Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 14244–14249, November 1998
Evolution

Explosive invasion of plant mitochondria by a group I intron

YANGRAE CHO, YIN-LONG QIU*, PETER KUHLMAN†, AND JEFFREY D. PALMER‡

Department of Biology, Indiana University, Bloomington, IN 47405

Communicated by Susan R. Wessler, University of Georgia, Athens, GA, September 24, 1998 (received for review July 28, 1998)

ABSTRACT **Group I introns are mobile, self-splicing genetic elements found principally in organellar genomes and nuclear rRNA genes. The only group I intron known from mitochondrial genomes of vascular plants is located in the *cox1* gene of *Peperomia*, where it is thought to have been recently acquired by lateral transfer from a fungal donor. Southern-blot surveys of 335 diverse genera of land plants now show that this intron is in fact widespread among angiosperm *cox1* genes, but with an exceptionally patchy phylogenetic distribution. Four lines of evidence—the intron's highly disjunct distribution, many incongruencies between intron and organismal phylogenies, and two sources of evidence from exonic coconversion tracts—lead us to conclude that the 48 angiosperm genera found to contain this *cox1* intron acquired it by 32 separate horizontal transfer events. Extrapolating to the over 13,500 genera of angiosperms, we estimate that this intron has invaded *cox1* genes by cross-species horizontal transfer over 1,000 times during angiosperm evolution. This massive wave of lateral transfers is of entirely recent occurrence, perhaps triggered by some key shift in the intron's invasiveness within angiosperms.**

Many group I introns encode site-specific endonucleases that catalyze their efficient spread from intron-containing to intronless alleles of the same gene in genetic crosses (1–3). This process, termed intron "homing," has been observed for introns located in a variety of mitochondrial (mt) and chloroplast genes (4–7), in nuclear rRNA genes of the slime mold *Physarum* (8), and in protein genes of T-even phage (9). Homing is initiated by the intron-encoded endonuclease, which makes a staggered double-strand break at its target site within a recipient intronless allele, and is thought to then proceed by the double-strand-break repair pathway (10).

The evolutionary importance of intron homing to the spread of group I introns across species barriers has been unclear, as relatively few cases of the horizontal transfer of group I introns between identical genomic sites of nonmating organisms are documented (11–17). Most of these cases involve the same genome and species belonging to the same phylum, usually fungi (11–13). Two notable exceptions are the transfer of two group I introns between identical sites of rRNA genes located in the chloroplast of a *Chlamydomonas*-type green alga and the mitochondrion of an *Acanthamoeba*-like ameboid (15).

The only group I intron known from vascular plant mt genomes (which contain many group II introns) is also thought to have been acquired by homing-mediated horizontal transfer from a distantly related organism. This intron is present in the *cox1* (cytochrome oxidase subunit 1) gene of the angiosperm *Peperomia* (16, 17) at the same location as related introns in the nonvascular plant *Marchantia*, the green alga *Prototheca*, the slime mold *Dictyostelium*, and several diverse fungi (see ref. 18 and references therein). This *cox1* intron is thought to have been recently acquired by *Peperomia*, most likely from a fungal donor, based on (*i*) its singular presence in *Peperomia* among 25 genera of vascular plants examined, (*ii*) its closer phylogenetic relationship to fungal introns than to those of the green "plants" *Marchantia* and *Prototheca*, and (*iii*) the presence of exonic signatures of homing-mediated coconversion immediately downstream of the *Peperomia* intron (16, 17).

We now show that *Peperomia* is only the tip of a large iceberg: there has been an explosive and recent wave of horizontal transfers of this intron into *cox1* genes of many different lineages of flowering plants. We surveyed over 300 diverse land plants and infer that, based on phylogenetic and molecular criteria, 32 separate transfers account for the intron's presence in 48 disparate genera of angiosperms. From this sampling, we estimate that the intron has been separately acquired over 1,000 times during angiosperm evolution.

MATERIALS AND METHODS

Latin names and voucher information for the 341 species of land plants examined in this study are available at http://www.bio.indiana.edu/~palmerlab. Total cellular DNA was extracted by using a modified cetyltrimethylammonium bromide procedure (19) and further purified by banding in a CsCl/ethidium bromide gradient. Southern transfers used Immobilon nylon membranes (Millipore). Probes were prepared by random-priming using ^{32}P. Hybridizations were carried out at 60°C for 18 hr in 5× SSC, 50 mM Tris (pH 8.0), 0.1% SDS, 10 mM EDTA, and 2× Denhardt's solution. Filters were twice washed for 30 min at 60°C in 2× SSC/0.1% SDS.

All but the first 165 bp and the last 77 bp of the 1,590-bp *cox1* coding sequence and the entirety of the gene's single, 953–1,008-bp intron were amplified from intron-containing taxa by using three pairs of primers: cox42F (GGATCTTCTCCACTAACCACAAA) and cox657R (GCGGGATCAGAAAAGGTTGTA), IP53 (GGAGGAGTTGATTTAGC) and IP56 (GAGCAATGTCTAGCCC), and INT1.2KF (AGCATGGCTAGCTTTCCTAGA) and *cox1*.6KR (AAGGCTGGAGGGCTTTGTAC). These primers amplified a ≈600-bp region of the 5′ exon, a ≈1,650-bp region containing the entire intron and flanking exonic sequences, and a ≈950-bp region containing part of the intron and part of the 3′ exon, respectively. For intron-lacking species, primer pairs cox42F/cox657R (≈600-bp product) and IP53/*cox1*.6KR (≈1,000-bp product) were used to amplify the same aggregate length of coding region as above. Annealing reactions were performed at 50–55°C by using 20–50 ng of total cellular DNA in a 10-μl reaction with 1 mM MgCl2 and 5% acetamide for 40 cycles

Abbreviations: mt, mitochondrial; *cox1*, cytochrome oxidase subunit 1.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AJ223411–AJ223439).

