J Mol Evol (2004) 59:80–89
DOI: 10.1007/s00239-004-2606-y

JOURNAL OF MOLECULAR EVOLUTION

Many Independent Origins of *trans* Splicing of a Plant Mitochondrial Group II Intron

Yin-Long Qiu,[1,2] Jeffrey D. Palmer[1]

[1] Department of Biology, Indiana University, Bloomington, IN 47405, USA
[2] Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA

Received: 1 October 2003 / Accepted: 6 February 2004

Abstract. We examined the *cis*- vs. *trans*-splicing status of the mitochondrial group II intron *nad1i728* in 439 species (427 genera) of land plants, using both Southern hybridization results (for 416 species) and intron sequence data from the literature. A total of 164 species (157 genera), all angiosperms, was found to have a *trans*-spliced form of the intron. Using a multigene land plant phylogeny, we infer that the intron underwent a transition from *cis* to *trans* splicing 15 times among the sampled angiosperms. In 10 cases, the intron was fractured between its 5′ end and the intron-encoded *matR* gene, while in the other 5 cases the fracture occurred between *matR* and the 3′ end of the intron. The 15 intron fractures took place at different time depths during the evolution of angiosperms, with those in Nymphaeales, Austrobaileyales, Chloranthaceae, and eumonocots occurring early in angiosperm evolution and those in *Syringodium filiforme*, *Hydrocharis morsus-ranae*, *Najas*, and *Erodium* relatively recently. The *trans*-splicing events uncovered in Austrobaileyales, eumonocots, Polygonales, Caryophyllales, Sapindales, and core Rosales reinforce the naturalness of these major clades of angiosperms, some of which have been identified solely on the basis of recent DNA sequence analyses.

Key words: Genomic structural characters — Group II intron — Intron evolution — Mitochondrial *nad1* — Plant phylogeny — *trans* splicing

Correspondence to: Yin-Long Qiu; *email:* ylqiu@umich.edu

Introduction

The *trans* splicing of group II introns, i.e., fragmentation of an intron into at least two separate pieces that belong to different transcriptional units and splicing of these intron pieces that are reunited at the RNA level via intermolecular base pairing, was first discovered in the chloroplast genes *rps12* (Fukuzawa et al. 1986; Koller et al. 1987; Zaita et al. 1987) and *psaA* (Kück et al. 1987). A few years later, several research groups reported a total of six *trans*-spliced introns in the angiosperm mitochondrial genes *nad1* (Chapdelaine and Bonen 1991; Conklin et al. 1991; Wissinger et al. 1991), *nad2* (Binder et al. 1992), and *nad5* (de Souza et al. 1991; Knoop et al. 1991). Recently, two *trans*-spliced introns were found in mitochondrial *nad3* of the green alga *Mesostigma viride* (Turmel et al. 2002a). Most *trans*-spliced introns are bipartite, but two tripartite group II introns are known, in the chloroplast *psaA* gene of *Chlamydomonas reinhardtii* (Goldschmidt-Clermont et al. 1991) and the mitochondrial *nad5* gene of *Oenothera berteriana* (Knoop et al. 1997).

cis-spliced orthologs of all *trans*-spliced introns known from angiosperm mitochondrial genomes have been isolated from either bryophytes or pteridophytes (Malek et al. 1997; Malek and Knoop 1998; Qiu et al. 1998a), suggesting that the *trans*-spliced form represents an evolutionarily derived condition of a *cis*-spliced intron. For all these introns except one, there is no evidence that the *cis*- to *trans*-splicing transition has occurred more than once, although in all cases relatively few taxa have been examined. The

exceptional case is the *nad1i728* (formerly called *nad1.i4* in angiosperms; here we follow a new intron nomenclature proposed by Dombrovska and Qiu 2004), which has evolved *trans*-splicing two or three times independently during angiosperm evolution, in the monocot family Poaceae (wheat, maize, and rice [Chapdelaine and Bonen 1991; Thomson et al. 1994; Notsu et al. 2002]) and in the eudicot families Solanaceae (petunia, Conklin et al. 1991) and Chenopodiaceae (sugar beet [Kubo et al. 2000]). We infer at least two and quite possibly three separate *trans*-splicing evolutionary events for this intron (which is *cis*-spliced in a moss and a hornwort [Qiu et al. 1998a]), because several rosids (which also are eudicots) have been investigated (Wahleithner et al. 1990; Thomson et al. 1994; Wissinger et al. 1991; Unseld et al. 1997) and all have the *cis*-spliced condition (the uncertainty here, two vs. three fractures, reflects the unresolved branching order of rosids, Solanaceae, and Chenopodiaceae in current estimates of angiosperm phylogeny [Soltis et al. 2000]), and because the Poaceae fracture (located 5′ of the intron-encoded *matR* gene) is physically distinct from and therefore evolutionarily independent of the two eudicot fractures (located 3′ of *matR*).

The evolution of *trans*-splicing involves the physical separation of intron fragments and adjacent exons, and represents a substantial genomic structural change. So far, it seems that this type of genomic event occurs at a rather low frequency. Rare genomic structural changes have been used in a number of cases to help resolve difficult issues in plant phylogeny (e.g., Jansen and Palmer 1987; Manhart and Palmer 1990; Raubeson and Jansen 1992; Qiu et al. 1998a). Hence, it is worth exploring the utility of *cis*- to *trans*-splicing changes for reconstructing plant phylogeny. Despite the recent success of multigene analyses in reconstructing angiosperm phylogeny (Parkinson et al. 1999; Qiu et al. 1999, 2000; Chase et al. 2000; Graham and Olmstead 2000; Soltis et al. 2000), certain important relationships remain unresolved or poorly supported. Furthermore, most major clades identified in these multigene phylogenies lack independent corroboration from other evidence. Additional characters, such as genomic structural changes, may help resolve some of the most difficult phylogenetic patterns in angiosperms and provide further support to the clades already suggested by the sequence analyses.

In this study, we investigated the evolution of *trans* splicing of *nad1i728* across land plants by carrying out a large-scale Southern hybridization survey of vascular plants. In addition, intron splicing status in bryophytes, pteridophytes, and several angiosperms was assessed from sequence data compiled from the literature (Wahleithner et al. 1990; Chapdelaine and Bonen 1991; Conklin et al. 1991; Wissinger et al. 1991; Thomson et al. 1994; Unseld et al. 1997; Kubo et al. 2000; Notsu et al. 2002; Qiu et al. 1998a; Dombrovska and Qiu 2004). The goals of our study were (1) to determine whether additional cases of *trans* splicing have occurred besides those two reported early on in Poaceae and Solanaceae [(Chapdelaine and Bonen 1991; Conklin et al. 1991; Thomson et al. 1994)—our project was initiated well before *nad1i728 trans* splicing was discovered in sugar beet (Kubo et al. 2000) and rice (Notsu et al. 2002; see Qiu and Palmer 1997); (2) to assess the evolutionary depth of *trans*-splicing events; and (3) to use these events as a new set of characters to evaluate plant phylogeny reconstructed from the multigene analyses.

