Systematic Botany (1996), 21(2): pp. 135–151

Multiple Independent Losses of the *rpoC1* Intron in Angiosperm Chloroplast DNA's

STEPHEN R. DOWNIE[1], ESMERALDA LLANAS, and DEBORAH S. KATZ-DOWNIE

Department of Plant Biology, University of Illinois, Urbana, Illinois 61801
[1]Author for correspondence

Communicating Editor: Elizabeth A. Kellogg

ABSTRACT. Previous studies have shown that in *Marchantia*, black pine, and the vast majority of angiosperms examined to date, the chloroplast gene *rpoC1* is interrupted by an intron (of about 750 base pairs), but that in the grasses and one of three subfamilies of cacti (Cactoideae) this intron has been lost. DNA's of the complete *rpoC1* intron region and portions of the flanking exon regions from 107 species (representing 54 families) of angiosperms were amplified with the polymerase chain reaction (PCR) in order to uncover other instances of intron loss. In addition to grass and Cactoideae chloroplast DNA's, we report that the *rpoC1* intron is missing from the chloroplast genomes of *Scaevola* and *Goodenia* (the only two representatives of Goodeniaceae examined), four of ten species of *Passiflora* (Passifloraceae), two of four genera of Aizoaceae (*Delosperma* and *Faucaria* but not *Monilaria* or *Tetragonia*), and one of the eight taxa of *Medicago* examined (*M. suffruticosa* subsp. *leiocarpa*; Fabaceae). Mirroring other instances of intron loss from the chloroplast genome, sequence analysis reveals that, for these taxa, the *rpoC1* gene has undergone a precise deletion of the intron. For those taxa with the intron, minimal size variation is apparent within the region and, in all taxa investigated, the intron lies in precisely the same position in the gene. Parsimony analyses of partial exon and intron *rpoC1* nucleotide sequences reveal that the latter, owing to considerable conservation of secondary structure in that region of the intron sequenced, can supply useful characters for phylogenetic analysis at high taxonomic levels.

The chloroplast *rpoC* region, homologous to the β' subunit of *Escherichia coli* (Migula) Castellani & Chalmers RNA polymerase, is divided into two genes, *rpoC1* and *rpoC2*, and is located in the large single-copy region of the plastid genome in most angiosperms (Cozens and Walker 1986; Hudson et al. 1988). These genes, along with *rpoB*, are cotranscribed as a single operon and encode three subunits of the chloroplast RNA polymerase (Hudson et al. 1988). The gene *rpoC1* is of systematic interest because it is known to be interrupted by a single intron in most, but not all, land plants. DNA sequencing has revealed that this intron is present in the liverwort *Marchantia polymorpha* L. (Ohyama et al. 1986), black pine (*Pinus thunbergii* Parl.; Wakasugi et al. 1994), tobacco (*Nicotiana tabacum* L. 'Bright Yellow 4'; Shinozaki et al. 1986), spinach (*Spinacia oleracea* L.; Hudson et al. 1988), and numerous species of Apiaceae and relatives (Downie et al. 1996, and unpubl. data) but absent in rice (*Oryza sativa* L. 'Nihonbare'; Hiratsuka et al. 1989; Shimada et al. 1990), maize (*Zea mays* L.; Igloi et al. 1990), and from all examined representatives of one subfamily of Cactaceae (Cactoideae; Wallace and Cota 1996). The absence of an intron in *rpoC1*, based on heterologous filter hybridization and polymerase chain reaction (PCR) experiments, has also been noted in several other grasses (*Agropyron*, *Avena*, *Bambusa*, *Cynodon*, *Danthonia*, *Dendrocalymus*, *Lolium*, *Secale*, and *Triticum*; Katayama and Ogihara 1993; S. Downie and J. Palmer, unpubl. data). Thus, it is clear that the intron was present in the common ancestor of angiosperms and subsequently lost independently in the lineages leading to the grasses and Cactoideae.

Owing to the conservative nature of chloroplast genome evolution among photosynthetic angiosperms, particularly with regard to its gene and intron content, major structural rearrangements were thought to be relatively infrequent events (Palmer et al. 1988; Palmer 1991; Downie and Palmer 1992). However, as broad groups of angiosperms are surveyed for specific structural mutations, an increasing number of these characters are being detected. Indeed, what was once considered to be a relatively infallible class of phylogenetic character because of its rarity, some major structural alterations of the chloroplast genome, such as the loss of the *rpl2* intron, are gaining notoriety as being highly homoplastic (Downie et al. 1991; Downie and Palmer 1992; Downie et al. 1994; Doyle et al. 1995). Other homoplastic chloroplast intron losses include the *cis*-spliced *rps12* intron in Ranunculaceae and Cuscutaceae (Hoot and Palmer

1994; Freyer et al. 1995), and the *rpl16* intron in Geraniaceae and Plumbaginaceae (Downie and Palmer 1992). Nevertheless, despite their homoplasy, intron loss characters are still powerful indicators of relationship within the different lineages in which they have occurred (Downie et al. 1991; Doyle et al. 1995).

Here we present the results of a PCR survey constructed to detect the presence/absence of the intron in the chloroplast gene *rpoC1* across a broad representation of angiosperm species. Sequence analysis was used to confirm both the absence of the intron and, when present, its conserved location within the gene. Additionally, the availability of partial *rpoC1* exon and intron nucleotide sequences for some of these taxa permitted us to use these data in a cladistic analysis. With this information, the relative utility of exon and intron sequences in providing characters to infer relationships across a comparable set of taxa can be assessed.

Materials and Methods

One hundred and seven species from 54 angiosperm families, representing members of all six subclasses of dicots and three of five subclasses of monocots sensu Cronquist (1981), were surveyed for the presence or absence of the *rpoC1* intron (Appendix 1). Included here, as controls, were tobacco, previously reported to contain the intron in *rpoC1* (Shinozaki et al. 1986), and wheat (*Triticum aestivum* L.), previously reported not to contain the intron (Katayama and Ogihara 1993). Total genomic DNA's were isolated from fresh leaf material of one or, rarely, more individual plants using the modified CTAB procedure of Doyle and Doyle (1987) and further purified by centrifugation in cesium chloride-ethidium bromide gradients. Leaves or seeds were either collected directly from the field or obtained from the various institutions and individuals cited in Appendix 1. For a few species, DNA's were supplied to us directly.

The complete *rpoC1* intron region and portions of the flanking exon regions in each genomic DNA were PCR-amplified using forward primer "5′*rpoC1* exon" and reverse primer "3′*rpoC1* exon" in an equimolar ratio. In tobacco chloroplast DNA (cpDNA), the *rpoC1* intron falls between sequence coordinates 23,094 and 23,834 and is 738 base pairs (bp) in size (Shinozaki et al. 1986; Shimada et al. 1990). Primers were designed by comparing the *rpoC1* exon sequences from tobacco, rice, and *Marchantia* and choosing regions highly conserved among the taxa. In tobacco, the 3′ end of primer "5′*rpoC1* exon" is 52 bp upstream from the exon 1/intron junction, whereas the 3′ end of primer "3′*rpoC1* exon" is 354 bp downstream from the exon 2/intron junction (Shinozaki et al. 1986). Primer sequences, written 5′ to 3′, are as follows: 5′*rpoC1* exon—ACGTCTTCCTAGYTAYATHGC; 3′*rpoC1* exon—AATAGACAYAANACCATCCA. The primers were synthesized by Operon Technologies, Inc. (Alameda, CA). Details of the amplification reactions, and the DNA purification and sequencing strategies used, were the same as outlined in Downie and Katz-Downie (1996) with the only exception being a reduction in the volume of each reaction (25 µl instead of 100 µl) in the PCR survey. The ensuing PCR fragments were separated by electrophoresis in 1% agarose gels, stained with ethidium bromide, and sized against *Eco*RI/*Hin*dIII-digested lambda DNA standards. Each set of reactions was monitored by the inclusion of positive (tobacco and wheat DNA's) and negative (no template) controls. Successful PCR amplifications resulted in a single DNA band of about 1,100 bp when the intron was present, or about 400 bp when the intron was absent.

To confirm the suspected loss and precise excision of the *rpoC1* intron, all but two of the small-sized (i.e., 400 bp) PCR fragments were sequenced using the same pair of primers identified above. Both strands of the DNA were completely sequenced. Additionally, for some of those taxa producing the larger (i.e., 1.1 kilobase) PCR product, the presence of the intron and its conserved position within the gene were confirmed by sequencing portions of the exons and adjacent intron regions. These sequence data were also used in the estimation of relationships (described below). The large size of these PCR-fragments and the use of only two external sequencing primers precluded the sequencing of both DNA strands in their entirety. However, due to the lack of any detectable sequence ambiguity (i.e., compressions or hard stops) and the presence, in the exon regions, of an intact reading frame undisturbed by insertions or deletions, we believe that the quality of our readings has not been jeopardized by sequencing only one strand. Furthermore, each set of sequences was read critically from the autoradiographs by three independent readers, entered into the computer, and then rechecked against the autoradiographs for errors. Alignment of these partial exon and intron sequences was done manually (and, for the latter, these sequences were compared to

conserved regions documented for *rpoC1* and other group II introns and used to guide the alignment when appropriate; Michel et al. 1989). Boundaries of the exon and intron regions were determined by comparing the DNA sequences to the corresponding boundaries in tobacco (Shinozaki et al. 1986) and consensus splice sites in other plants for group II introns (Michel et al. 1989). Pairwise nucleotide differences of all positions were determined using the distance matrix option in PAUP version 3.1.1 (Swofford 1993). Thus, these divergence values were calculated simply as a proportion of divergent sites in each pairwise comparison with no provision made to account for superimposed events (multiple hits) which must have occurred at many positions.

The resulting (partial exon and intron) data matrices were analyzed separately and together by assuming unordered character states (i.e., Fitch parsimony) using PAUP run on a Power Macintosh 8100/100 AV computer. All heuristic searches were replicated 500 times with random addition sequence and tree bisection-reconnection branch swapping. The options mulpars, steepest descent, collapse, and acctran optimization were selected. All searches were performed using equal character weighting. Bootstrap values (Felsenstein 1985) were calculated from 100 replicate analyses using the heuristic search strategy and simple addition sequence of the taxa. "Decay" analyses (Bremer 1988) were also used to test the robustness of particular clades. However, owing to the large number of trees obtained upon the relaxation of parsimony and the memory limitations of the computer, the analyses of trees three, four, and seven steps longer than those most parsimonious (derived from the cladistic analyses of the exon, intron, and combined data matrices, respectively) could not be done. Transition/transversion ratios were calculated using MacClade version 3.01 (Maddison and Maddison 1992).

RESULTS

All 107 species of angiosperms surveyed for the presence or absence of the *rpoC1* intron produced a single, bright DNA band after amplification. These taxa, representing both monocots and dicots, belong to lineages that may have diverged as much as 200 million years ago (Wolfe et al. 1989). As these primers were constructed from *rpoC1* exon sequence regions highly conserved among *Marchantia*, tobacco, and rice cpDNA's, it is not surprising that these primers should anneal to a multitude of different angiosperm species.

Our survey revealed two major size categories of PCR products—the first corresponding to the presence of the intron and the second corresponding to its loss. Major length variants in *rpoC1* intron sequences across these taxa were non-existent, as the vast majority of products surveyed were similar in size to each other and to that observed for tobacco. Detectable size variation was only evident in a dozen species, differing in size from the tobacco PCR fragment by about 100 bp or less. As discussed in Doyle et al. (1995), most of the group II introns present in cpDNA are within a few hundred bp of the minimum size (ca. 500 bp) required for intron splicing. Thus, other than the loss of the entire intron itself, major deletions are not expected.