A Commentary on this article begins on page 14003.

*Present address: Institute of Systematic Botany, University of Zurich, Zollikerstrasse 107, 8008 Zurich, Switzerland.
†Present address: Department of Chemistry and Biochemistry, Denison University, Granville, OH 43023.
‡To whom reprint requests should be addressed. e-mail: jpalmer@bio.indiana.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

with a 1-min extension time. Products were purified from agarose gels and cloned by using a TA cloning kit (Invitrogen). Nucleotide sequences were determined for both strands of (usually) a single clone of each species by using LiCor automated DNA sequencers.

Nucleotide sequences were initially aligned by using the program PILEUP (Genetics Computer Group, Madison, WI); alignments were then adjusted by eye and are available on request from J.D.P. Gaps were excluded from all phylogenetic analyses, as was the 3′ exonic coconversion region. The global intron phylogenetic analyses were carried out by using PAUP*d56 (from D. L. Swofford, Smithsonian Institution, Washington, DC) on an 1,153-character alignment of the *cox1* and related introns. Maximum-likelihood analysis used the HKY85 model with empirical base frequencies and an empirical transition/transversion ratio of 0.46. Seven random-addition heuristic searches yielded nine trees of equally low log likelihood, one of which is shown (these trees differ only within angiosperms). Bootstrapping involved 100 replicates, each with 1 random addition sequence. Parsimony analysis used all characters unordered and unweighted, steepest descent, tree bisection and resection, and 200 bootstrap replicates, each of one heuristic search with random taxon addition. Neighbor-joining analysis used Kimura two-parameter distances and 100 bootstrap replicates.

Angiosperm intron and organismal maximum-likelihood analyses were performed by using the F84 model in PHYLIP version 3.5 (from J. Felsenstein, University of Washington, Seattle) and FASTDNAML version 1.06 (20). Four different transition/transversion ratios (1.0, 1.5, 2.0, 2.5), empirical base frequencies, and two addition sequences under global swapping conditions were used during preliminary analyses. The ratio that produced the lowest log-likelihood tree for each data set was selected for further analyses by using multiple randomized addition sequences and global swapping of up to 28 branches at each step. Bootstrapping was performed with FASTDNAML using 100 random data sets, generated by SEQBOOT using the same swapping and sequence-addition conditions as described above.

RESULTS

Intron Distribution. Of 25 genera of vascular plants previously examined (16, 17), this intron was known to be present in the mt *cox1* gene only in *Peperomia*. We were therefore surprised to encounter, in a comparative sequencing study of mutation-rate variation in plant mtDNA, an intron of highly similar length (966 vs. 953 bp) and sequence (92% identity) located at the same position within *cox1* in the distantly related angiosperm *Veronica*. The highly disjunct distribution of these two introns suggested that they might have arisen by separate insertions and caused us to ask how frequently and how recently this intron had been acquired during plant evolution.

Taking advantage of the generally very low mutation rate in plant mtDNAs (21, 22), we used a Southern-hybridization approach to rapidly survey the intron's distribution among 335 diverse genera of land plants. As a control, a probe for the *cox1* coding region hybridized to all of the DNAs tested (e.g., Fig. 1). In contrast, a probe for the *cox1* intron showed a very patchy pattern of hybridization, hybridizing to 49 largely phylogenetically disjunct DNAs, all from angiosperms (Figs. 1 and 2). In all but one of the 49 positive cases, the intron hybridized with proportionally equivalent strength as the *cox1* exon and to at least one band in common with it (Fig. 1, *Top* and *Middle*). We therefore conclude that these 48 DNAs probably contain an intron in their (presumptively mt) *cox1* genes that is closely related to the *Veronica* probe intron. In the 49th case, the intron hybridized weakly and to a different band than the exon probe; this probably represents a small and/or divergent region of intron homology located in some non-*cox1* region in this plant (*Brasenia*; Fig. 1).

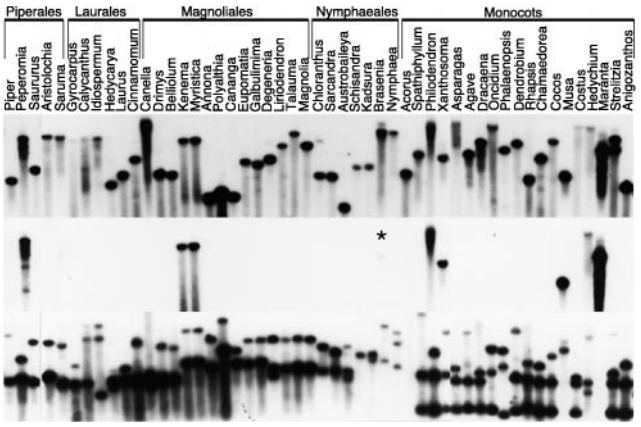

FIG. 1. Southern-blot hybridizations showing presence or absence of two mt introns among 51 of the 341 land plants examined in this study. *Bam*HI-cut DNAs were arranged according to presumptive phylogenetic relationship and hybridized with probes internal to the *cox1* coding sequence from *Beta vulgaris* (*Top*), the single *cox1* group I intron from *Veronica ugrestis* (*Middle*), and the single *cox2* group II intron from *Zea mays* (*Bottom*). * indicates weak, non-*cox1* hybridization in *Brasenia schreberi* (see text).