Materials and Methods

DNA Preparation

Plant tissues were collected fresh from the field, botanical gardens, or greenhouses. In a few cases, silica gel-dried leaves were used. Total cellular DNAs were extracted using a modified CTAB (hexadecyltrimethylammonium bromide) method (Doyle and Doyle 1987) and further purified by CsCl/ethidium bromide gradient ultracentrifugation. Almost all DNA samples were vouchered, with a few exceptions for which insufficient material was available or for which the species were commercial crops. Voucher information is available from the corresponding author upon request.

Identification of trans*-Spliced Introns by Southern Hybridization*

To identify *trans*-spliced introns, we designed the following strategy. Total cellular DNAs were digested by a restriction enzyme, separated by agarose gel electrophoresis, and transferred onto nylon membranes. Using PCR, three probes were made for Southern hybridizations (Fig. 1); these correspond to intron positions 16–427 (5′-end of the intron; probe nad1i728 h), 709–2394 (*matR*; probe nad1i728matR), and 2680–3088 (3′-end of the intron; probe nad1i728t) of *nad1i728* in *Oenothera berteriana* (Wissinger et al. 1991). The probes were generated using template DNA from *Oenothera organensis*, because this material was readily available. If a species has a *cis*-spliced intron, all three probes will generally hybridize to the same restriction fragment, assuming that there is no restriction site for the enzyme used within the intron (see below for our approach to tackle this problem when there is a site). If a species has a *trans*-spliced intron that is broken between the 5′-end of the intron and *matR*, as in wheat (Chapdelaine and Bonen 1991), then probe nad1i728h will hybridize to a different restriction fragment than probes nad1i728matR and nad1i728t, which should hybridize to the same fragment. Likewise, if a species has a *trans*-spliced intron that is broken between *matR* and the 3′-end of the intron, as in the case of petunia (Conklin et al. 1991), probes nad1i728 h and nad1i728matR will hybridize to a different restriction fragment than probe nad1i728t. Figure 2 presents a set of taxa that have either a *cis*- or a *trans*-spliced intron and demonstrates how this method works.

Two rounds of hybridization surveys were carried out. In the first round, the total cellular DNAs were digested with *Bam*HI or *Hin*dIII (New England Biolabs). We used these two enzymes to reduce the chance of both enzymes cutting in the same region, since

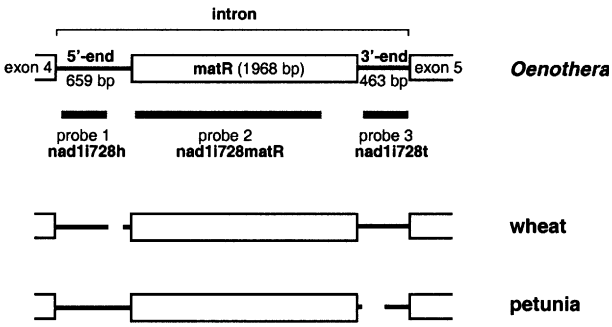

Fig. 1. Structure of the mitochondrial intron *nad1i728* in *Oenothera berteriana* (*cis*-spliced), wheat (*trans*-spliced; fracture occurs between exon 4 and *matR*), and petunia (*trans*-spliced; fracture occurs between *matR* and exon 5). The three probes used in this study to detect *trans*-spliced introns are shown by thick gray lines. The drawing is to scale for *Oenothera* only.

the former has a GC-rich recognition site and the latter an AT-rich site. Digestion conditions followed those provided by the enzyme manufacturer. The digested DNAs were separated on 0.8% agarose gels for a running distance of 6 cm (for the bromophenol blue tracking dye), so that we could include 140 (35 × 4) samples on a piece of 20 × 25-cm Kodak X-ray film. Southern transfer used Immobilon-Ny+ transfer membranes (Millipore; we used a test version), which allowed at least 16 successive probe hybridizations (these membranes were used for surveying many genes and introns [e.g., Cho et al. 1998; Qiu et al. 1998a; Adams et al. 2000, 2002]). Probes were labeled with ^{32}P using a homemade random-priming kit and the Klenow DNA polymerase (3′ → 5′ exo$^-$) from New England Biolabs. The membranes were prehybridized for 2 h at 60°C in a solution containing 5× SSC, 50 m*M* Tris (pH 8.0), 0.2% SDS, 10 m*M* EDTA, and 2× Denhardt's and hybridized for 18 h at 60°C in the above solution plus 5% dextran sulfate (Pharmacia). After hybridization, the membranes were washed twice in 2× SSC for 5 min at room temperature, followed by twice in 2× SSC and 0.5% SDS for 30 min at 60°C. Autoradiography was carried out using intensifying screens at −80°C for 18–48 h. The same set of membranes was hybridized sequentially with the three probes described above, allowing direct overlay of autoradiographs with adjacent probes to determine whether they share the same restriction fragment(s) and thereby enabling scoring of *cis*- vs. *trans*-splicing status as outlined in the preceding paragraph.

The occurrence of restriction sites in or near interprobe regions, or the short electrophoresis distance used (which may result in insufficient separation of fragments of similar sizes), could generate either false-positive or false-negative diagnosis of the *trans*-spliced condition. Therefore, we carried out a second round of Southern blot survey for almost all the species tentatively identified as containing a *trans*-spliced intron in the first round; only the very few taxa that do not have enough DNA were not included in this round. Three measures were taken to eliminate false positives in this round of hybridization. First, we used four more enzymes, *Bcl*I, *Eco*RV, *Nde*I, and *Xho*I; these are infrequent cutters based on examining restriction sites in *nad1i728* of the six diverse angiosperms with then-known sequences (wheat, maize, *Oenothera berteriana*, *Arabidopsis thaliana*, soybean, and broad bean). The chances of all six enzymes having sites in the interprobe regions (or very near the end of a probe) should be very low. Second, to achieve better separation of restriction fragments, we ran the gels twice as far (12 cm) as in the first round. Third, we mapped the distribution of species that contain a putatively *trans*-spliced intron onto a phylogenetic tree of seed plants that was reconstructed by analyzing sequences of the chloroplast gene *rbcL* ([Chase et al. 1993]; the three-gene angiosperm phylogeny by Soltis et al. [2000] was not available at this point in the study). Using this and other

