Down the Slippery Slope: Plastid Genome Evolution in Convolvulaceae

Saša Stefanović,[1,2] Richard G. Olmstead[1]

[1] Department of Biology, University of Washington, Box 355325, Seattle, WA 98195-5325, USA
[2] Department of Biology, University of Toronto at Mississauga, Mississauga, Ontario, L5L 1C6, Canada

Abstract. *Cuscuta* (dodder) is the only parasitic genus found in Convolvulaceae (morning-glory family). We used long PCR approach to obtain large portions of plastid genome sequence from *Cuscuta sandwichiana* in order to determine the size, structure, gene content, and synteny in the plastid genome of this *Cuscuta* species belonging to the poorly investigated holoparasitic subgenus *Grammica*. These new sequences are compared with the tobacco chloroplast genome, and, where data are available, with corresponding regions from taxa in the other *Cuscuta* subgenera. When all known plastid genome structural rearrangements in parasitic and nonparasitic Convolvulaceae are considered in a molecular phylogenetic framework, three categories of rearrangements in *Cuscuta* are revealed: plesiomorphic, autapomorphic, and synapomorphic. Many of the changes in *Cuscuta*, previously attributed to its parasitic mode of life, are better explained either as plesiomorphic conditions within the family, i.e., conditions shared with the rest of the Convolvulaceae, or, in most cases, autapomorphies of particular *Cuscuta* taxa, not shared with the rest of the species in the genus. The synapomorphic rearrangements are most likely to correlate with the parasitic lifestyle, because they represent changes found in *Cuscuta* exclusively. However, it appears that most of the affected regions, belonging to all of these three categories, have probably no function (e.g., introns) or are of unknown function (a number of open reading frames, the function of which, if any, has yet to be discovered).

Correspondence to: Saša Stefanović; *email:* sstefano@utm.utoronto.ca

Introduction

The plastid genome of land plants is highly conserved in size, structure, gene content, and synteny (Palmer 1985, 1991; Palmer and Stein 1986; Downie et al. 1991, 1994; Downie and Palmer 1992). Any changes may be important phylogenetic markers (Palmer 1985, 1990), but perhaps more importantly, they may also be used to study molecular evolutionary and genetic processes involving the structure, function, and evolution of the plant plastid genome (Palmer 1985, 1990, 1991; dePamphilis and Palmer 1990).

The morning-glory family, Convolvulaceae, is one of the few angiosperm families exhibiting substantial plastid genome rearrangement (Downie and Palmer 1992). There is evidence from a small number of sampled species indicating structural differences in the Convolvulaceae plastid DNA (ptDNA) relative to other flowering plants. For example, the *rpl2* gene is interrupted by an intron in most but not all angiosperms. Based on plastid genome sequences of diverse land plants, including one liverwort (Ohyama et al. 1986), one gymnosperm (Wakasugi et al. 1994), two basal angiosperms (Goremykin et al. 2003, 2004), five eudicots (Shinozaki et al. 1986a; Spielmann et al. 1988; Sato et al. 1999; Hupfer et al. 2000; Kato et al. 2000), and five monocots (Posno et al. 1986; Hiratsuka et al. 1989; Maier et al. 1995; Ogihara et al.

2002; Stefanović et al. 2004), it became evident that this intron was present in the common ancestor of flowering plants (and probably land plants), but was subsequently lost in some angiosperm lineages (e.g., *Spinacia oleracea*; Zurawski et al. 1984; Sehmitz-Linneweber et al. 1999). In their large-scale filter hybridization survey across angiosperms, Downie et al. (1991) reported at least six independent losses of the *rpl2* intron, including in the ancestor of two representatives of green Convolvulaceae and one *Cuscuta* species. Outside of Convolvulaceae, no other member of the Asteridae was found to lack this intron. In another example, Downie and Palmer (1992) reported the loss of an open reading frame (ORF). *ycf1* (= ORF1244 in *Nicotiana tabacum*), of unknown function, but common to plastid genomes of land plants in three nonparasitic Convolvulaceae and one *Cuscuta* species (as well as in Poaceae, Fabaceae, and some other families not closely related to Convolvulaceae). Given the conserved nature and distribution of this ORF across the land plants, it is most parsimonious to conclude that *ycf1* was lost several times independently within angiosperms, similar to the *rpl2* intron. A third example is the *atpB* gene which, like *rbcL*, is very conserved in its length across all angiosperms (except at the very end of the gene). However, in the two members of green Convolvulaceae for which sequences were known, an in-frame deletion encoding two amino acids has been reported at the 5′ end of this gene (Savolainen et al. 2000).

Cuscuta, the only parasitic genus associated with Convolvulaceae, has been the subject of more extensive molecular analyses than the rest of this family (Machado and Zetche 1990; Haberhausen et al. 1992; Bömmer et al. 1993; Haberhausen and Zetsche 1994; van der Kooij et al. 2000; Krause et al. 2003). Most major synoptic works on flowering plants (e.g., Cronquist 1988; Takhtajan 1997) accept *Cuscuta* at the family level, thus implying that it is only distantly related to other species in Convolvulaceae, but recent phylogenetic results indicate it belongs within Convolvulaceae (Stefanović et al. 2002; Stefanović and Olmstead 2004). Cytological, anatomical, and morphological characters indicate the presence of three well-defined groups in *Cuscuta* and classification in three subgenera (*Cuscuta*, *Grammica*, and *Monogyna*) was proposed by Engelmann (1859; formalized by Yuncker 1932). Members of this genus are characterized by twining, slender, pale stems, with reduced, scale-like leaves, no roots, and are attached to the host by haustoria, therefore depending entirely or almost entirely on their hosts to supply water and nutrients (Kuijt 1969; Dawson et al. 1994). Many *Cuscuta* species are characterized also by the reduced amounts or absence of chlorophylls (van der Kooij et al. 2000). However, some species produce significant amounts of chlorophylls, especially in the tips of

seedlings not attached to the host and in fruiting sepals and ovaries (Panda and Choudhury 1992; Dawson et al. 1994). Thus, both holoparasitic and hemiparasitic species occur in this genus. In holoparasites, found mostly in subgenera *Cuscuta* and *Grammica*, all carbohydrates are provided by the host. In herniparasites, confined predominantly to subgenus *Monogyna*, carbohydrates supplied by hosts are supplemented, to a small extent, by the plant's own photosynthesis (Hibberd et al. 1998). This diversity of photosynthetic ability among *Cuscuta* species initiated several physiological studies of photosynthetic enzymes (Machado and Zetche 1990; van der Kooij et al. 2000) and molecular evolution studies of the plastid genome (Haberhausen et al. 1992; Bömmer et al. 1993; Haberhausen and Zetsche 1994; Krause et al. 2003).

Even though the plastids of *Cuscuta reflexa* (subgenus *Monogyna*) have no visible grana and the number of thylakoids is highly reduced compared to normal plastids, Machado and Zetche (1990) showed the presence of both chlorophyll a and b, detected residual light- and CO_2-dependent photosynthetic activity, and demonstrated incorporation of $^{14}CO_2$ into carbohydrates. In contrast, the plastids of *C. europaea* (subgenus *Cuscuta*) lack both grana and thylakoids and appear to lack chlorophylls completely. Consequently, neither ribulose-1,5-bisphosphate carboxylase-oxygenase (Rubisco) activity nor light-dependent CO_2 fixation could be detected (Machado and Zetche 1990). Further studies, focusing mainly on species of subgenus *Grammica*, demonstrated that in most cases the *rbcL* gene, thylakoids, and chlorophylls are present, and low amounts of the large subunit of Rubisco could be detected immunologically (van der Kooij et al. 2000). However, two species from the same subgenus have been shown to lack thylakoids and chlorophylls, and neither the *rbcL* gene nor its protein product, Rubisco large subunit, could be detected (van der Kooij et al. 2000).

The ptDNA sequence of *Cuscuta reflexa*, a member of the predominantly hemiparasitic subgenus *Monogyna*, has been studied in most detail from a molecular evolutionary standpoint. Results indicate that this species retains an affected, yet functional, plastid genome (Haberhausen et al. 1992; Bömmer et al. 1993; Haberhausen and Zetsche 1994). *Epifagus virginiana* (Orobanchaceae, Lamiales) is the only parasitic plant species for which the entire plastid genome has been mapped (dePamphilis and Palmer 1990) and subsequently sequenced (Wolfe et al. 1992). In contrast to this holoparasitic species, which lacks all photosynthetic and most putatively chlororespiratory *ndh* genes (dePamphilis and Palmer 1990). *C. reflexa* retains most of the plastid genes generally found in autotrophic land plants, including both

those involved in photosynthesis and "house-keeping" functions (Haberhausen et al. 1992). However, chlororespiratory (*ndh*) genes seem to be either altered to the point of becoming pseudogenes (e.g., *ndhB*) or are lost from the plastid genome (Haberhausen and Zetsche 1994). In addition, a large deletion (~6.5 kb) at the junction between the inverted repeat (IR) and the large single-copy (LSC) region, comprising two ribosomal protein genes, *rpl2* and *rpl23*, one tRNA (*trnI*-CAA), as well as a large portion of the *ycf2* gene (= ORF2280 in tobacco), has been reported in *C. reflexa* (Bömmer et al. 1993). Given this sequence information and the absence of detectable Southern hybridization to total *C. reflexa* cellular DNA using tobacco probes for affected genes, Bömmer et al. (1993) postulated the total loss of *rpl2* and *rpl23* and hypothesized further that the whole translation apparatus might be nonfunctional in this *Cuscuta* species.