The extremely patchy distribution of the *cox1* intron (Figs. 1 and 2) contrasts with the nearly universal hybridization of a group II intron from the *cox2* gene (Fig. 1, *Bottom*) and each of 10 other mt group II introns examined (ref. 23; Y.-L.Q. and J.D.P., unpublished results). These data strongly indicate that these latter 11 introns were ancestrally present in the mt genome of, minimally, all angiosperms, with occasional intron losses (e.g., two blank lanes in Fig. 1, *Bottom*). In contrast, the unusually sporadic distribution of the *cox1* intron suggests that lateral transfer has played a major role in its evolution. Indeed, an all-gain model (Fig. 2) to account for the intron's distribution would postulate 30 independent acquisitions among angiosperms, which barely exceeds the number of events in the most parsimonious models (e.g., 22 gains and 7 losses; not shown in Fig. 2) and is much lower than the 85 events under a model of ancestral presence followed by frequent loss (also not shown).

Discordant Intron and Organismal Phylogenies. To assess the relative contributions of horizontal and vertical genetic transmission to the intron's phylogenetic history, we sequenced the *cox1* intron and coding region from 29 of the 48 hybridizing angiosperms, and compared phylogenies of the intron with those of the organisms in which it resides. These 29 introns, plus the *Peperomia* intron (16), are highly similar in length (953–1,008 bp) and sequence (≥92% identity) and are located at the same position within the *cox1* gene. All 30 introns contain a ≈270-bp core region typical of group I introns (1–3) interrupted by and partially overlapping with a ≈834-bp ORF. The inferred protein from this ORF is about 52% identical over 229 residues with the yeast *cox1* aI4 intronic protein, which encodes site-specific DNA endonuclease and RNA maturase activities (4, 5, 24).

The similarity of the angiosperm *cox1* introns is vividly illustrated in the global phylogeny of Fig. 3*A*. Twenty representative angiosperm introns cluster tightly relative to the long branches separating this intron clade and all other members of this intron family from each other. As in the initial study of the *Peperomia* intron (16), the angiosperm intron clade is more closely related to a group of fungal mt introns than to the identically positioned *cox1* intron of the nonvascular plant *Marchantia*. This suggests that angiosperm *cox1* genes acquired their introns either in a single transfer from a fungal source or