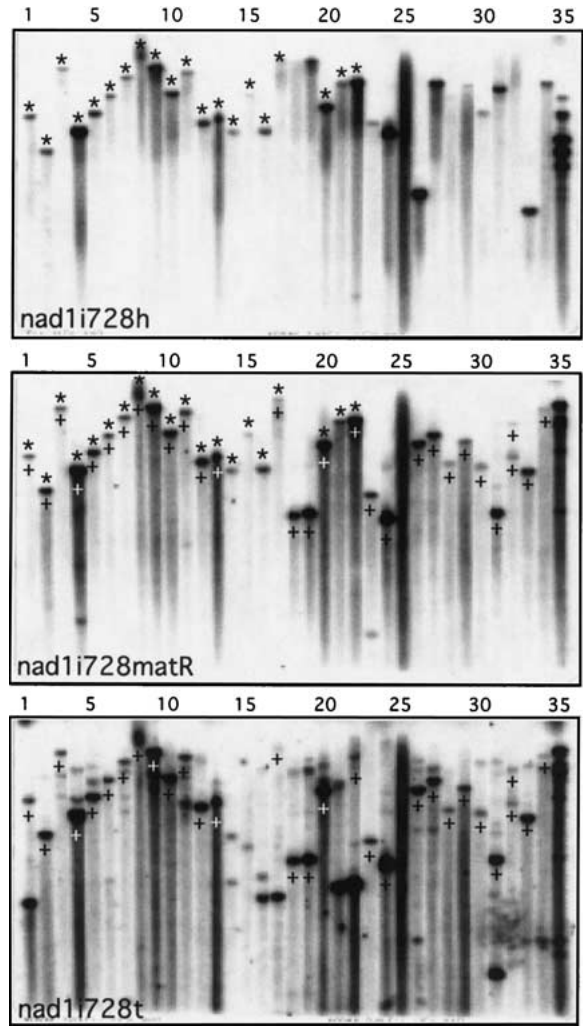

Fig. 2. Autoradiographs showing hybridization results of three *nad1i728* probes to *Eco*RV-digested DNAs from 35 monocots. Species 1–17 and 20–22 have a *cis*-spliced intron, and species 18, 19, and 23–35 have a *trans*-spliced intron with fracture between exon 4 and *matR*. Fragments below asterisks are shared between probes nad1i728 h and nad1i728matR. Fragments above plus signs are shared between probes nad1i728matR and nad1i728t. White color is used for "asterisks" or "plus" signs in places where the hybridization background or smear is too dark. When no fragment is shared between probes nad1i728 h and nad1i728matR, a *trans*-spliced intron is inferred. A few species (Nos. 14, 15, 16, and 21) lack any shared *Eco*RV fragment for probes nad1i728matR and nad1i728t due to the presence of *Eco*RV restriction sites in or near the interprobe region (other enzyme digests did produce shared fragments). Species 25, 32, and 35 have partially digested DNAs, thus producing a hybridization smear or multiple bands. In the nad1i728t autoradiograph, more than one fragment was visualized, likely due to cross-hybridization of the probe with nonorthologous, related introns in domains V and VI. Species shown: 1—*Acorus calamus*, 2—*Pleea tenuifolia*, 3—*Amorphophallus rivieri*, 4—*Anthurium scherzerianum*, 5—*Arisaema triphyllum*, 6—*Dieffenbachia* sp., 7—*Peltandra virginica*, 8—*Philodendron oxycardium*, 9—*Scindapsus aureus*, *10*—*Spathiphyllum clevelandii*, *11*—*Xanthosoma mafatta*, *12*— *Zamioculcas zamiifolia*, *13*—*Zantedeschia aethiopica*, *14*—*Alisma* sp., *15*—*Echinodorus radicans*, *16*—*Sagittaria* sp., *17*—*Hydrocleys* sp., *18*—*Najas gracillima*, *19*—*Najas minor*, *20*—*Anacharis* sp., *21*—*Potamogeton berchtoldii*, *22*—*Potamogeton crispus*, *23*—*Burmannia capitata*, *24*—*Dioscorea* sp., *25*—*Pandanus veitchii*, *26*—*Alstroemeria* sp., *27*—*Clintonia uniflora*, *28*—*Disporum hookeri*, *29*—*Streptopus amplexaca*, *30*—*Zigadenus glaberrimus*, *31*—*Trillium* sp., *32*—*Lilium* sp., *33*—*Smilax* sp., *34*—*Asparagus officinalis*, *35*—*Cordyline terminalis*.

relevant phylogenetic information, we then sampled more species for each case of putative *trans* splicing. This increase in species sampling should help corroborate, or call into question, the original observation of *trans* splicing in each plant lineage and should also reduce potential technical errors such as misidentification of species and mixup or contamination of DNA samples. Altogether, 416 species from 402 genera of vascular plants were investigated in the hybridization surveys (see Fig. 3).

Inference of Evolutionary Frequency and Depth of trans *Splicing by Phylogenetic Mapping*

To infer the number of independent occurrences of *trans* splicing of *nad1i728* and the evolutionary depth of each event, we mapped the taxa with *trans*-spliced introns onto a land plant phylogeny. There have been several studies with moderate to extensive gene and taxon sampling in pteridophytes (Pryer et al. 2001), gymnosperms (Chaw et al. 2000), basal angiosperms (Qiu et al. 1999), monocots (Chase et al. 2000), and angiosperms (Soltis et al. 2000). Hence, we constructed a synthetic land plant phylogeny by taking the following approach. First, we downloaded *rbcL* sequences from GenBank, which covered about 95% of our taxa, and generated a rough phylogeny by parsimony analysis. This phylogeny matched the results of the above-mentioned studies quite well. We then modified this phylogeny according to these studies as well as several others that presented detailed phylogenies for the groups that we sampled intensively: Araceae (French et al.1995), Alismatidae (Les et al. 1997), Asparagales (Chase et al. 1995a), Zingiberales (Kress et al. 2001), Caryophyllidae (Cuenoud et al. 2002), and Asteridae (Albach et al. 2001). The bryophytes were arranged based on our unpublished results. Finally, the few taxa without *rbcL* sequences were added onto the tree by consulting Mabberley (1987). The resulting synthetic phylogeny is shown in Fig. 3.

Those taxa that have a putative *trans*-spliced intron according to Southern hybridization or sequencing studies are highlighted in Fig. 3. The frequency of evolution of *trans* splicing is inferred by simply counting the number of phylogenetically unrelated lineages in which the condition has been observed. The depth of each *trans*-splicing event is assessed by examining its breadth of phylogenetic distribution as well as the position of the clade within the plant phylogeny. We avoided assigning absolute time to each *trans*-splicing evolutionary event, as we believe that current age estimates for various land plant clades using molecular clocks and fossil evidence (e.g., Wikstrom et al. 2001) do not provide a level of resolution and confidence to warrant such a practice. On the contrary, the state of plant phylogeny reconstruction does allow us to estimate the frequency and phylogenetic depth of evolution of *trans* splicing.

Results and Discussion

Evolution of nad1i728

Among the 439 species (427 genera) examined (416 species [402 genera] in our hybridization survey and 39 species [38 genera] in previous sequencing studies [Wahleithner et al. 1990; Chapdelaine and Bonen 1991; Conklin et al. 1991; Wissinger et al. 1991; Thomson et al. 1994; Unseld et al. 1997; Kubo et al. 2000; Notsu et al. 2002; Qiu et al. 1998a, Dombrovska and Qiu 2004]; a small number of species were overlapped in the two types of studies), a total of 164 species (157 genera) was found to have a *trans*-spliced intron. Of these, 130 species had the intron broken between the 5′ end and *matR* (the wheat type [Chapdelaine and Bonen 1991]), whereas 34 species had the intron broken between *matR* and the 3′ end (the petunia type [Conklin et al. 1991]). By mapping these data onto a land plant phylogeny (Fig. 3), we estimate that this intron has evolved the *trans*-splicing condition 15 times independently. In 10 cases, the wheat type evolved, in Nymphaeales, Austrobaileyales, Chloranthaceae, Winteraceae, *Syringodium filiforme* in Cymdoceaceae, *Hydrocharis morsus-ranae* in Hydrocharitaceae, *Najas* (in both species examined) in Najadaceae, eumonocots, *Erodium* (in all three species investigated) in Geraniaceae, and core Rosales. The five cases of the petunia type are found in Sapindales, *Pterocarya stenoptera* and *Juglans cinera* in Juglandaceae, part of Polygonales, Caryophyllales, and Solanaceae.