The PCR experiments indicated that most of the angiosperm cpDNA's examined contain an intron in *rpoC1*. On the basis of these same experiments, the intron was inferred to be absent in three of five species of Aizoaceae (*Delosperma cooperi*, *D. nubigenum*, and *Faucaria tigrina*), both examined representatives of Goodeniaceae (*Goodenia ovata* and *Scaevola sericea*), four of ten species of *Passiflora* (Passifloraceae), and one of the eight taxa of *Medicago* examined (*M. suffruticosa* subsp. *leiocarpa*; Fabaceae; Table 1). DNA sequencing of the *rpoC1* region in *Delosperma cooperi* and *Faucaria*, both species of Goodeniaceae, *Passiflora helleri*, *P. morifolia*, and *P. suberosa*, and *Medicago suffruticosa* subsp. *leiocarpa* confirmed that the intron is indeed missing from this region in these taxa. Owing to similarly-sized PCR fragments, it is expected that *Passiflora capsularis* and *Delosperma nubigenum* also lack the intron. Furthermore, these sequence data show that the *rpoC1* gene has undergone a precise deletion of the intron with the two remaining exons juxtaposed into a single, uninterrupted gene. This excision coincides precisely with established splice sites for group II introns (Michel et al. 1989) and leads to the union of the two coding regions without disrupting the reading frame. This precise excision of an intron has been reported from other chloroplast genes (Hiratsuka et al. 1989; Igloi et al. 1990; Downie et al. 1991; Freyer et al. 1995; Doyle et al. 1995; Wallace and Cota 1996). The mechanism of intron loss might involve reverse transcription of the spliced RNA followed by reintegration into the genome at precisely the same location (Hiratsuka et al. 1989; Igloi et al. 1990; Downie et al. 1991), but other undocumented processes theoretically could effect intron removal, such as random DNA-level dele-

TABLE 1. Distribution of the chloroplast *rpoC1* intron in angiosperm families Aizoaceae, Goodeniaceae, Passifloraceae, and Fabaceae. [The intron is also absent from all examined Poaceae and Cactoideae (Cactaceae) cpDNA's; see text for further discussion.] Subfamilial classification of Aizoaceae based on Bittrich and Hartmann (1988); infrageneric classifications of *Passiflora* and *Medicago* based on Killip (1938) and Small and Jomphe (1989), respectively. Asterisks denote those species whose intron absence, or presence, has been confirmed by DNA sequencing.

	Subfamily	
AIZOACEAE		
RpoC1 Intron Absent		
Delosperma cooperi (Hook. f.) L.*	Ruschioideae	
Delosperma nubigenum (Schltr.) L.	Ruschioideae	
Faucaria tigrina (Haw.) Schwantes*	Ruschioideae	
RpoC1 Intron Present		
Monilaria moniliformis (Thunb.) H. D. Ihlenf. & S. Jörg.	Ruschioideae	
Tetragonia tetragonioides (Pall.) Kuntze*	Tetragonioideae	
GOODENIACEAE		
RpoC1 Intron Absent		
Goodenia ovata Smith*		
Scaevola sericea Vahl*		
	Subgenus	**Section**
PASSIFLORACEAE		
RpoC1 Intron Absent		
Passiflora capsularis L.	*Plectostemma*	*Xerogona*
P. helleri Peyr.*	*Plectostemma*	*Decaloba*
P. morifolia Mast.*	*Plectostemma*	*Cieca*
P. suberosa L.*	*Plectostemma*	*Cieca*
RpoC1 Intron Present		
Passiflora caerulea L.	*Granadilla*	
P. coccinea Aubl.	*Distephana*	
P. eichleriana Mast.	*Granadilla*	
P. foetida L.*	*Dysosmia*	
P. incarnata L.*	*Granadilla*	
P. quadrangularis L.	*Granadilla*	
		Section
FABACEAE		
RpoC1 Intron Absent		
Medicago suffruticosa Ramond ex DC. subsp. *leiocarpa* (Benth.) P. Fourn.*		*Medicago*
RpoC1 Intron Present		
M. polymorpha L.		*Spirocarpos*
M. rigidula (L.) All.		*Spirocarpos*
M. ruthenica (L.) Ledebour		*Platycarpae*
M. sativa L.*		*Medicago*
M. sauvagei Nègre		*Spirocarpos*
M. truncatula Gaertn.		*Spirocarpos*
M. turbinata (L.) All.		*Spirocarpos*
(Representatives from 13 other genera of Fabaceae also have the intron; see Appendix 1)		

tion and gene conversion between an intron-containing gene and its spliced transcript (Palmer 1991).

For those DNA's where an intron was inferred to be present using PCR, sequencing of the *rpoC1* intron region in 29 taxa (Appendix 1) confirmed its presence and showed that the intron in each of these DNA's lies in precisely the same position in the gene. This position is identical to that reported for *Marchantia*, black pine, tobacco, and spinach *rpoC1* sequences (Ohyama et al. 1986; Shinozaki et al. 1986; Hudson et al. 1988; Wakasugi et al. 1994).

The six families in which at least some members lack an intron in chloroplast gene *rpoC1* [i.e., Aizoaceae (Caryophyllidae), Goodeniaceae (Asteridae), Passifloraceae (Dilleniidae), Poaceae (Com-

melinidae), Cactaceae (Caryophyllidae), and Fabaceae (Rosidae)] represent a diverse array of plants assigned to five subclasses of angiosperms that include both monocots and dicots. Other than Aizoaceae and Cactaceae, which have been inferred by some authors to be closely related (discussed below), these families are distantly related to one another in systems of classification based largely on morphology (e.g., Cronquist 1981) and in phylogenies based on cpDNA *rbcL* or *ndhF* gene sequences (e.g., Chase et al. 1993; Kim and Jansen 1995, and unpubl. data). Thus, we infer that the *rpoC1* intron has been lost at least six times during the evolution of the angiosperms.

The primary goal of this study was to survey for the presence/absence of the *rpoC1* intron across a broad representation of angiosperms. DNA sequencing was used to confirm the intron's conserved position within the gene or its loss, where appropriate. However, the availability of the partial *rpoC1* exon and intron sequences permitted us to use these data in a parsimony analysis. This study was not designed to be an exhaustive comparative analysis of *rpoC1* sequences among angiosperms—the selection of taxa was based primarily on those DNA's we had on hand, the availability of published sequences, and those species thought to be appropriate outgroups to those taxa lacking the intron. Poor sampling density, the small sizes of the regions, and the presence of missing data (see Figs. 1 and 2), are limitations to obtaining the best interpretation of relationship. Our purpose is not to attempt to resolve issues of relationships among these taxa but to compare the relative utility of *rpoC1* exon and, especially, intron sequences in providing characters to infer relationships across a comparable set of taxa.

Alignment of partial *rpoC1* exon and intron nucleotide sequences for all but a few of the taxa resulted in matrices of 342 and 238 positions, respectively (Figs. 1 and 2). The exon portions sequenced flank the intron and comprise 21 bp from exon 1 and 321 bp from exon 2, or about 17% of the entire gene excluding the intron. For those taxa lacking the intron, this region is contiguous; for those taxa with the intron, the two exon portions have been placed end to end. The portion of the intron sequenced ranges in size from 195 to 233 bp, averages 212 bp, and comprises about one-third of the length of the entire intron. Only the sequence data at the 5′ end of the intron (i.e., proximal to exon 1) were used; data from the 3′ end were minimal (or missing for several taxa). The *Marchantia rpoC1* intron was too divergent to include. No gaps were necessary to align the exon sequences, whereas 22 gaps (of one to sixteen bp, relative to the outgroup spinach) were required to align the intron sequences. For the latter, all positions were included in the phylogenetic analysis even though two regions within the matrix exist where alternative alignments are possible (i.e., positions 161–172, and 206–219; Fig. 2). Deleting these 26 positions did little to affect the ensuing tree topology.

Of the 342 aligned exon positions for 38 taxa, 137 (40%) were unvarying, 59 (17%) were autapomorphic, and the remaining 146 (43%) were potentially informative phylogenetically. Of the 238 aligned positions from the intron region for 27 taxa, 128 (54%) were unvarying, 38 (16%) were autapomorphic, and 72 (30%) were potentially informative. To make these comparisons equitable, the 38-taxon exon matrix (which includes *Marchantia* and two grass species) was pruned to the same 27 taxa as contained in the intron data matrix. Subsequently, the number of conserved positions in this reduced exon data set increased to 199 (58.2%), the number of potentially informative positions decreased to 83 (24.3%), and the number of autapomorphies remained about the same (60 or 17.5%). In direct pairwise comparisons of all dicot exon positions, sequence divergence values ranged from 0.3 (between *Villarsia* and *Nymphoides*, both Menyanthaceae) to 21.3% of nucleotides (between *Goodenia* and *Passiflora suberosa*). In the 38-taxon exon matrix the greatest divergence values, reaching a maximum of 33%, were observed between *Marchantia* and the two grasses. For the intron sequence comparisons (excluding gaps), divergence values were similar to those obtained for the exons, ranging from identity (between *Gamocarpha* and *Calycera*, both Calyceraceae) to 23.3% of nucleotides (between *Coriandrum* and *Bixa*).

Parsimony analysis of the exon sequences for 38 taxa resulted in four minimal-length trees, whose strict consensus tree with accompanying bootstrap values and decay indices is shown in Fig. 3a. Each of these four trees had a length of 514 steps, a consistency index (CI) of 0.586 (all characters) and 0.525 (excluding uninformative characters), and a retention index (RI) of 0.691. Steps calculated over all trees by codon position indicated that most changes occurred, as expected, in the third position (271–272) followed by first (138) and second (104–105) position. Analysis of the intron data generated three maximally-parsimonious topologies of 232 steps (CI = 0.685 and 0.605, with and