Utilizing a combination of heterologous and homologous Southern hybridization, in conjunction with PCR amplification using internal primers, Krause et al. (2003) showed evidence for parallel loss of the *rpoA* and *rpoB* genes coding for the plastid-encoded RNA polymerase (PEP) in three holoparasitic *Cuscuta* species from subgenus *Grammica*. These genes were, however, retained as ORFs in hemiparasitic *C. reflexa*, and one examined species of autotrophic Convolvulaceae showed strong hybridization signal for both genes under investigation (Krause et al. 2003).

Most of these findings of plastid genome structural rearrangements in *Cuscuta* were attributed to its parasitic lifestyle, but without comparison to related nonparasitic members of the family. Also, sequence analysis of *Cuscuta* species other than *C. reflexa* are quite limited (but see Freyer et al. 1995; Krause et al. 2003). The plastid genome of *C. europaea* (subgenus *Cuscuta*) was shown to retain an *rbcL* ORF, as well as plastid-encoded *psbA* and nucleus-encoded *rbcS* genes, although this species contains no chlorophyll and is unable to photosynthesize (Machado and Zetsche 1990). In addition, more extensive ptDNA sequences from species belonging to the subgenus *Grammica* which could have potentially the most affected plastid genome (van der Kooij et al. 2000; Krause et al. 2003), have not been reported.

Collectively, all of these observations make Convolvulaceae one of the very few angiosperm families showing substantial structural variation in the plastid genome. This is even more intriguing considering that the sister family. Solanaceae, has virtually no rearrangements and represents a "typical" chloroplast genome (Olmstead and Palmer 1992; Sugiura 1992). However, in order to accurately trace the sequence of events in plastid molecular evolution within Convolvulaceae, a reconstruction of phylogenetic rela-

tionships among parasitic and autotrophic members of the family is necessary. Convolvulaceae have been the subject of only two family-wide molecular phylogenetic studies (Stefanović et al. 2002; Stefanović and Olmstead 2004). In aggregate, these two studies used seven genes drawn from all three plant genomes (five chloroplast, one mitochondrial, and one nuclear) with taxonomic sampling ranging from 32 to 109 species, representing the diversity of both parasitic and nonparasitic Convolvulaceae. *Cuscuta* was shown to be nested well within Convolvulaceae with at least two autotrophic lineages diverging before *Cuscuta*. However, the exact sister group of *Cuscuta* could not be ascertained, even though many alternatives were tested and rejected with confidence (Stefanović et al. 2002; Stefanović and Olmstead 2004).

The purpose of the present study is four-pronged: (1) to introduce new ptDNA sequence data for the underinvestigated *Cuscuta* species belonging to the holoparasitic subgenus *Grammica*, (2) to summarize the patterns of ptDNA variation observed in sequence data drawn from a large sample of heterotrophic and autotrophic species of Convolvulaceae, (3) to offer an interpretation for newly obtained data and for data obtained in previous studies in light of the current estimate of phylogenetic relationships within Convolvulaceae, and (4) to investigate implications of observed ptDNA variation on the molecular processes of plastid genome evolution in Convolvulaceae.

Materials and Methods

Seed of *Cuscuta sandwichiana* Choisy (subgenus *Grammica*) was obtained from a herbarium specimen deposited at the WTU herbarium (*Degener & Degener 36596*, collected in 1984 in Hawaii, USA), and were germinated in the greenhouse. The seedlings were grown on *Beta vulgaris* or *Coleus* species as host plants. Total genomic DNA was isolated from fresh *C. sandwichiana* tissue by the modified CTAB procedure (Doyle and Doyle 1987), followed by ultracentrifugation in CsCl-ethidium bromide gradient. DNA samples from other *Cuscuta* species and from other Convolvulaceae, obtained for phylogenetic studies, were also used (for extraction and voucher information see Stefanović et al. 2002 and Stefanović and Olmstead 2004).

A number of long polymerase chain reaction (PCR) amplifications were designed to amplify extensive portions of *C. sandwichiana* plastid genome. Also, a series of PCR experiments was designed to assay for two types of plastid rearrangements previously reported in *Cuscuta* and/or Convolvulaceae, namely the loss of genes and/or introns and the presence of large inversions. Presence/absence of introns was determined by comparing the size of amplified products of organisms under investigation with those of a reference species (in most cases tobacco). Some representative products were also sequenced to find out whether the intron loss was complete or partial, and to examine its boundaries. To assay for the inversion, two pairs of forward and reverse primers were used, each pair spanning one of the putative inversion endpoints, resulting in four diagnostic amplifications (Wolfe and Liston 1998).

The presence/absence of inversions is determined from the resulting pattern of successful and failed amplifications. Tobacco was used as a control.

Long PCRs were conducted using the Expand™ Long Template PCR System kit (Roche Applied Science), following instructions provided by the manufacturer. Other PCRs were conducted using *Taq* polymerase (Promega). Initially, sets of primers designed by Stefanović et al. (2004), which cover a large portion of the plastid genome, were used for amplifications and/or sequencing. Based on these initial sequences, a number of *Cuscuta*-specific sequencing primers were designed and used for chromosome walking with long PCR products. Primer sequences are available upon request from SS. PCR products were separated by electrophoresis using 1% agarose gels and were visualized with ethidium-bromide. Amplified PCR products from both long and regular amplifications, intended for sequencing, were cleaned using Qiagen columns. Cleaned products were then directly sequenced using the BigDye™ Terminator cycle sequencing kit (Applied Biosystems) on an ABI 377 DNA automated sequencer (Applied Biosystems). Sequence data were edited and assembled using Sequencher™ 4.1 (Gene Codes Corporation). All newly obtained sequences reported in this study are deposited in GenBank (accession numbers AY936335-AY9336360).

Slot-blot DNA hybridizations followed standard procedures described in Doyle et al. (1995) and Adams et al. (1999). Immobilon-Ny+ membranes (Millipore) were prehybridized and hybridized at 60–62°C and filters were washed at the hybridization temperature. PCR-derived probes were labeled with ^{32}P using random oligonucleotide primers. Autoradiography was carried out using intensifying screens at room temperature overnight (for positive control probe) or −80°C for 18–48 h (for other probes).

Results and Discussion

PtDNA Variation Within Cuscuta and Its Comparison With Tobacco Plastid Genome

Newly obtained sequence data from *Cuscuta* subgenus *Grammica* (this study) and existing sequences from *Cuscuta* subgenus *Monogyna* (Haberhausen et al. 1992; Bömmer et al. 1993; Haberhausen and Zetsche 1994) were compiled and compared with corresponding regions of *Nicotiana tabacum* cpDNA (Shinozaki et al. 1986a). These sequence results, together with the results from a series of experiments involving only PCR amplification, using *Cuscuta* species from all three subgenera, are summarized in Fig. 1.

It was demonstrated by Southern hybridization that the *rpl2* intron is missing in five species belonging to tribes Ipomoeeae and Convolvuleae as well as in one species of *Cuscuta* (Downie et al. 1991). We have documented further a family-wide loss of this intron, using primers located in the conserved regions of the surrounding gene. Intron absence is confirmed for 30 representatives selected from throughout the Convolvulaceae (see Fig. 3), including *Humbertia*, for all of which the PCR bands are ∼700 bp shorter than the control (tobacco). This PCR-based determination was investigated further by DNA sequencing in order to determine the nature of the intron loss in *rpl2* (i.e., partial or complete deletion, precise or imprecise in-

tron excision). This approach confirmed the complete loss of the *rpl2* intron in Convolvulaceae (see Fig. 3; see below), resulting in combination of the two exons into a single ORF, similar to the intron-deletion cases encountered in previous studies (e.g., *rpoC1, clpP, rpl16, trnI*; Hiratsuka et al. 1989; Zurawski et al. 1984; Downie et al. 1991).

The situation with *Cuscuta* is more complicated. Bömer et al. (1993) reported the complete loss of the *rpl2* gene, along with *rpl23, trnI*, and a big portion of the *ycf2* homologue in *C. reflexa* (=ORF2280 in tobacco). These authors argued that the translation apparatus must, therefore, be nonfunctional in this parasitic genus. However, we were able to amplify the *rpl2* gene, with the same set of primers used for the nonparasitic taxa, in several *Cuscuta* species representing all three subgenera including *C. japonica*, a close relative of *C. reflexa* (subgenus *Monogyna*; Fig. 1; see also Fig. 3). These conflicting results can be reconciled in two ways: (1) either this large deletion may be a taxon-specific loss in *C. reflexa*, thus an autapomorphy for this species rather than a genus-wide feature linked to the parasitic lifestyle, or more likely (2) this represents a contraction of the inverted repeat (IR) in the subgenus *Monogyna* and possibly in the entire genus (see also Plunkett and Downie 2000). The reported deletion for *C. reflexa* occurs precisely at the tobacco IR_A-LSC junction (i.e., J_{LA}) and could also be explained by a contraction of the IR_A in this species. If the J_{LA} in *C. reflexa* were to fall within *ycf2*, the region containing *rpl2, rpl23, trnI*, and a 5′ end of the *ycf2* would still be found in the LSC region adjacent to the other IR (i.e., IR_B). Given that only one junction was sequenced (J_{LA}, as indicated by the presence of *trnH*, which is always neighboring the IR_A in land plants), this possibility might have been overlooked. Bömer et al. (1993) reported negative Southern hybridization results of *C. reflexa* total cellular DNA hybridized against *rpl2/rpl23* and 5′-*ycf2* probes from tobacco, thus implying the total absence of these regions from all *C. reflexa* genomes. However, the authors failed to provide a positive control demonstrating the presence of detectable levels of *C. reflexa* ptDNA on these blots. Our result, demonstrating the presence of *rpl2* in closely related *C. japonica* (subgenus *Monogyna*), further supports the notion of the IR contraction and retention of only one copy of *rpl2* (and probably also one copy of *rpl23, trnI*, and *ycf2* for which sequence confirmations are not available).