All cases of *trans* splicing were found in angiosperms. The fracture events that created these *trans*-spliced introns occurred over a variety of taxonomic ranges and time depths (Fig. 3). The most ancient events are in Nymphaeales, Austrobaileyales, Chloranthaceae, and eumonocots, while the most recent ones are restricted to one or two genera of Cymdoceaceae, Hydrocharitaceae, Najadaceae, Geraniaceae, and Juglandaceae. These data indicate that *trans* splicing in this intron has evolved repeatedly throughout the history of angiosperm evolution. The process started soon after angiosperms originated, as Nymphaeales and Austrobaileyales are very close to the root of angiosperm phylogeny (Mathews and Donoghue 1999; Parkinson et al. 1999; Qiu et al. 1999; Soltis et al. 2000; Graham and Olmstead 2000) and both have ancient fossils dated to the Early Cretaceous (Upchurch 1984; Friis et al. 2001).

In only 3 of the 439 species examined do we regard the splicing status of *nad1i728* as ambiguous. One case involves *Amborella trichopoda*, the putatively earliest angiosperm (Mathews and Donoghue 1999; Parkinson et al. 1999; Qiu et al. 1999; Graham and Olmstead, 2000; Soltis et al. 2000). This species may possess both *cis*- and *trans*-spliced forms of the intron, because multiple fragments were detected in hybridizations of all three probes in all enzyme digests and because there are shared fragments between hybridizations (Fig. 4). Given this complexity of hybridization results, DNA sequencing is required to properly elucidate the *cis*- and/or *trans*-splicing status of *Amborella*. Similarly, *Hedychium coronarium* (Zingiberaceae) may also possess more than one copy of the intron, with one in the *cis*-spliced form, as multiple bands were observed in both nad1i728h and nad1i728matR probe hybridizations, and some shared fragments were observed. In the case of *Pontederia cordata* (Pontederiaceae), the hybridization data strongly indicate either the presence of only a

cis-spliced intron or, if still *trans*-spliced, that fortuitous rearrangement has brought the two parts of the intron very close together. These two last species are both deeply imbedded in a large monocot clade that

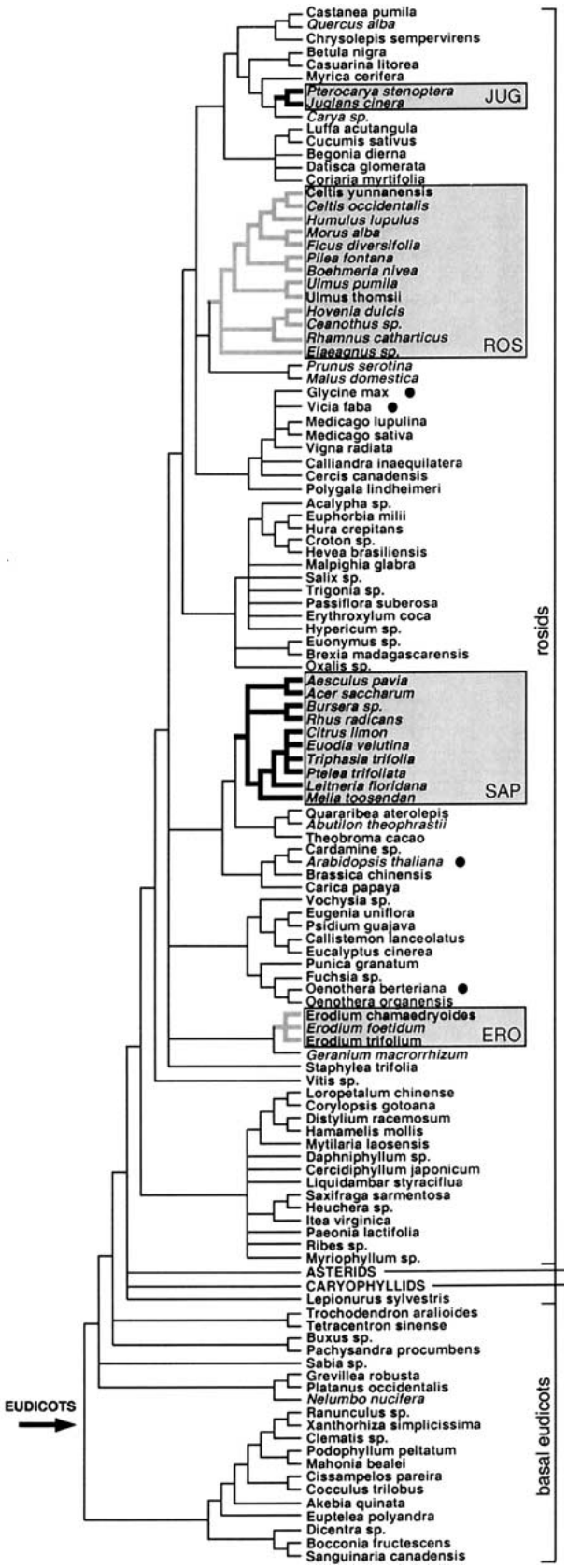

Fig. 3. Phylogenetic distribution of *trans* splicing of *nad1i728* in land plants. The phylogeny shown was generated as described in Materials and Methods. The species with *trans*-spliced introns are shown in gray boxes (except for *Hedychium coronarium* and *Pontederia cordata*, which may have secondarily derived *cis*-spliced introns and are indicated by asterisks, and *Amborella trichopoda*, whose intron status is even less clear and is therefore marked by a question mark). The lineages whose *trans*-spliced introns are broken between the 5′-end of the intron and *matR* are shown by thick gray lines, while those with breakage between *matR* and the 3′-end of the intron are shown by thick black lines. For taxa with a black dot, the splicing status of the intron was determined by sequencing (see text for citations in which the sequences were reported). For taxa in italics, the splicing status of the intron was inferred using Southern hybridizations with four single-enzyme digests and 12-cm electrophoretic separation (many of these taxa were also investigated in the first round of hybridizations using two other, single-enzyme digests and 6-cm electrophoretic separation). All other taxa (in plain text) have a *cis*-spliced intron according to the results of Southern hybridizations using two single-enzyme digests and 6 cm electrophoretic separation. NYM—Nymphaeales, AUS—Austrobaileyales, CHL—Chloranthaceae, WIN—Winteraceae, CYM—Cymdoceaceae, HYD—Hydrocharitaceae, NAJ—Najadaceae, EUM—Eumonocots, POL—Polygonales, CAR—Caryophyllales, ERO—*Erodium*, SAP—Sapindales, ROS—Rosales, JUG—Juglandaceae, SOL—Solanaceae.

otherwise has only *trans*-spliced introns (this includes the other species of Pontederiaceae examined, *Eichhornia crassipes*). As part of a separate study, we are now sequencing and studying the expression of the *nad1* gene from these two and various related monocots to distinguish among several possible explanations for these potential reversals from *trans* to *cis* splicing.