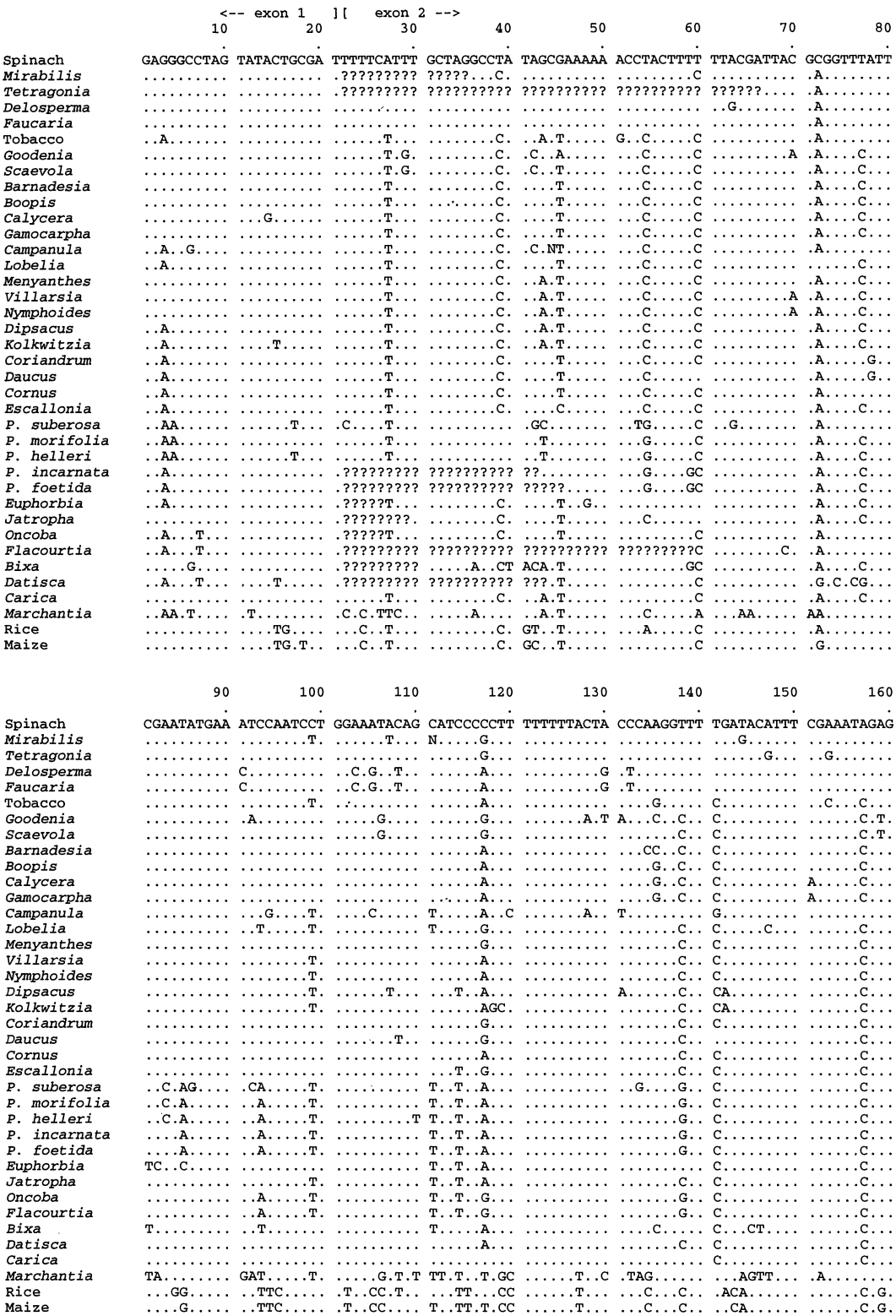


FIG. 1. Aligned nucleotide sequences from the cpDNA *rpoC1* exon regions. For those taxa with an intron, these sequences represent 21 bp from exon 1 (positions 1–21) and 321 bp from exon 2 (positions 22–342); these exon regions flank the intron and have been, for the purpose of this figure, placed end to end. For those taxa lacking the intron (Table 1), the coding region is contiguous and appears as illustrated. Nucleotide identity with the first sequence is indicated by a dot; missing data are indicated by question marks; N = nucleotide of uncertain state. Complete taxon names are provided in Appendix 1.

```
                    170        180        190        200        210        220        230        240
                     .          .          .          .          .          .          .          .
Spinach     AAATATCTAC TGGAGCAGGT GCTATTCGAG AACAATTAGC CGATTTGGAT TTGCGTACTA TTATAGATTA TTCATTCGCA
Mirabilis   .......... .........C .......... .......... .......... .....A.... .......... ..........
Tetragonia  .......G.. .......... .......... .......... .......... .....A.... .......... ..........
Delosperma  .......... .......... .......... .......... .......... .....A...G ........C ..........
Faucaria    .......... .......... .......... .......... .......... .....A...G .........C ..........
Tobacco     ....C..... ...C..G... .....C.... .......... ....C.A... ..A..A.T.. .......GA. ...G..G.T.
Goodenia    .G..C..... C..TT..... .....AA... .G........ ....C.A... ..A..A.T.C ........G. ......G.T.
Scaevola    .G..C..... C..GT..... .....CA... .......... ....C.A... ..A..A.G.C .......... ......GTT.
Barnadesia  .G..C..... C..G...A.. .....CA... .......... ....C.A... ..A..A.T.. .......... ......G.T.
Boopis      .G..C..... C..G...A.. .....CA... .......... ......A... ..A..A.T.. .......... ...G..G.T.
Calycera    .G..C..G.. C..G...A.. .....CA... .......... ......A... ..A..A.T.. .......... ...G..G.T.
Gamocarpha  .G..C..G.. C..G...AA. .....CA... .......... ......A... ..A..A.T.. .......... ...G..G.T.
Campanula   .G..CG.C.. .......... T....C.... .......... G...C.A... ..A..A.AA. ...G.C...C ........TT.
Lobelia     .G..C..C.. .......... .....C.... .......G.. ....C.A... ..A..C.GA. ..C..A...C ......GTT.
Menyanthes  .G..C..... ...G...... .....C.... .......... T...C.A... ..A..A.TG. .......... ......G.T.
Villarsia   .G..C..... ...G...... .....C.... .......... T...C.A... ..A..A.T.. ..C....... ......G.T.
Nymphoides  .G..C..... ...G...... .....C.... .......... T...C.A... ..A..A.T.. ..C....... ......G.T.
Dipsacus    .G..C..... .......... .....C.... .......... ....C.A... ..AG.A.T.C .........C ...G..G.T.
Kolkwitzia  .G..C..... .......... .....C.... .......... ....C.A... ..A..A.T.G ..G....... ...G..G.T.
Coriandrum  .G..C..G.. C......... .....C.... .......... T...C.A... ..A..G.T.. .........C ...G..G.T.
Daucus      .G..C..... .......... .....C.... .......... ....C.C... ..A..G.T.. .........C ......G.T.
Cornus      .G..C..... .......... .....C.... .......... ....C.A... ..A..A.T.. .......... ...C..G.T.
Escallonia  .G..N..... .......... .......... .......... ....C.A... ..A..A.T.. .......... ......G.T.
P. suberosa ....C..C.. ......C... .....C.... .......... .A..C..... ....AA.T.. .....A..G. .....CG...
P. morifolia ....C..... ......C... .....C.... .......... ....C..... ....AA.T.. .......C.. .....CG.T.
P. helleri  ....C..... ......C... .....C.... .......... ....C..... ....AA.T.. .......C.. .....CG.T.
P. incarnata ....C..... ......C... .....C.... .......... ....C..... ...AAA.T.. .......... .....CG.T.
P. foetida  ....C..... ......C... .....C.... .......... ....C..... ...AAA.T.. .......... .....CG.T.
Euphorbia   ....T..... A.....T..C .....C.... .......... T...C..... .....A.T.. .........G .....CG.T.
Jatropha    ....T..... A.....T... .....C.... .......... ....C..... .....A.T.. .......... .....CG.T.
Oncoba      ....T..... ......C... .....C.... ...T...... ....C..... .....A.T.. ..C....... ...G.CG.T.
Flacourtia  ....T..... ......C... .....C.... ...T...... ....C..... .....A.T.. ..C....... ...G.CG.T.
Bixa        ....T..... .......... .....C.... .......... ....C.A... .....A.T.. .........T ...G.GG.T.
Datisca     ....T..... A......... .....C.... .......... T...C..... .....ACT.. ..C....... ...G..G.T.
Carica      ....T..... C..G..G... .....C.... .......... ....C.A... ....AA.T.. .......... ...G...TT.
Marchantia  ....CG.A.. .....G..A. .......A.A ........A. TA....AA.C ..A.AA.A.G .A...A..CT .G..CA.TT.
Rice        ....TG.G.. A......... ......A... .......... A......... .....A.T.. ..T....GA. ...C.CG.TC
Maize       ....TG.A.. A........C ......A... .......... A.....A... .....A.T.. .......GA. ...T..G.TT

                    250        260        270        280        290        300        310        320
                     .          .          .          .          .          .          .          .
Spinach     GAATGGAAAG AGTTAGGGGA AGAAGGGTCT ACAGGAAATG AATGGGAAGA CCGAAAAGTT GGAAGAAGAA AGGATTTTTT
Mirabilis   .......... .......... G......C.. .....G.... .......... .......... .......... ..........
Tetragonia  .......... .......... .......A.. .....G.... .......... .......... .......... ..........
Delosperma  .......... .......... .....A.C.. .....G.... .......... .......... A......... ..........
Faucaria    .????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
Tobacco     ......G... .A..G..... .......CAC .....G.... .......... T......... .......... ....C.....
Goodenia    .......... .A..G..A.. ......ACTC .....C.... .G........ T....TC... .A........ .AT.......
Scaevola    .......... .A..G..A.. ......ACTC .....G.... .G........ T....TC... .A........ .A......C.
Barnadesia  .......... .A..G..... ...G..AC.C .....G.... .......... T......... .......... .A........
Boopis      .......... .A..G..... ...G..AC.C .....G.... .......... T......... .......... .A........
Calycera    .......... .A..G..... ...G..AC.C .....G.... .......... T......... .......... .A........
Gamocarpha  .......... .A..G..... ...G..AC.C .....G.... .......... T......... .......... .A........
Campanula   ........GA .A..G.C.A. .A....TC.C ..G..G...A ......C... T.A...G... C......... .A........
Lobelia     .......... .A..G..... ......TC.C .....G.... .......... T.....G... A..C...... .A...C....
Menyanthes  .......... .A..G..... .......C.C .....G.... .......... T......... .......... .A........
Villarsia   .......... .A..G..... .......C.C .....G.... .......... T......... .......... .A........
Nymphoides  .......... .A..G..... ...G...C.C .....G.... .......... T......... .......... .A........
Dipsacus    .......... .A..G..... ..CG.A.CT. .....G.... .......... T......... .......... ...C......
Kolkwitzia  .......... .A..G..... ...G...CT. .....G.... .......... T......... .......... ..........
Coriandrum  .......... ....G..... ...C..NN.C .....G.... .G........ T.....G... ......C... ..........
Daucus      .......... ....G..... ...C..NN.C .....G.... .G........ T......... ......C... ..........
Cornus      .......... .A........ .......C.C .....G.... .......... T......A.. ........G. ..........
Escallonia  .......... .A..G..... .......C.C .......... .G........ T......... .......... ..........
P. suberosa .......... .A....AA.. .AGG...... ..G.A..... .....A.... T.T....... A......... ...C......
P. morifolia .......... .A.....A.. .A.G...... ..G..G.... .....A.... T.T....... A......... ..........
P. helleri  .......... .A.....A.. .A.G...... ..G..G.... .....A.... T.T....... A......... ..........
P. incarnata .......... .A.....A.. .A.G...... ..G..G.... .......... T......... A......... ..........
P. foetida  .......... .A.....A.. .A.GA..... ..G..G.... .......... T......... A......... ..........
Euphorbia   .......... .A........ .......C.C .....G.... .......... T.....G... .......... .A........
Jatropha    .......... .A........ .......C.. .....G.... .......... T......... .......... ..........
Oncoba      .......... .A........ ......AC.C ..C..G.... .......... T......... .......... ..........
Flacourtia  .......... .A..G..... ......AC.C ..C..G.... .......... T......... .......... ..........
Bixa        .......... .T.....A.. .......CTC ..G..T.... .......... T......A.. .......... ..........
Datisca     .......... .A........ .........C G.G..T.... .......... T........G .......... ..........
Carica      .......... .A.....A.. ...G..TC.C .....G.... .......... T......A.. .......... ..........
Marchantia  .......... ....T.CT.. .C..AAA..A ..T....... .......... TA.....A.. CAGC...... .A...C....
Rice        ........G. .A....AA.. C..G....A. .GT...G... .......... TA.....AGA C...T..... .A.T......
Maize       ........G. .A....AA.. C..G....A. .GT...G... .......... TA.....AGA C...T..... .A.T......
```

```
                    330         340
                     .           .
Spinach         AGTTAGACGC ATGGAATTAG TT
Mirabilis       .....A.... ..A..... C.
Tetragonia      .......... .......... C.
Delosperma      ......G... ....G..... C.
Faucaria        ?????????? ?????????? ??
Tobacco         G......... G.......G. C.
Goodenia        G......... ..AC...... C.
Scaevola        G......... ..CC...... C.
Barnadesia      G......... .......... C.
Boopis          G......... ........G. C.
Calycera        G......... ........GT C.
Gamocarpha      G......... ........GT C.
Campanula       GA........ ....G...G. C.
Lobelia         G......... ....G...G. C.
Menyanthes      G......... ........G. C.
Villarsia       G......... ........G. C.
Nymphoides      G......... ........G. C.
Dipsacus        GTC....... .......... C.
Kolkwitzia      GTC......G .......... C.
Coriandrum      G..C.....T .......... CG
Daucus          G..C.....T .......... CG
Cornus          G......... .......... C.
Escallonia      G......... .......... C.
P. suberosa     G......... G......... C.
P. morifolia    G......... G......... C.
P. helleri      G......... G......... C.
P. incarnata    G......... .......... C.
P. foetida      G......... G......... C.
Euphorbia       G...C..... G......... C.
Jatropha        G......... G......... C.
Oncoba          .......... G......... C.
Flacourtia      .......... G......... C.
Bixa            G......... ........G. C.
Datisca         G........T .......... C.
Carica          .......... .......... C.
Marchantia      .........A ..AA..C... C.
Rice            GA........ ...C....G. CG
Maize           GA........ ...C....G. CG
```

without uninformative characters, respectively; RI = 0.749); the strict consensus of these trees, along with bootstrap estimates and decay indices, is presented in Fig. 3b. Parsimony analysis of the reduced (i.e., 27-taxon) exon matrix produced only one minimal-length tree (not shown) with a CI (excluding uninformative characters) of 0.535 and a RI of 0.688. These CI and RI values are only slightly lower than those obtained from the analysis of the 27 intron sequences, indicating that the amount of homoplasy in both data sets is approximately the same. Average transition/transversion ratios, as calculated by MacClade over the most parsimonious tree(s) for each data set, were 1.75 for the 38 exon sequences and 0.97 for the 27 intron sequences. As the ratio of transitions to transversions approaches 1.0, as it does here for the intron sequence comparisons, it is often assumed that transitions are becoming largely saturated (Mindell and Honeycutt 1990, but see Reeder 1995, for an exception). Considering both intron and exon data for these 27 taxa together in a single matrix resulted in two most-parsimonious cladograms; their strict consensus is depicted in Fig. 4 (length, 527 steps; CI, excluding uninformative characters, 0.560; RI, 0.708).