Haberhausen and Zetsche (1994) showed that the *ndhB* gene is reduced to a pseudogene in *C. reflexa* due to many frameshift mutations, while the other ten *ndh* genes are either lost or strongly altered in this species, as suggested by lack of hybridization signals using heterologous gene probes derived from tobacco. We have documented, using PCR amplifica-
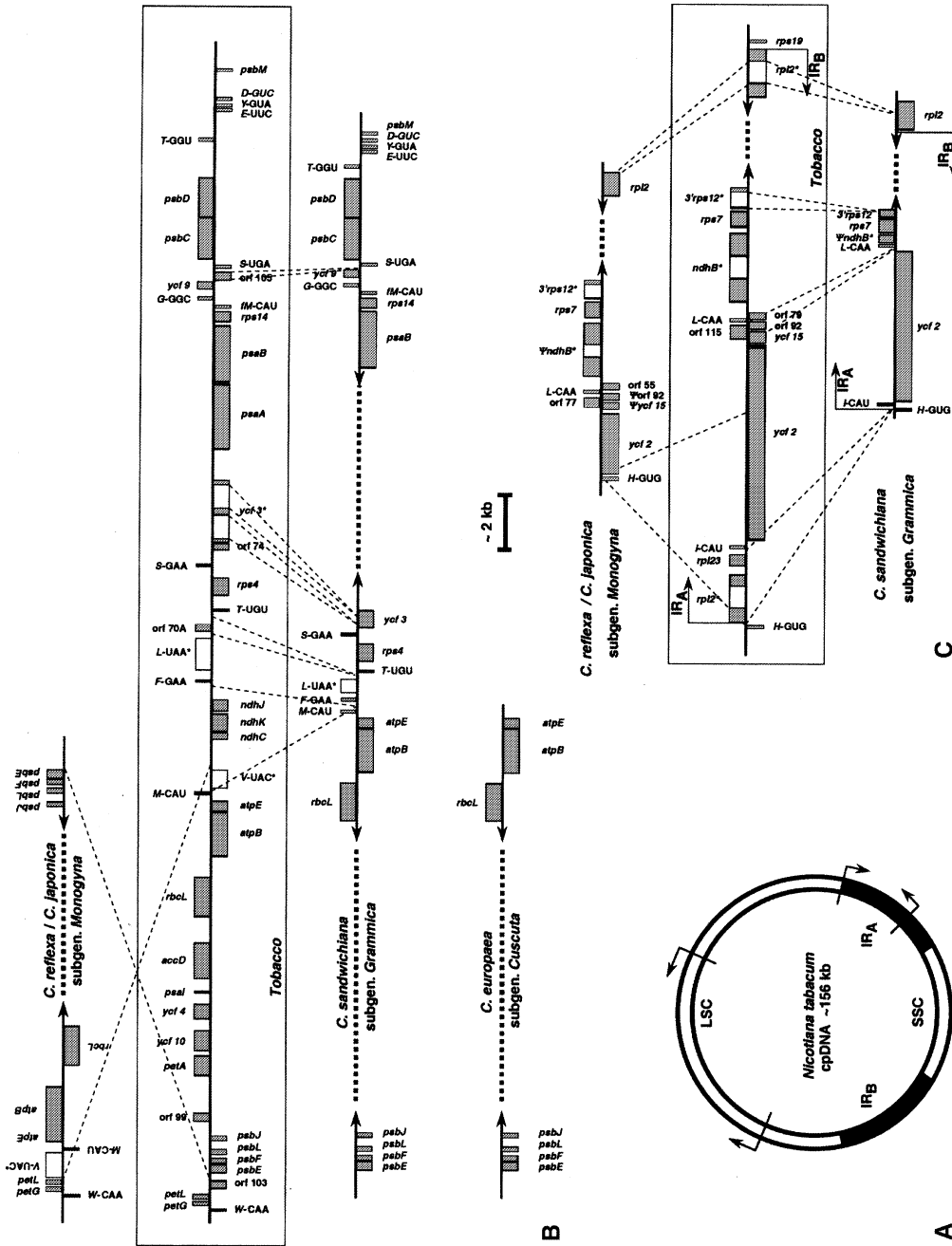
Fig. 1. Comparison of the size, structure, gene content, and synteny in the plastid genomes of tobacco and different species of the parasitic genus *Cuscuta*. The newly obtained sequence data from *C. sandwichiana* (subgenus *Grammica*) and existing sequences from *C. reflexa* (subgenus *Monogyna*; Haberhausen et al. 1992; Bömmer et al. 1993; Haberhausen and Zetsche 1994) are compiled and compared with corresponding regions of tobacco chloroplast DNA (cpDNA) shown in the box between the *Cuscuta* genome maps. Shaded boxes depict coding regions; open boxes indicate introns (asterisks additionally denote genes that have introns), and thick dotted lines depict unsequenced regions. Genes shown above the line are transcribed from right to left; genes below the line are transcribed from left to right. Gene nomenclature follows Wakasugi et al. (1998). (**A**) Structural organization of tobacco cpDNA according to Wakasugi et al. (1998). The inverted repeat regions (IR$_A$ and IR$_B$) divide the rest of the genome into large single copy (LSC) and small single copy (SSC). Arrows indicate location of the LSC and the IR regions used for detailed comparison with parasitic ptDNA as shown in B and C. (**B**) Comparison of *C. sandwichiana* sequences with the corresponding LSC regions of tobacco (~40 kb). The ~15-kb inversion in *Cuscuta* subgenus *Monogyna* is depicted with dashed lines as well as the lack of it in the other two *Cuscuta* subgenera (as determined by the PCR experiments and/or sequencing; see text for full explanation). (**C**) Comparison of *C. sandwichiana* and *C. reflexa* sequences with the corresponding IR$_A$ regions of tobacco (~15 kb).
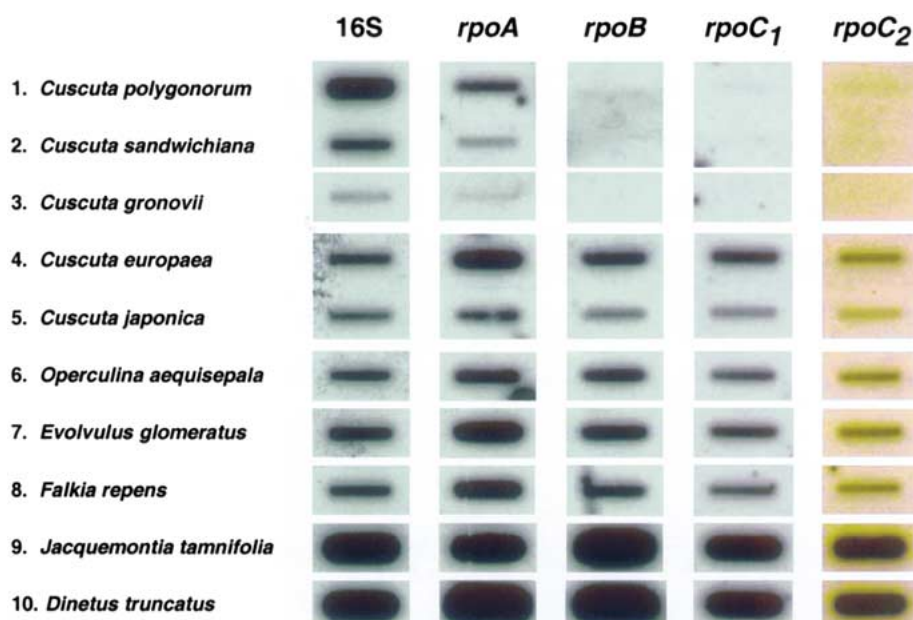
Fig. 2. Autoradiographs showing slot-blot hybridization results of four probes derived from *rpo* genes to DNAs from ten Convolvulaceae species. Small plastid ribosomal subunit (16S rRNA) was used as positive control. Species 1–5 are from parasitic genus *Cuscuta* (1–3 belong to subgenus *Grammica*, 4 to subgenus *Cuscuta*, and 5 to subgenus *Monogyna*), whereas species 6–10 represent diverse autotrophic members of the family. Note that the absence of *rpoB*, *rpoC₁*, and *rpoC₂* hybridization as well as significant signal intensity diminution for *rpoA* are restricted to three *Cuscuta* subgenus *Grammica* representatives.

tion, the absence of at least one cluster of *ndh* genes (*ndhCJK*) in several other *Cuscuta* species, sampled from all three subgenera. The same cluster is present, however, in five representatives sampled throughout nonparasitic Convolvulaceae (*Operculina aequisepala, Falkia repens, Evolvulus glomeratus, Jacquemontia tamnifolia*, and *Dinetus truncatus*) and sequencing confirmed that these three *ndh* genes are present as ORFs in all five cases. In addition, the amplification of the *ndhF* gene, using primers both within the coding and flanking regions, failed in multiple *Cuscuta* species examined (using the same DNA samples that yielded PCR products with other primer combinations), while *ndhF* was easily amplifiable and is present as ORF in the same five above-mentioned nonparasitic Convolvulaceae species. Even though negative PCR results (i.e., the lack of amplification) cannot be explained unequivocally, this result is consistent with the one obtained by Southern hybridization (Haberhausen and Zetsche 1994).