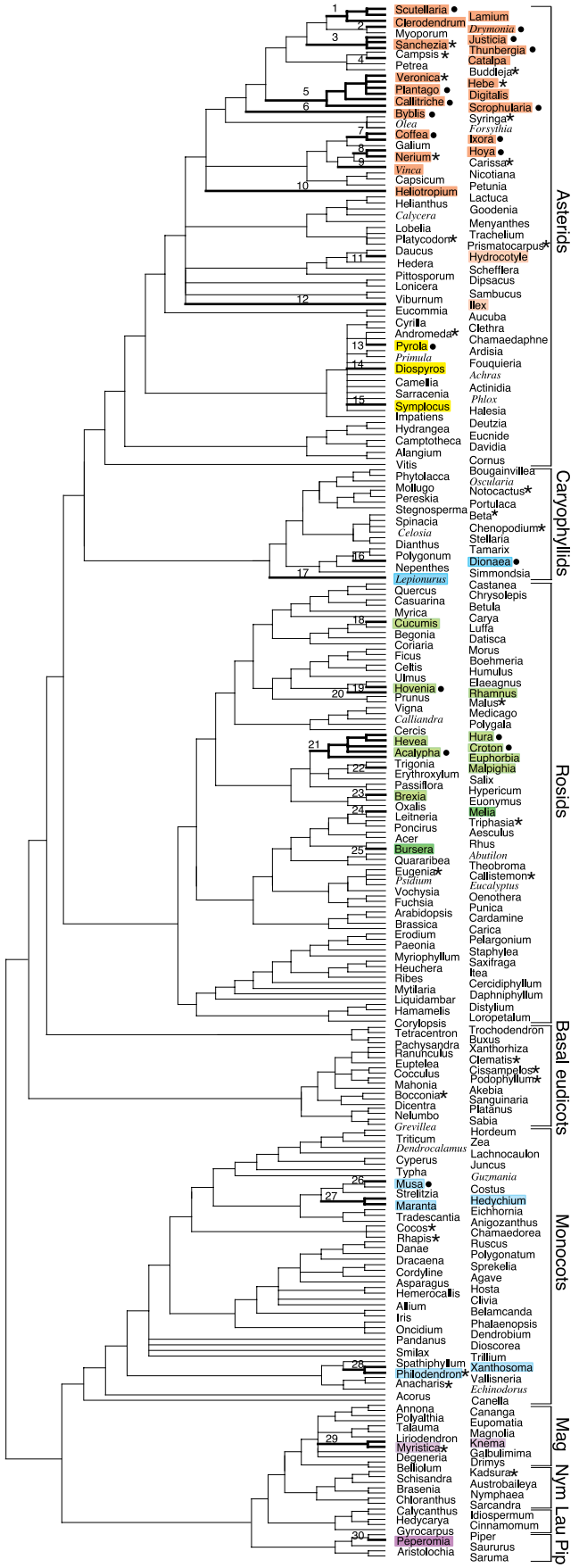

FIG. 2. Sporadic distribution of the *cox1* intron among 281 examined species of angiosperms. The cladogram is rooted on gymnosperms and is from a parsimony analysis (Y.-L.Q., unpublished results) of a 1,428-bp region of the chloroplast rbcL gene. The 48 taxa that hybridized strongly to the *cox1* intron probe are shaded in color; the colors designate nine major groups of angiosperms (cf. Fig. 3*C*). Heavy branches mark the 30 monophyletic intron-containing clades (numbered 1–30) under an all-gain/no-loss model of intron evolution. ● marks the 18 intron-hybridizing taxa whose *cox1* genes were not sequenced (cf. Fig. 3). Names of the 19 taxa for which phylogenetic substitutes were used in the rbcL analyses are italicized (cf. Fig. 3*C*). ∗ marks the 25 taxa for which rbcL sequences are not available and which were positioned based on other evidence. Mag, Magnoliales; Nym, Nymphaeales; Lau, Laurales; Pip, Piperales.

in multiple transfers from a group of closely related fungal donors.

Phylogenies for the 30 angiosperms whose *cox1* introns have been sequenced were constructed separately from the coding sequences of the mt *cox1* gene and the chloroplast rbcL gene. The *cox1* tree, although less well resolved than the rbcL tree owing to the generally very low rate of plant mt substitutions (21, 22), is nonetheless highly congruent with the rbcL tree [data not shown; the two trees showed 89% congruency in a formal compatibility test (25)]. This congruence indicates that there is no reason to suspect any lateral exchange of *cox1* genes during angiosperm evolution and justifies combining the *cox1* and rbcL coding sequences to produce a bigenomic estimate of "organismal" phylogeny for these 30 angiosperms. This organismal tree resolves the taxa into nine groups, which are shown color-coded and named in Fig. 3*C* and which mirror those recovered in an analysis of some 500 rbcL sequences (26).

The interspersion of colors in the *cox1* intron phylogeny (Fig. 3*B*) illustrates substantial incongruence with the organismal phylogeny (Fig. 3*C*) and indicates extensive lateral transfer during the evolution of the *cox1* intron. There are but four small groups that are identically recovered, or nearly so, in both trees (bold branches in Fig. 3*B*), indicative of vertical transmission of the intron in these lineages. All other intron clades are, with varying levels of bootstrap support, composed of distantly related taxa. To point out just a few examples, note the three pairs of taxa (*Ilex*/*Hydrocotyle*, *Symplocus*/*Diospyros*, and *Maranta*/*Hedychium*) that each receive 100% bootstrap support in the organismal tree (Fig. 3*C*), but whose members are separated by multiple robust nodes in the intron tree (Fig. 3*B*). The log likelihood of an intron tree constrained to match the topology of the organismal tree (Fig. 3*C*) is −4,042 (compared with −3568 for Fig. 3*B*), and the KH test (27) rejects these two data sets producing the same topology with $P < 0.0001$.

Coconversion-Tract Evidence for Multiple Intron Gains. Additional evidence, of two kinds, for many separate acquisitions of this intron comes from analysis of an exonic coconversion tract (Fig. 4). Group I intron homing is known in genetic crosses to lead to coconversion of recipient exonic sequences flanking the acquired intron by donor exonic sequences (1–3). An 18-bp region 3′ to the intron is virtually unchanged in the 24 diverse intronless vascular plants whose sequences are shown in Fig. 4*A*, whereas 29 of the 30 intron-containing angiosperms show one or more variations in this region and 28 show three or more variations (Fig. 4*B*). Moreover, the variations all are identical at a given site and extend in a 3′ gradient away from the intron insertion site. It thus appears that a short 3′ tract of at least 3–18 bp has been coconverted in all but one of the intron-containing plants. Because the mutation rate in plant mtDNA is generally extremely low (ref. 21 and 22; Fig. 4*A*), because there is no apparent selective pressure for back-mutation (all six sites changed by coconversion are silent sites), and because there is no evidence for back-mutation (which would abolish the 3′ coconversion gradient seen), the incidence of back-mutation at sites changed by coconversion must be very low. We therefore conclude that taxa such as *Xanthosoma* and *Philodendron*,