Apparent in Fig. 3a is a sister group relationship between *Barnadesia* (Asteraceae) and a clade comprising three genera of Calyceraceae (*Boopis*, *Calycera*, and *Gamocarpha*). *Goodenia* and *Scaevola*, both Goodeniaceae, fall as sister-group to Asteraceae+ Calyceraceae. Due to the absence of the *rpoC1* intron in Goodeniaceae, *Goodenia* and *Scaevola* could not be included in Figs. 3b and 4. Nevertheless, in all analyses, Asteraceae and Calyceraceae are sister groups, and Menyanthaceae (*Menyanthes*, *Villarsia*, and *Nymphoides*) then Campanulaceae/ Lobeliaceae fall as successive sister taxa to the Asteraceae+Calyceraceae(+Goodeniaceae, if present) clade. The union of *Coriandrum* with *Daucus* (both Apiaceae) and of *Dipsacus* with *Kolkwitzia* (Dipsacaceae and Caprifoliaceae, respectively) is also evident in all trees. However, the relative positions of these two clades to each other and to tobacco (Solanaceae), *Escallonia* (Grossulariaceae), and *Cornus* (Cornaceae) are different in the three trees. The three intron-absent and two intron-containing species of *Passiflora* constitute a clade (Fig. 3a), and Passifloraceae are sister group to Flacourtiaceae (*Oncoba* and *Flacourtia*) in both the exon (Fig. 3a) and combined (Fig. 4) analyses. In contrast, when only the intron sequences are considered (Fig. 3b), Passifloraceae is sister group to a clade consisting of Flacourtiaceae+ Euphorbiaceae (*Euphorbia* and *Jatropha*). The three species of *Passiflora* lacking the intron form a clade, as do *Goodenia* and *Scaevola* (Goodeniaceae), *Delosperma* and *Faucaria* (Aizoaceae), and rice and maize (Poaceae; Fig. 3a).

Parsimony analyses of separate exon, intron, and combined data sets yield highly-resolved and moderately to well-supported phylogenies that are also generally supported, in part, by parsimony analyses of complete *rbcL* and *ndhF* chloroplast gene sequences (e.g., Chase et al. 1993; Olmstead et al. 1993; Cosner et al. 1994; Kim and Jansen 1995, and unpubl. data) and/or morphological and chemical characters (e.g., Gustafsson and Bremer 1995). Many of the depicted relationships are also similar to those implied by conventional classification systems (e.g., Killip 1938; Cronquist 1981). Discordance of relationship among our analyses and those of others occurs in those regions of the trees where sampling is inadequate.

Discussion

Distribution of the rpoC1 *Intron Loss in Angiosperms.* The phylogenetic implications of the four instances of *rpoC1* intron loss detected here are

discussed below. Although the loss of the *rpoC1* intron has likely occurred independently six times during angiosperm evolution (when the grasses and Cactoideae are also considered), like the homoplastic distribution of the *rpl2* intron loss (Downie et al. 1991; Doyle et al. 1995), its absence can still be a powerful indicator of relationship within the different lineages in which it has occurred. Obviously, the distribution of this mutation in each of these four groups needs to be assessed critically by examining additional taxa. We welcome these investigations and, to this end, would gladly supply researchers with the necessary primers for PCR amplification and DNA sequencing. Of the 54 families that have been examined for the presence or absence of the *rpoC1* intron, six have members whose *rpoC1* genes are not split by introns. It is expected that as representatives of other families are surveyed, additional losses of the *rpoC1* intron will be detected. Moreover, as we have found that some families and even genera are polymorphic for the intron, it should not be surprising to uncover other cases of intron loss as additional representatives from these same genera are sampled.

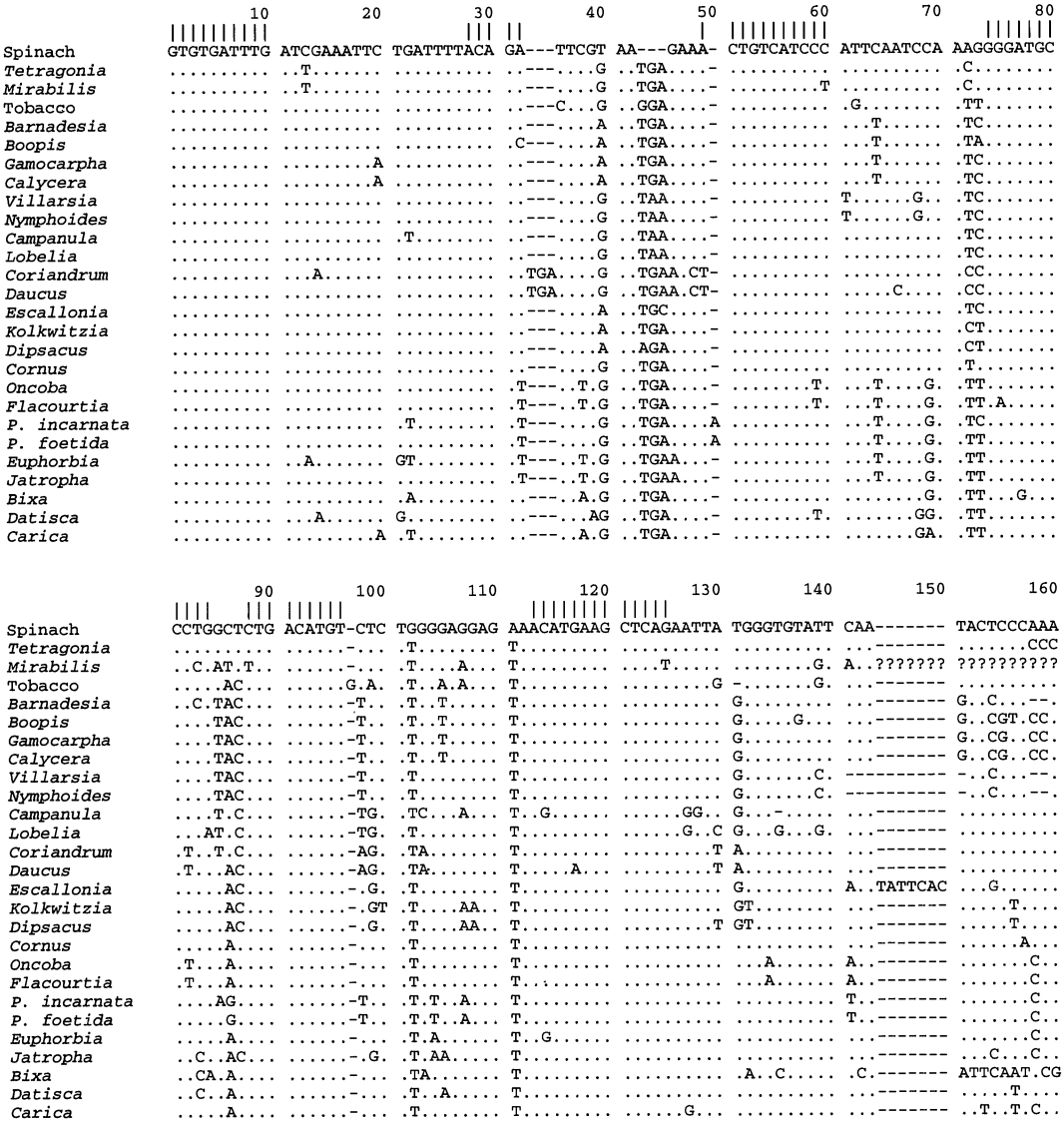
GOODENIACEAE. The family Goodeniaceae, of tropical and subtropical distribution, consists of about 14 genera and 300 species, with the two largest genera being *Goodenia* and *Scaevola* (Cronquist 1981). Recent cladistic analyses of morphological characters and *rbcL* sequences from nine genera of Goodeniaceae have shown that *Goodenia* and *Scaevola* may be sister taxa (Gustafsson 1995). Additional species within Goodeniaceae must be investigated to circumscribe the distribution of this intron loss. The traditional close association of this family with Asteraceae and Calyceraceae is supported by recent morphological and molecular studies (Michaels et al. 1993; Gustafsson and Bremer 1995; Kim and Jansen 1995), and the analysis of *rpoC1* exon sequences presented herein. Representatives from these latter two families, however, both contain an intron in *rpoC1*.

PASSIFLORACEAE. The family Passifloraceae, defined by Cronquist (1981) as consisting of about 16 genera and 650 species, is widespread in tropical and warm-temperate regions. With about 400 species, *Passiflora* is by far the largest and most important genus in the family. In the most comprehensive study of New World passionflowers to date, Killip (1938) recognized 22 subgenera and 13 sections, of which species from four subgenera were sampled during the course of this investigation. All four species of *Passiflora* lacking the *rpoC1* intron belong to one subgenus (*Plectostemma*), whereas those species having the intron have been classified into subgenera *Distephana*, *Dysosmia*, and *Granadilla* (Table 1). A recent survey of chromosome numbers in *Passiflora* has revealed that the basic chromosome number of $x = 6$, characteristic of subgenus *Plectostemma*, is likely derived within the genus (Snow and MacDougal 1993). Moreover, the analysis of nucleotide substitutions from the *rpoC1* exon regions in five species of *Passiflora* (Fig. 3a) suggests monophyly of *Plectostemma*. The absence of the intron in *Plectostemma* further distinguishes this subgenus in Passifloraceae.

Passifloraceae are regarded as being related to Flacourtiaceae (Cronquist 1981). Cronquist placed these two families, along with several others, such as Bixaceae, Caricaceae, Violaceae, Cucurbitaceae, Datiscaceae, and Begoniaceae, in his order Violales. In a phylogeny based on *rbcL* sequences, Chase et al. (1993) showed *Euphorbia* (Euphorbiaceae) as sister-taxon to *Passiflora* with *Datisca* closely related. *Flacourtia*, *Oncoba*, and *Bixa* were not included in their analyses, and *Carica* was placed elsewhere. Representatives from all of these putative outgroups have an intron in *rpoC1*.

AIZOACEAE. The family Aizoaceae, circumscribed by Bittrich and Hartmann (1988) on the shared presence of unique bladder idioblasts in the epidermis, comprises about 2,500 species that are primarily South African and Australian in distribution (Cronquist 1981). The presence of the *rpoC1* intron in *Monilaria* and *Tetragonia* but not *Delosperma* or *Faucaria* cpDNA's (Table 1) may indicate a major dichotomy within Aizoaceae. Of the five Aizoaceae subfamilies recognized by Bittrich and Hartmann (1988), the largest is Ruschioideae, with about 120 genera, whereas the smallest is Tetragonioideae, with only two genera. The absence of the *rpoC1* intron in *Delosperma* and *Faucaria* but not *Monilaria* (all representatives of subfamily Ruschioideae; Table 1), further suggests that *Monilaria* may be basal within the subfamily if Ruschioideae are indeed monophyletic.