The 15 independent origins of *trans* splicing inferred for *nad1i728* make this the only group II intron for which multiple evolutionarily separate transitions from *cis* to *trans* splicing have been discovered. Other *trans*-spliced group II introns fall into two general categories with respect to the extent of survey of their splicing status and also the timing of their transition to *trans* splicing. One set of introns has been studied in a wide variety of plants and can be inferred to have evolved *trans* splicing at an early stage in their existence. These include *rps12i114* (formerly known as rps12.il; we are using new intron names according to the nomenclature proposed by Dombrovska and Qiu 2004) in the chloroplast genomes of streptophytes (charophytes + land plants) (Fukuzawa et al. 1986; Zaita et al. 1987; Lew and Manhart 1993; Turmel et al. 2002b), and *nad1i394* (formerly nad1Ti1), *nad1i669* (formerly nad1Ti3), *nad2i542* (formerly nad2Ti2), *nad5i1455* (formerly nad5Ti2), and *nad5i1477* (formerly nad5Ti3) in the mitochondrial genomes of land plants (Chapdelaine and Bonen 1991; Conklin et al. 1991; de Souza et al. 1991; Knoop et al. 1991; Wissinger et al. 1991; Binder et al. 1992; Malek et al. 1997; Malek and Knoop 1998). The *rps12i114* is *trans* spliced in all streptophytes that have been investigated, including *Spirogyra* and *Chaetosphaeridium* (both charophytic green algae); those in the *nad* genes are

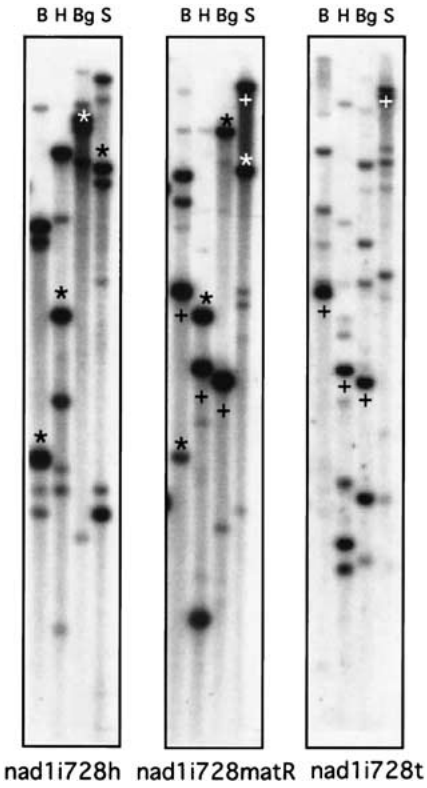

Fig. 4. Autoradiographs showing hybridization results of three intron probes to *Bam*HI (B)-, *Hin*dIII (H)-, *Bgl*II (Bg)-, and *Sac*I (S)-digested DNA from *Amborella trichopoda*, which may have both *cis*- and *trans*-spliced introns. The fragments below asterisks are shared between probes nad1i728 h and nad1i728matR. The fragments above plus signs are shared between probes nad1i728-matR and nad1i728t. White color is used for the asterisks or plus signs in places where the hybridization background or smear is too dark. Note the multiple fragments detected by each of the three probes.

trans spliced in at least monocots and eudicots. The other set of *trans*-spliced introns has been studied in only a small number of taxa and appears to be fairly restricted in their phylogenetic distribution. These include *psaAi80* and *psaAi260* (no formal names were ever used previously for these two introns) in the chloroplast of *Chlamydomonas reinhardtii* (Kück et al. 1987; Goldschmidt-Clermont et al. 1991) and *nad3i87* (formerly *Mvnad3.1*) and *nad3i301* (formerly *Mvnad3.2*) in the mitochondrion of *Mesostigma viride* (Turmel et al. 2002a).

No *trans* splicing of *nad1i728* was detected in gymnosperms, pteridophytes, hornworts, and mosses. A reasonable number of species that cover most major lineages of these groups were sampled in the Southern hybridization surveys (Fig. 3). In the PCR-sequencing studies by Qiu et al. (1998a) and Dombrovska and Qiu 2004, only *cis*-spliced introns could have been isolated, and so there was some inherent technical bias there. Indeed, a possibility exists for *trans*-spliced introns in *Psilotum nudum* and *Ophioglossum lusitanicum*, for which repeated PCR attempts by Dombrovska and Qiu 2004 failed to amplify the entire intron and instead recovered only the 5′-end or 3′-end (but not *matR*) of the intron. Given how many times *trans* splicing of this intron has arisen in angiosperms, it would not be surprising if more extensive survey of mosses and pteridophytes revealed cases of *trans* splicing within these groups, both of which are relatively large and diverse. Furthermore, all 24 group II introns known variously in one or another angiosperm mitochondrial genome are present broadly across angiosperms (Unseld et al. 1997; Kubo et al. 2000; Notsu et al. 2002), and most of these introns are present throughout vascular plants or even extend into bryophytes (Malek and Knoop 1998; Qiu et al. 1998a; Beckert et al. 2001; Gugerli et al., 2001; Hashimoto and Sato 2001; Pruchner et al. 2001; Dombrovska and Qiu 2004). Few of these introns have been surveyed very extensively for their splicing status, and therefore some may ultimately join *nad1i728* in the category of frequent transitions to *trans*-splicing. However, *cox2i373* (formerly cox2.i1, in Kudla et al. [2002], or cox2.i3, in Qiu et al. [1988a]) and *nad1i477* (formerly nad1.i2, in Gugerli et al. [2001]) have been investigated in a relatively large number of plants, and *trans* splicing has not been reported (Rabbi and Wilson 1993; Gugerli et al., 2001; Joly et al. 2001; Kudla et al. 2002). Hence, it would appear that certain introns (*nad1i728*) are more prone to evolve *trans* splicing than others (*cox2i373* and *nad1i477*).