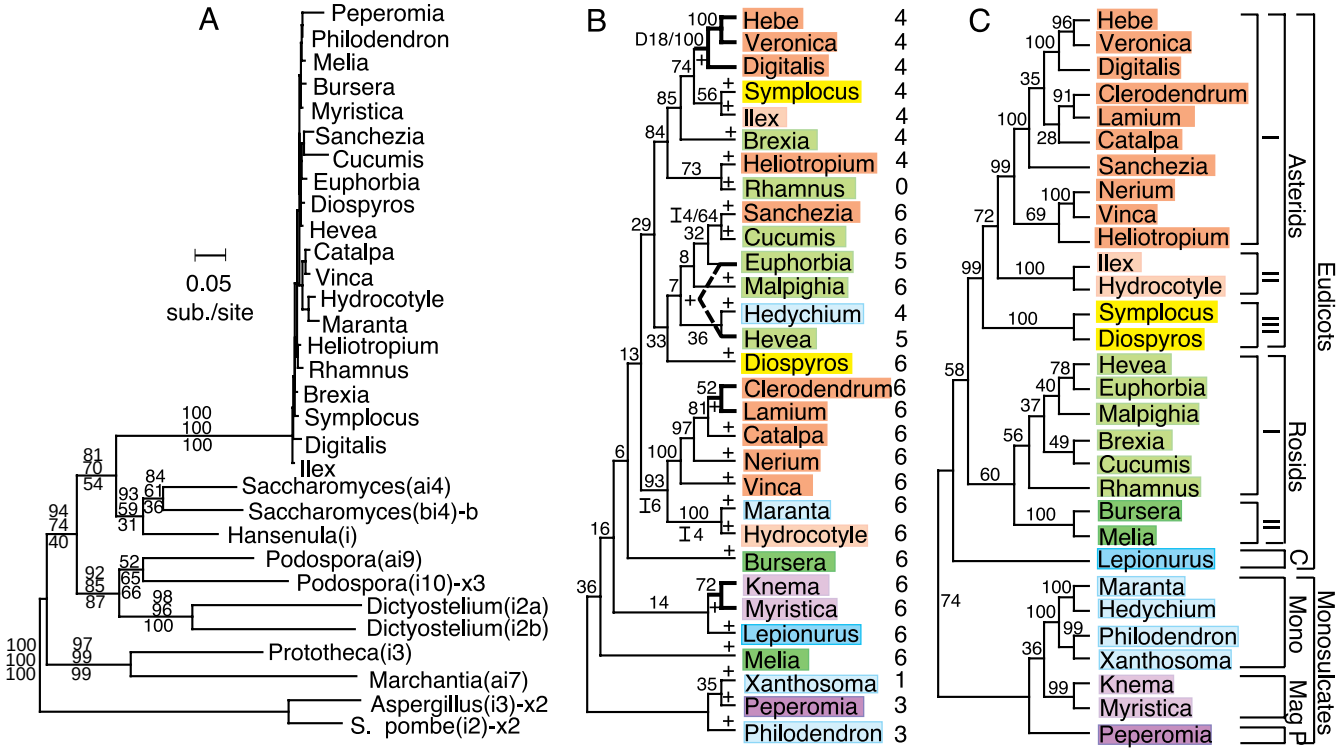

FIG. 3. Phylogenetic history of the *cox1* intron. (*A*) Maximum-likelihood tree of the *cox1* and all related introns. Bootstrap values for each node are shown in a column above and below the corresponding branch and are, from top to bottom, from neighbor-joining, parsimony, and likelihood analyses, respectively. Introns located at the same *cox1* position as in this study are unmarked; introns at two other positions in *cox1* are marked "x2" or "x3"; and an intron in the cob gene is marked with a "b." (*B*) Maximum-likelihood tree of 30 angiosperm *cox1* introns. Numbers on the tree are bootstrap values. The four synapomorphic intron gaps (which were not used to build this tree) are marked by "I" or "D" (for insertion or deletion relative to the *Peperomia* intron) followed by the gap's length in bp. + signs on the tree mark 25 inferred gains of the intron among these 30 taxa. Color-coding is as in Fig. 3*C*. Numbers at right indicate number of 3′-flanking nucleotides changed by coconversion (see Fig. 4*B*). Bold branches mark four small clades of introns thought to have originated from the same intron gain event. (*C*) Organismal tree from a maximum-likelihood analysis of a combined data set of chloroplast rbcL and mt *cox1* coding sequences, excluding the coconversion region (see Fig. 4). Numbers are bootstrap values. Color-coding is as described in the text. C, caryophyllids; Mono, monocots; Mag, Magnoliales; P, Piperales.

whose coconversion tracts differ in length, most likely acquired their introns separately, despite the fact that their relationships in the intron (Fig. 3*B*) and organismal (Fig. 3*C*) trees are not significantly incongruent. By the same logic, the four different coconversion tract lengths observed among the six Rosidae I taxa imply at least four separate acquisition events within this group (Figs. 3*B* and 4*B*).

Inspection of flanking exonic sequences in intronless taxa closely related to intron-containing plants reveals evidence for still additional cases of separate intron gain. Consider *Clerodendrum*, *Lamium*, and *Catalpa*, whose intron phylogeny exactly mirrors their organismal phylogeny (Fig. 3 *B* and *C*) and which have indistinguishable coconversion tracts (Fig. 4*B*). Nonetheless, two separate intron gains are marked for this group in Fig. 3*B* rather than a single gain. This is because *Petrea* and *Campsis*, which are specific within-family relatives of *Clerodendrum*/*Lamium* (Verbenaceae/Lamiaceae) and *Catalpa* (Bignoniaceae), respectively, lack the intron and also show no coconversion evidence (Fig. 4*A*) of ever having had it [importantly, intronless plant *cox1* genes that once possessed this intron do retain their coconversion tracts (29)]. By the same logic, we conclude that the introns in *Bursera* and *Melia*, whose phylogenetic separation in Fig. 3*B* lacks significant statistical support and which have identical coconversion tracts, were acquired separately. This is because the six Rosid II taxa between *Bursera* and *Melia* in Fig. 2 all lack the intron and any coconversion signatures. In total, intronless taxa have been sequenced from 10 different families known to include intron-containing members (nine from this study and one from ref. 17), and in all cases the intronless taxa show no evidence of coconversion. This again implies that they probably never possessed the intron and that in all cases intron acquisition occurred recently, within the family's evolution.

