Curr Genet (1997) 31: 419–429

ORIGINAL PAPER

Mary E. Cosner · Robert K. Jansen · Jeffrey D. Palmer
Stephen R. Downie

The highly rearranged chloroplast genome of *Trachelium caeruleum* (Campanulaceae): multiple inversions, inverted repeat expansion and contraction, transposition, insertions/deletions, and several repeat families

Received: 10 July 1996 / 28 January 1997

Abstract Comprehensive gene mapping reveals that the chloroplast genome of *Trachelium caeruleum* is highly rearranged relative to those of other land plants. Evolutionary scenarios that consist of seven to ten inversions, one or two transpositions, both expansion and contraction of the typically size-conserved inverted repeat, a presumed gene loss, deletions within two large open reading frames and several insertions, are sufficient to derive the *Trachelium* arrangement from the ancestral angiosperm chloroplast DNA arrangement. Two of the rearrangements disrupt transcriptional units that are otherwise conserved among land plants. At least five families of small dispersed repeats exist in the *Trachelium* chloroplast genome. Most of the repeats are associated with inversion endpoints and may have facilitated inversions through recombination across homologous repeats.

Key words Chloroplast DNA rearrangements · *Trachelium* · Inversions · Transposition · Gene loss

Introduction

The chloroplast genome of photosynthetic land plants consists of a circular chromosome ranging in size from about 120 to 217 kilobase pairs (kb). Chloroplast DNA (cpDNA) consists of a large inverted repeat (IR) separated by a large single-copy (LSC) region and a small single-copy (SSC) region. Exceptions include conifers (Lidholm et al. 1988; Strauss et al. 1988; Raubeson and Jansen 1992) and certain taxa in a few angiosperm families, the most notable of which are six tribes of the Fabaceae (legumes) (Palmer and Thompson 1981; Lavin et al. 1990; Palmer 1991), which have lost one segment of the IR. The IR is generally 20–30 kb in length, but ranges in size from about 10 to 76 kb (Palmer 1991). Expansion or contraction of the IR occurs infrequently; the most drastic case of expansion is found in *Pelargonium* (geranium), in which the IR has grown to 76 kb (Palmer et al. 1987). Gene content, gene order, and nucleotide sequence are highly conserved in cpDNA (Palmer 1991), particularly in the IR (Wolfe et al. 1987; Downie and Palmer 1992 a). Few cases of chloroplast gene loss have been reported in land plants (Palmer 1991) and rearrangement of genes is relatively rare. Most cases in which regions of cpDNA have been inverted involve one or only a few inversions. For example, a 22-kb derived inversion separates most members of the angiosperm family Asteraceae from the subfamily Barnadesioideae and other angiosperms (typified by tobacco) (Jansen and Palmer 1987 a,b). Three inversions characterize all examined members of the grass family, Poaceae (Howe 1985; Quigley and Weil 1985; Hiratsuka et al. 1989; Doyle et al. 1992). More complicated genomes exhibiting several to many inversions include those of conifers (Lidholm et al. 1988; Strauss et al. 1988; Raubeson and Jansen 1992), some legumes (Palmer and Thompson 1981; Palmer et al. 1988; Milligan et al. 1989), *Pelargonium* (Palmer et al. 1987), *Lobelia* (Knox et al. 1993), and some members of the Ranunculaceae (Hoot and Palmer 1994).

The existence of multiple cpDNA rearrangements in two closely related angiosperm families, Campanulaceae and Lobeliaceae, was noted several years ago following a survey of restriction-site variation (Palmer 1985; Jansen and Palmer 1988; Downie and Palmer 1992 b). We have found that the chloroplast genome of *Trachelium caeruleum* (Campanulaceae) is highly rearranged relative to that of the majority of land plants. At least seven inversions,

M. E. Cosner[1] · R. K. Jansen
Department of Botany and Institute of Cell and Molecular Biology, University of Texas, Austin, TX 78713, USA

J. D. Palmer
Department of Biology, Indiana University, Bloomington, IN 47405, USA

S. R. Downie
Department of Plant Biology, University of Illinois, Urbana, IL 61801, USA

[1] Deceased

Communicated by B. B. Sears

IR contraction/expansion, two major presumed deletions, several insertions, and a putative transposition have transformed the chloroplast genome of *Trachelium* from a tobacco-like ancestor.

Materials and methods

Chloroplast-enriched DNA was isolated from *T. caeruleum* by the sucrose gradient method (Palmer 1986). The cpDNA was digested with the restriction endonuclease *Hin*dIII, resulting in 26 fragments larger than 0.5 kb, ranging in size from 1.1 to approximately 18 kb. The *Hin*dIII-digested cpDNA was shotgun-cloned into the plasmid vector pUC19. The ligation mixture was then used to transform *Escherichia coli* strain DH5alpha. Plasmid DNA was recovered using the alkaline-lysis method of Birnboim and Doly (1979), and the sizes of plasmid inserts were determined by *Hin*dIII restriction digestion followed by the visualization of fragments in agarose gels. Certain fragments that were not successfully shotgun-cloned were cut out of a 1.0% SeaPlaque (FMC) agarose gel and ligated into pUC19 in agarose. Total DNA from *T. caeruleum* was prepared according to the CTAB isolation method of Doyle and Doyle (1987). Agarose-gel electrophoresis, bidirectional transfer of DNA onto Zetabind (AMF CUNO) nylon membranes, labeling of recombinant plasmids by nick translation, filter hybridizations, and autoradiography were all carried out essentially as described by Palmer (1986). Restriction-enzyme digests of *Trachelium* total DNA were electrophoresed in 1.2% agarose gels. Hybridization probes consisted of cloned *Hin*dIII cpDNA fragments from *T. caeruleum*. Fragments that were not cloned were cut out of a 1.0% SeaPlaque (FMC) agarose gel and nick-translated directly in the agarose. In addition, 109 cloned tobacco cpDNA fragments were used as probes in hybridization experiments. These probes (Palmer et al. 1994) were subcloned from larger cloned cpDNA fragments (Sugiura et al. 1986; Olmstead and Palmer 1992) and average about 1.2 kb in size. The polymerase chain reaction (PCR) was employed to determine if introns are present in the gene *clpP* in *Trachelium*. One forward and one reverse primer each (for a total of eight 14- to 21-mer primers) for exon 3, intron 2, exon 2, and intron 1 were constructed using available sequences from tobacco (Shinozaki et al. 1986), *Epifagus virginiana* (Wolfe et al. 1992), or a consensus. Fifty microliter PCR reactions consisted of 0.2 mM dNTPs, 0.5 units of Tfl polymerase (Epicentre Technologies), 1×Tfl polymerase buffer (includes 1.5 mM of MgCl), approximately 1 μM of primers, and a 0.5-μl aliquot of unquantified DNA. The thermocycler was programmed for one initial denaturation cycle consisting of 3-min denaturation at 95°C, 1-min primer annealing at 55–60°C, and 1-min extension at 72°C, followed by 30 cycles of 1-min denaturation at 95°C, 1-min annealing at 55–60°C, and 1-min extension at 72°C. A final extension period of 5 min at 72°C terminated the PCR reactions.

Results and discussion

Clone bank of *T. caeruleum* cpDNA

The shotgun cloning strategy together with the cloning of gel-isolated fragments of approximately 18, 11.5, 9.6, and 7.4 kb resulted in a clone bank consisting of 21 of the 26 *Hin*dIII cpDNA fragments from *Trachelium*, representing 82% of the genome (see closed circles in Fig. 1). Uncloned fragments include those of 12.5, 8.3, 4.4, 3.8, and 1.5 kb in size.