The shared absence of the *rpoC1* intron in some Aizoaceae and all examined members of Cactaceae subfamily Cactoideae (Wallace and Cota 1996) provides tantalizing evidence for their close relationship. Although several authors have illustrated such a relationship (e.g., Burret et al. 1981; Cronquist 1981, Rodman et al. 1984), the characters used to support this inference have been subject to


10 20 30 40 50 60 70 80
Spinach GTGTGATTTG ATCGAAATTC TGATTTTACA GA---TTCGT AA---GAAA- CTGTCATCCC ATTCAATCCA AAGGGGATGC
Tetragonia ....T..... ....---....G ..TGA....- .... .C.....
Mirabilis ....T..... ....---....G ..TGA....- .....T ..... .C.....
Tobacco ....---C...G ..GGA....- ..... .G..... .TT.....
Barnadesia ....---....A ..TGA....- ....T..... .TC.....
Boopis .... .C---....A ..TGA....- ....T..... .TA.....
Gamocarpha ....A ....---....A ..TGA....- ....T..... .TC.....
Calycera ....A ....---....A ..TGA....- ....T..... .TC.....
Villarsia ....---....G ..TAA....- ..... T......G.. .TC.....
Nymphoides ....---....G ..TAA....- ..... T......G.. .TC.....
Campanula .... .T..... ..---....G ..TAA....- .... .TC.....
Lobelia ....---....G ..TAA....- .... .TC.....
Coriandrum ....A..... ....TGA....G ..TGAA.CT- .... .CC.....
Daucus ....TGA....G ..TGAA.CT- ....C.... .CC.....
Escallonia ....---....A ..TGC....- .... .TC.....
Kolkwitzia ....---....A ..TGA....- .... .CT.....
Dipsacus ....---....A ..AGA....- .... .CT.....
Cornus ....---....G ..TGA....- .... .T.....
Oncoba .... .T---..T.G ..TGA....- .....T. ...T....G. .TT.....
Flacourtia .... .T---..T.G ..TGA....- .....T. ...T....G. .TT.A....
P. incarnata .... .T..... .T---....G ..TGA....A ....T....G. .TC.....
P. foetida .... .T---....G ..TGA....A ....T....G. .TT.....
Euphorbia ....A..... GT..... .T---..T.G ..TGAA...- ....T....G. .TT.....
Jatropha .... .T---..T.G ..TGAA...- ....T....G. .TT.....
Bixa .... .A..... ..---..A.G ..TGA....- ....G. .TT...G...
Datisca ....A..... G..... ..---...AG ..TGA....- .....T. .....GG. .TT.....
Carica ....A .T..... ..---..A.G ..TGA....- ....GA. .TT.....
90 100 110 120 130 140 150 160
Spinach CCTGGCTCTG ACATGT-CTC TGGGGAGGAG AAACATGAAG CTCAGAATTA TGGGTGTATT CAA------- TACTCCCAAA
Tetragonia ....-... .T..... T..... ....------- CCC
Mirabilis ..C.AT.T.. .....-... .T....A... T..... ....T..... .....G. A..??????? ??????????
Tobacco .....AC... .....G.A. .T..A.A... T..... .....G -.....G. ...------- .....
Barnadesia ..C.TAC... .....-T.. .T..T..... T..... ..... G..... ...------- G..C...--.
Boopis ....TAC... .....-T.. .T..T..... T..... ..... G.....G... ...------- G..CGT.CC.
Gamocarpha ....TAC... .....-T.. .T..T..... T..... ..... G..... ...------- G..CG..CC.
Calycera ....TAC... .....-T.. .T..T..... T..... ..... G..... ...------- G..CG..CC.
Villarsia ....TAC... .....-T.. .T..... T..... ..... G.....C. ---------- -..C...--.
Nymphoides ....TAC... .....-T.. .T..... T..... ..... G.....C. ---------- -..C...--.
Campanula ....T.C... .....-TG. .TC...A... T..G..... .....GG.. G...-..... ...------- .....
Lobelia ...AT.C... .....-TG. .T..... T..... .....G..C G...G...G. ...------- .....
Coriandrum .T..T.C... .....-AG. .TA..... T..... .....T A..... ...------- .....
Daucus .T...AC... .....-AG. .TA..... T.....A... .....T A..... ...------- .....
Escallonia .....AC... .....-.G. .T..... T..... ..... G..... A..TATTCAC ...G.....
Kolkwitzia .....AC... .....-.GT .T....AA.. T..... ..... GT..... ...------- .....T....
Dipsacus .....AC... .....-.G. .T....AA.. T..... .....T GT..... ...------- .....T....
Cornus .....A.... .....-.... .T..... T..... ....------- .....A...
Oncoba .T...A.... .....-.... .T..... T..... ....A..... A..------- .....C..
Flacourtia .T...A.... .....-.... .T..... T..... ....A..... A..------- .....C..
P. incarnata ....AG.... .....-T.. .T.T..A... T..... .... T..------- .....C..
P. foetida .....G.... .....-T.. .T.T..A... T..... .... T..------- .....C..
Euphorbia .....A.... .....-.... .T.A..... T..G..... ....------- .....C..
Jatropha ..C..AC... .....-.G. .T.AA..... T..... ....------- ...C...C..
Bixa ..CA.A.... .....-.... .TA..... T..... ..... .A..C..... .C.------- ATTCAAT.CG
Datisca ..C..A.... .....-.... .T..A..... T..... ....------- .....T....
Carica .....A.... .....-.... .T..... T..... .....G... ....------- ..T..T.C..
170 180 190 200 210 220 230 238
Spinach T-----GAAA A-GGGAATTG ATCCATGGTC GATTTCGTGA CAAAG-AGAT ATA----AAT TTCTAGTTGT ACC-TCGT
Tetragonia C-----A... .-..... ....A. .....????? ?????????? ?????????? ????????
Mirabilis ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ????????
Tobacco .-----A... .-..... ...T..... ....A. ....TC.A.. .A.GAGG... .....T.... ...-....
Barnadesia A--------- --..... C..T..... ......A.A. G..------- ---------- ....A. ...-....
Boopis A-----A... --..... ...T..... ......A.A. G..------- ---------- ----....A. ...-....
Gamocarpha A-----A... --..... ...T..... ......A.A. G..------- ---------- ----....A. ...C....
Calycera A-----A... --..... ...T..... ......A.A. G..------- ---------- ----....A. ...C....
Villarsia A-----A..- --..... ...T..... ......A.AG G..------- ---------. ...G....A. ...-....
Nymphoides A-----A... --..... ...T..... ......A.AG G..------- ---------. ...G....A. ..T-....
Campanula .-----A... GG..... ...T..... ......A.A. T...T-.A.. ----AGG... ....A. ...-....
Lobelia .-----C... GG..... ...T..... ......A.A. .....-.A.. ----AGGC.. ....A. ...-....
Coriandrum .-----A.C. .G...G...T ...T..... ....A. ....A-.A.. ----AGGC.. ..AA....A. ...-....
Daucus .-----A.C. .G.....T ...T..... ....A. ....A-.A.. ----AGGC.. ..T.....A. ...-....
Escallonia .-----A... .G..... ...T..... ....A. ....T-.A.. ----AGG... ..A.....A. ...-....
Kolkwitzia .-----A.C. .A..... ...T..... ....CA. ....T-.A.. ----AGG... ..A.....A. ...-....
Dipsacus .-----A.C. .A..... ...T..... ....A. ....T-.A.. ----AGG... ..T.....A. ...-....
Cornus .-----A..C GG..A..... ...T.....T .....TA.A. ....T-.A.. ----AGGC.. ....-....

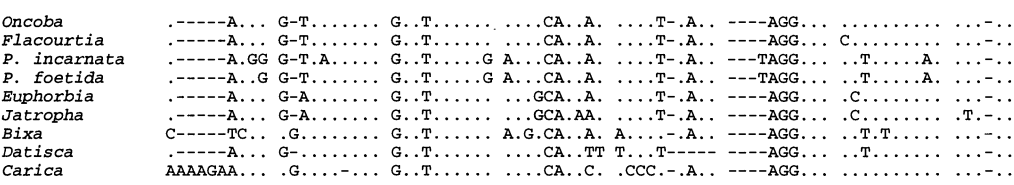


FIG. 2. Aligned nucleotide sequences from the cpDNA *rpoC1* intron. Nucleotide sites are arranged 5′ to 3′ and begin at the exon 1/intron boundary. These sequences comprise about one-third of the entire intron. Nucleotide identity with the first sequence is indicated by a dot; gaps required for alignment are indicated by hyphens; missing data are indicated by question marks; N = nucleotide of uncertain state. Complete taxon names are provided in Appendix 1. The nucleotides participating in highly conserved base-paired stem formation or single-stranded sequence in all organelle group II introns (Michel et al. 1989) are indicated by vertical bars above the alignment. The identification and locations of the four subdomains of domain I are as follows (numbers refer to coordinates in alignment): IA (28–54); IB (55–79); IC (80–174); ID (181–238).

much criticism (see discussions in Levin 1985, and Bittrich and Hartmann 1988). Furthermore, within Cactaceae, the putatively basal subfamily Pereskioideae [*Pereskia*; Cronquist (1981)], contains a split *rpoC1* gene. Representatives from all other Caryophyllalean families examined (i.e., Amaranthaceae, Basellaceae, Caryophyllaceae, Chenopodiaceae, Didiereaceae, Molluginaceae, Nyctaginaceae, Phytolaccaceae, and Portulacaceae) also all have the intron (Wallace and Cota 1996; this study).

MEDICAGO. The genus *Medicago* (Fabaceae; Trifolieae) consists of about 85 annual and perennial species divided into 12 sections (Small and Jomphe 1989) of which three sections were represented in this survey (Table 1). Only one (*M. suffruticosa* subsp. *leiocarpa*) of eight species examined lack the *rpoC1* intron. The relationship of *M. suffruticosa* and *M. hybrida* (Pourret) Trautv. (which together make up section *Suffruticosae*; Lesins and Lesins 1979) to other taxa within the genus are obscure. The absence of the *rpoC1* intron in *M. suffruticosa* subsp. *leiocarpa* further distinguishes this taxon from all other *Medicago* examined but is neutral with respect to its proper sectional placement. All other Fabaceae sampled, including putatively allied *Trifolium* (also Trifolieae), have an intron in *rpoC1* (Appendix 1).

Chloroplast Group II Intron Secondary Structure. Plastid introns in land plants have been classified into two groups, based on conserved secondary structural features and boundary sequences (Michel et al. 1989). Only the tRNA gene *trnL*(UAA) is interrupted by a group I intron; all remaining split genes, including *rpoC1*, have group II introns.

Group II introns exhibit considerable evolutionary conservation of secondary structure and are characterized by six centrally-radiating structural components (designated as domains I-VI; Michel et al. 1989). Domain I, important in intron excision, is composed of four subdomains [designated A, B, C, and D, with the latter two subdomains further divided into C1 and C2, and D1, D2, and D3 regions, respectively (Michel et al. 1989)]. Each of these structural regions includes highly conserved, double-stranded stem portions (helices) and, generally, less conserved, single-stranded loop portions (Michel and Dujon 1983; Michel et al. 1989). Of the 238 aligned positions in the intron data matrix (Fig. 2), 81—or about one-third of all positions—are involved in highly conserved base-paired stem formation or single-stranded sequence (such as the 5′ terminal nucleotides). These positions are indicated above the alignment in Fig. 2. The portion of the intron sequenced coincides with conserved group II intron regions IA, IB, IC1, IC2 (position 28–174; Fig. 2), and one strand of the region identified as ID1$^{(i)}$ to ID1$^{(iv)}$ (position 181–238; Michel et al. 1989). These sequences comprise about 54% of the total length of domain I, the largest domain in group II introns.