What factors might promote the evolution of *trans* splicing? Angiosperm mitochondrial genomes are well known to rearrange at very high rates (Palmer and Herbon 1988; Fauron et al. 1995) and to harbor many families of short dispersed repeat sequences that can promote recombination and rearrangement (Andre et al. 1992). Recombination between repeat elements located internally to group II introns and also elsewhere in the genome could certainly fracture an intron and create a *trans*-spliced gene. Worth noting here is that the two pieces of the *trans*-spliced *nad1i728* in wheat both carry an 11-nucleotide purine-rich direct repeat, which is also present as one copy in the *cis*-spliced broad bean *nad1i728* (Chapdelaine and Bonen 1991). An eight-nucleotide-long repeat was found in the two pieces of the *trans*-spliced maize *nad1i728* (Thomson et al. 1994). The tripartite *nad5i1455* (formerly nad5Ti3, in Malek and Knoop [1998]) in *Oenothera berteriana* also carries an 11-nucleotide repeat in the first two pieces of the intron (Knoop et al. 1997). Recently, Dai and Zimmerly (2002) have found that in fragmented group II intron pieces in eubacterial genomes, repeat sequences, in one case a transposon, were often attached to the intron fragments. It is even possible that a *trans*-splicing intron could be generated by recombination between intron-encoded ORFs (such as *matR* in *nad1i728*) and related retroelement ORFs located

elsewhere in the mitochondrial genome. Several retroelements have been identified in plant mitochondrial DNAs (Knoop et al. 1996; Unseld et al. 1997; Notsu et al. 2002). Consistent with this idea, the ORF is always located in the domain IV of the intron and all *trans*-spliced introns have been found to be broken in that same domain.

Given the strong propensity of *nad1i728* to evolve *trans* splicing on both sides of *matR*, we thought that this intron would be an excellent candidate to look for more cases of tripartite *trans* splicing (Goldschmidt-Clermont et al. 1991; Knoop et al. 1997), particularly in such large *trans*-splicing clades as the eumonocots and the Caryophyllales. Nonetheless, we uncovered no cases of apparent tripartite *trans* splicing. Perhaps selection disfavors *matR* as a free-standing unit (for instance, its transcription may need to be coordinated with one of the two intron pieces or adjacent exons). A more likely reason why we did not find any cases of tripartite *trans* splicing is our experimental strategy. In both known cases of tripartite *trans* splicing, the ''second'' site of intron breakage occurs in domain I, in addition to the normal *trans*-splicing breakpoint in domain IV (Goldschmidt-Clermont et al. 1991; Knoop et al. 1997). Because our probe covering the 5′-end of the intron spans this domain I breakage point (Fig. 1), we probably would have missed any case of intron breakage within this domain. Future study with more numerous and specific smaller intron probes should better uncover any cases of tripartite *trans* splicing in *nad1i728* and other *trans*-spliced introns.

Phylogenetic Implications of Evolution of trans Splicing of nad1i728

The discovery of *nad1i728 trans* splicing in all examined Austrobaileyales provides the first non-sequence-based evidence for this only recently erected clade of basal angiosperms. This clade was first recognized in the 1993 *rbcL* analyses of Qiu et al. and has since been recovered in several multigene analyses (Parkinson et al. 1999; Qiu et al. 1999, 2000; Soltis et al. 2000). Several morphological characters, such as sieve-element plastid features (Behnke 1988) and gynoecial traits (Endress and Igersheim 1997), also support the monophyly of this group, although these characters have sufficient homoplasy that on their own they never led to the diagnosis of this clade as a natural group.

Trans splicing is universal among the two basalmost groups (excepting *Amborella*) of angiosperms, Nymphaeales and Austrobaileyales. This raises the possibility that, contrary to what is shown in Fig. 3, these two lineages actually are sister clades, forming a clade right after *Amborella* at the base of angiosperm phylogeny. Although the recent report of four-celled female gametophyte and diploid endosperm in these lineages makes this hypothesis tantalizing (Williams and Friedman 2002), none of the many sequence-based analyses (see above references) has recovered this topology. Furthermore, a Kishino–Hasegawa test similar to those reported by Qiu et al. (2000) also found that the topology of Nymphaeales and Austrobaileyales as sister clades has a much lower likelihood than the topology in which they form a grade (Qiu, data not shown). Therefore, we think that *trans* splicing probably evolved independently in Nymphaeales and Austrobaileyales.

One of the phylogenetically most significant *trans*-splicing events uncovered in this study is the one that marks the so-called ''eumonocot'' clade (Fig. 3). The term eumonocot was first used by Stevenson and Loconte (1995) to refer to a different group of monocots. Because their eumonocot clade is not recovered in phylogenetic analyses of molecular and combined molecular and morphological data sets (Chase et al. 1995b; 2000), we hereby reappropriate this term to refer to a clade comprising all monocots except for *Acorus* and the Alismatales, which *has* been consistently identified in all molecular analyses. This eumonocot clade received a support of decay index 3 in Chase et al.'s (1995b) combined *rbcL* and morphological analysis of monocots, 78–95% bootstrap support in Chase et al.'s (2000) combined *rbcL*, *atpB*, and 18S analysis of monocots, and 99% jackknife support in Soltis et al.'s (2000) same three-gene analysis of angiosperms. Like many other major clades identified in sequence-based phylogenies, the eumonocot clade currently lacks any nonsequence synapomorphy to define it.

Trans splicing also serves as an important phylogenetic marker for defining Polygonales among the caryophyllids. This clade (importantly, including *Tamarix*) was recovered with only 61% bootstrap in Cuenoud et al.'s (2002) recent four-gene analysis and was not recovered at all in the three-gene study of Soltis et al. (2000). Thus this putatively synapomorphic *trans*-splicing event provides much-needed additional evidence for the monophyly of Polygonales. In contrast, the other clade, the Caryophyllales, that is marked by *trans* splicing has long enjoyed ample support from many lines of evidence (Cuenoud et al. 2002 and references therein).

Another phylogenetically informative *trans*-splicing event was discovered in Sapindales (Fig. 3). Even though this clade has been consistently recovered in molecular phylogenetic studies (Qiu et al. 1998b; Soltis et al. 2000), it is still desirable to have a genomic structural character to add to the growing body of evidence that defines this clade. The relationship among Elaeagnaceae, Rhamnaceae, and traditional Urticales in the newly established Rosales was only recognized starting with DNA sequence analyses (Chase et al. 1993; Qiu et al. 1998b; Soltis et

al. 2000). The *trans* splicing we report here in these taxa again adds to the evidence that helps to establish their affinities.

The distribution of *trans* splicing in Nymphaeales, Chloranthaceae, Winteraceae, *Syringodium filiforme*, *Erodium*, Juglandaceae, and Solanaceae corresponds to well-defined clades in all cases and, thus, are not discussed any further. According to a current molecular phylogeny (Les et al. 1997), *trans* splicing evolved independently in *Hydrocharis* and *Najas* (Fig. 3). However, because relationships in the Hydrocharitaceae–Najadaceae are still only poorly resolved, it is wise to wait until more evidence is available before a firm conclusion can be reached.

Acknowledgments. We thank Keming Song for technical help, Joe Peters for making a test version of the Immobilon-Ny+ transfer membranes available to us, and Olena Dombrovska for providing us a preprint. This work was supported by an NIH postdoctoral fellowship (GM17923-01), an NSF Career Award (DEB 0093012) to Y.-L.Q., and an NIH research grant to J.D.P. (GM-35087).