Physical and gene mapping

Total DNA from *Trachelium* was digested with the restriction endonucleases *Bam*HI, *Bgl*II, *Eco*RI, *Eco*RV, *Hin*dIII, *Sst*I, and *Xba*I, and double digests were carried out using *Hin*dIII and the other six enzymes. The cloned *Hin*dIII fragments from *Trachelium* and gel isolations of uncloned fragments were used as homologous probes in order to construct restriction-site maps for the seven enzymes by the identification of overlapping fragments (Fig. 1). Summation of the restriction-fragment sizes of the seven enzymes indicates that *Trachelium* has a chloroplast genome of about 162 kb, consisting of an IR of about 27 kb, and large and small single-copy regions of approximately 100 and 8 kb, respectively.

Hybridizations using 109 small cloned tobacco cpDNA fragments (Palmer et al. 1994) were used to map the genome organization of *Trachelium* relative to tobacco and to determine the locations of chloroplast genes (Fig. 1). The differential hybridization of clones containing the 5′ and 3′ ends of certain chloroplast genes made it possible to determine the direction of transcription for these genes (Fig. 1). The direction of transcription of other genes was inferred either from their co-transcription in tobacco with genes for which 5′ and 3′ differential hybridization was possible, or by comparisons of their known locations in polycistronic transcription units in tobacco with their similar locations in *Trachelium* (Fig. 1). Because many of the tobacco cpDNA probes contain more than one gene, it was not always possible to determine which genes define inversion endpoints. A conservative approach was taken that kept operons known from other plants intact, but it should be noted that the locations of some genes were inferred.

Inverted repeat

Although the size of the IR in *Trachelium* is similar to that of most other angiosperms, its gene content is significantly different. Several genes that are single copy in tobacco are duplicated (i.e., are located in the IR) in *Trachelium* and vice versa (Fig. 1). The difference in extent of the *Trachelium* IR has resulted from spreading of the repeat into what is normally the SSC region and shrinkage at the LSC end. Genes between and including *ndhE* and ORF1901 in the IR of *Trachelium* are located in the SSC region in tobacco (Fig. 1; a small portion of ORF1901 is in the IR in tobacco, whereas the entire gene is duplicated in *Trachelium*), whereas genes between and including 3′-*rps12* and ORF2280 in the LSC of *Trachelium* are in tobacco's IR (Fig. 1). In addition, 5′-*rps12* and *clpP*, which reside in the IR of *Trachelium*, are in the LSC of tobacco and most other land plants. The rRNA operon, a conserved feature of the IR in all IR-containing plants, remains in the *Trachelium* IR despite the disruptive changes in IR structure. However, the rRNA operon is located close to the LSC region in *Trachelium*, whereas the operon is close to the SSC end of the IR in other angiosperms.

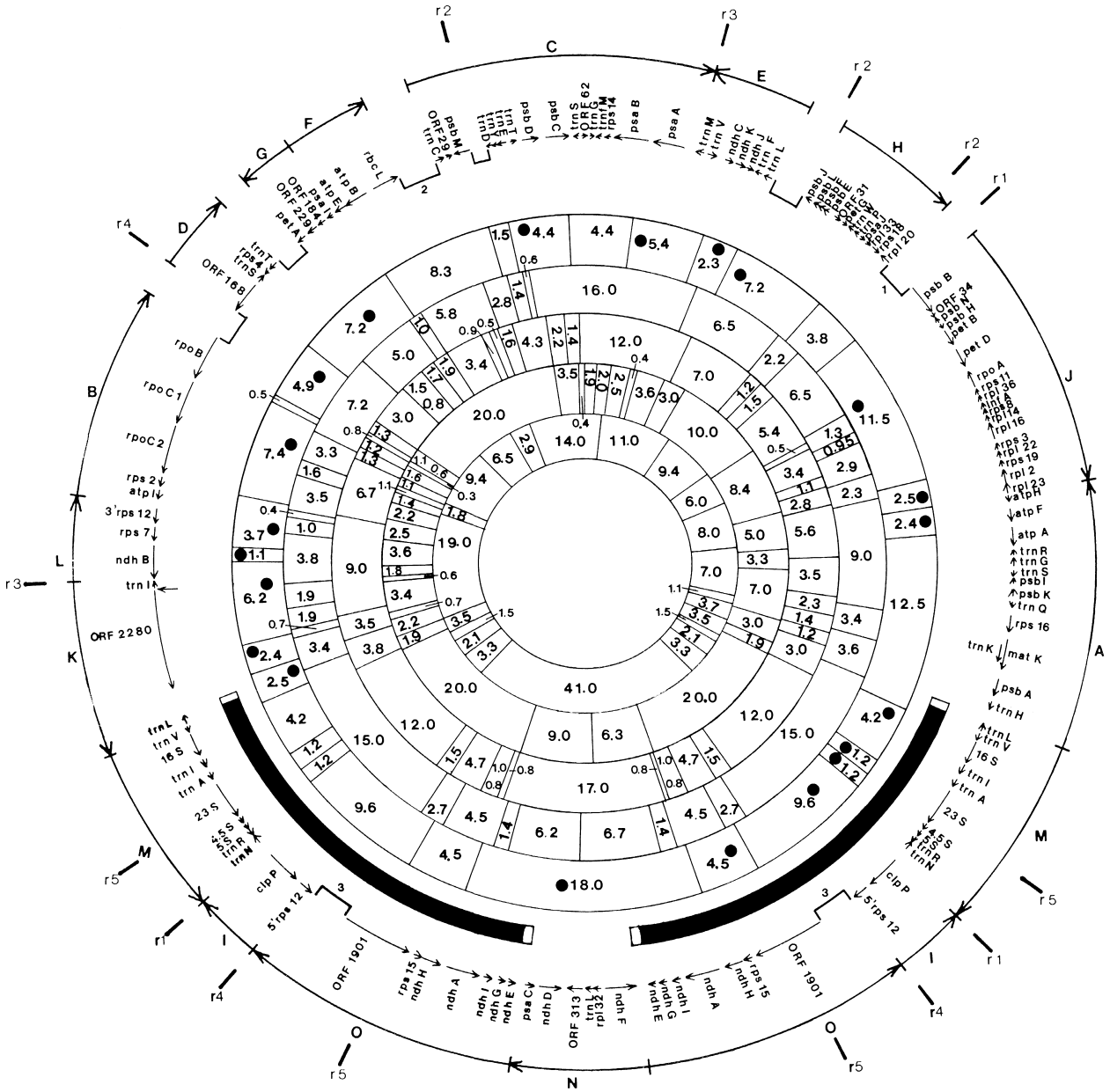

Fig. 1 Physical and gene map of *T. caeruleum* cpDNA. *Complete circles* from outside to inside show restriction sites and restriction fragment sizes for *Hin*dIII, *Eco*RV, *Xba*I, *Bam*HI, and *Sst*I in that order. Enzymes mapped but not pictured include *Bgl*II and *Eco*RI. The two *black bars* represent the minimum extent of the inverted repeat and the *open extensions* of these bars its maximum possible extent. Locations, orientations and approximate lengths of genes are shown by *arrows* labeled with gene names; sizes of ORFs are those of tobacco. *Outermost arrows* labeled A–O indicate blocks of genes and their orientations relative to tobacco. *Numbered lines* indicate the positions of elements belonging to five families of dispersed repeats (*r*). Repeat 5 is a small segment of the 23*s* rRNA gene located in both the *23s rDNA* and *ORF1901*. *Brackets* indicate regions that did not hybridize with tobacco cpDNA probes; *numbered brackets* designate regions that may contain pseudogenes of *clpP* and 5′-*rps12* (*1*), *accD* (*2*), and part of ORF1901 (*3*). The *small arrow* shows the location of a 213-bp tobacco cpDNA probe that does not hybridize to *Trachelium* DNA. *Closed circles* indicate cloned *Hin*dIII fragments

