of the intron (Delahodde et al. 1989; Wenzlau et al. 1989; Henke, Butow, and Perlman 1995) presumably imposes very strong selection on maintenance of the ORF as an intact and expressed reading frame. Selection to maintain active maturase is probably much stronger than selection to maintain active endonuclease. Therefore, depending on the extent to which the two activities of this bifunctional protein are inseparable (Henke, Butow, and Perlman 1995), selection to maintain maturase may actually drive maintenance of endonuclease activity. This predicts that the intron should be lost preferentially under circumstances in which selection is lost on its RNA maturase function (e.g., if a second copy of the intron is present elsewhere in the genome to provide this function, or if some other *trans*-acting factor happens to supplant its role). Consistent with this prediction, the only case of loss of this intron yet observed in angiosperms occurs in the only genus (*Plantago*) of 35 intron-containing angiosperm genera examined in which the intron's ORF is a pseudogene (Cho et al. 1998*a*).

Recent Explosion of *cox1* Intron Transfers in Araceae and Other Angiosperms

In our recent study (Cho et al. 1998*b*), we inferred 32 separate gains of this intron among nearly 300 diverse angiosperms examined. Extrapolating across all angiosperms, we estimated that this intron had been horizontally introduced into angiosperm *cox1* genes ''over a thousand times'' and that ''all of these events seem to be recent.'' The present study reinforces both of these conclusions. Our inference of three to five separate acquisitions of this intron within a mere 14 genera from a single family gives us even greater confidence that many hundreds of gains of this intron among angiosperms remain to be identified. Furthermore, except for the shared gain in *Arisaema* and *Pistia,* all of the Araceae gains are potentially genus- or even species-specific, thus highlighting the recency of this intron's spread among angiosperms (see also Adams, Clements, and Vaughn 1998). We refer the reader to Cho et al. (1998*b*) for a discussion of several factors that could account for such a recent explosion of intron invasions, as well as two opposing models for the history of horizontal transfer of this intron.

We face a conundrum of why, if they truly arose by three to five separate events, the six Araceae *cox1* introns cluster (albeit weakly) as a monophyletic group in the intron phylogeny (fig. 4). Does this reflect some underlying biological factor? The Araceae are a cosmopolitan, mainly tropical, family, found in a variety of habitats; they thus lack the tight geographical and ecological clustering that might render them prone to infection with a similarly clustered set of donor or vectoring agents. Phylogenetic history may be the key here. All Araceae may share a genetic propensity to be infected by a particular vectoring agent which is transmitting *cox1* introns from another preferred group of hosts. Another possibility is that the intron was transmitted from one member of the Araceae to another across true mating barriers by cross-pollination unaccompanied by full fertilization and production of fertile hybrid offspring. This assumes a phylogenetic basis for such wide cross-pollination and also that this would be accompanied by sufficient pollen tube growth to occasionally—on an evolutionary timescale—allow introduction of this invasive intron into a fertilized egg. Or perhaps mere chance is at play here; perhaps these Araceae acquired their introns from very closely related donors merely by random chance, in which case further sampling of angiosperms should cause this apparent patterning to break down.

Acknowledgments

We thank Jim French for providing his cp restriction site data set of the Araceae for use in phylogenetic analyses, two anonymous reviewers for exceptionally helpful reviews, Yin-Long Qiu for providing Araceae DNAs and helpful discussion about Araceae phylogeny, Claude dePamphilis for providing the cox657R primer, and Keith Adams and Patrick Keeling for critical reading of the manuscript. This study was supported by NIH grant GM35087 to J.D.P.

LITERATURE CITED

ADAMS, K. L., M. J. CLEMENTS, and J. C. VAUGHN. 1998. The *Peperomia* mitochondrial *coxI* group I intron: timing of horizontal transfer and subsequent evolution of the intron. J. Mol. Evol. **46**:689–696.

BELFORT, M., and P. S. PERLMAN. 1995. Mechanisms of intron mobility. J. Biol. Chem. **270**:30237–30240.

BHATTACHARYA, D., T. FRIEDL, and S. DAMBERGER. 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.

BINISZKIEWICZ, D., E. CESNAVICIENE, and D. A. SHUB. 1994. Self-splicing group I intron in cyanobacterial initiator methionine tRNA: evidence for lateral transfer of introns in bacteria. EMBO J. **13**:4629–4635.

BLANCHARD, J. L., and G. W. SCHMIDT. 1995. Pervasive migration of organellar DNA to the nucleus in plants. J. Mol. Evol. **41**:397–406.