References

Adams KL, Daley DO, Qiu Y-L, Whelan J, Palmer JD (2000) Repeated, recent and diverse transfers of a mitochondrial gene to the nucleus in flowering plants. Nature 408:354–357

Adams KL, Qiu Y-L, Stoutemyer M, Palmer JD (2002) Punctuated evolution of mitochondrial gene content: High and variable rates of mitochondrial gene loss and transfer during angiosperm evolution. Proc Natl Acad Sci USA 99:9905–9912

Albach DC, Soltis PS, Soltis DE, Olmstead RG (2001) Phylogenetic analysis of asterids based on sequences of four genes. Ann Mo Bot Gard 88:163–212

Andre C, Levy A, Walbot V (1992) Small repeated sequences and the structure of plant mitochondrial genomes. Trends Genet 8:128–132

Beckert S, Muhle H, Pruchner D, Knoop V (2001) The mitochondrial *nad2* gene as a novel marker locus for phylogenetic analysis among early land plants: A comparative analysis in mosses. Mol Phylogenet Evol 18:117–126

Behnke H-D (1988) Sieve-element plastids, phloem protein, and evolution of flowering plants: III. Magnoliidae. Taxon 37:699–732

Binder S, Marchfelder A, Brennicke A, Wissinger B (1992) RNA editing in *trans*-splicing intron sequences of *nad2* mRNAs in *Oenothera* mitochondria. J Biol Chem 267:7615–7623

Chapdelaine Y, Bonen L (1991) The wheat mitochondrial gene for subunit I of the NADH dehydrogenase complex: A trans-splicing model for this gene-in-pieces. Cell 65:465–472

Chase MW, Soltis DE, Olmstead RG et al. (39 co-authors) (1993) Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene rbcL. Ann Mo Bot Gard 80:528–580

Chase MW, Duvall MR, Hills HG, et al. (1995a) Molecular phylogenetics of Lilianae. In: Rudall PJ, Cribb PJ, Cutler DF, Humphries CJ (eds) Monocotyledons: Systematics and evolution. Royal Botanic Gardens, Kew, pp 109–137

Chase MW, Stevenson DW, Wilkin P, Rudall PJ (1995b) Monocot systematics: A combined analysis. In: Rudall PJ, Cribb PJ, Cutler DF, Humphries CJ (eds) Monocotyledons: Systematics and evolution. Royal Botanic Gardens, Kew, pp 685–730

Chase MW, Soltis DE, Soltis PS, et al. (2000) Higher-level systematics of the monocotyledons: An assessment of current knowledge and a new classification. In: Wilson KL, Morrison DA (eds) Monocots: Systematics and evolution. CSIRO, Melbourne, pp 3–16

Chaw S-M, Parkinson CL, Cheng Y, Vincent TM, Palmer JD (2000) Seed plant phylogeny inferred from all three plant genomes: Monophyly of extant gymnosperms and origin of Gnetales from conifers. Proc Natl Acad Sci USA 97:4086–4091

Cho Y, Qiu Y-L, Kuhlman P, Palmer JD (1998) Explosive invasion of plant mitochondria by a group I intron. Proc Natl Acad Sci USA 95:14244–14249

Conklin PL, Wilson RK, Hanson MR (1991) Multiple *trans*-splicing events are required to produce a mature *nad1* transcript in a plant mitochondrion. Genes Dev 5:1407–1415

Cuenoud P, Savolainen V, Chatrou LW, Powell M, Grayer RJ, Chase MW (2002) Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, *atpB*, and *matK* DNA sequences. Am J Bot 89:132–144

Dai L, Zimmerly S (2002) Compilation and analysis of group II intron insertions in bacterial genomes: Evidence for retroelement behavior. Nucleic Acids Res 30:1091–1102

deSouza AP, Jubier M-F, Delcher E, Lancelin D, Lejeune B (1991) A trans-splicing model for the expression of the tripartite *nad5* gene in wheat and maize mitochondria. Plant Cell 3:1363–1378

Dombrovska O, Qiu Y-L (2004) Distribution of introns in the mitochondrial gene *nad1* in land plants: Phylogenetic and molecular evolutionary implications. Mol Phylogenet Evol (in press)

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

Endress PK, Igersheim A (1997) Gynoecium diversity and systematics of the Laurales. Bot J Linn Soc 125:93–168

Fauron C, Casper M, Gao Y, Moore B (1995) The maize mitochondrial genome: dynamic, yet functional. Trends Genet 11:228–235

French JC, Chung MG, Hur YK (1995) Chloroplast DNA phylogeny of the Ariflorae. In: Rudall PJ, Cribb PJ, Cutler DF, Humphries CJ (eds) Monocotyledons: Systematics and evolution. Royal Botanic Gardens, Kew, pp 255–275

Friis EM, Pedersen K, Crane PR (2001) Fossil evidence of water lilies (Nymphaeales) in the Early Cretaceous. Nature 410:357–360

Fukuzawa H, Kohchi T, Shirai H, Ohyama K, Umesono K, Inokuchi H, Ozeki H (1986) Coding sequences for chloroplast ribosomal protein S-12 from the liverwort *Marchantia polymorpha* are separated far apart on the different DNA strands. FEBS Lett 198:11–15

Goldschmidt-Clermont M, Choquet Y, Girard-Bascou J, Michel F, Schirmer-Rahire M, Rochaix J-D (1991) A small chloroplast RNA may be required for *trans*-splicing in *Chlamydomonas reinhardtii*. Cell 65:135–143

Graham SW, Olmstead RG (2000) Utility of 17 chloroplast genes for inferring the phylogeny of the basal angiosperms. Am J Bot 87:1712–1730

Gugerli F, Sperisen C, Büchler U, Brunner I, Brodbeck S, Palmer JD, Qiu Y-L (2001) The evolutionary split of Pinaceae from other conifers: Evidence from an intron loss and a multigene phylogeny. Mol Phylog Evol 21:167–175

Hashimoto K, Sato N (2001) Characterization of the mitochondrial *nad7* gene in *Physcomitrella patens:* Similarity with angiosperm *nad7* genes. Plant Sci 160:807–815

Jansen RK, Palmer JD (1987) A chloroplast DNA inversion marks an ancient evolutionary split in the sunflower family (Asteraceae). Proc Natl Acad Sci USA 84:5818–5822

Joly S, Brouillet L, Bruneau A (2001) Phylogenetic implications of the multiple losses of the mitochondrial *coxII.i3* intron in the angiosperms. Int J Plant Sci 162:359–373

Knoop V, Schuster W, Wissinger B, Brennicke A (1991) *Trans*-splicing integrates an exon of 22 nucleotides into the *nad5* mRNA in higher plant mitochondria. EMBO J 10:3483–3493

Knoop V, Unseld M, Marienfeld J, Brandt P, Sunkel S, Ullrich H, Brennicke A (1996) Copia-, gypsy and LINE-like retrotransposon fragments in the mitochondrial genome of *Arabidopsis thaliana*. Genetics 142:579–585

Knoop V, Altwasser M, Brennicke A (1997) A tripartite group II intron in mitochondria of an angiosperm plant. Mol Gen Genet 255:269–276

Koller B, Fromm H, Galun E, Edelman M (1987) Evidence for *in vivo trans* splicing of premessenger-RNAs in tobacco chloroplasts. Cell 48:111–119