Putative deletions

Most of the 109 small tobacco cpDNA probes hybridized to *Trachelium* cpDNA, but lack of hybridization of four tobacco clones suggests possible deletions. Two probes that span a 2.8 kb region of tobacco cpDNA and contain the entire *accD* gene (Li and Cronan 1992) and flanking nontranscribed sequences showed no hybridization, suggesting that the gene has been deleted from the *Trachelium* chloroplast genome. It cannot be ruled out that the sequence has simply diverged significantly from that of tobacco, because mapping in the region of the putative deletion indicates sufficient space (approximately 2.0 kb) to contain a gene of this size (bracket 2 in Fig. 1). It is possible that *accD* is a pseudogene in *Trachelium*. It is also conceivable that the unidentified 2.0-kb segment was in-

serted before or after the deletion of *accD*. Regardless of whether deletion, insertion, or gene divergence has occurred, an inversion endpoint is associated with the location (Fig. 1).

Only two other tobacco cpDNA probes did not hybridize to *Trachelium*. One is a small clone of 213 base pairs (bp), about half of which consists of the 5′ 108 bp of tobacco's large ORF2280. Although this open reading frame is noted for sustaining fairly sizable deletions while remaining an intact ORF (Blasko et al. 1988; Zhou et al. 1988; Downie et al. 1994), it is possible that this gene is not functional in *Trachelium*. The putative deletion may be fairly recent, because the 213-bp tobacco probe hybridizes to cpDNA in some other members of Campanulaceae (M. Cosner and R. Jansen, unpublished). The second probe is located in another large open reading frame of tobacco, ORF1901. No hybridization was detected for a 1415-bp segment of this gene, although, as seen with *accD*, mapping indicates a space sufficiently large to accommodate the segment, suggesting the possibility of divergence rather than deletion (bracket 3 in Fig. 1). The insertion of the genes *clpP* and 5′-*rps12* near the 5′ end of the ORF and the apparent duplicative transposition of a portion of *23s rDNA* (both discussed below) provide further evidence that this gene is non-functional in *Trachelium*.

Insertions

Several apparent insertions (as many as seven) are present in *Trachelium* cpDNA (Fig. 1). These appear to consist of DNA unique to *Trachelium*, as no hybridization to tobacco in these regions was detected. Three of these may not be insertions, but may instead consist of non-functional genes (putative pseudogenes) that have diverged sufficiently as to no longer hybridize to tobacco cpDNA. One of these is the approximately 2.0-kb segment where *accD* normally resides, and the second is located where the 1415-bp tobacco ORF1901 probe should hybridize (both discussed above). The third is a 3.0-kb segment located between the genes *psbB* and *rpl20*, within an otherwise unrearranged stretch of DNA (bracket 1 in Fig. 1). This corresponds to the region from which *clpP* and 5′-*rps12* were either transposed or inverted (see below). It is possible, however, that all three of these DNA segments are insertions of unknown content and origin, rather than pseudogenes. Four other insertions are present, three of which average about 2.0 kb in size and occur at inversion endpoints (Fig. 1). The fifth insertion is about 1.0 kb and is not located at an inversion endpoint (Fig. 1).

Structural organization of the *Trachelium* chloroplast genome

Results from restriction-site and gene mapping show that the chloroplast genome of *Trachelium* is radically rearranged compared to that of other land plants (Fig. 1). Genome modifications include changes in gene order and orientation, the extent of the IR at both large single-copy and small single-copy ends, and gene content. Figure 2 diagrams three possible series of evolutionary events involving the fewest numbers of rearrangements necessary to produce the *Trachelium* genome structure from a tobacco-like ancestor. Rearrangements in *Trachelium* consist of no fewer than seven inversions, IR expansion/contraction, one putative gene loss, putative deletions within two large ORFs, several apparent insertions, and a possible transposition (see below). However, complex hybridization patterns indicate that a more complicated evolutionary scenario consisting of a greater number of rearrangements may actually have occurred. In many instances, individual tobacco probes hybridized to several locations in the cpDNA. Many of these hybridization patterns are readily explained by the obvious splitting of the probe sequence by a rearrangement endpoint. In some cases, the complex hybridization patterns were not as easily interpreted, and were presumably due either to "endpoint remnants" left by previous inversions, to small dispersed repeats within the genome, or to hybridization with non-chloroplast sequences. Although Fig. 2 portrays a step-wise series, it must be emphasized that the order of events is not known.

Evolutionary models

The most complicating factor in postulating evolutionary scenarios for the *Trachelium* chloroplast genome is the relocation of *clpP* and 5′-*rps12* (sequence block I in Fig. 2) from the LSC region to the IR. For the simplicity of modelling, cpDNA insertions and deletions in *Trachelium* are not shown in Fig. 2.

The first model (Fig. 2 A) proposes a series of steps involving seven inversions, one transposition, one IR contraction, and one IR expansion. A constraint on ordering these events is that inversion 3 (Fig. 2) must occur following the IR contraction, since one inversion endpoint lies within what is typically the IR in most angiosperms. This proposed constraint is based on the assumption that inversions between the IR and either single-copy region would be highly disruptive and unlikely to occur. It is likely that two other inversions (1 and 2) involving sequences (segments L and K) found almost entirely in the LSC region of *Trachelium*, but in the IR of most angiosperms, occurred before the IR shrinkage, because the gene *trnL* is the only portion of sequence block K (Fig. 1) remaining in the IR. The remaining four inversions lie in what is typically the LSC region in most land plants. Besides the constraints mentioned above, the rearrangements could have occurred in any sequence.

The second evolutionary model (Fig. 2 B) consists of ten inversions, one IR loss, and one IR re-growth. This scenario is constrained in that loss of the IR segment must occur before inversion 3, inversion 3 must occur before inversions a and b, and inversions a and b must occur before the IR re-growth. A fourth constraint involves the relative orientations of single-copy regions in the ancestral genome. Chloroplast DNA exists in two orientations in equi-