DISCUSSION

We infer fully 25 separate intron gains to account for the presence of this intron among the 30 angiosperms whose *cox1* introns have been sequenced. This inference rests on four lines of evidence: (*i*) many incongruencies between intron and organismal phylogenies (Fig. 3 *B* and *C*), (*ii*) the highly disjunct distribution of intron-containing plants in the rbcL phylogeny shown in Fig. 2, (*iii*) different lengths of coconversion among otherwise related introns (Figs. 3*B* and 4), and (*iv*) the existence of ancestrally intron-lacking taxa within families containing the intron. Furthermore, by criterion *ii*, we infer 7 additional gains among the 18 intron-containing taxa whose introns were not sequenced (Fig. 2). Remarkably, this total of 32 inferred intron gains actually exceeds the 30 gains postulated under an all-gain model based solely on the intron's distribution across the angiosperm phylogeny of Fig. 2. This discrepancy reflects the two pairs of intron-containing sister taxa in Fig. 2 which by either incongruence (*Maranta* and *Hedychium*; Fig. 3) or coconversion (*Philodendron* and *Xanthosoma*; Fig. 4) evidence acquired their introns separately.

Extrapolating from these ≈32 separate cases of inferred intron acquisition among the 278 genera and 281 species of angiosperms examined by Southern blots in this study, we estimate that this intron has invaded the *cox1* gene over 1,000 times among the >13,500 genera and >300,000 species of extant angiosperms. Moreover, all of these events seem to be recent; many, possibly all, of the characterized gains have

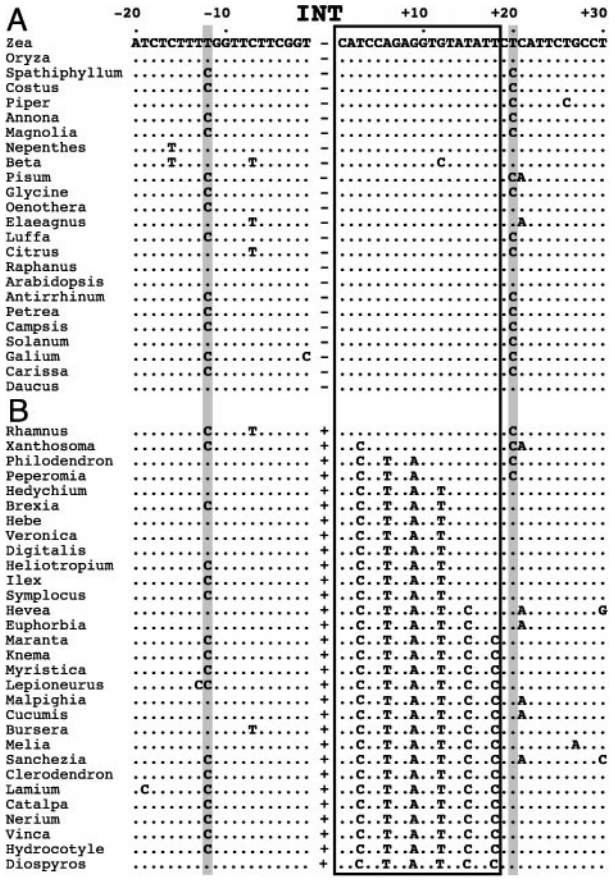

FIG. 4. Coconversion of *cox1* exonic sequences immediately 3′ (boxed region) to the intron insertion site (marked "INT"). Dots indicate identity to the *Zea* reference sequence. Taxa are arranged in phylogenetic order, with the intron-containing angiosperms first grouped according to length of coconversion tract. Shaded columns indicate positions of C-to-U RNA editing (A. Shirk and J.D.P., unpublished results); editing sites are known to evolve rapidly (28). (*A*) Angiosperms lacking the intron. (*B*) Angiosperms containing the intron.

occurred within the evolution of a particular family of flowering plants. Consistent with these conclusions, more intensive study of a single family of flowering plants indicates 5 separate intron gains among the 6 taxa (of only 14 examined) found to contain the intron (30). Among mobile genetic elements of any type, this rampancy of lateral transfer seems to be approached only by the mariner transposable elements of insects and other animals (31–33), whereas such recently emergent and massive promiscuity seems without precedent.

The close relationships (Fig. 3*A*) of members of this family of extraordinarily invasive introns, together with their identical (in sequence, irrespective of length) tracts of 3′ coconversion (Fig. 4*B*), suggest two opposing models for the history of horizontal transmission of the intron. Many or all of the donors of the intron might have been a nonplant (perhaps a fungus; Fig. 5*A*), in which case the donors themselves must all be closely related. Alternatively, a single or a few fungal donations might have been followed by hundreds or thousands of recent plant-to-plant lateral transfers (Fig. 5*C*).

These two models make testably distinct phylogenetic predictions on further sampling of the intron in plants and nonplants. The all-nonplant-to-plant model (Fig. 5*A*) predicts the existence of a clade of closely related intron-containing nonplants, various members of which are sister taxa in intron phylogenetic analyses to each clade of plants that has separately acquired the intron. That is, the plant introns would be phylogenetically interspersed with the donor introns (Fig. 5*B*).