Kress WJ, Prince LM, Hahn WJ, Zimmer EA (2001) Unraveling the evolutionary radiation of the families of the Zingiberales using morphological and molecular evidence. Syst Biol 50:926–944

Kubo T, Nishizawa S, Sugawara A, Itchoda N, Estiati A, Mikami T (2000) The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNACys (GCA). Nucleic Acids Res 28:2571–2576

Kück U, Choquet Y, Schneider M, Dron M, Bennoun P (1987) Structural and transcriptional analysis of two homologous genes for the P700 chlorophyll alpha apoproteins in *Chlamydomonas reinhardtii*: Evidence for in vivo trans-splicing. EMBO J 6:2185–2196

Kudla J, Albertazzi FJ, Blazevic D, Hermann M, Bock R (2002) Loss of the mitochondrial *cox2* intron 1 in a family of monocotyledonous plants and utilization of mitochondrial intron sequences for the construction of a nuclear intron. Mol Genet Genom 267:223–230

Les DH, Cleland MA, Waycott M (1997) Phylogenetic studies in Alismatidae. II: Evolution of marine angiosperms (seagrasses) and hydrophily. Syst Bot 22:443–463

Lew KA, Manhart JR (1993) They *rps12* gene in *Spirogyra maxima* (Chlorophyta) and its evolutionary significance. J Phycol 29:500–505

Mabberley DJ (1987) The plant book. Cambridge University Press, Cambridge

Malek O, Knoop V (1998) Trans-splicing group II introns in plant mitochondria: The complete set of *cis*-arranged homologs in ferns, fern allies, and a hornwort. RNA 4:1599–1609

Malek O, Brennicke A, Knoop V (1997) Evolution of *trans*-splicing plant mitochondrial introns in pre-Permian times. Proc Natl Acad Sci USA 94:553–558

Manhart JR, Palmer JD (1990) The gain of two chloroplast tRNA introns marks the green algal ancestors of land plants. Nature 345:268–270

Mathews S, Donoghue MJ (1999) The root of angiosperm phylogeny inferred from duplicate phytochrome genes. Science 286:947–950

Notsu Y, Masood S, Nishikawa T, Kubo N, Akiduki G, Nakazono M, Hirai A, Kadowaki K (2002) The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. Mol Genet Genom 268:434–445

Palmer JD, Herbon LA (1988) Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. J Mol Evol 28:87–97

Parkinson CL, Adams KL, Palmer JD (1999) Multigene analyses identify the three earliest lineages of extant flowering plants. Curr Biol 9:1481–1485

Pruchner D, Nassal B, Schindler M, Knoop V (2001) Mosses share mitochondrial group II introns with flowering plants, not with liverworts. Mol Genet Genom 266:608–613

Pryer KM, Schneider H, Smith AR, Cranfill R, Wolf PG, Hunt JS, Sipes SD (2001) Horsetails and ferns are a monophyletic group and the closest relatives to seed plants. Nature 409:618–622

Qiu Y-L, Palmer JD (1997) Mitochondrial genome evolution and land plant phylogeny. In: Wessler SR, Freeling M, Meyerowitz EM (eds) Proceedings of the Keystone Symposia on Molecular and Cellular Biology: Evolution of Plant Development.

Qiu Y-L, Chase MW, Les DH, Parks CR (1993) Molecular phylogenetics of the Magnoliidae: cladistic analyses of nucleotide sequences of the plastid gene rbcL. Ann Mo Bot Gard 80:587–606

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

Qiu Y-L, Chase MW, Hoot SB, Conti E, Crane PR, Sytsma KJ, Parks CR (1998b) Phylogenetics of the Hamamelidae and their allies: parsimony analyses of nucleotide sequences of the plastid gene rbcL. Int J Plant Sci 159:891–905

Qiu Y-L, Lee J, Bernasconi-Quadroni F, Soltis DE, Soltis PS, Zanis M, Zimmer EA, Chen Z, Savolainen V, Chase MW (1999) The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. Nature 402:404–407

Qiu Y-L, Lee J, Bernasconi-Quadroni F, Soltis DE, Soltis PS, Zanis M, Zimmer EA, Chen Z, Savolainen V, Chase MW (2000) Phylogeny of basal angiosperms: Analyses of five genes from three genomes. Int J Plant Sci 161:S3–S27

Rabbi MF, Wilson KG (1993) The mitochondrial *coxII* intron has been lost in two different lineages of dicots and altered in others. Am J Bot 80:1216–1223

Raubeson LA, Jansen RK (1992) Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. Science 255:1697–1699

Soltis DE, Soltis PS, Chase MW, et al. (2000) Angiosperm phylogeny inferred from 18S rDNA, *rbcL* and *atpB* sequences. Bot J Linn Soc 133:381–461

Stevenson DW, Loconte H (1995) Cladistic analysis of monocot families. In: Rudall PJ, Cribb PJ, Cutler DF, Humphries CJ (eds) Monocotyledons: Systematics and evolution. Royal Botanic Gardens, Kew, pp 543–578

Thomson MC, MacFarlane JL, Beagley CT, Wolstenholme DR (1994) RNA editing of *mat-r* transcripts in maize and soybean increases similarity of the encoded protein to fungal and bryophyte group II intron maturases: evidence that *mat-r* encodes a functional protein. Nucleic Acids Res 22:5745–5752

Turmel M, Otis C, Lemieux C (2002a) The complete mitochondrial DNA sequence of *Mesostigma viride* identifies this green alga as the earliest green plant divergence and a highly compact mitochondrial genome in the ancestor of all green plants. Mol Biol Evol 19:24–38

Turmel M, Otis C, Lemieux C (2002b) The chloroplast and mitochondrial genome sequences of the charophyte *Chaetosphaeridium globosum*: Insights into the timing of the events that restructured organelle DNAs within the green algal lineage that led to land plants. Proc Natl Acad Sci USA 99:11275–11280

Unseld M, Marienfeld JR, Brand P, Brennicke A (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. Nature Genet 15:57–61

Upchurch GRJ (1984) Cuticle evolution in Early Cretaceous angiosperms from the Potomac Group of Virginia and Maryland. Ann Mo Bot Gard 71:522–550

Wahleithner JA, MacFarlane JL, Wolstenholme DR (1990) A sequence encoding a maturase-related protein in a group II intron of a plant mitochondrial nad1 gene. Proc Natl Acad Sci USA 87:548–552

Wikstrom N, Savolainen V, Chase MW (2001) Evolution of the angiosperms: Calibrating the family tree. Proc Roy Soc Lond B 268:2211–2220

Williams JH, Friedman WE (2002) Identification of diploid endosperm in an early angiosperm lineage. Nature 415:522–526

Wissinger B, Schuster W, Brennicke A (1991) Trans-splicing in *Oenothera* mitochondria: *nad1* mRNA are edited in exon and trans-splicing group II intron sequences. Cell 65:473–482

Zaita N, Torazawa K, Shinozaki K, Sugiura M (1987) *Trans* splicing *in vivo* joining of transcripts from the divided gene for ribosomal protein S12 in the chloroplasts of tobacco. FEBS Lett 210:153–156