BONITZ, S. G., G. CORUZZI, B. E. THALENFELD, A. TZAGOLOFF, and G. MACINO. 1980. Assembly of the mitochondrial membrane system. Structure and nucleotide sequence of the gene coding for subunit I of yeast cytochrome oxidase. J. Biol. Chem. **255**:11927–11941.

CHASE, M. W., D. E. SOLTIS, R. G. OLMSTEAD et al. (35 coauthors). 1993. Phylogenetics of seed plants: an analysis of

nucleotide sequences from the plastid gene *rbcL*. Ann. Mo. Bot. Gard. **80**:528–580.

CHO, Y., K. L. ADAMS, Y.-L. QIU, P. KUHLMAN, J. C. VAUGHN, and J. D. PALMER. 1998*a*. A highly invasive group I intron in the mitochondrial *cox1* gene. Pp. 19–23 *in* I.-M. MOLLER, P. GARDESTROM, K. GLIMELIUS, and E. GLASER, eds. Plant mitochondria: from gene to function. Backhuys Publishers, Leiden, The Netherlands.

CHO, Y., Y.-L. QIU, P. KUHLMAN, and J. D. PALMER. 1998*b*. Explosive invasion of plant mitochondria by a group I intron. Proc. Natl. Acad. Sci. USA **95**:14244–14249.

DELAHODDE, A., V. GOGUEL, A. M. BECAM, F. CREUSOT, J. PEREA, J. BANROQUES, and C. JACQ. 1989. Site-specific DNA endonuclease and RNA maturase activities of two homologous intron-encoded proteins from yeast mitochondria. Cell **56**:431–441.

DOYLE, J. J., and J. S. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. **19**:11–15.

DUJON, B. 1989. Group I intron as mobile genetic elements: facts and mechanistic speculations—a review. Gene **82**:91–114.

DUVALL, M. R., M. T. CLEGG, M. W. CHASE et al. (11 coauthors). 1993. Phylogenetic hypotheses from the monocotyledons constructed from *rbcL* sequence data. Ann. Mo. Bot. Gard. **80**:607–619.

EYRE-WALKER, A., and B. S. GAUT. 1997. Correlated rates of synonymous site evolution across plant genomes. Mol. Biol. Evol. **14**:445–460.

FRENCH, J. C., M. G. CHUNG, and Y. K. HUR. 1995. Chloroplast DNA phylogeny of the Ariflorae. Pp. 255–275 *in* P. J. RUDALL, P. J. CRIBBE, D. F. CUTLE, and C. J. HUMPHRIES, eds. Monocotyledons: systematics and evolution. Royal Botanic Gardens, Kew, England.

GENETICS COMPUTER GROUP. 1991. Wisconsin sequence analysis package. Version 6. Genetics Computer Group, Madison, Wis.

HENKE, R. M., R. A. BUTOW, and P. S. PERLMAN. 1995. Maturase and endonuclease functions depend on separate conserved domains of the bifunctional protein encoded by the group I intron aI4α of yeast mitochondrial DNA. EMBO J. **14**:5094–5099.

HIBBETT, D. S. 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.

ITADANI, H., T. WAKASUGI, M. SUGITA, M. SUGIURA, M. NAKAZONO, and A. HIRAI. 1994. Nucleotide sequence of a 28-kbp portion of rice mitochondrial DNA: the existence of many sequences that correspond to parts of mitochondrial genes in intergenic regions. Plant Cell Physiol. **35**:1239–1244.

JACQUIER, A., and B. DUJON. 1985. An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell **41**:383–394.

KISHINO, H., and M. HASEGAWA. 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.

LAMBOWITZ, A. M., and M. BELFORT. 1993. Introns as mobile genetic elements. Annu. Rev. Biochem. **62**:587–622.

LANG, B. F. 1984. 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.

LAROCHE, J., P. LI, L. MAGGIA, and J. BOUSQUET. 1997. Molecular evolution of angiosperm mitochondrial introns and exons. Proc. Natl. Acad. Sci. USA **94**:5722–5727.

LEMIEUX, C., and R. W. LEE. 1987. Nonreciprocal recombination between alleles of the chloroplast 23S rRNA gene in interspecific *Chlamydomonas* crosses. Proc. Natl. Acad. Sci. USA **84**:4166–4170.

LONSDALE, D. M. 1987. The molecular biology and genetic manipulation of the cytoplasm of higher plants. Pp. 47–102 *in* P. W. J. RIGBY, ed. Genetic engineering. Vol. 6. Academic Press, London.

———. 1989. The plant mitochondrial genome. Pp. 229–295 *in* A. MARCUS, ed. The biochemistry of plants, Vol. 15. Molecular biology. Academic Press, San Diego.