The function of the *ndh* genes in the plant plastid genome was controversial in the past, but their involvement in chlororespiration now appears to be well established (dePamphilis and Palmer 1990; Sugiura 1992; Guedeney et al. 1996; Casano et al. 2000; Nixon 2000; reviewed in Pettier and Cournac 2002). The conserved nature of these genes suggests that a functional constraint is present throughout the angiosperms. Based on the genome map for the *Epifagus virginiana* plastid, which shows the complete loss of photosynthetic and chlororespiratory genes in this holoparasitic flowering plant, dePamphilis and Palmer (1990) proposed that the *ndh* genes are involved in a metabolic pathway associated with photosynthesis in autotrophic plants. However, in all *Cuscuta* species under investigation, including both hemi- and holoparasitic ones, most, if not all, photosynthetic genes are present in a more or less functional form, while all *ndh* genes are either lost or significantly altered, suggesting the decoupling of photosynthetic and chlororespiratory functions (see also Bungard 2004).

Loss of the *rpoA* and *rpoB* genes (coding for the plastid-encoded RNA polymerase; PEP) in three holoparasitic *Cuscuta* species from subgenus *Grammica* as well as retention of these genes as ORFs in one hemiparasitic species from subgenus *Monogyna* was demonstrated recently by Krause et al. (2003). In order to further investigate the phylogenetic extent of these findings, we conducted slot-blot hybridization on five *Cuscuta* species from all three subgenera (including subgenus *Cuscuta*, previously not assayed), as well as five autotrophic members of Convolvulaceae. In addition, our survey included probes for two remaining *rpo* genes, *rpoC₁* and *rpoC₂*. Results of the hybridization survey are presented in Fig. 2. In five nonparasitic species, hybridization signal is of similar intensity among all four *rpo* probes and the 16S rRNA positive control indicating a high level of similarity with tobacco *rpo* sequences. Likewise, *Cuscuta* species from subgenera *Cuscuta* and *Monogyna* (*C. europaea* and *C. japonica*, respectively) show little, if any, evidence of signal diminution. The presence of intact ORFs deduced from hybridization in these two *Cuscuta* species was further corroborated by complete or partial sequencing of all four *rpo* genes, and no evidence of frameshift deletion or premature stop codons was found. In contrast, three remaining parasitic species belonging to subgenus *Grammica* show a highly diminished signal for *rpoA* and complete absence of hybridization in three other genes (Fig. 2). PCR amplification and sequencing of

rpoA region using primes in adjacent plastid genes (*petD* and *rps11*) revealed the presence of *rpoA* pseudogenes in *C. sandwichiana* and *C. gronovii*, while multiple primer combinations in each of three remaining *rpo* genes failed to produce amplicons in these species, consistent with the negative hybridization results. Together with previously published data (Krause et al. 2003), our hybridization survey of *rpo* genes indicates that loss of these genes is concerted and within Convolvulaceae confined only to *Cuscuta* species belonging to subgenus *Grammica*.

A large inversion in the *C. reflexa* plastid genome, ~15 kb in length (Fig. 1B), was hypothesized based on the position of the *petG* gene relative to the *rbcL* gene and its transcriptional orientation (Haberhausen et al. 1992). Our PCR assay indicates that the same inversion reported for *C. reflexa* exists in *C. japonica* (both are species of subgenus *Monogyna*). However, this inversion is absent from the other species of *Cuscuta* belonging to subgenera *Cuscuta* and *Grammica* (Fig. 1B), as well as from all nonparasitic Convolvulaceae investigated. The arrangement of genes in ptDNAs from these species are syntenic with that of tobacco ptDNA, indicating that the 15-kb inversion appears to be a synapomorphy for *Cuscuta* subgenus *Monogyna*.

Long PCR amplifications with *C. sandwichiana* (subgenus *Grammica*) resulted in sequences from three ptDNA regions (Fig. 1B, C). The first sequenced area contains the *rbcL* gene, *atpB-E* operon, a series of tRNAs (*trnM*, *trnF*, *trnL*, and *trnT*), the *rps4* gene, *trnS*, and *ycf3* (= ORF168 in tobacco). This region, approximately 7 kb long, is collinear with a 15-kb region of tobacco ptDNA located in the large single copy (LSC) region (Fig. 1B). The difference in size is due to the combined effect of several deletions involving genes, intergenic spacer (IGS) regions, and introns. The most conspicuous deletion in this area of *C. sandwichiana* ptDNA is the portion that corresponds to 3.5 kb in tobacco containing *trnV* and the *ndhC-J* operon. Also lacking is one open reading frame (ORF70A), found in tobacco but not common in other plastid genomes (Wakasugi et al. 1998), and two introns usually present in *ycf3*, an ORF of unknown function that is found commonly throughout land plants. The juxtaposition of what are, in most plants, three exons into a single, uninterrupted *ycf3* gene shows that the introns have been removed precisely from this gene, keeping the ORF intact, similar to the situation found in the *rpl2* that was described above. No intermediate cases, i.e., those cases in which reduced intron(s) would be present, were found. This kind of "clean" intron deletion is generally thought to involve through a reverse-transcriptase-mediated mechanism, where the reverse transcription of a spliced transcript is followed by homologous recombination between the intronless cDNA and the original gene, as opposed to illegitimate recombination at a random site in the genome (Fink 1987; Dujon 1989, Hiratsuka et al. 1989; Downie et al. 1991).

The second sequenced region of *C. sandwichiana*, ~7 kb long, is collinear with a 10.5-kb region of tobacco ptDNA located also in the LSC (Fig. 1B). It contains, in the following order, the *psaB* and *rps14* genes, *trnfM* and *trnG*, *ycf9* (= ORF62 in tobacco), *trnS*, *psbC-D* operon, a cluster of tRNAs (*trnT*, *trnE*, *trnY*, and *trnD*) and the *psbM* gene. Almost all of the ~3-kb difference in size can be explained by deletions in the IGS regions. Only one ORF, tobacco-specific ORF105, is lacking in this portion of *C. sandwichiana* ptDNA compared to the tobacco sequence.

Finally, the third region of *C. sandwichiana* sequenced in this study (Fig. 1C) has similar linear gene arrangement as that reported for the inverted repeat A (IR$_A$) region of tobacco and *C. reflexa* (subgenus *Monogyna*; Bömmer et al. 1993). The portion of IR$_A$, which covers about 15 kb in tobacco, is reduced in this *Cuscuta* species to approximately 7 kb. The twofold size reduction is due to a large number of reductions/deletions including the following: (1) deletion of the *rpl2* gene and its intron as well as loss of the *rpl23* gene, (2) significant reduction in length of the largest plastid gene *ycf2* (= ORF2280 in tobacco), which has no known function, (3) a series of ORFs of unknown function (*ycf15*, ORF115, ORF92, and ORF79) are missing, (4) an intron usually found in the in the 3'-*rps12* gene is excised precisely, and (5) the *ndhB* gene (and its intron) is not only greatly reduced in length but also includes a number of frameshift mutations. This transformation of *ndhB* to a pseudogene and the length reduction, albeit to a somewhat lesser extent, is also reported from *C. reflexa* (Bömmer et al. 1993; Fig. 1C).