Each of the six domains of group II introns, and their characteristic structural features, can vary markedly in their acceptance of mutational change (Clegg et al. 1986, 1994; Learn et al. 1992; Downie et al. 1996). In a recent comparative analysis of *rpoC1* intron sequences from the Apiaceae and relatives (Downie et al. 1996), domain II was deemed to be the most variable. Similarly, in a detailed investigation of cpDNA *trnV*(UAC) intron sequences, Learn et al. (1992) showed that domain II had the highest rate of sequence divergence, approaching the synonymous substitution rate reported for some protein coding genes. In contrast, domain I had regions with some of the slowest rates of sequence evolution within the intron. In this study, extensive

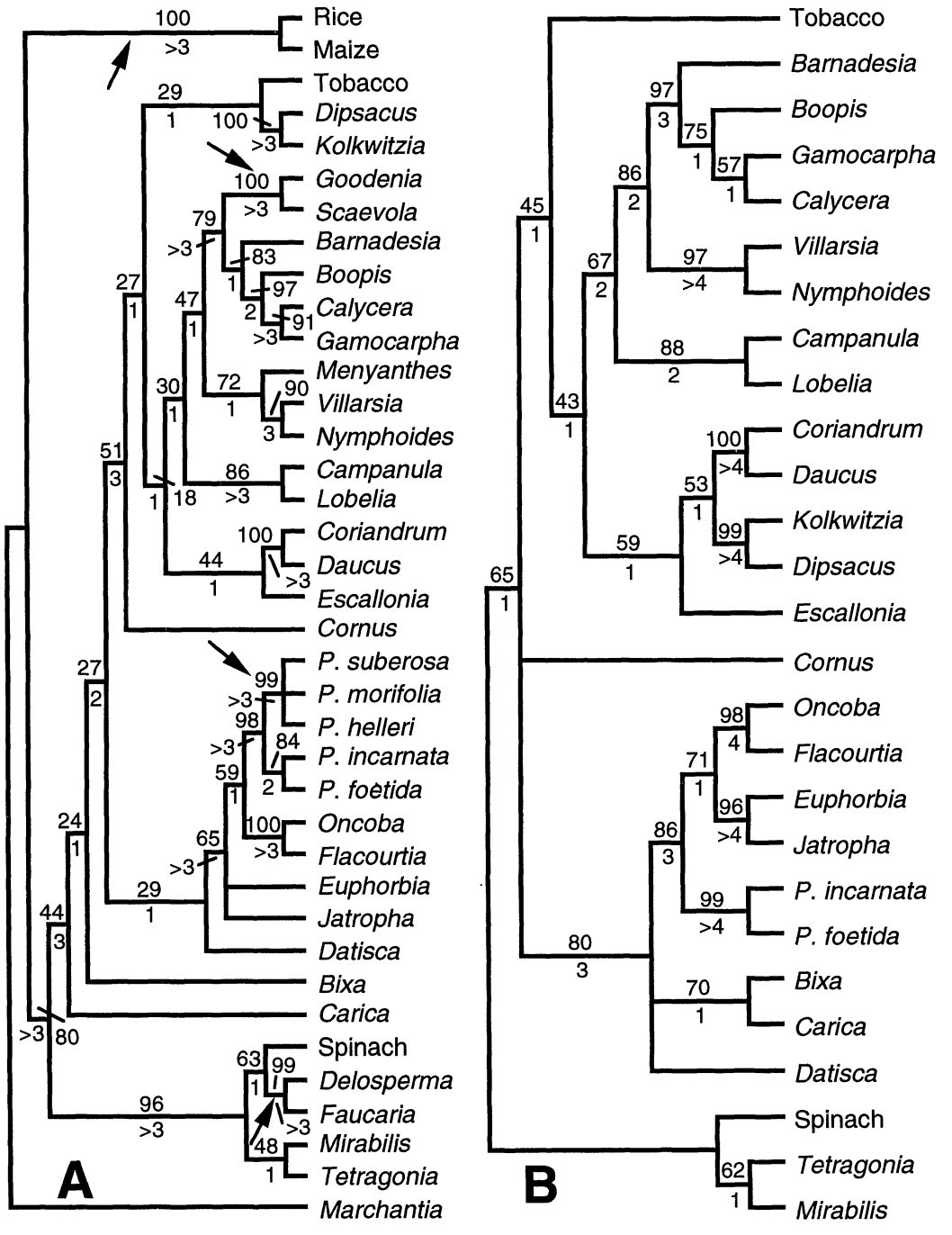


FIG. 3. Strict consensus trees derived from unweighted parsimony analysis of cpDNA *rpoC1* sequence data. Numbers above the nodes indicate the number of times the group occurred in 100 bootstrap replicates. Numbers below the nodes indicate the number of additional steps required before the node collapses. *P* = *Passiflora*; complete taxon names are provided in Appendix 1. A) Strict consensus of four maximally-parsimonious 514-step trees derived from partial *rpoC1* exon sequences (CI excluding uninformative characters = 0.525; RI = 0.691). The four lineages lacking the *rpoC1* intron are indicated by arrows. B) Strict consensus of three maximally-parsimonious 232-step trees derived from the analysis of partial *rpoC1* intron sequences (CI excluding uninformative characters = 0.605; RI = 0.749).

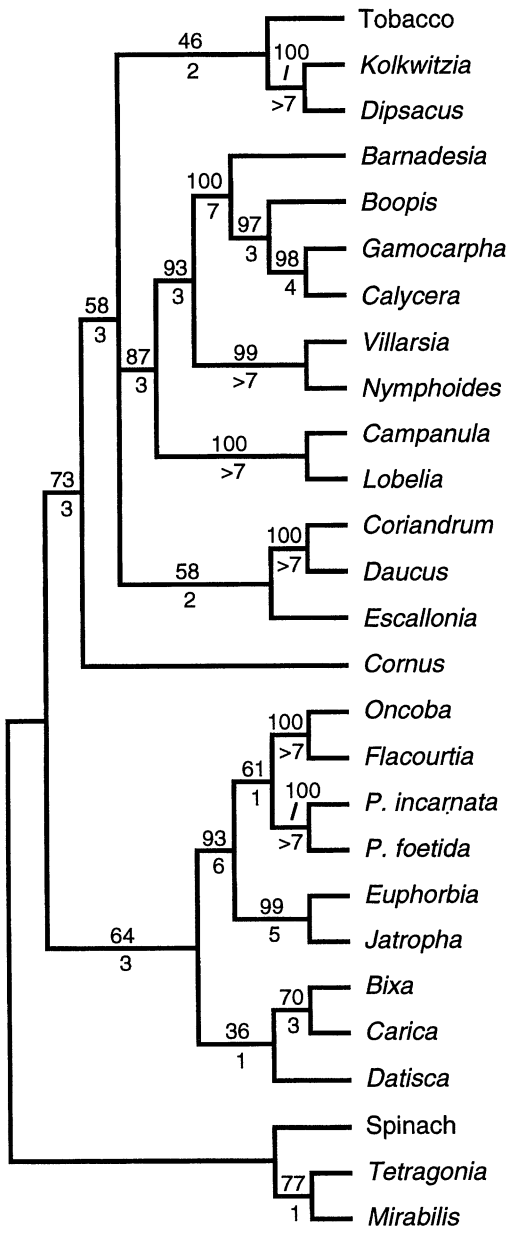


FIG. 4. Strict consensus tree of two minimal-length cladograms, each of 527 steps, derived from unweighted parsimony analysis of combined *rpoC1* exon and intron DNA sequences (CI excluding uninformative characters = 0.560; RI = 0.708). Numbers above the nodes are bootstrap estimates for 100 replicate analyses. Numbers below the nodes indicate the number of additional steps required before the node collapses. *P* = *Passiflora*; complete taxon names are provided in Appendix 1.

variation, both in sequence and in length, occurred in both the single-stranded loop regions at the tip of the helical stems and in the stretches of nucleotides connecting the subdomains. All but 23 of the 110 variable positions (38 autapomorphies and 72 informative sites) detected in the intron data matrix are located in these terminal loop portions or other non-conserved regions of the intron. Similarly, all insertion and deletion events (and, therefore, all potentially ambiguous positions) also occurred in these same regions. Consequently, these mutations do not disrupt the secondary structure of the intron and are likely evolving independently.

Phylogenetic Utility of* rpoC1 *Intron Sequences. It is generally acknowledged that chloroplast intron sequences, like noncoding intergenic spacers, are rapidly evolving, both in their high rates of nucleotide substitution and in their propensity to accumulate many insertions and deletions (Curtis and Clegg 1984; Zurawski and Clegg 1987; Palmer 1991; Golenberg et al. 1993; Gielly and Taberlet 1994; Fangan et al. 1994). In this study, however, both the intron and exon portions sequenced for the same set of 27 taxa had similar percentages of conserved sites (58 and 54%, respectively). Also, for each data set, the range in nucleotide sequence divergence values for all pairwise comparisons was approximately the same. The conservatism observed in intron sequence comparisons relative to exon comparisons is undoubtedly due to the high percentage of positions participating in base-paired secondary structure in that portion of the intron sequenced. It is expected that data from other and more variable regions of the intron, such as domain II, will show greater divergence among pairwise intron comparisons. For example, sequence divergence values using full-length exon and intron sequences between tobacco and spinach (Shinozaki et al. 1986; Hudson et al. 1988) differed two-fold (9.3 and 18.4%, respectively), whereas nucleotide differences of partial exon and intron sequences between these same two taxa, excluding gaps, were similar (12.3 and 10.0%, respectively). Because the pattern of intron evolution is constrained by a high degree of secondary structure, compensatory substitutions in stem regions of helices and, thus, nonindependence of characters should be considered in phylogenetic analyses.

Single-nucleotide insertions or deletions were the most frequent (10 of 22 events) in the multiple alignment of intron sequences. In contrast, no gaps were needed to align the exon sequences. Although the homology of small length mutations can be

difficult to ascertain, particularly if they fall within regions where frequent insertions and deletions occur (e.g., Palmer et al. 1985), such characters can potentially be useful phylogenetically if their homology can be confirmed (Olmstead and Palmer 1994). Of the 22 indels inferred here in the alignment of intron sequences, 11 are potentially informative and, of these, eight are congruent (i.e., they are not homoplastic) with the most parsimonious trees based on nucleotide substitutions only (Fig. 3b).

The presence of an intron in *rpoC1* is not a liability. Its location within the gene is conserved and readily identified (as its boundary sequences are diagnostic), and major length variants within the intron are non-existent. Moreover, the availability of an additional 750 bp or so of sequence, containing both highly conserved (stem) and variable (loop) regions, can provide additional phylogenetic information at various levels of genetic divergence. Among distantly related taxa, however, the presence of numerous small length mutations can confound alignment and, concomitantly, phylogenetic analyses. Similarly, the likely saturation of transitions in non-conserved positions by multiple substitutions at this level suggests that this category of mutation might be best downweighted or ignored in a phylogenetic analysis. Precise information on the phylogenetic utility of *rpoC1* exon and intron sequences requires complete sequences for these regions. Once these data are available, comparisons of evolutionary rates, patterns of base substitution variability and conservation, and phylogenetic relationships can be made.

ACKNOWLEDGMENTS. The authors thank the various institutions and individuals cited in Appendix 1 for providing us with seeds or leaf material, D. Seigler for supplying material of *Passiflora* and J. MacDougal for determining their identifications, J. Palmer for providing some of the DNA's, L. Paulius and R. Patel for laboratory assistance, and E. Kellogg, G. Plunkett, K. Potgieter, and two anonymous reviewers for helpful comments on an earlier draft of the manuscript. This work was supported by laboratory start-up funds from the University of Illinois to SRD, and, in part, by a grant from the National Science Foundation (DEB-9407712).