LONSDALE, D. M., T. BREARS, T. P. HODGE, S. E. MELVILLE, and W. H. ROTTMANN. 1988. The plant mitochondrial genome: homologous recombination as a mechanism for generating heterogeneity. Philos. Trans. R. Soc. Lond. B Biol. Sci. **319**:149–163.

MACFARLANE, J. L., J. A. WAHLEITHNER, and D. R. WOLSTENHOLME. 1990. A broad bean mitochondrial *atp6* gene with an unusually simple, non-conserved 5′ region. Curr. Genet. **18**:87–91.

MADDISON, W. P., and D. R. MADDISON. 1992. MacClade: analysis of phylogeny and character evolution. Version 3.0. Sinauer, Sunderland, Mass.

MICHEL, F., and B. DUJON. 1986. Genetic exchanges between bacteriophage T4 and filamentous fungi? Cell **46**:323.

MUELLER, J. E., D. SMITH, and M. BELFORT. 1996. Exon coconversion biases accompanying intron homing: battle of the nucleases. Genes Dev. **10**:2158–2166.

MUSCARELLA, D. E., and V. M. VOGT. 1989. A mobile group I intron in the nuclear rDNA of *Physarum polycephalum*. Cell **56**:443–454.

MUSE, S. V., and B. S. GAUT. 1994. A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates with application to the chloroplast genome. Mol. Biol. Evol. **11**:715–724.

NISHIDA, H., Y. TAJIRI, and J. SUGIYAMA. 1998. Multiple origins of fungal group I introns located in the same position of nuclear SSU rRNA gene. J. Mol. Evol. **46**:442–448.

PALMER, J. D., and L. A. HERBON. 1988. Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. J. Mol. Evol. **28**:87–97.

PERLMAN, P. S., and R. BUTOW. 1989. Mobile introns and intron-encoded proteins. Science **246**:1106–1109.

QUIRK, S. M., D. BELL-PEDERSEN, and M. BELFORT. 1989. Intron mobility in the T-even phages: high frequency inheritance of group I introns promoted by intron open reading frames. Cell **56**:455–465.

STOCKEY, R. A., G. L. HOFFMAN, and G. W. ROTHWELL. 1997. The fossil monocot *Limnobiophyllum scutatum*: resolving the phylogeny of Lemnaceae. Am. J. Bot. **84**:355–368.

SWOFFORD, D. L. 1998. PAUP*. Version 4.0. Sinauer, Sunderland, Mass.

SZOSTAK, J. W., T. L. ORR-WEAVER, R. J. ROTHSTEIN, and F. W. STAHL. 1983. The double strand break repair model for recombination. Cell **33**:25–35.

TURMEL, M., V. COTE, C. OTIS, J.-P. MERCIER, M. W. GRAY, K. M. LONERGAN, and C. LEMIEUX. 1995. Evolutionary transfer of ORF-containing group I introns between different subcellular compartments (chloroplast and mitochondrion). Mol. Biol. Evol. **12**:533–545.

TURMEL, M., J.-P. MERCIER, and M.-J. COTE. 1993. Group I introns interrupt the chloroplast *psbA* and *psbC* and the mitochondrial *rrnL* gene in *Chlamydomonas*. Nucleic Acids Res. **21**:5242–5250.

UNSELD, M., J. R. MARIENFELD, P. BRANDT, and A. BRENNICKE. 1997. The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. Nat. Genet. **15**:57–61.

VAUGHN, J. C., M. T. MASON, G. L. SPER-WHITIS, P. KUHLMAN, and J. D. PALMER. 1995. Fungal origin by horizontal transfer of a plant mitochondrial group I intron in the chimeric *coxI* gene of *Peperomia.* J. Mol. Evol. **41**:563–572.

WATANABE, K. I., M. EHARA, Y. INAGAKI, and T. OHAMA. 1998. Distinctive origins of group I introns found in the COXI genes of three green algae. Gene **213**:1–7.

WENZLAU, J. M., R. J. SALDANHA, R. A. BUTOW, and P. S. PERLMAN. 1989. A latent intron-encoded maturase is also an endonuclease needed for intron mobility. Cell **56**:421–430.

WOLFE, K. H. 1996. Molecular evolution of plants: more genomes, fewer generalities. Pp. 45–57 *in* B. ANDERSSON, A. H. SALTER, and J. BARBER, eds. Molecular genetics of photosynthesis. IRL Press, Oxford.

ZINN, A. R., and R. A. BUTOW. 1985. Nonreciprocal exchange between alleles of the yeast mitochondrial 21S rRNA gene; kinetics and the involvement of a double-strand break. Cell **40**:887–895.

THOMAS EICKBUSH, reviewing editor

Accepted May 25, 1999