PtDNA Variation Within Convolvulaceae

Because the phylogenetic framework for the morning-glory family, obtained through molecular phylogenetic analyses (Stefanović et al. 2002; Stefanović and Olmstead 2004), shows *Cuscuta* nested within Convolvulaceae, and because the autotrophic members of this family also exhibit significant ptDNA rearrangements, it became evident that a more appropriate comparison of ptDNA variation observed in the parasitic *Cuscuta* would be with other members of Convolvulaceae, rather than with the more distantly related tobacco (Solanaceae). Hence, additional insight into the structural mutations in the plastid genome of the Convolvulaceae was provided by the ptDNA sequence data drawn from the large sample of autotrophic species of this family as well as a number of *Cuscuta* species from all three subgenera,
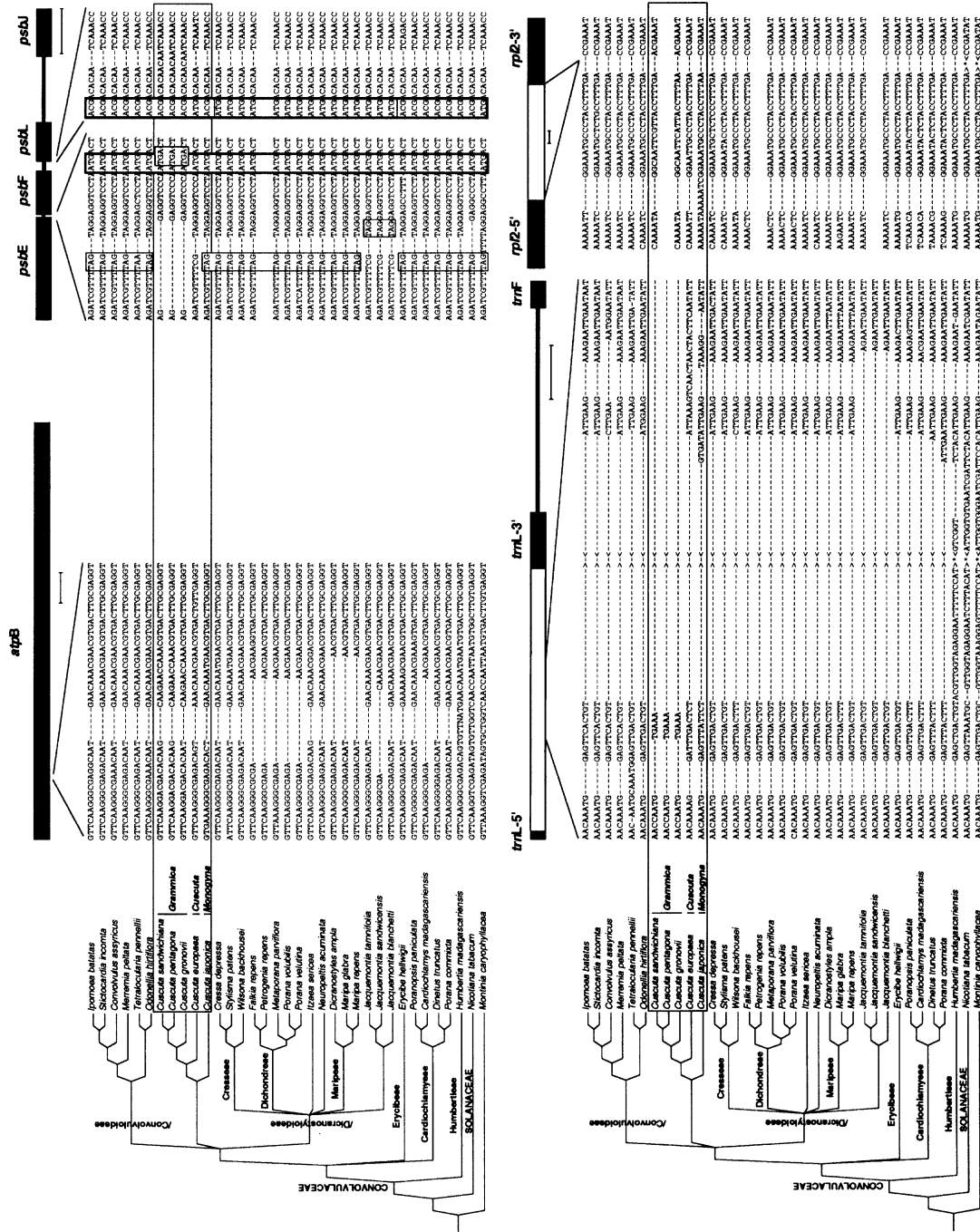
Fig. 3. Partial alignments from four plastid regions (*atpB* gene, *trnL-F* intron/spacer, *psbE-J* operon, and *rpl2* gene) showing portions relevant for comparison of plastid DNA structural mutations between parasitic and nonparasitic members of Convolvulaceae. These plastid regions were parts of molecular data matrices used originally for phylogeny reconstructions of Convolvulaceae, results of which are summarized in trees on the far left side (Stefanović et al. 2002; Stefanović and Olmstead 2004). Approximate position of the alignment within the corresponding gene/region is indicated above each alignment segment. Tobacco plastid DNA is used as the reference species (solid boxes depict coding sequence, open boxes depict introns, and thick lines depict intergenic spacers). Alignments derived from *Cuscuta* are shaded. Stop codons (thin-lined boxes) and start codons (thick-lined boxes) are indicated. Asterisks and dashes in between arrows represent discontinuations in the *trnL* intron alignment. Scale bars correspond to ~100 bp.

used originally for the phylogenetic reconstructions. The relevant portions of the alignments for four such plastid regions (*atpB* gene, psbE-J operon, *trnL-F* intron/spacer, and *rpl2* gene) are summarized in Fig. 3 and are accompanied by the current phylogenetic hypothesis for Convolvulaceae.

The *atpB* gene was found to be missing at least two codons near the 5′ end (position 163–168 in tobacco *atpB*) for all members of Convolvulaceae, including *Cuscuta* species (Fig. 3) The only exception to this is *Humbertia madagascariensis*, which has a full-length gene, as do all other angiosperms known (Savolainen et al. 2000). In the same region, some nonparasitic species, belonging mainly to the Dicranostyloideae clade (Fig. 3), are lacking three additional amino acids (AA). Given their position in the alignment, it seems that at least four additional and mutually independent deletions occurred (in Dichondreae, Maripeae, one *Jacquemontia* species within Dicranostyloideae, and in *Cardiochlamis* within the Cardiochlamyeae clade; Fig. 3) in what appears to be a hotspot for deletion mutations in the *atpB* gene in Convolvulaceae. *Cuscuta* species from all three subgenera are missing only two AA in this region as is the case with some 80 additional nonparasitic Convolvulaceae for which *atpB* sequences are known (Stefanović et al. 2002).

The absence of a 169-bp portion of the *trnL* intron (as compared to tobacco) in the Convolvulaceae was first reported by Stefanović et al. (2002). This large deletion has a distribution pattern similar to the one described for the 5′ *atpB* deletion. It is found in all Convolvulaceae, including *Cuscuta* species, with the exception of *Humbertia*. *Cuscuta japonia* (subgenus *Monogyna*) and *C. europaea* (subgenus *Cuscuta*), two *Cuscuta* species showing the least amount of nucleotide divergence compared to the green members of the family, were fully alignable throughout the *trn*L-F region. Also, the *trnL* intron deletion boundaries in these two subgenera followed closely those of other Convolvulaceae (Fig. 3). This was not the case with any member of the more highly divergent *Cuscuta* subgenus *Grammica*. In this subgenus, additional deletion(s) occurred resulting in an even larger gap (Fig. 3). The only other group that had a substantial departure with respect to the intron gap boundaries is *Jacquemontia*, the genus that shows the greatest sequence divergence of any nonparasitic Convolvulaceae (Stefanović et al. 2002). The *trnL-F* spacer region is evolving more rapidly than the *trnL* intron, in both point mutations and indels, as noted in previous studies dealing with this type of data (e.g., Gielly and Taberlet 1994; McDade and Moody 1999). This region was almost entirely missing in *Cuscuta* subgenus *Grammica* (data not shown; see Stefanović et al. 2002). In addition, the small remaining part of the spacer showed great nucleotide divergence and

could not be unambiguously aligned with the members of the other *Cuscuta* subgenera, nor with green Convolvulaceae species.

Most of the length variation of the *psbE-J* operon is found in its three IGS regions. No size variation was found in the *psbF* and *psbJ* genes. However, the *psbE* and *psbL* genes exhibited some length variation, while always maintaining the open reading frame. The *psbE* gene has three independent occurrences of gaps with respect to tobacco, two of them found in green Convolvulaceae (*Humbertia* and *Jacquemontia*; not shown) and one in *Cuscuta* subgenus *Grammica* (Fig. 3). Due to this deletion in *Cuscuta* subgenus *Grammica*, the coding sequence for the *psbE* gene extends into the *psbE-F* IGS region and its stop codon overlaps with the start codon for *psbF*. The *psbE* gene of *C. europaea* (subgenus *Cuscuta*) is truncated due to an early stop codon (not shown), while this gene in *C. japonica* (subgenus *Monogyna*) is identical in length to that of nonparasitic relatives. *Jacquemontia* is the only genus of autotrophic Convolvulaceae that shows length variation due to an A to C substitution in the stop codon (Fig. 3), which is followed immediately by another stop codon found in all Convolvulaceae *psbE* sequences, in this otherwise very conserved region across angiosperms (Graham and Olmstead 2000). In most flowering plants (with the exception of monocots) for which the sequences are known, the start codon of *psbL* has an edit site (Kudla et al. 1992; Bock et al. 1993; Graham and Olmstead 2000); editing of C to U is necessary to produce a functional translation initiation codon for this gene in these taxa. This putative edit site is found also in most of the Convolvulaceae, except in a subset of taxa in which the editing is not necessary, and in the outgroup, *Montinia*. This subset corresponds to the "bifid style" clade (Dicranostyloideae; Fig. 3), first explicitly identified by Stefanović et al. (2002). All examined *Cuscuta* species have this inferred edited start codon except for *C. europaea*, which has the nonedited codon (Fig. 3).

As already mentioned, an intron usually found in the *rpl2* gene of angiosperms is deleted in all Convolvulaceae, including *Humbertia* and *Cuscuta* (Fig. 3), representing a unique event within Asteridae, and a synapomorphy for Convolvulaceae (Stefanović et al. 2002). The *rpl2* alignment within Convolvulaceae required only one gap, an in-frame insertion of two codons in *Cuscuta japonica* (subgenus *Monogyna*; Fig. 3). A comparison of the Convolvulaceae *rpl2* sequences homologous to those immediately flanking the 5′ intron terminus in the outgroups (Solanaceae, Montiniaceae) reveals that this 6-bp region, the intron-binding sequence 1 (IBSI; Michel et al. 1989), which is essential for intron splicing, is conserved among all taxa examined,

including some *Cuscuta* species (Fig. 3). Alignment of the uninterrupted *rpl2* genes from Convolvulaceae with the exons of outgroup taxa demonstrates that these sequences are maintained even after deletion of the intron and are similar to the sequences of other group II introns (Shinozaki et al. 1986b; Michel et al. 1989).