LITERATURE CITED

BITTRICH, V. and H. E. K. HARTMANN. 1988. The Aizoaceae—a new approach. Botanical Journal of the Linnaean Society 97: 239–254.

BREMER, B. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42: 795–803.

BURRET, F., Z. RABESA, P. ZANDONELLA, and B. VOIRIN. 1981. Contribution biochimique à la systématique de l'ordre des Centrospermales. Biochemical Systematics and Ecology 9: 257–262.

CHASE, M. W., D. E. SOLTIS, R. G. OLMSTEAD, D. MORGAN, D. H. LES, B. D. MISHLER, M. R. DUVALL, R. A. PRICE, H. G. HILLS, Y.-L. QIU, K. A. KRON, J. H. RETTIG, E. CONTI, J. D. PALMER, J. R. MANHART, K. J. SYTSMA, H. J. MICHAELS, W. J. KRESS, K. G. KAROL, W. D. CLARK, M. HEDREN, B. S. GAUT, R. K. JANSEN, K.-J. KIM, C. F. WIMPEE, J. F. SMITH, G. R. FURNIER, S. H. STRAUSS, Q.-Y. XIANG, G. M. PLUNKETT, P. S. SOLTIS, S. M. SWENSEN, S. E. WILLIAMS, P. A. GADEK, C. J. QUINN, L. E. EGUIARTE, E. GOLENBERG, G. H. LEARN, JR., S. W. GRAHAM, S. C. H. BARRETT, S. DAYANANDAN, and V. A. ALBERT. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. Annals of the Missouri Botanical Garden 80: 528–580.

CLEGG, M. T., B. S. GAUT, G. H. LEARN, JR., and B. R. MORTON. 1994. Rates and patterns of chloroplast DNA evolution. Proceedings of the National Academy of Sciences U.S.A. 91: 6795–6801.

———, K. RITLAND, and G. ZURAWSKI. 1986. Processes of chloroplast DNA evolution. Pp. 275–294 in *Evolutionary processes and theory*, eds. S. Karlin and E. Nevo. New York: Chapman and Hall.

COSNER, M. E., R. K. JANSEN, and T. G. LAMMERS. 1994. Phylogenetic relationships in the Campanulales based on *rbcL* sequences. Plant Systematics and Evolution 190: 79–95.

COZENS, A. L. and J. E. WALKER. 1986. Pea chloroplast DNA encodes homologues of *Escherichia coli* ribosomal subunit S2 and the β′ subunit of RNA polymerase. Biochemical Journal 236: 453–460.

CRONQUIST, A. 1981. *An integrated system of classification of flowering plants*. New York: Columbia Univ. Press.

CURTIS, S. E. and M. T. CLEGG. 1984. Molecular evolution of chloroplast DNA sequences. Molecular Biology and Evolution 1: 291–301.

DOWNIE, S.R. and J. D. PALMER. 1992. Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. Pp. 14–35 in *Molecular systematics of plants*, eds. P. S. Soltis, D. E. Soltis, and J. J. Doyle. New York: Chapman and Hall.

——— and D. S. KATZ-DOWNIE. 1996. A molecular phylogeny of Apiaceae subfamily Apioideae: evidence from nuclear ribosomal DNA internal transcribed spacer sequences. American Journal of Botany 83: 234–251.

———, ———, K. H. WOLFE, P. J. CALIE, and J. D. PALMER. 1994. Structure and evolution of the largest chloroplast gene (ORF2280): internal plasticity and multiple gene loss during angiosperm evolution. Current Genetics 25: 367–378.

———, ———, and K.-J. CHO. 1996. Phylogenetic analysis of Apiaceae subfamily Apioideae using nucleotide sequences from the chloroplast *rpoC1* intron. Molecular Phylogenetics and Evolution. 6:1–18.

———, R. G. OLMSTEAD, G. ZURAWSKI, D. E. SOLTIS, P. S. SOLTIS, J. C. WATSON, and J. D. PALMER. 1991. Six independent losses of the chloroplast DNA *rpl2* intron in dicotyledons: molecular and phylogenetic implications. Evolution 45: 1245–1259.

DOYLE, J. J. and J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19: 11–15.

———, ———, and J. D. PALMER. 1995. Multiple independent losses of two genes and one intron from legume chloroplast genomes. Systematic Botany 20: 272–294.

FANGAN, B. M., B. STEDJE, O. E. STABBETORP, E. S. JENSEN, and K. S. JAKOBSEN. 1994. A general approach for PCR-amplification and sequencing of chloroplast DNA from crude vascular plant and algal tissue. BioTechniques 16: 484–494.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791.

FREYER, R., K. NECKERMANN, R. A. MAIER, and H. KOSSEL. 1995. Structural and functional analysis of plastid genomes from parasitic plants: loss of an intron within the genus *Cuscuta*. Current Genetics 27: 580–586.

GIELLY, L. and P. TABERLET. 1994. The use of chloroplast DNA to resolve plant phylogenies: noncoding versus *rbcL* sequences. Molecular Biology and Evolution 11: 769–777.

GOLENBERG, E. M., M. T. CLEGG, M. L. DURBIN, J. DOEBLEY, and D. P. MA. 1993. Evolution of a noncoding region of the chloroplast genome. Molecular Phylogenetics and Evolution 2: 52–64.

GUSTAFSSON, M. H. G. 1995. Intra- and interfamilial relationships of the Goodeniaceae as suggested by *rbcL* sequences and morphology. American Journal of Botany 82: S133. (abstract)

——— and K. BREMER. 1995. Morphology and phylogenetic interrelationships of the Asteraceae, Calyceraceae, Campanulaceae, Goodeniaceae, and related families (Asterales). American Journal of Botany 82: 250–265.

HIRATSUKA, J., H. SHIMADA, R. WHITTIER, T. ISHIBASHI, M. SAKAMOTO, M. MORI, C. KONDO, Y. HONJI, C.-R. SUN, B.-Y. MENG, Y.-Q. LI, A. KANNO, Y. NISHIZAWA, A. HIRAI, K. SHINOZAKI, and M. SUGIURA. 1989. The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distant tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. Molecular and General Genetics 217: 185–194.

HOOT, S. B. and J. D. PALMER. 1994. Structural rearrangements, including parallel inversions, within the chloroplast genome of *Anemone* and related genera. Journal of Molecular Evolution 38: 274–281.

HUDSON, G. S., T. A. HOLTON, P. R. WHITFELD, and W. BOTTOMLEY. 1988. Spinach chloroplast *rpoBC* genes encode three subunits of the chloroplast RNA polymerase. Journal of Molecular Biology 200: 639–654.

IGLOI, G. L., A. MEINKE, I. DORY, and H. KOSSEL. 1990. Nucleotide sequence of the maize chloroplast *rpo* $B/C_1/C_2$ operon: comparison between the derived protein primary structures from various organisms with respect to functional domains. Molecular and General Genetics 221: 379–394.

KATAYAMA, H. and Y. OGIHARA. 1993. Structural alterations of the chloroplast genome found in grasses are not common in monocots. Current Genetics 23: 160–165.

KILLIP, E. P. 1938. The American species of Passifloraceae. Publications of the Field Museum of Natural History, Botanical Series 19: 1–613.

KIM, K.-J. and R. K. JANSEN. 1995. *ndhF* sequence evolution and the major clades in the sunflower family. Proceedings of the National Academy of Sciences U.S.A. 92: 10379–10383.

LEARN, G. H., JR., J. S. SHORE, G. R. FURNIER, G. ZURAWSKI, and M. T. CLEGG. 1992. Constraints on the evolution of plastid introns: the group II intron in the gene encoding tRNA-Val(UAC). Molecular Biology and Evolution 9: 856–871.

LESINS, K. A. and I. LESINS. 1979. Genus *Medicago* (Leguminosae). A taxogenetic study. Dr. W. Junk, Publishers. The Hague, Netherlands.

LEVIN, G. A. 1985. Character analysis and cladistics: a response to Rodman et al. Systematic Botany 10: 496–503.

MADDISON, W. P. and D. R. MADDISON. 1992. *MacClade, version 3.0: Analysis of phylogeny and character evolution*. Sunderland, Massachusetts: Sinauer Associates, Inc.

MICHAELS, H. J., K. M. SCOTT, R. G. OLMSTEAD, T. SZARO, R. K. JANSEN, and J. D. PALMER. 1993. Interfamilial relationships of the Asteraceae: insights from *rbcL* sequence variation. Annals of the Missouri Botanical Garden 80: 742–751.

MICHEL, F. and B. DUJON. 1983. Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast- and nuclear-encoded members. EMBO Journal 2: 33–38.

———, K. UMESONO, and H. OZEKI. 1989. Comparative and functional anatomy of group II catalytic introns—a review. Gene 82: 5–30.

MINDELL, D. P. and R. L. HONEYCUTT. 1990. Ribosomal RNA in vertebrates: evolution and phylogenetic applications. Annual Review of Ecology and Systematics 21: 541–566.

OHYAMA, K., H. FUKUZAWA, T. KOHCHI, H. SHIRAI, T. SANO, S. SANO, K. UMESONO, Y. SHIKI, M. TAKEUCHI, Z. CHANG, S. AOTA, H. INOKUCHI, and H. OZEKI. 1986. Complete nucleotide sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Plant Molecular Biology Reporter 4: 148–175.

OLMSTEAD, R. G. and J. D. PALMER. 1994. Chloroplast DNA systematics: a review of methods and data analysis. American Journal of Botany 81: 1205–1224.

———, B. BREMER, K. M. SCOTT, and J. D. PALMER. 1993. A parsimony analysis of the Asteridae sensu lato based

on *rbcL* sequences. Annals of the Missouri Botanical Garden 80: 700–722.

PALMER, J. D. 1991. Plastid chromosomes: structure and evolution. Pp. 5–53 in *The molecular biology of plastids*, eds. L. Bogorad and I. K. Vasil, vol. 7A of *Cell culture and somatic cell genetics of plants*. San Diego: Academic Press.

———, R. K. JANSEN, H. J. MICHAELS, M. W. CHASE, and J. R. MANHART. 1988. Chloroplast DNA variation and plant phylogeny. Annals of the Missouri Botanical Garden 75: 1180–1206.

———, R. A. JORGENSEN, and W. F. THOMPSON. 1985. Chloroplast DNA variation and evolution in *Pisum*: patterns of change and phylogenetic analysis. Genetics 109: 195–213.

REEDER, T. W. 1995. Phylogenetic relationships among Phrynosomatid lizards as inferred from mitochondrial ribosomal DNA sequences: substitutional bias and information content of transitions relative to transversions. Molecular Phylogenetics and Evolution 4: 203–222.

RODMAN, J. E., M. K. OLIVER, R. R. NAKAMURA, J. U. MCCLAMMER, JR., and A. H. BLEDSOE. 1984. A taxonomic analysis and revised classification of Centrospermae. Systematic Botany 9: 297–323.

SHIMADA, H., M. FUKUTA, M. ISHIKAWA, and M. SUGIURA. 1990. Rice chloroplast RNA polymerase genes: the absence of an intron in *rpoC1* and the presence of an extra sequence in *rpoC2*. Molecular and General Genetics 221: 395–402.

SHINOZAKI, K., M. OHME, M. TANAKA, T. WAKASUGI, N. HAYASHIDA, T. MATSUBAYASHI, N. ZAITA, J. CHUNWONGSE, J. OBOKATA, K. YAMAGUCHI-SHINOZAKI, C. OHTO, K. TORAZAWA, B.-Y. MENG, M. SUGITA, H. DENO, T. KAMAGOSHIRA, K. YAMADA, J. KUSUDA, F. TAKAIWA, A. KATO, N. TOHDOH, H. SHIMADA, and M. SUGIURA. 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its organization and expression. EMBO Journal 5: 2043–2049.

SMALL, E. and M. JOMPHE. 1989. A synopsis of the genus *Medicago* (Leguminosae). Canadian Journal of Botany 67: 3260–3294.