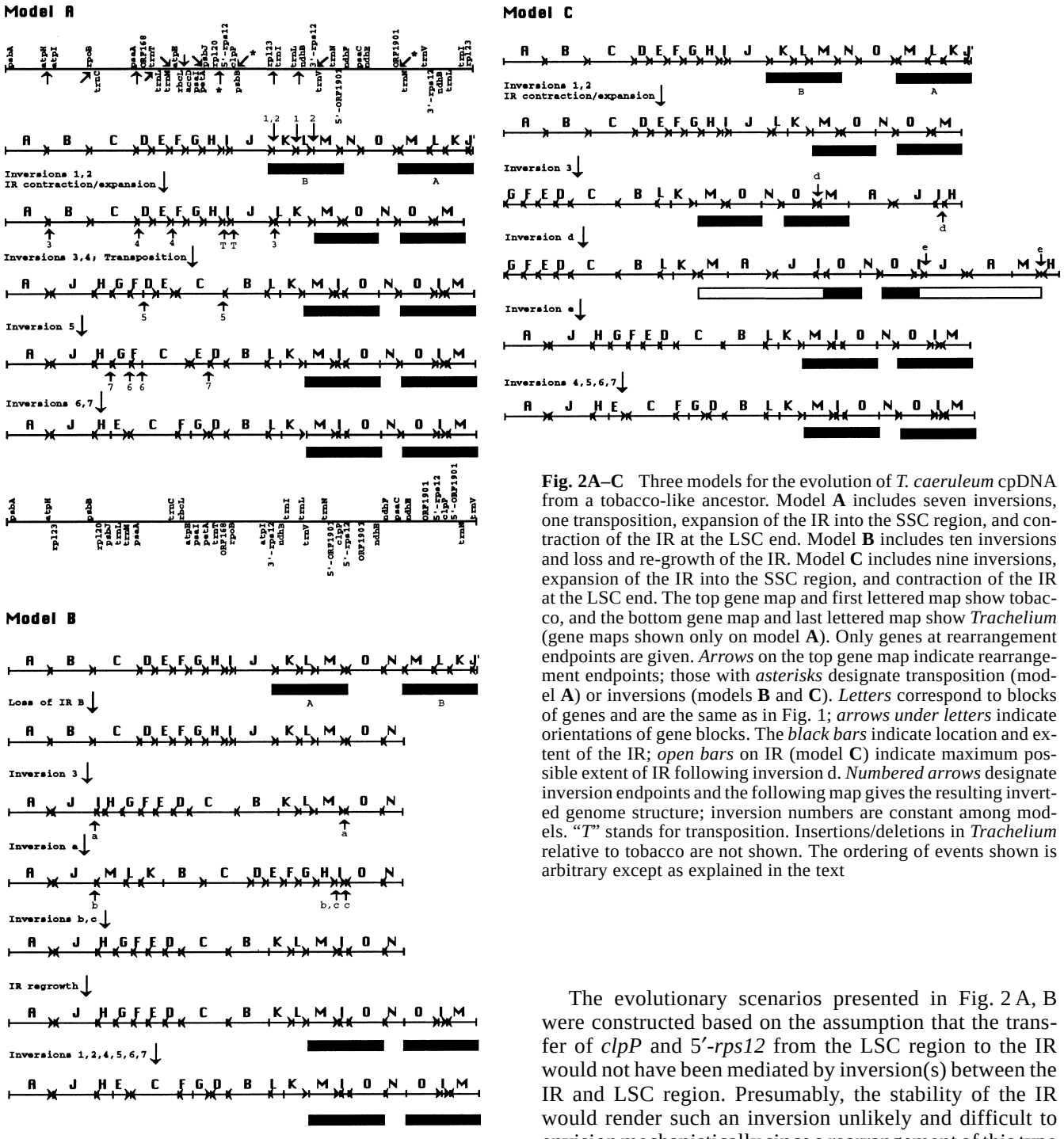

Fig. 2A–C Three models for the evolution of *T. caeruleum* cpDNA from a tobacco-like ancestor. Model **A** includes seven inversions, one transposition, expansion of the IR into the SSC region, and contraction of the IR at the LSC end. Model **B** includes ten inversions and loss and re-growth of the IR. Model **C** includes nine inversions, expansion of the IR into the SSC region, and contraction of the IR at the LSC end. The top gene map and first lettered map show tobacco, and the bottom gene map and last lettered map show *Trachelium* (gene maps shown only on model **A**). Only genes at rearrangement endpoints are given. *Arrows* on the top gene map indicate rearrangement endpoints; those with *asterisks* designate transposition (model **A**) or inversions (models **B** and **C**). *Letters* correspond to blocks of genes and are the same as in Fig. 1; *arrows under letters* indicate orientations of gene blocks. The *black bars* indicate location and extent of the IR; *open bars* on IR (model **C**) indicate maximum possible extent of IR following inversion d. *Numbered arrows* designate inversion endpoints and the following map gives the resulting inverted genome structure; inversion numbers are constant among models. "*T*" stands for transposition. Insertions/deletions in *Trachelium* relative to tobacco are not shown. The ordering of events shown is arbitrary except as explained in the text

molar quantities due to intramolecular recombination between the paired segments of the IR (Palmer 1983; Stein et al. 1986). This "flip-flop" recombination results in two molecules whose only difference is in the relative orientations of their single-copy regions. For model B to operate, the single copy segments must exist in the orientations shown (Fig. 2 B); for comparison, the opposite orientation is diagrammed in Fig. 2 A. All other rearrangements could have occurred in any order.

The evolutionary scenarios presented in Fig. 2 A, B were constructed based on the assumption that the transfer of *clpP* and 5′-*rps12* from the LSC region to the IR would not have been mediated by inversion(s) between the IR and LSC region. Presumably, the stability of the IR would render such an inversion unlikely and difficult to envision mechanistically since a rearrangement of this type would be disruptive to the IR. However, given the complexity of the *Trachelium* chloroplast genome, it is conceivable that inversion between the IR and the LSC region may have occurred, and this is proposed as a mechanism in model C (Fig. 2). Model C invokes expansion/contraction of the IR and nine inversions, at least one of which must occur between the IR and LSC region, accompanied by temporary growth and/or shrinkage of the IR. If inversions d and e of model C occurred with the IR intact, both the transposition of model A and the IR loss/re-growth and one inversion of model B would be eliminated. The effect

of an inversion between the IR and LSC region on the IR is unknown, but it seems likely that inversion d would be accompanied by either severe shrinkage or extreme growth of the IR; model C shows the presumed minimum and maximum extent of the IR following the inversion. This model depicts the maximum extent of the IR following presumed copy correction (Palmer 1991; Birky and Walsh 1992) of the IR copy not involved in inversion d. The minimum extent of the IR following inversion d would contain only sequence block O. In this case, sequence blocks I, J, A, and M would presumably be made single copy by gene conversion or copy correction following loss of one copy of segment M. Alternatively, inversion d could result in a tandem repeat of segment M (data not shown) and the subsequent re-inversion e would then be accompanied by a slight expansion of the IR in its middle (segment I). Inversion e reverses most of inversion d, and it is possible that these two events occurred in concert rather than in an ordered fashion, thus reducing disruption to the IR. A constraint on the ordering of events is that inversion 3 must occur before inversions d and e. In addition, the IR contraction is diagrammed as occurring before inversion 3 to eliminate an additional inversion between the IR and LSC region, since one endpoint of inversion 3 lies in what is typically the IR in angiosperms. For proper orientation of sequence block I, inversions d and e must occur between the LSC region and IR A, adding a third constraint. As in model A, inversions 1 and 2 must occur prior to the IR contraction at its LSC end.

We favor the model that involves transposition (Fig. 2 A) for three reasons. First, model A postulates two fewer steps than model B, and three fewer than model C, by eliminating three and two inversions (and associated growth/shrinkage of the IR in model C), respectively, but adding one transposition. Second, it invokes considerably less radical and complex change to the IR. In particular, although this model proposes extensive contraction of the IR, the rDNA core, a feature of virtually all cpDNA IRs, remains intact. The rDNA-containing IR is temporarily lost in model B, and although the effect of inversion d of model C on the IR is unknown, the temporary exclusion of the rDNA from the IR is possible. Third, model A is less constrained temporally, having only two constraints compared to four each in models B and C.

All three scenarios propose events that do not typically occur in chloroplast genomes. Complete loss and subsequent re-growth of the IR has never been suggested as an evolutionary mechanism in the chloroplast genome, transposition has only rarely been postulated (for further discussion on cpDNA transpositions, see below), and inversion between the IR and a single-copy region is unknown. The numbers of steps and constraints of the various models may not be as important in determining which scenario may be more acceptable as is the actual likelihood of the events. The relative simplicity of the events in model A compared to those in models B-C also favors a model involving transposition.