Interpretation of the Observed ptDNA Variation in a Phylogenetic Context

Figure 4 summarizes all of the available data on plastid genome variation in *Cuscuta* as well as in the rest of the Convolvulaceae, mapped on a simplified family phytogeny. In the context of a rooted phylogenetic hypothesis for the Convolvulaceae, all of the ptDNA structural mutations discussed in this report can be considered either plesiomorphic, autapomorphic, or synapomorphic relative to *Cuscuta*. The first group consists of those changes that are found in all or almost all members of Convolvulaceae, regardless of whether they are parasitic or not (Fig. 4). For example, the loss of the *rpl2* intron, unique among the Asteridae, is confirmed in all Convolvulaceae, indicating that this deletion probably predated the diversification of the family and provides additional evidence for its monophyly, including *Humbertia*. Also, the absence of the deletions in *Humbertia* for both the *atpB* gene (at the 5′ and 3′ ends) and the *trnL* intron underlines the isolated position of this genus as a sister to the rest of the family. Given the available Southern hybridization data (Downie and Palmer 1992), the absence (or major alternation) of the *ycf1* gene cannot be unambiguously mapped, but it is most likely to be missing in *Cuscuta* species as well as in some other Convolvulaceae. These findings indicate that the plastid genome of the nonparasitic taxa in this family has undergone a number of structural changes similar to those observed in parasitic plants. Hence, not all of the changes in *Cuscuta* are correlated with its parasitic habit, and a significant proportion of those, at least three given the presently available data, can be better explained as plesiomorphic characters of the family.

The autapomorphic group consists of events that have occurred in some but not all *Cuscuta* species/subgenera. For example, the large 15-kb inversion in the LSC is restricted to *Cuscuta* subgenus *Monogyna*. None of the other Convolvulaceae investigated, including members of the two other *Cuscuta* subgenera, have this inversion. However, inversions sometimes occur in photosynthetic plants as well (e.g., large inversions in the LSC of *Oryza sativa* and *Pinus thunbergiana*; Hiratsuka et al. 1989; Wakasugi et al. 1994), and there is no reason to suspect a priori an association of this type of structural rearrangement with parasitism. The *trnV*-UAC gene and its
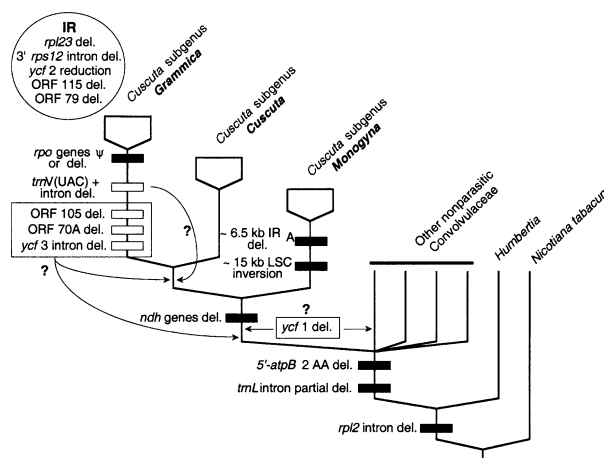

Fig. 4. Plastid genome structural rearrangements observed in three *Cuscuta* subgenera and the rest of Convolvulaceae mapped on a simplified family phylogeny. Solid boxes depict rearrangements that can be unambiguously assigned to different nodes. Open boxes indicate rearrangements inferred from *C. sandwichiana* that could be potentially shared with other taxa, but cannot be mapped unambiguously at present due to the lack of comparative data (see text). Arrows indicate alternative positions for this type of rearrangement. A number of deletions found in the inverted repeat (IR) of *C. sandwichiana*, which could potentially be explained by the IR contraction, are encircled. Indels that have occurred in some *Cuscuta* species and in some, but not all, nonparasitic Convolvulaceae (examples can be found in Fig. 3) are not depicted on the tree.

intron, usually found in the LSC of angiosperms, are absent in at least one species of *Cuscuta* subgenus *Grammica* (*C. sandwichiana*). On the other hand, this gene is present in *Cuscuta* subgenus *Monogyna*, albeit inverted with respect to its orientation in tobacco (Fig. 1B). The data concerning this tRNA gene are not available for the third subgenus, *Cuscuta*. However, given the phylogenetic relationships within *Cuscuta*, and the results for *Cuscuta* subgenus *Monogyna*, it is most likely that this deletion represents either a taxon-specific event restricted to subgenus *Grammica* or a feature shared between subgenera *Grammica* and *Cuscuta*. In either case, this tRNA deletion, like the 15-kb inversion in *Monogyna*, represents an apomorphy for some lineages within the parasitic clade, and cannot be explained entirely by parasitism. Some autapomorphic changes, however, may be related to parasitism. For the most part genomic changes associated with the transition to parasitic habit are expected to be progressive, i.e., found in some clades and not in others. Such an evolutionary transition series cannot be rejected in *Cuscuta*. Based on ptDNA analyses (Stefanović et al. 2002), the hypothesis of progressive transition to parasitism in this genus is supported by the resulting phylogenetic relationships as well as progressive increase in sequence divergence, as deduced from the branch lengths (Stefanović et al. 2002). The hemiparasitic subgenus *Monogyna*, showing the least amount of sequence divergence compared to the green members of the family, is found to be the sister

to holoparasitic subgenera *Cuscuta* and *Grammica*. The branch leading to subgenus *Cuscuta* is longer, compared to that of subgenus *Monogyna*, and is followed by even more highly divergent sequences in subgenus *Grammica* (Stefanović et al. 2002). This stands in contrast with the case of parasitic plants in the Orobanchaceae, where several independent losses of photosynthetic ability were inferred from the phylogenetic relationships among the holo- and hemiparasites, and consequently the hypothesis of evolutionary transition series was rejected (Young et al. 1999; Olmstead et al. 2001).

The third group concerns changes that are not found in any of the nonparasitic Convolvulaceae but are found in all *Cuscuta* species investigated and represent synapomorphies for this genus. Those are the changes that help define the genus, and could, in some case, be directly implicated in ability to become a parasitic plant. Presently, only the loss or significant modification of all *ndh* genes is identified as a ptDNA synapomorphy for all *Cuscuta* members. A number of additional structural changes, mainly deletions and size reductions, are identified in *C. sandwichiana* (Figs. 1, 3). However, due to the lack of comparative data from other *Cuscuta* subgenera, and representative samples of the green family members, it is not possible to infer at which point in the evolution of Convolvulaceae these changes occurred (Fig. 4). This is also complicated for the changes observed in the IR region. To distinguish the loss of genes from simple IR contraction, which is frequent in angiosperms (e.g., Goulding et al. 1996; Plunkett and Downie, 2000, and references therein), comparative data are needed to locate the junctions between both IRs and the LSC.

Molecular Implications of the Observed ptDNA Variation

In addition to conclusions regarding the pattern of plastid genome evolution in Convolvulaceae that can be drawn from a phylogenetic perspective, these inferences on structural rearrangements in Convolvulaceae in general, and *Cuscuta* in particular, have some molecular implications. Regardless of how the changes observed in *Cuscuta sandwichiana* will be distributed with respect to the two other *Cuscuta* subgenera and autotrophic taxa, it is already clear that most of the affected regions are those with no function or with unknown function.

Changes that probably have little or no functional consequences include a number of deleted introns, such as those usually found in *rpl2*, *ycf3*, and 3′-*rps12* genes. Introns are highly stable components of land plant plastid genomes, with no cases of intron gain known during land plant evolution. The lack of these regions, however, does not necessarily affect the functioning of plastids, as evidenced by examples found throughout angiosperms where different introns are lacking in many successfully photosynthetic organisms, e.g., the *rpoC1* intron and two *clpP* introns in *Oryza sativa* (Poaceae; Hiratsuka et al. 1989), the *rpl2* intron in all Caryophyllales (Zurawski et al. 1984; Downie et al. 1991), Menyanthaceae, Saxifragaceae, and Convolvulaceae (Downie et al. 1991; this study), the *rpl16* intron in Geraniaceae, and *trnI* intron in Campanulaceae (Downie et al. 1991).

Changes with unknown functional consequences include a variety of lost *ycfs* and ORFs. The *ycfs* abbreviation is reserved usually for hypothetical plastid open reading frames that have no known function but whose homologs are found throughout land plants. The open reading frames, also without known function, but found only in tobacco ptDNA (i.e., there are no homologs in the other known plastid genomes) are referred to with the ORF abbreviation. The *ycf1* is reported missing or significantly altered in *Cuscuta* and in some photosynthetic Convolvulaceae (Downie and Palmer 1992; Fig. 4). The homologs of this hypothetical protein also are lacking from Poaceae (Hiratsuka et al. 1989; Downie and Palmer 1992) and Campanulaceae (Downie and Palmer 1992) without affecting their photosynthetic ability. The largest plastid gene, *ycf2*, may or may not be significantly reduced in *Cuscuta*, depending on whether the documented reduction affected only one of the IRs in this genus or both (see above).

Major length mutations *ycf2* were detected in several photosynthetic representatives of Passifloraceae, Geraniaceae, Campanulaceae, and Poaceae (Downie et al. 1994), while at the same time the homolog of this gene is present in the otherwise significantly reduced plastid genome of holoparasitic *Epifagus virginiana* (dePamphilis and Palmer 1990). When combined, these results imply that *ycf2* is not involved in photosynthetic metabolism as already suggested by Wolfe et al. (1992). Its function remains unknown, even though a possible proteolytic ATPase activity (Wolfe 1994) and/or a chromoplast-specific function (Richards et al. 1991) have been proposed. Finally, because their homologs are not found in other land plants, photosynthetic or not, the suite of tobacco-specific ORFs (ORF105, ORF70A, ORF115, and ORF79) that are absent in *C. sandwichiana* are most likely not connected with parasitism.