SNOW, N. and J. M. MACDOUGAL. 1993. New chromosome reports in *Passiflora* (Passifloraceae). Systematic Botany 18: 261–273.

SWOFFORD, D. L. 1993. *PAUP: Phylogenetic analysis using parsimony*, version 3.1.1 Champaign: Illinois Natural History Survey.

WAKASUGI, T., J. TSUDZUKI, S. ITO, M. SHIBATA, and M. SUGIURA. 1994. A physical map and clone bank of the black pine (*Pinus thunbergii*) chloroplast genome. Plant Molecular Biology Reporter 12: 227–241.

WALLACE, R. S. and J. H. COTA. 1996. An intron loss in the chloroplast gene *rpoC1* supports a monophyletic origin for the subfamily Cactoideae of the Cactaceae. Current Genetics 29: 275–281.

WOLFE, K. H., M. GOUY, Y. W. YANG, P. M. SHARP, and W.-H. LI. 1989. Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. Proceedings of the National Academy of Sciences U.S.A. 86: 6201–6205.

ZURAWSKI, G. and M. T. CLEGG. 1987. Evolution of higher-plant chloroplast DNA-encoded genes: implications for structure-function and phylogenetic studies. Annual Review of Plant Physiology 38: 391–418.

APPENDIX 1. Taxa surveyed for presence/absence of the cpDNA *rpoC1* intron. Those taxa for which sequence data have been obtained are indicated by asterisks. Source of material and voucher, if available, are given after each species (Downie's vouchers are, or will be, deposited at ILL). Abbreviations are as follows: BBG = Brooklyn Botanical Garden; MBG = Matthaei Botanical Garden, University of Michigan; MoBG = Missouri Botanical Garden; MSU = Michigan State University W. J. Beal Botanical Garden; UIUC = University of Illinois at Urbana-Champaign Plant Science Laboratory; USDA = United States Department of Agriculture, Washington State University.

Aizoaceae: *Delosperma cooperi* (Hook. f.) L.* (BBG; *Downie 1060*); *Delosperma nubigenum* (Schltr.) L. (BBG; *Downie 1057*); *Faucaria tigrina* (Haw.) Schwantes* (MSU; *Downie 1023*); *Monilaria moniliformis* (Thunb.) H. D. Ihlenf. & S. Jörg. (UIUC 89203; *Downie 716*); *Tetragonia tetragonioides* (Pall.) Kuntze* (MSU 89B423; *Downie 1070*). Alismataceae: *Sagittaria* sp. (Palmer Lab DNA no. 772). Amaranthaceae: *Amaranthus tricolor* L. (MSU; *Downie 756*); *Celosia argentea* L. 'Pink Tassles' (BBG; *Downie 1053*). Apiaceae: *Coriandrum sativum* L.* (cult. UIUC; *Downie s.n.*); *Daucus carota* L.* (*Downie 741*). Apocynaceae: *Aspidosperma myristicifolium* Markgr. (*Potgieter 5*). Aristolochiaceae: *Aristolochia durior* Hill (Palmer Lab DNA no. 17). Asteraceae: *Barnadesia caryophylla* (Vell.) S. F. Blake* (MBG; *Olmstead 16;* Palmer Lab DNA no. 154); *Lactuca sativa* L. (cult. UIUC; *Downie s.n.*). Basellaceae: *Anredera cordifolia* (Ten.) Steenis (MBG 840353; *Olmstead 51;* Palmer Lab DNA no. 163). Begoniaceae: *Begonia* × *semperflorens-cultorum* (UIUC 85070; *Downie 656*). Bixaceae: *Bixa orellana* L.* (UIUC 85414; *Downie 687*). Cactaceae: *Pereskia grandifolia* Haw. (MBG; *Olmstead 46;* Palmer Lab DNA no. 158). Calyceraceae: *Boopis* sp.* (*A. Hunziker;* Palmer Lab DNA no. 817); *Calycera sympaganthera* Kuntze* (*M. Devore;* Palmer Lab DNA no. 906); *Gamocarpha poeppigii* DC.* (*M. Devore;* Palmer Lab DNA no. 905). Campanulaceae: *Campanula garganica* Ten.* (*T. Ayers;* Palmer Lab DNA no. 480). Caprifoliaceae: *Kolkwitzia amabilis* Graebn.* (cult. UIUC; *Downie s.n.*); *Sambucus canadensis* L. (cult. UIUC; *Downie s.n.*); *Symphoricarpos albus* (L.) S. F. Blake (cult. UIUC; *Downie s.n.*). Caricaceae: *Carica papaya* L. 'Solo' * (UIUC 90321; *Downie 666*). Caryophyllaceae: *Lychnis coronaria* (L.) Desr. (BBG; *Downie 1048*); *Petrorhagia saxifraga* (L.) Link (MSU; *Downie 1069*). Chenopodiaceae: *Archiatriplex nanpinensis* Chu (*Downie 759*); *Sarcobatus vermiculatus* (Hook.) Torr. (*Downie 726*). Cornaceae: *Aucuba japonica* Thunb. (*L. Hufford;* Palmer Lab DNA no. 947); *Cornus florida* L.* (cult. UIUC; *Downie s.n.*). Cucurbitaceae: *Cucurbita moschata* (Duchesne) Poir. (cult. UIUC; *Downie 667*). Datiscaceae: *Datisca glomerata* Baill.* (UIUC 90054; *Downie 694*). Didiereaceae: *Alluaudia montagnacii* Rauh (BBG; *Downie 1055*). Dipsacaceae: *Dipsacus sylvestris* Huds.* (*Downie s.n.*). Droseraceae: *Drosera* sp. (cult. UIUC; *Downie s.n.*). Euphorbiaceae: *Acalypha hispida* Burm. f. (UIUC 85465; *Downie 668*); *Euphorbia milii* Des Moul.* (UIUC 85165; *Downie 655*); *Jatropha podagrica* Hook.* (UIUC 85204; *Downie 659*); *Trewia nudiflora* L. (UIUC 85596; *Downie 685*). Fabaceae: *Acacia farnesiana* L. (*Clarke 1*); *Albizia julibrissin* Durazz. (*Clarke 19*); *Caesalpinia pulcherrima* (L.) Sw. (*Clarke 99*); *Calliandra* sp. (*Seigler 13822*); *Enterolobium cyclocarpum* Griseb. (*Seigler 13821*); *Glycine max* (L.) Merr.* (cult. UIUC, *Downie s.n.*); *Inga edulis* Mart. (*Clarke 154*); *Lablab purpureus* L. (*Clarke 102*); *Lysiloma* sp. (*Seigler 13823*); *Medicago polymorpha* L. (USDA 577440; *Downie 894*); *M. rigidula* (L.) All. (USDA 537228; *Downie 895*); *M. ruthenica* (L.) Ledeb. (USDA 568103; *Downie 896*); *M. sativa* L.* (*Downie 862*); *M. sauvagei* Nègre (USDA 499153; *Downie 904*); *M. suffruticosa* Ramond ex DC. subsp. *leiocarpa* (Benth.) P. Fourn.* (USDA 516912; *Downie 907*); *M. truncatula* Gaertn. (USDA 577634; *Downie 905*); *M. turbinata* (L.) All. (USDA 577644; *Downie 906*); *Phaseolus vulgaris* L. (*Downie 866*); *Pisum sativum* L.* (*Downie 814*); *Pithecellobium saman* (Jacq.) Benth. (*Seigler 13921*); *Trifolium repens* L. (*Downie 864*). Flacourtiaceae: *Flacourtia inermis* Roxb.* (UIUC 85482; *Downie 684*). *Oncoba spinosa* Forssk.* (UIUC 85535; *Downie 686*). Goodeniaceae: *Goodenia ovata* Smith* (Palmer Lab DNA no. 982); *Scaevola sericea* Vahl* (Palmer Lab DNA no. 978). Grossulariaceae: *Escallonia rubra* (Ruiz & Pav.) Pers.* (*L. Hufford;* Palmer Lab DNA no. 949). Hydrangeaceae: *Hydrangea* sp. (*L. Hufford;* Palmer Lab DNA no. 945). Hydrophyllaceae: *Hydrophyllum virginianum* L. (*Downie s.n.*). Liliaceae: *Allium cepa* L. (cult. UIUC; *Downie s.n.*). Lobeliaceae: *Lobelia mildbraedii* Engl.* (*E. Knox;* Palmer Lab DNA no. 742). Malpighiaceae: *Banisteriopsis caapi* (Griseb.) Morton (UIUC 90319; *Downie 712*). Malvaceae: *Gossypium hirsutum* L. (*N. Ayres;* Palmer Lab DNA no. 186). Menyanthaceae: *Menyanthes trifoliata* L.* (*R. Olmstead;* Palmer Lab DNA no. 301); *Nymphoides thunbergiana* Kuntze* (*Pagels s.n.; Downie 752*); *Villarsia exaltata* F. Muell. (*Ornduff s.n.; Downie 704*); *Villarsia violifolia* F. Muell.* (*Ornduff s.n.; Downie 709*). Molluginaceae: *Mollugo verticilliata* L. (MSU; *Downie 1068*). Nepenthaceae: *Nepenthes* sp. (*M. Chase;* Palmer Lab DNA no. 918). Nyctaginaceae: *Mirabilis nyctaginea* (Michx.) MacMill.* (MSU B87137; *Downie 1067*). Oleaceae: *Forsythia* sp. (Palmer Lab DNA no. 781). Paeoniaceae: *Paeonia lactiflora* Pall. (*R. Olmstead;* Palmer Lab DNA no. 253). Passifloraceae: *Passiflora caerulea* L. (UIUC 85650; *Downie 592*); *P. capsularis* L. (UIUC 89247; *Downie 604*); *P. coccinea* Aubl. (*Downie 594*); *P. eichleriana* Mast. (UIUC 89262; *Downie 642*); *P. foetida* L.* (UIUC 85653; *Downie 597*); *P. helleri* Peyr.* (MBG; Palmer Lab DNA no. 707); *P. incarnata* L.* (UIUC 87042; *Downie 598*); *P. morifolia* Mast.* (UIUC 89216; *Downie 591;* UIUC 89256; *Downie 647*); *P. suberosa* L.* (UIUC 85655; *Downie 603;* UIUC 89144; *Downie 599;* UIUC 89264; *Downie 670*); *P. quadrangularis* L. (UIUC 88137; *Downie 650*). Petiveriaceae: *Petiveria alliacea* L. (UIUC 85543; *Downie 714*). Phytolaccaceae: *Phytolacca americana* L. (BBG; *Downie 1054*). Pittosporaceae: *Pittosporum tobira* (Thunb.) Ait. (*Downie 534*). Plumbaginaceae: *Goniolimon tataricum* (L.) Boiss. (MoBG; *Downie 546*). Poaceae: *Triticum aestivum* L. (*Downie s.n.*). Portulacaceae: *Calandrinia ciliata* (Ruiz & Pav.) DC. (MSU; *Downie 1066*); *Claytonia perfoliata* Donn (MSU 90B1242W; *Downie 1031*); *Portulaca umbraticola* Kunth (cult. UIUC; *Downie 717*). Punicaceae: *Punica granatum* L. (UIUC 85565; *Downie 710*). Rosaceae: *Prunus serotina* Ehrh. (*R. Olmstead;* Palmer Lab DNA no. 249). Rubiaceae: *Rubia tinctorum* L. (UIUC 85573; *Downie 711*). Solanaceae: *Nicotiana tabacum* L. (cult. UIUC; *Downie s.n.*). Stegnospermataceae: *Stegnosperma halimifolium* Benth. (MoBG 720287; *Downie 1061*). Ulmaceae: *Ulmus americana* L. (*R. Olmstead;* Palmer Lab DNA no. 254). Violaceae: *Viola dissecta* Ledeb. var. *chaerophylloides* (Regel) Makino (UIUC 85348; *Downie 658*).