In a chloroplast genome as rearranged as *Trachelium*, considerations of parsimony or the likelihood of specific rearrangements may be inadequate for determining which of various alternative models may most closely represent actual evolutionary events. Instead, it may be best to use comparative information by examining related chloroplast genomes. Chloroplast DNA mapping data from other genera in the Campanulaceae supply strong evidence against the IR loss/re-growth model (Fig. 2 B). Essentially all taxa examined in this family, including three putatively basal genera in which *clpP* and 5′-*rps12* have not been transferred from the LSC region to the IR, share the same IR/SSC junction (M. Cosner and R. Jansen, unpublished). If the IR was lost in the ancestor of *Trachelium* and most other members of the Campanulaceae (but not in the basal genera) prior to the inversional transfer of these two genes to the IR, an independent IR growth corresponding to the same IR/SSC junction in these two different lineages in the Campanulaceae would be extremely unlikely. A modification (data not shown) of the IR loss/re-growth model compatible with this evidence would include shrinkage of the IR at the LSC end severe enough that the inversion of *clpP* and 5′-*rps12* could occur entirely within the LSC region, thus avoiding inversion between the IR and LSC region. However, this modification would also require part of the IR to re-grow.

Transposition

Since we favor the model (Fig. 2 A) involving transposition of *clpP* and 5′-*rps12*, further discussion of transposition in cpDNA is warranted. Variations in gene order in land-plant chloroplast genomes are most often the result of inversions or gene deletions. Only recently has evidence pointed to a possible role for transposition events in cpDNA evolution. Milligan et al. (1989) reported evidence for the occurrence of at least two transpositions in subclover cpDNA. One of these is a duplicative transposition of part of the ribosomal protein gene *rpl23*, and the second is an apparent transposition of *rpoB* and *rpoC1*. In addition, small segments of certain ribosomal protein genes have shifted positions in the chloroplast genomes of some wheat species, presumably by transposition (Bowman et al. 1988; Ogihara et al. 1988).

The genes *clpP* and 5′-*rps12*, found in the LSC region in most land plants, have been transferred to the IR in *Trachelium*, presumably either by transposition (Fig. 2 A) or inversion (Fig. 2 B, C). The third gene of the *clpP* operon, *rpl20*, has remained in the LSC region. The new location of the genes in the IR is more simply explained by transposition (Fig. 2 A), particularly since the removal of the two genes left an otherwise unrearranged segment of DNA. If transposition was indeed the mechanism, it is unclear whether the event was DNA- or RNA-mediated.

If the *clpP*/5′-*rps12* transposition was RNA-mediated, it would be necessary for these two genes to acquire proper promoter sequences to function in their new location. The two genes are now located within a large open reading frame (ORF1901 of tobacco), very near its 5′ end (Fig. 3). It is possible that the genes have remained functional

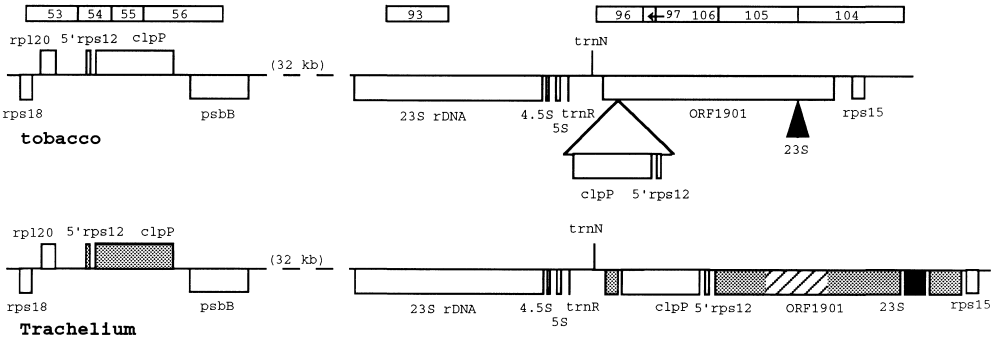

Fig. 3 Model for duplicative transposition of 5′-*rps12* and *clpP* from the LSC region to the IR and a small segment of *23s rDNA* within the IR of *T. caeruleum* cpDNA. *Numbered boxes* show locations of certain tobacco cpDNA probes (Palmer et al. 1994). *Open boxes* indicate presumably functional genes, *shaded boxes* designate putative pseudogenes, the *striped box* shows the region corresponding to a 1415-bp tobacco ORF1901 cpDNA clone that does not hybridize to *Trachelium*, and the *black box* represents a partial copy of the 23*s* rRNA gene. The *dashed line* indicates the DNA segment separating the regions pictured. *Triangles* represent the transposition events

through the cannibalization of the presumably non-functional ORF's promoter. It also is possible that *rpl20* has retained the promoter sequences typically shared with 5′-*rps12* and *clpP*.

PCR amplification of *clpP* using primers for exons 2 and 3, and introns 1 and 2, indicates that both introns still exist in *clpP* in *Trachelium*. (The 9.6-kb *Hin*dIII *Trachelium* clone, i.e., the IR clone that contains *clpP* and 5′-*rps12*, was used as the template in PCR reactions.) If retrotransposition did play a role in the insertion of the two genes in the IR, the presence of introns in *clpP* indicates that the cDNA was made from a partially processed transcript from which *rpl20* was removed, but from which *clpP* introns were not yet spliced. Absence of introns would provide evidence for the involvement of an RNA intermediate, but intron presence is merely inconclusive. Kohchi et al. (1988) found that multiple mRNAs are transcribed from the *clpP* operon (*clpP*, 5′-*rps12*, *rpl20*), and they also described the occurrence of ordered processing and splicing of the transcripts that involved processing prior to splicing of the *clpP* introns. This is consistent with a retrotransposition hypothesis involving an RNA species containing both 5′-*rps12* and *clpP* (including introns) but not *rpl20*.

Our data are also consistent with a DNA-mediated transposition of these two genes. In a DNA-mediated transposition, the necessity of acquiring proper promoter sequences is readily avoided, as it is conceivable that the promoter could be transferred as well. However, it would then be necessary for *rpl20* to acquire a promoter, as it would be left orphaned by such an event unless it were duplicative.

It also appears that a duplicative transposition of a segment of the 23*s* rRNA gene has taken place in *Trachelium*. A 1680-bp tobacco cpDNA probe consisting entirely of coding sequence (the 5′ and 3′ ends of the gene are located in two other probes) from within this gene hybridizes both to the presumably functional 23*s* rRNA locus and to a region approximately 14 kb downstream in *Trachelium* (Fig. 3). The exact size of the duplication is unknown, but the smallest restriction fragment containing the duplicated DNA is an *Xba*I fragment of approximately 1.0 kb. The partially duplicated *23s rDNA* is located within the same large open reading frame (*ORF1901*) to which part of the *clpP* operon was also putatively transposed (Fig. 3).