A third category consists of lost genes, the function of which is either known or suggested. For example, the lack of *trnI*-CAU and *rpl23* in the IR of some *Cuscuta* species and of *trnV*-UAC in the LSC of *C. sandwichiana* (and possibly further in some other *Cuscuta* species, but definitely present in the subgenus *Monogyna*; Fig. 1B) might indicate that the translational apparatus is altered in some *Cuscuta* plastids.

The IR deletions could be explained by the IR contraction and retention of only one copy of those genes, which would be sufficient to maintain the function of the plastid translational machinery. This IR-contraction hypothesis is supported by the demonstrated presence of the *rpl2* gene in different *Cuscuta* plastid genomes, previously thought to be completely absent (Bömmer et al. 1993; see above). The lack of *trnV*-UAC, even if confirmed for the rest of the plastid genome in *C. sandwichiana*, does not necessarily affect the entire translational mechanism, because it could be compensated by the other valine tRNA (*trnV*-GAC) located in the IR. Deletion of *ndh* genes seems to be correlated with evolution of parasitism in Convolvulaceae. Due to the significant amino-acid similarity with mitochondrial genes encoding mitochondrial respiratory chain NADH dehydrogenase, it has been suggested that these genes may be involved in respiratory processes in the plastid (i.e., chlororespiration; Sugiura 1992; Peltier and Cournac 2002). The best evidence for the existence a of chlororespiratory chain comes from investigation of the unicellular green alga *Chlamydomonas reinhardtii* (Bennuon 1982; Peltier and Schmidt 1991). Additionally, these genes are expressed (Matsubayashi et al. 1987) and conserved across angiosperms (Olmstead and Palmer 1994; Olmstead et al. 2000) supporting the idea that such a plastid respiratory chain also exists in land plants. Because all of the *ndh* genes are lacking (or altered) in concert with all photosynthetic genes in holoparasitic *Epifagus virginiana*, it has been hypothesized also that these genes are involved in a metabolism closely connected with photosynthesis in land plants (dePamphilis and Palmer 1990). As pointed out by Haberhausen and Zetsche (1994), this is probably not the case in *Cuscuta*, which retains all of its photosynthetic genes in unaltered or nearly unaltered form, indicating the lack of a direct connection between photosynthesis and any putative chlororespiration mechanism. This lack of connection is further supported by the loss or significant alteration of all 11 *ndh* genes encountered in the black pine (*Pinus thunbergii*; Wakasugi et al. 1994).

In sum, our data indicate that a major trend in *Cuscuta* plastid evolution, and to a certain extent in other Convolvulaceae as well, is overall genome reduction. This decrease in ptDNA size is achieved through gene/ORF loss, intron loss, and strong reduction in IGS lengths. According to the presently available data, no plastid genes of a known function (with the exception of chlororespiratory genes) are missing in the parasitic genus *Cuscuta* that could not be potentially replaced by duplicate genes (e.g., in the IR) or genes with similar function (tRNA genes specifying the same AA). Many of the changes in *Cuscuta*, previously attributed to its parasitic mode of life, could be better explained either as retention of ancestral conditions within the family (i.e., plesiomorphies) or, in most cases, autapomorphies of some *Cuscuta* species not shared with the rest of the genus. The autapomorphic changes, given their exact phylogenetic distribution and extent, which is yet to be established, could lend further support for the hypothesis of progressive transition to parasitism in *Cuscuta*, as already has been indicated by the ptDNA sequence analyses. However, the third, synapomorphic, group is most likely to be explained by the parasitic lifestyle alone, because it represents changes found in *Cuscuta* exclusively.

These mostly unaltered plastid genomes of *Cuscuta* species stand in sharp contrast with the enormous morphological and physiological modifications that the ancestors of this parasitic genus must have undergone. The unexpectedly conservative nature of the *Cuscuta* plastids, especially when compared to those of its closest nonparasitic relatives, could be accounted for either by relatively recent origin of this lineage, or, more likely, by relaxed but still present natural selection for photosynthetic capacity, presumably because of the need for such an ability in some stages of *Cuscuta*'s life cycle, for example during the post-germination/pre-host-attachment period or during endosperm formation.

References

Adams KL, Song K, Roessler PG, Nugent JM, Doyle JL, Doyle JJ, Palmer JD (1999) Intracellular gene transfer in action: dual transcription and multiple silencing of nuclear and mitochondrial cox2 genes in legumes. Proc Natl Acad Sci USA 96:13863–13868

Bennoun P (1982) Evidence for a respiratory chain in the chloroplast. Proc Natl Acad Sci USA 79:4352–4356

Bock R, Hagemann R, Kössel H, Kudla J (1993) Tissue- and stage-specific modulation of RNA editing of *psbF* and *psbL* transcript from spinach plastid: a new regulatory mechanism? Mol Gen Genet 240:238–244

Bömmer D, Haberhausen G, Zetsche K (1993) A large deletion in the plastid DNA of the holoparasitic flowering plant *Cuscuta reflexa* concerning two ribosomal proteins (*rpl2*, *rpl23*), one transfer RNA (*trnI*) and an ORF 2280 homologue. Curr Genet 24:171–176

Bungard RA (2004) Photosynthetic evolution in parasitic plants: insight from the chloroplast genome. Bioessays 26:235–247

Casano LM, Zapata JM, Martin M, Sabater B (2000) Chlororespiration and poising of cyclic electron transport: plastoquinone

as electron transporter between tylakoid NADH dehydrogenase and peroxidase. J Biol Chem 275:942–948

Cronquist A (1988) The evolution and classification of flower plants. The New York Botanical Garden, Bronx, NY

Dawson JH, Musselman LJ, Wolswinkel P, Dörr I (1994) Biology and control of *Cuscuta*. Rev Weed Sci 6:265–317

dePamphilis CW, Palmer JD (1990) Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. Nature 348:337–339

Downie SR, Palmer JD (1992) Use of chloroplast DNA rearrangements in reconstructing plant phytogeny. In: Soltis PS, Soltis DE, Doyle JJ (eds) Molecular systematics of plants. Chapman and Hall, London, UK, pp 14–35

Downie SR, Olmstead RG, Zurawski G, Soltis DE, Soltis PS, Watson JC, Palmer JD (1991) Six independent losses of the chloroplast DNA *rpl2* intron in Dicotyledons: molecular and phylogenetic implications. Evolution 45:1245–1259

Downie SR, Katz-Downie DS, Wolfe KH, Calie PJ, Palmer JD (1994) Structure and evolution of the largest chloroplast genome (ORF2280): internal plasticity and multiple gene loss during angiosperm evolution. Curr Genet 25:367–378

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

Doyle JJ, Doyle JL, Palmer JD (1995) Multiple independent losses of two genes and one intron from legume chloroplast genomes. Syst Bot 20:272–294

Dujon B (1989) Group I introns as mobile genetic elements: facts and mechanistic speculations: a review. Gene 82:91–114

Engelmann G (1859) Systematic arrangement of the species of the genus *Cuscuta* with critical remarks on old species and descriptions of new ones. Trans Acad Sci St Louis 1:453–523

Fink GR (1987) Pseudogenes in yeast? Cell 49:5–6

Freyer R, Neckermann K, Maier RM, Kössel H (1995) Structural and functional analysis of plastid genomes from parasitic plants: loss of an intron within the genus *Cuscuta*. Curr Genet 27:580–586

Gielly L, Taberlet P (1994) The use of chloroplast DNA to resolve plant phytogenies: noncoding versus *rbcL* sequences. Mol Biol Evol 11:769–777

Goremykin VV, Hirsch-Ernst KI, Wölf S, Hellwig FH (2003) Analysis of *Amborella trichopoda* chloroplast genome sequence suggests that *Amborella* is not a basal angiosperm. Mol Biol Evol 20:14499–14505

Goremykin VV, Hirsch-Ernst KI, Wölf S, Hellwig FH (2004) The chloroplast genome of *Nymphaea alba*: whole-genome analyses and the problem of identifying the most basal angiosperm. Mol Biol Evol 21:1445–1454

Goulding SE, Olmstead RG, Morden CW, Wolfe KH (1996) Ebb and flow of the chloroplast inverted repeat. Mol Gen Genet 252:195–206

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

Guedeney G, Corneille S, Cuine S, Peltier G (1996) Evidence for an association of *ndhB, ndhJ* gene products and feredoxin-NADP-reductase as components of a chloroplastic NAD(P)H dehydrogenase complex. FEBS Lett 378:277–280

Haberhausen G, Zetsche K (1994) Functional loss of all *ndh* genes in an otherwise relatively unaltered plastid genome of the holoparasitic flowering plant *Cuscuta reflexa*. Plant Mol Biol 24:217–222

Haberhausen G, Valentin K, Zetsche K (1992) Organization and sequence of photosynthetic genes from the plastid genome of the holoparasitic flowering plant *Cuscuta reflexa*. Mol Gen Genet 232:154–161

Hibberd JM, Bungard RA, Press MC, Jeschke WD, Scholes JD, Quick WP (1998) Localization of photosynthetic metabolism in the parasitic angiosperm *Cuscuta reflexa*. Planta 205:506–513