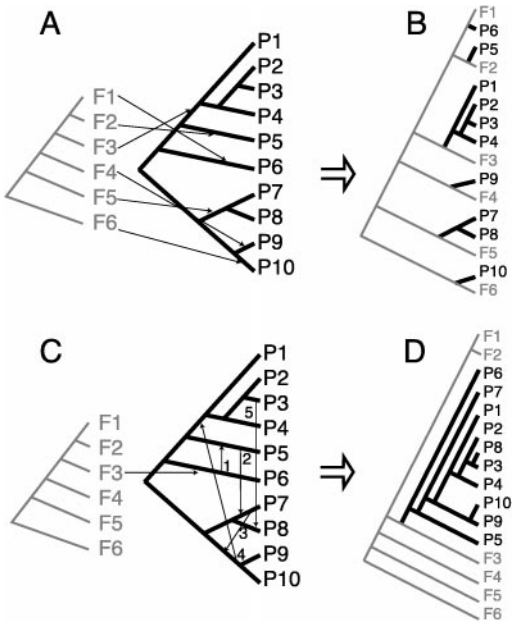

FIG. 5. Two extreme models for the pathway of *cox1* intron transfer in plants and nonplants. (*A*) An all nonplant-to-plant transfer model. *Left* cladogram shows six donor organisms, all nonplants (F1–6). Arrows show donor-recipient relationships for six separate intron transfers to plants (P1–10 in *Right* cladogram). (*B*) Intron phylogeny based on the transfers diagrammed in *A*, showing phylogenetic interspersion of donor (F) and recipient (P) sequences. (*C*) An all-plant-to-plant transfer model. A single initiating transfer from the nonplant F3 lineage to plants is shown, followed by five successive plant-to-plant transfers. (*D*) Intron phylogeny based on *C*, showing a clade of 10 plant introns whose phylogeny is incongruent with that of the same plants in *C*. Branch lengths in these cladograms are not proportional to time.

By contrast, the extreme form of the plant-to-plant model (one nonplant transfer, the rest all plant-to-plant; Fig. 5*C*) predicts a phylogenetic hierarchy of plant *cox1* introns in which only one clade of plant introns (*P6* in Fig. 5*D*) derives directly from nonplant-to-plant transfer, with all subsequent recipient introns nested within it (Fig. 5*D*), regardless of the true phylogeny of the host plants (Fig. 5*C*). Each subsequent plant-to-plant transfer will thus appear as a further nested hierarchy of paraphyletic donor introns from which emerges an organismally unrelated group of recipient introns (e.g., note nesting of *P8* within the *P1–P4* intron clade in Fig. 5*D*). Unfortunately, correct deciphering of donor-recipient identities for even a single case of horizontal transfer will in essence require working out the phylogeny of this intron across the >1,000 lineages of angiosperms estimated to have separately acquired it, or else otherwise potential bridging taxa will be missed. This makes determining the timing of transfer and any biogeographic, ecological, or phylogenetic determinants of donor-recipient relationships a daunting task. Nonetheless, an answer to the basic question of whether there are few or many plant-to-plant transfers should emerge with relatively modest but judicious further sampling of angiosperms.

The intron phylogeny in Fig. 3*B* reveals one case of apparent plant-to-plant transfer; however, this collapses under closer scrutiny. This case involves a strongly supported (100% bootstrap support) clade of five introns (from *Clerodendron* through *Vinca*) whose members all have the same coconversion tract length. The two basal members of this clade—*Vinca* and *Nerium*—both belong to the Apocynaceae. Their introns have precisely the paraphyletic relationship with respect to the other three introns in this clade (each of which belongs to a different other family of plants) that is expected if the Apocynaceae had first acquired its intron by a single ancestral gain and then donated its *cox1* intron to each of the other three

families. This scenario collapses because *Carissa*, an apocynaceous genus that is more closely related to *Nerium* than either is to *Vinca* (Fig. 2), both lacks the intron and, based on the absence of any coconconversion signatures (Fig. 4*A*), never possessed it. For this reason, *Nerium* and *Vinca* are inferred to have acquired their introns by separate events (Fig. 3*B*).

If transmission has been largely plant-to-plant, then exchange of genes between disparate plants may, at least on an evolutionary time scale, be more prevalent than is generally recognized. However, whether these exchanges occur so frequently as to be relevant to present concerns over the likelihood of genetically engineered crop genes spreading laterally to wild plants is unclear and will require extensive survey at lower taxonomic levels than studied herein. In any event, the exquisitely powerful homing mechanism of group I introns (see *Introduction* and refs. 1–9), together with their protection from genomic deletional forces (as opposed to cDNA- or RNA-mediated forces) by being sheltered within genes, makes them in many ways the perfect molecular parasites and thus perhaps the most sensitive monitors of gene flow across breeding barriers.

Regardless of the historical pathways of intron transfers, vectoring agents are probably involved (e.g., viruses, bacteria, aphids, mycorrhizal fungi, etc.). These could ferry this intron either as transiently ingested DNA or in a genetically integrated form. It is thought that a semiparasitic mite may act as a vector for P-element transposons across species boundaries of drosophilid flies (34).

The recency of this massive wave of intron gains is striking. Does it reflect the recent emergence of a widely promiscuous donor or vectoring agent, the former fortuitously containing the intron in its mt *cox1* gene? Or perhaps key is some recently evolved special property of a particular clade of introns? Such properties could include an extremely active homing endonuclease, splicing that is either unusually independent of host factors or else dependent on ones that are highly conserved and ubiquitious (in either case, enabling the intron to spread readily without regard to host), or perhaps unusually short coconversion, yielding intron insertions that are silent with respect to *COX1* function. This last possibility is attractive given that the coconversion tracts observed here are in fact much shorter than those typically observed in group I intron homing (refs. 35 and 36, and references therein).

Note Added in Proof. Watanabe *et al.* (37) recently discovered in the green alga *Chlorella vulgaris* strain NIES 227, a related, endonuclease-encoding form of this intron located at the same position in the *cox1* gene. Their phylogenetic analysis places the *Chlorella* intron together with that of another green alga, *Protheca*, whereas the intron from the angiosperm *Peperomia* groups with those of fungi (specifically yeasts), in agreement with the results of Fig. 3*A* and ref. 16.