Disruption of cpDNA operons

At least two operons have been disrupted by rearrangements in *Trachelium*. Genes within both the *atpA* and *clpP* operons have been separated by different rearrangements. One endpoint of inversion 3 (Fig. 2) lies between two genes for subunits of the H^+-ATPase complex, *atpH* and *atpI*, reported to be transcriptionally linked in spinach (Westhoff et al. 1985; Hennig and Herrmann 1986; Hudson et al. 1987), pea (Cozens et al. 1986; Hudson et al. 1987), rice (Kanno and Hirai 1993), and wheat (Hoglund et al. 1990). Another rearrangement has placed *clpP* and 5′-*rps12* in the *Trachelium* IR, while *rpl20* has remained in the LSC region (Fig. 3). These three genes were shown to be co-transcribed in the liverwort *Marchantia* (Kohchi et al. 1988), *Pinus* (Clarke et al. 1994), rice (Kanno and Hirai 1993), and maize (Weglohner and Subramanian 1992). The *clpP* operon was disrupted either by inversion b or d (Fig. 2 B and 2 C, respectively) or by the transposition postulated in Fig. 2 A. Only two other studies report disruption of cpDNA transcription units by rearrangements in land plants. The mung bean *rpl23* operon is split by an inversion (Palmer et al. 1988), whereas a transposition has been implicated in the disruption of subclover's *rpoB* operon (Milligan et al. 1989). It is possible that new transcription linkages have been formed in *Trachelium*, as has been reported for the rearranged chloroplast genome of pea (Nagano et al. 1991).

Dispersed repeats

Comparisons among chloroplast genomes of highly divergent plant lineages show a remarkable conservation of

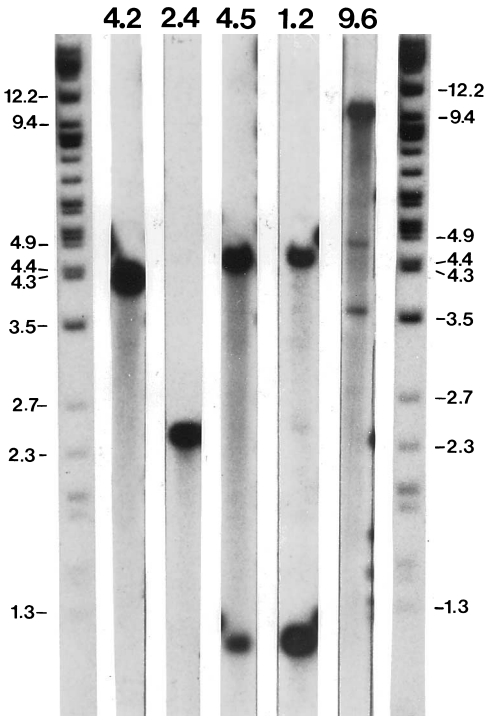

Fig. 4 Short dispersed repeats in the chloroplast genome of *T. caeruleum*. Filters containing *Trachelium* total DNA digested with *Hin*dIII and electrophoresed on a 1.2% agarose gel were hybridized with each of five *Trachelium* cpDNA *Hin*dIII clones. Sizes in kb of the *Hin*dIII inserts are given above each lane. The *outside lanes* contain lambda DNA molecular-weight standards with fragment sizes in kb indicated. The 1.2-kb and 4.5-kb fragments contain the *23s rDNA* and its duplicated segment, respectively

chromosome structure, gene content, and gene sequence. In particular, the large IR is especially conserved, showing an even slower rate of sequence evolution than either single-copy region (Wolfe et al. 1987; Downie and Palmer 1992 a). The IR is believed to confer an inherent structural stability on cpDNA by limiting recombination between other segments of the chromosome (Palmer 1985). With the exception of the large inverted repeat, chloroplast genomes seldom contain repeated sequences, possibly related to a constraint on the size of the cpDNA molecule (Palmer 1985). Exceptions include the cpDNAs of geranium (Palmer et al. 1987), wheat (Howe 1985; Quigley and Weil 1985; Bowman and Dyer 1986; Bowman et al. 1988; Ogihara et al. 1988), rice (Shimada and Sugiura 1989), subclover (Milligan et al. 1989), broad bean (Bonnard et al. 1985) and Douglas fir (Tsai and Strauss 1989), whose chloroplast genomes all contain a series of short dispersed repeats. In these species, dispersed repeats have been associated with inversion endpoints and are thought to have facilitated inversions through recombination across homologous repeats.

Certain cloned *Trachelium Hin*dIII cpDNA fragments hybridize only to themselves (e.g., fragments of 2.4 and 4.2 kb; Fig. 4) when hybridized to *Hin*dIII digests of total *Trachelium* DNA, but others hybridize to one or more additional *Hin*dIII fragments. For example, a 4.5-kb *Hin*dIII clone hybridizes to itself and to a 1.2 kb fragment (Fig. 4). The reciprocal hybridization of the 1.2-kb clone to the 4.5-kb fragment indicates the existence of a region of common sequence between the two fragments (Fig. 4). In another case, a 9.6-kb clone hybridizes to itself and to fragments of 4.9 and 3.8 kb (Fig. 4). Reciprocal hybridization indicates that the 4.9-and 3.8-kb fragments do not hybridize to each other, and thus share different sequence homologies with the 9.6-kb fragment. There are at least five clear sets of dispersed repeats within the *Trachelium* chloroplast genome (r1–r5 in Fig. 1). Many other *Trachelium Hin*dIII clones hybridize faintly to additional *Hin*dIII fragments, but these data are more difficult to interpret. They may be due to the hybridization of probes with non-chloroplast DNA fragments.

The *Trachelium* chloroplast genome contains at least five families of dispersed repeats (Figs. 1 and 4). All of these are associated with various rearrangements, including inversions, insertions/deletions and a possible transposition, thereby suggesting a mechanism for at least some of the rearrangements. One of the repeated sequences consists of part of the 23*s* rRNA gene, presumably the result of a duplicative transposition (see above) (Fig. 3), whereas the remainder are of unidentified content. It is not known how these repeats may have originated nor why they are present in so few plant species. Although dispersed repeats may play a role in the de-stabilization of the chloroplast genome of *Trachelium*, not all inversion endpoints are associated with dispersed repeats. It is possible that in some instances the small repeats were lost as the result of inversion, but it seems likely that additional inversion mechanisms were involved in *Trachelium* as well.

In addition to small dispersed repeats, insertions and deletions are frequently associated with inversion endpoints in *Trachelium* (Fig. 1), and may be part of the inversion process. However, it is not known if the insertions and deletions are a consequence or a cause of the inversions.

Endpoint re-use

Regardless of which evolutionary model is accepted, reuse of inversion endpoints has occurred in *Trachelium*. According to Fig. 2 B, there are ten inversions and 13 endpoints. If all inversions used different endpoints, the maximum number of endpoints would be 20. In Fig. 2 A, seven inversions use ten endpoints; maximally 14 endpoints could exist. The nine inversions of Fig. 2 C use 13 endpoints rather than the maximum 18. Without knowing the precise order of events, it is not possible to determine the exact nature of the endpoint re-use. The cpDNA of the closely related family Lobeliaceae is also rearranged, but not as radically as in the Campanulaceae. More extensive endpoint re-use has apparently occurred in the Lobeliaceae. Among 18 taxa from three genera examined, 11 inversions sharing ten endpoints were found (Knox et al. 1993).

The re-use of endpoints led Knox et al. (1993) to postulate the existence of recombinational “hot spots” (Kung

et al. 1982; Ogihara et al. 1988, 1991) in the Lobeliaceae. The tRNA genes have been associated with cpDNA inversion endpoints in several families of angiosperms including the Asteraceae (Jansen and Palmer 1987 b), the Fabaceae (Michalowski et al. 1987; Palmer et al. 1988), the Onagraceae (Herrmann et al. 1983) and the Poaceae (Howe 1985; Quigley and Weil 1985; Rodermel et al. 1987; Howe et al. 1988; Hiratsuka et al. 1989), as well as in some conifers (Strauss et al. 1988; Lidholm and Gustafsson 1991; Tsudzuki et al. 1992), a fern (Hasebe and Iwatsuki 1992), a moss (Calie and Hughes 1987), and a liverwort (Ohyama et al. 1986). The inversions may have been mediated through recombination across short repeated sequences found within or near the tRNA genes (Bonnard et al. 1985; Howe 1985; Howe et al. 1988; Hiratsuka et al. 1989; Shimada and Sugiura 1989; Rodermel 1992), but the exact mechanism by which the repeats may facilitate recombination is unclear. Most of the endpoints in the Lobeliaceae are associated with tRNA genes (Knox et al. 1993). In *Trachelium*, six or seven endpoints (depending on which model is accepted) occur near one or more tRNA genes (Figs. 1 and 2). Although tRNA genes and associated repeats may play a role in cpDNA inversion, not all inversion endpoints lie near these genes, indicating that other mechanistic factors are also involved.