Hiratsuka J, Shimada H, Whitter R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun CR, Meng BY, Li YQ, Kanno A, Nishizawa Y, Hirai A, Shinozaki K, Sugiura M (1989) The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes account for a major plastid DNA inversion during the evolution of the cereals. Mol Gen Genet 217:185–194

Hupfer H, Swaitek M, Hornung S, Hermann RG, Maier RM, Chiu WL, Sears B (2000) Complete nucleotide sequence of the *Oenothera elata* plastid chromosome, representing plastome I of the five distinguishable *Euoenothera* plastomes. Mol Gen Genet 263:581–585

Kato T, Kaneko T, Sato S, Nakamura Y, Tabata S (2000) Complete structure of the chloroplast genome of a legume, *Lotus japonicus*. DNA Res 7:323–330

Krause K, Berg S, Krupinska K (2003) Plastid transcription in the holoparasitic plant genus *Cuscuta*: parallel loss of the *rrn16* PEP-promoter and of the *rpoA* and *rpoB* genes coding for the plastid-encoded RNA polymerase. Planta 216:815–823

Kudla JG, Igloi GL, Metzlaf M, Hagemann R, Kössel H (1992) RNA editing in tobacco chloroplasts leads to the formation of a translatable *psbL* mRNA by C to U substitution within the initiation codon. EMBO J 11:1099–1103

Kuijt J (1969) The biology of parasitic flowering plants. University of California Press, Berkeley, CA

Machado MA, Zetsche K (1990) A structural, functional, and molecular analysis of plastids of the holoparasites *Cuscuta reflexa* and *Cuscuta europaea*. Planta 181:91–96

Maier RM, Neckermann K, Igloi GL, Kössel H (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence, and fine tuning of genetic information by transcript editing. J Mol Biol 251:614–628

Matsubayashi T, Wakasugi T, Shinozaki K, Yamagushi-Shinozaki K, Zaita N, Hidaka T, Meng BY, Ohto C, Tanaka M, Kato A, Maruyama T, Sugiura M (1987) Six chloroplast genes (*ndhA-F*) homologous to human mitochondrial genes encoding components of the respiratory chain NADH dehydrogenase are actively expressed: determination of splice sites in *ndhA* and *ndhB* pre-mRNAs. Mol Gen Genet 210:385–393

McDade LA, Moody ML (1999) Phylogenetic relationships among Acanthaceae: evidence from noncoding *trnL-trnF* chloroplast DNA sequences. Am J Bot 86:70–80

Michel F, Umesono K, Ozeki H (1989) Comparative and functional anatomy of group II catalytic introns: a review. Gene 82:5–30

Nixon PJ (2000) Chlororespiration. Philos Trans R Soc Lond Ser B 355:1541–1547

Ogihara Y, Isono K, Kojima T, Endo A, Hanaoka M, Shiina T, Terachi T, Utsugi S, Murata M, Mori N, Takumi S, Ikeo K, Gojobori T, Murai R, Murai K, Matsuoka Y, Ohnishi Y, Tajiri H, Tsunewaki K (2002) Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA. Mol Genet Genom 266:740–746

Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota S, Inokuchi H, Ozekj H (1986) Chloroplast gene organization deduced from complete sequence of liverworth *Marchantia polymorpha* chloroplast DNA. Nature 322:572–574

Olmstead RG, Palmer JD (1992) A chloroplast DNA phylogeny of the Solanaceae: subfamilial relationships and character evolution. Ann Mo Bot Gard 79:346–360

Olmstead RG, Palmer JD (1994) Chloroplast DNA systematics: a review of methods and data analysis. Am J Bot 81:1205–1224

Olmstead RG, Kim KJ, Jansen RK, Wagstaff SJ (2000) The phylogeny of the Asteridae sensu lato based on chloroplast *ndhF* gene sequences. Mol Phylogenet Evol 16:96–112

Olmstead RG, dePamphilis CW, Wolfe AD, Young ND, Elisons WJ, Reeves PA (2001) Disintegration of the Scrophulariaceae. Am J Bot 88:348–361

Palmer JD (1985) Comparative organization of chloroplast genomes. Ann Rev Genet 19:325–354

Palmer JD (1990) Contrasting modes and tempos of genome evolution in land plants organelles. Trends Genet 6:115–120

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

Palmer JD, Stein DB (1986) Conservation of chloroplast genome structure among vascular plants. Curr Genet 10:823–833

Panda MM, Choudhury NK (1992) Effect of irradiance and nutrients on chlorophyll and carotenoid content and Hill reaction activity of Cuscuta reflexa. Photosynthetica 26:585–592

Peltier G, Cournac L (2002) Chlororespiration. Annu Rev Plant Biol 53:523–550

Peltier G, Schmidt GW (1991) Chlororespiration: an adaptation to nitrogen deficiency in Chlamydomonas reinhardtii. Proc Natl Acad Sci USA 88:4791–4795

Plunkett GM, Downie SR (2000) Expansion and contraction of the chloroplast inverted repeat in Apiaceae subfamily Apioideae. Syst Bot 25:648–667

Posno M, van Vliet A, Groot GSP (1986) Localization of chloroplast ribosomal protein genes on Spirodela oligorhiza chloroplast DNA. Curr Genet 10:923–930

Richards CM, Hinman SB, Boyer CD, Hardison RC (1991) Survey of plastid RNA abundance during tomatofruit ripening: the amounts of rna from the ORF2280 region increase in chromoplasts. Plant Mol Biol 17:1179–1188

Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S (1999) Complete structure of the chloroplast genome of Arabidopsis thaliana. DNA Res 6:283–290

Savolainen V, Chase MW, Hoot SB, Morton CM, Soltis DE, Bayer C, Fay MF, deBruijn AJ, Sullivan S, Qiu YL (2000) Phylogenetics of flowering plants based upon a combined analysis of plastid atpB and rbcL gene sequences. Syst Biol 49:306–362

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986a) The complete nucleotide sequence of tobacco chloroplast genome: its gene organization and expression. EMBO J 5:2043–2049

Shinozaki K, Deno H, Sugita M, Kuramitsu S, Sugiura M (1986b) Intron in the gene for the ribosomal protein S16 of tobacco chloroplast and its conserved boundary sequences. Mol Gen Genet 202:1–5

Schmitz-Linneweber C, Maier RM, Alcaraz JP, Cottet A, Herrmann RG, Mache R (2001) The plastide chromosome of spinach (Spinacia oleracea): complete nucleotide sequence and gene organization. Plant Mol Biol 45:307–315

Spielman A, Roux E, von Allmen JM, Stutz E (1988) The soybean chloroplast genome: complete sequence of the rps19 gene, including flanking parts containing exon 2 of rpl2 (upstream), but lacking rpl22 (downstream). Nucleic Acids Res 16:1199

Stefanović S, Olmstead RG (2004) Testing the phylogenetic position of a parasitic plant (Cuscuta, Convolvulaceae, Asteridae): Bayesian inference and the parametric bootstrap on data drawn from three genomes. Syst Biol 53:384–399

Stefanović S, Krueger L, Olmstead RG (2002) Monophyly of the Convolvulaceae and circumscription of their major lineages based on DNA sequences of multiple chloroplast loci. Am J Bot 89:1510–1522

Stefanović S, Austin DF, Olmstead RG (2003) Classification of Convolvulaceae: a phylogenetic approach. Syst Bot 28:791–806

Stefanović S, Rice DW, Palmer JD (2004) Long branch attraction, taxon sampling, and the earliest angiosperms: Amborella or monocots? BMC Evol Biol 4:35

Sugiura M (1992) The chloroplast genome. Plant Mol Biol 19:149–168

Takhtajan A (1997) Diversity and classification of flowering plants. Columbia University Press, New York, New York, USA

van der Kooij TAW, Krause K, Dörr I, Krupinska K (2000) Molecular, functional and ultrastructural characterization of plastids from six species of the parasitic flowering plant genus Cuscuta. Planta 210:701–707

Wakasugi T, Tsudzuki T, Ito S, Nakashima K, Tsudzuk T, Sugiura M (1994) Loss of all ndh genes as determined by sequencing the entire chloroplast genome of the black pine Pinns thunbergii. Proc Natl Acad Sci USA 91:9794–9798

Wakasugi T, Sugita M, Tsudzuki T, Sugiura M (1998) Updated gene map of tobacco chloroplast DNA. Plant Mol Biol Reporter 16:231–241

Wolfe AD, Liston A (1998) Contribution of PCR-based methods to plant systematics and evolutionary biology. In: Soltis PS, Soltis DE, Doyle JJ (eds) Molecular systematics of plants, II, DNA sequencing. Kluwer Academic Publishers, Boston, MA, USA, pp 43–86

Wolfe KH (1994) Similarity of plastid ORF2280 to the CDC48 family: a proteolytic ATPase? Curr Genet 25:379–383

Wolfe KH, Morden CW, Palmer JD (1992) Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. Proc Natl Acad Sci USA 89:10648–10652

Young ND, Steiner KE, dePamphilis CW (1999) The evolution of parasitism in Scrophulariaceae/Orobanchaceae: plastid gene sequences refute an evolutionary transition series. Ann Mo Bot Gard 86:876–893

Yuncker TG (1932) The genus Cuscuta. Mem Torrey Bot Club 18:113–331

Zurawski G, Bottomley W, Whitfeld PR (1984) Junction of the large single copy region and the inverted repeat in Spinacea oleracea and Nicotiana debneyi chloroplst DNA: sequence of the genes for tRNAHis and the ribosomal proteins S19 and L2. Nucleic Acids Res 12:6547–6558