We thank K. Adams and P. Keeling for critical reading of the manuscript and D. Burton, D. Campbell, M. W. Chase, S. Kroken, K. A. Kron, J. Lemon, C. R. Parks, L. Rieseberg, K. Song, S. Swensen, and R. Wallace for plant material. This work was supported by National Institutes of Health fellowship GM-17923 to Y.-L.Q. and National Institutes of Health Research Grant GM-35087 to J.D.P.

1. Dujon, B. (1989) *Gene* **82,** 91–114.
2. Lambowitz, A. M. & Belfort, M. (1993) *Annu. Rev. Biochem.* **62,** 587–622.
3. Belfort, M. & Perlman, P. S. (1995) *J. Biol. Chem.* **270,** 30237–30240.
4. Delahodde, A., Goguel, V., Becam, A. M., Creusot, F., Perea, J., Banroques, J. & Jacq, C. (1989) *Cell* **56,** 431–441.
5. Wenzlau, J. M., Saldanha, R. J., Butow, R. A. & Perlman, P. S. (1989) *Cell* **56,** 421–430.
6. Jacquier, A. & Dujon, B. (1985) *Cell* **41,** 383–394.
7. Lemieux, C. & Lee, R. W. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 4166–4170.
8. Muscarella, D. E. & Vogt, V. M. (1989) *Cell* **56,** 443–454.
9. Quirk, S. M., Bell-Pederson, D. & Belfort, M. (1989) *Cell* **56,** 455–465.
10. Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J. & Stahl, F. W. (1983) *Cell* **33,** 25–35.
11. Lang, B. F. (1984) *EMBO J.* **3,** 2129–2136.
12. Hibbett, D. S. (1996) *Mol. Biol. Evol.* **13,** 903–917.
13. Nishida, H., Tajiri, Y. & Sugiyama, J. (1998) *J. Mol. Evol.* **46,** 448–448.
14. Bhattacharya, D., Friedl, T. & Damberger, S. (1996) *Mol. Biol. Evol.* **13,** 978–989.
15. Turmel, M., Côté, V., Otis, C., Mercier, J.-P., Gray, M. W., Lonergan, K. M. & Lemieux, C. (1995) *Mol. Biol. Evol.* **12,** 533–545.
16. Vaughn, J. C., Mason, M. T., Sper-Whitis, G. L., Kuhlman, P. & Palmer, J. D. (1995) *J. Mol. Evol.* **41,** 563–572.
17. Adams, K. L., Clements, M. J. & Vaughn, J. C. (1998) *J. Mol. Evol.* **46,** 689–696.
18. Ogawa, S., Matsuo, K., Angata, K., Yanagisawa, K. & Tanaka, Y. (1997) *Curr. Genet.* **31,** 80–88.
19. Doyle, J. J. & Doyle, J. S. (1987) *Phytochem. Bull.* **19,** 11–15.
20. Olsen, G. J., Matsuda, H., Hagstrom, R. & Overbeek, R. (1994) *Comp. Appl. Biosc.* **10,** 41–48.
21. Wolfe, K. H., Li, W.-H. & Sharp, P. M. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 9054–9058.
22. Palmer, J. D. & Herbon, L. A. (1988) *J. Mol. Evol.* **28,** 87–97.
23. Qiu, Y.-L., Cho, Y., Cox, J. C. & Palmer, J. D. (1998) *Nature (London)* **394,** 671–674.
24. Henke, R. M., Butow, R. A. & Perlman, P. S. (1995) *EMBO J.* **20,** 5094–5099.
25. Farris, J. S., Kallersjo, M., Kluge, A. G. & Bult, C. (1995) *Cladistics* **10,** 315–319.
26. Chase, M. W., Soltis, D. E., Olmstead, R. G., Morgan, D., Les, D. H., Mishler, B. D., Dulvall, M. R., Price, R. A., Hills, H. G., Qiu, Y.-L., *et al.* (1993) *Ann. Mo. Bot. Gard.* **80,** 528–580.
27. Kishino, H. & Hasegawa, M. (1989) *J. Mol. Evol.* **29,** 170–179.
28. Shields, D. C. & Wolfe, K. H. (1997) *Mol. Biol. Evol.* **14,** 344–349.
29. Cho, Y., Adams, K., Qiu, Y.-L., Kuhlman, P., Vaughn, J. C. & Palmer, J. D. (1998) in *Plant Mitochondria: from Gene to Function*, eds. Moller, I.-M., Gardestrom, P. & Glaser, E. (Backhuys, Leiden), pp. 19–23.
30. Cho, Y. & Palmer, J. D. (1999) *Mol. Biol. Evol.*, in press.
31. Robertson, H. M. (1993) *Nature (London)* **362,** 241–245.
32. Robertson, H. M. (1997) *J. Hered.* **88,** 195–201.
33. Robertson, H. M., Soto-Adames, F. N., Walden, K. K. O., Avancini, R. M. P. & Lampe, D. J. (1998) in *Horizontal Gene Transfer*, eds. Syvanen, M. & Kado, C. (Chapman & Hall, London), pp. 268–284.
34. Houck, M. A., Clark, J. B., Peterson, K. R. & Kidwell, M. G. (1991) *Science* **253,** 1125–1129.
35. Mueller, J. E., Smith, D. & Belfort, M. (1996) *Genes Dev.* **10,** 2158–2166.
36. Bussieres, J., Lemieux, C., Lee, R. W. & Turmel, M. (1996) *Curr. Genet.* **30,** 356–365.
37. Watanabe, K. I., Ehara, M., Inagaki, Y. & Ohama, T. (1998) *Gene* **213,** 1–7.