It has been suggested that the IR stabilizes cpDNA by decreasing intrachromosomal recombination (Palmer 1985). Conifers (Lidholm et al. 1988; Strauss et al. 1988; Raubeson and Jansen 1992) and some legumes (Palmer and Thompson 1981; Lavin et al. 1990; Palmer 1991), in which chloroplast genomes have become highly rearranged, have also lost one copy of the inverted repeat, whereas *Trachelium* retains its IR. It is clear that factors other than the loss of one copy of the IR have served to de-stabilize the cpDNA of *Trachelium*.

Conclusions

Chloroplast genome evolution in *Trachelium* is notable in five major respects. First, as the result of at least seven inversions, gene order in *Trachelium* is highly rearranged relative to most other land plants. Second, the IR has spread into the adjacent SSC region but has been reduced in size at its LSC end, resulting in larger and smaller LSC and SSC regions, respectively. Third, evidence suggests that transposition, a rarely reported event in cpDNA evolution (Palmer 1991), has also contributed to the scrambled gene order of *Trachelium*. Fourth, at least five families of dispersed repeats have been identified in *Trachelium*, whereas most chloroplast genomes lack small repeated sequences. Fifth, at least two rearrangements (either two inversions or one inversion and one transposition) have resulted in the disruption of each of two different transcriptional units.

Acknowledgements Financial support for this study was provided by a National Science Foundation doctoral dissertation improvement grant BSR-9101239 to M.E.C. (Daniel Crawford, principal investigator), an American Society of Plant Taxonomists research grant to M.E.C., and a National Science Foundation grant DEB-9318279 to R.K.J. Loni Walker provided valuable assistance with photography, and John Eary assisted in preparing Fig. 1. Thanks also to Daniel Crawford for support and guidance.

References

Birky CM Jr, Walsh JB (1992) Biased gene conversion, copy number, and apparent mutation-rate differences within chloroplast and bacterial genomes. Genetics 130:677–683

Birnboim HC, Doly JC (1979) A rapid alkaline-extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513–1523

Blasko K, Kaplan SA, Higgins KG, Wolfson R, Sears BB (1988) Variation in number of a 24-base pair tandem repeat in the chloroplast DNA of *Oenothera hookeri* strain Johansen. Curr Genet 14:287–292

Bonnard G, Weil J-H, Steinmetz A (1985) The intergenic region between the *Vicia faba* chloroplast $tRNA^{Leu}$(CAA) and $tRNA^{Leu}$(UAA) genes contains a partial copy of the split $tRNA^{Leu}$(UAA) gene. Curr Genet 9:417–422

Bowman CM, Dyer TA (1986) The location and possible evolutionary significance of small dispersed repeats in wheat ctDNA. Curr Genet 10:931–941

Bowman CM, Barker RF, Dyer TA (1988) In wheat ctDNA, segments of ribosomal protein genes are dispersed repeats, probably conserved by non-reciprocal recombination. Curr Genet 14:127–136

Calie PJ, Hughes KW (1987) The consensus land-plant chloroplast gene order is present, with two alterations, in the moss *Physcomitrella patens*. Mol Gen Genet 208:335–341

Clarke, AK, Gustafsson, P, Lidholm JA (1994) Identification and expression of the chloroplast *clpP* gene in the conifer *Pinus contorta*. Plant Mol Biol 26:851–862

Cozens AL, Walker JE, Phillips AL, Huttly AK, Gray JC (1986) A sixth subunit of ATP synthase, an F_0 component, is encoded in the pea chloroplast genome. EMBO J 5:217–222

Downie SR, Palmer JD (1992 a) Restriction-site mapping of the chloroplast DNA inverted repeat: a molecular phylogeny of the Asteridae. Ann Missouri Bot Gard 79:266–283

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

Downie SR, Katz-Downie DS, Wolfe KH, Calie PJ, Palmer JD. (1994) Structure and evolution of the largest chloroplast gene (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, Davis JI, Soreng RJ, Garvin D, Anderson MJ (1992) Chloroplast DNA inversions and the origin of the grass family (Poaceae). Proc Natl Acad Sci USA 89:7722–7726

Hasebe M, Iwatsuki K (1992) Gene localization on the chloroplast DNA of the maiden hair fern; *Adiantum capillus-veneris*. Bot Mag (Tokyo) 105:413–419

Hennig J, Herrmann RG (1986) Chloroplast ATP synthase of spinach contains nine non-identical subunit species, six of which are encoded by plastid chromosomes in two operons in a phylogenetically conserved arrangement. Mol Gen Genet 203:117–128

Herrmann RG, Westhoff P, Alt J, Winter P, Tittgen J, Bsanz C, Sears BB, Nelson N, Hurt E, Hauska G, Viebrock A, Sebald W (1983) Identification and characterization of genes for polypeptides of the thylakoid membrane. In: Ciferri O, Dure L (eds) Structure and function of plant genomes. Plenum Press, New York, pp 143–153

Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun C-R, Meng B-Y, Li Y-Q, 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 accounts for a major plastid DNA inversion during the evolution of the cereals. Mol Gen Genet 217:185–194
Hoglund A-S, Plant AL, Gray JC (1990) Expression of the wheat chloroplast gene for CF_0 subunit IV of ATP synthase. Curr Genet 18:471–476
Hoot SB, Palmer JD (1994) Structural rearrangements, including parallel inversions, within the chloroplast genome of *Anemone* and related genera. J Mol Evol 38:274–281
Howe CJ (1985) The endpoints of an inversion in wheat chloroplast DNA are associated with short repeated sequences containing homology to *att*-lambda. Curr Genet 10:139–145
Howe CJ, Barker RF, Bowman CM, Dyer TA (1988) Common features of three inversions in wheat chloroplast DNA. Curr Genet 13:343–349
Hudson GS, Mason JG, Holton TA, Koller B, Cox GB, Whitfield PR, Bottomley W (1987) A gene cluster in the spinach and pea chloroplast genomes encoding one CF_1 and three CF_0 subunits of the H^+-ATP synthase complex and the ribosomal protein S2. J Mol Biol 196:283–298
Jansen RK, Palmer JD (1987 a) A chloroplast DNA inversion marks an ancient evolutionary split in the sunflower family (Asteraceae). Proc Natl Acad Sci USA 84:5818–5822
Jansen RK, Palmer JD (1987 b) Chloroplast DNA from lettuce and *Barnadesia* (Asteraceae): structure, gene localization, and characterization of a large inversion. Curr Genet 11:553–564
Jansen RK, Palmer JD (1988) Phylogenetic implications of chloroplast DNA restriction-site variation in the Mutisieae (Asteraceae). Am J Bot 75:753–766
Kanno A, Hirai A (1993) A transcription map of the chloroplast genome from rice (*Oryza sativa*). Curr Genet 23:166–174
Knox EB, Downie SR, Palmer JD (1993) Chloroplast genome rearrangements and the evolution of giant lobelias from herbaceous ancestors. Mol Biol Evol 10:414–430
Kohchi T, Ogura Y, Umesono K, Yamada Y, Komano T, Ozeki H, Ohyama K (1988) Ordered processing and splicing in a polycistronic transcript in liverwort chloroplasts. Curr Genet 14:147–154
Kung SD, Zhu YS, Shen GF (1982) *Nicotiana* chloroplast genome. 3. Chloroplast DNA evolution. Theor Appl Genet 61:73–79
Lavin M, Doyle JJ, Palmer JD (1990) Evolutionary significance of the loss of the chloroplast DNA inverted repeat in the Leguminosae subfamily Papilionoideae. Evolution 44:390–402
Li S-J, Cronan JE Jr (1992) Putative zinc-finger protein encoded by a conserved chloroplast gene is very likely a subunit of a biotin-dependent carboxylase. Plant Mol Biol 20:759–761
Lidholm J, Gustafsson P (1991) A three-step model for the rearrangement of the chloroplast *trnK-psbA* region of the gymnosperm *Pinus contorta*. Nucleic Acids Res 19:2881–2887
Lidholm J, Szmidt AE, Hallgren J-E, Gustafsson P (1988) The chloroplast genomes of conifers lack one of the rDNA-encoding inverted repeats. Mol Gen Genet 212:6–10
Michalowski C, Breunig KD, Bohnert HJ (1987) Points of rearrangements between plastid chromosomes: location of protein coding regions on broad bean chloroplast DNA. Curr Genet 11:265–274
Milligan BG, Hampton JN, Palmer JD (1989) Dispersed repeats and structural re-organization in subclover chloroplast DNA. Mol Biol Evol 6:355–368
Nagano Y, Matsuno R, Sasaki Y (1991) Sequence and transcriptional analysis of the gene cluster *trnQ-zfpA-psaI*-ORF231-*petA* in pea chloroplasts. Curr Genet 20:431–436
Ogihara Y, Terachi T, Sasakuma T (1988) Intramolecular recombination of the chloroplast genome mediated by short direct-repeat sequences in wheat species. Proc Natl Acad Sci USA 85:8573–8577
Ogihara Y, Terachi T, Sasakuma T (1991) Molecular analysis of the hot-spot region related to length mutations in wheat chloroplast DNAs. I. Nucleotide divergence of genes and intergenic spacer regions located in the hot-spot region. Genetics 129:873–884
Olmstead RG, Palmer JD (1992) A chloroplast DNA phylogeny of the Solanaceae: subfamilial relationships and character evolution. Ann Missouri Bot Gard 79:346–360
Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota S, Inokuchi H, Ozeki H (1986) Chloroplast gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322:572–574
Palmer JD (1983) Chloroplast DNA exists in two orientations. Nature 301:92–93
Palmer JD (1985) Evolution of chloroplast and mitochondrial DNA in plants and algae. In: MacIntyre RJ (ed) Monographs in evolutionary biology: molecular evolutionary genetics. Plenum Press, New York, pp 131–240
Palmer JD (1986) Isolation and structural analysis of chloroplast DNA. Methods Enzymol 118:167–186
Palmer JD (1991) Plastid chromosomes: structure and evolution. In: Bogorad L, Vasil IK (eds) The molecular biology of plastids, vol. 7A in Vasil IK (ed) Cell culture and somatic cell genetics of plants. Academic Press, New York, pp 5–53
Palmer JD, Thompson WF (1981) Rearrangements in the chloroplast genomes of mung bean and pea. Proc Natl Acad Sci USA 78: 5533–5537
Palmer JD, Nugent JM, Herbon LA (1987) Unusual structure of geranium chloroplast DNA: a triple-sized inverted repeat, extensive gene duplications, multiple inversions and two repeat families. Proc Natl Acad Sci USA 84:769–773
Palmer JD, Osorio B, Thompson WF (1988) Evolutionary significance of inversions in legume chloroplast DNAs. Curr Genet 14:65–74
Palmer JD, Downie SR, Nugent JM, Brandt P, Unseld M, Klein M, Brennicke A, Schuster W, Borner T (1994) The chloroplast and mitochondrial DNAs of *Arabidopsis thaliana*: conventional genomes in an unconventional plant. In: Somerville C, Meyerowitz E (eds) *Arabidopsis*. Cold Spring Harbor Press, Cold Spring Harbor, New York, pp 37–62
Quigley F, Weil JH (1985) Organization and sequence of five tRNA genes and of an unidentified reading frame in the wheat chloroplast genome: evidence for gene rearrangements during the evolution of chloroplast genomes. Curr Genet 9:495–503
Raubeson LA, Jansen RK (1992) A rare chloroplast-DNA structural mutation is shared by all conifers. Biochem Syst Ecol 20:17–24
Rodermel SR (1992) Nucleotide sequence of a maize chloroplast DNA fragment containing an inversion breakpoint, *trnG* (GCC), *trnG* (UCC), *trnfM*, and a *trnG* pseudogene. Nucleic Acids Res 20:5844
Rodermel SR, Orlin P, Bogorad L (1987) The transcription termination region between two convergently transcribed photoregulated operons in the maize plastid chromosome contains *rps14*, *trnR* (UCU) and a putative *trnfM* pseudogene. Nucleic Acids Res 13:5493
Shimada H, Sugiura M (1989) Pseudogenes and short repeated sequences in the rice chloroplast genome. Curr Genet 16:293–301
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 B-Y, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5:2043–2049
Stein DB, Palmer JD, Thompson WF (1986) Structural evolution and flip-flop recombination of chloroplast DNA in the fern genus *Osmunda*. Curr Genet 10:835–841
Strauss SH, Palmer JD, Howe GT, Doerksen AH (1988) Chloroplast genomes of two conifers lack an inverted repeat and are extensively rearranged. Proc Natl Acad Sci USA 85:3898–3902
Sugiura M, Shinozaki K, Zaita N, Kusuda M, Kumano M (1986) Clone bank of the tobacco (*Nicotiana tabacum*) chloroplast genome as a set of overlapping restriction endonuclease fragments: mapping of 11 ribosomal protein genes. Plant Sci 44:211–216
Tsai C-H, Strauss SH (1989) Dispersed repetitive sequences in the chloroplast genome of Douglas-fir. Curr Genet 16:211–218
Tsudzuki J, Nakashima K, Tsudzuki T, Hiratsuka J, Shibata M, Wakasugi T, Sugiura M (1992) Chloroplast DNA of black pine

retains a residual inverted repeat lacking rRNA genes: nucleotide sequences of *trnQ*, *trnK*, *psbA*, *trnI* and *trnH* and the absence of *rps16*. Mol Gen Genet 232:206–214

Weglohner W, Subramanian AR (1992) Nucleotide sequence of a region of maize chloroplast DNA containing the 3′ end of *clpP*, exon 1 of *rps12* and *rpl20* and their co-transcription. Plant Mol Biol 18:415–418

Westhoff P, Alt J, Nelson N, Herrmann RG (1985) Genes and transcripts for the ATP synthase CF_0 subunits I and II from spinach thylakoid membranes. Mol Gen Genet 199:290–299

Wolfe KH, Li W-H, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054–9058

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

Zhou D-X, Massenet O, Quigley F, Marion MJ, Moneger F, Huber P, Mache R (1988) Characterization of a large inversion in the spinach chloroplast genome relative to *Marchantia*: a possible transposon-mediated origin. Curr Genet 13:433–439