

Structural Rearrangements, Including Parallel Inversions, Within the Chloroplast Genome of *Anemone* and Related Genera

Sara B. Hoot,¹ Jeffrey D. Palmer²

¹ University of Michigan Herbarium, North University Building, Ann Arbor, MI 48109, USA

² Department of Biology, Indiana University, Bloomington, IN 47405, USA

Received: 4 May 1993 / Revised: 20 September 1993

Abstract. Chloroplast DNA cleavage sites for 10 restriction enzymes were mapped for 46 species representing all sections of *Anemone*, four closely related genera (*Clematis*, *Pulsatilla*, *Hepatica*, and *Knowltonia*), and three more distantly related outgroups (*Caltha*, *Ranunculus*, and *Adonis*). Comparison of the maps revealed that the chloroplast genomes of *Anemone* and related genera have sustained an unusual number and variety of rearrangements. A single inversion of a 42-kb segment was found in the large single-copy region of *Adonis aestivalis*. Two types of rearrangements were found in the chloroplast genome of *Clematis*, *Anemone*, *Pulsatilla*, *Hepatica*, and *Knowltonia*: An approximately 4-kb expansion of the inverted repeat and four inversions within the large single-copy region. These rearrangements support the monophyletic status of these genera, clearly separating them from *Caltha*, *Ranunculus*, and *Adonis*. Two further inversions were found in two *Clematis* species and three *Anemone* species. While appearing to support a monophyletic grouping for these taxa, these two inversions conflict with data from both chloroplast restriction sites and morphology and are better interpreted as having occurred twice independently. These are the first two documented cases of homoplastic inversions in chloroplast DNA. Finally, the second intron of the chloroplast *rps12* gene was shown to have been lost in the common ancestor of the same three *Anemone* species that feature the two homoplastic inversions.

Key words: Chloroplast DNA — Rearrangements — Inversions — Intron loss — Homoplasy — Ranunculaceae — *Anemone* complex

Structural rearrangements such as inversions are rare in the chloroplast genome of the majority of land plants (Palmer et al. 1988b; Palmer 1991). Most angiosperm chloroplast DNAs are relatively uniform in size (135–160 kb pairs), containing a 20–30-kb-pair inverted repeat and two single-copy regions which are usually invariant in gene order (Palmer 1991). Prominent exceptions to this are the highly rearranged chloroplast DNAs (cpDNAs) found in conifers (Strauss et al. 1988), grasses (Quigley and Weil 1985; Howe et al. 1988; Hiratsuka et al. 1989), legumes (Palmer et al. 1988a; Milligan et al. 1989), geraniums (Palmer et al. 1987a), and Lobeliaceae (Knox et al. 1993).

Inversions to date appear devoid of homoplasy (parallel, convergent, and reverse evolution). For this reason, they have made excellent taxonomic markers in studies of the Asteraceae (Jansen and Palmer 1987), Phaseoleae (Palmer et al. 1988a; Bruneau et al. 1990), ferns (Stein et al. 1986), Poaceae (Doyle et al. 1992), vascular plants (Raubeson and Jansen 1992), and Lobeliaceae (Knox et al. 1993).

In a survey of cpDNA restriction-site variability of *Anemone* and related genera, we found multiple inversion events in the large single-copy region. Four of these inversions appear to have occurred at a time before the *Anemone* complex (consisting of *Anemone*, *Hepatica*, *Pulsatilla*, and *Knowltonia*) and *Clematis* had

Correspondence to: S. Hoot, Field Museum of Natural History, Roosevelt Road at Lake Shore Drive, Chicago, IL 60605, USA

diversified. Two additional inversions appear to have occurred twice independently. We discuss factors which may have contributed to these multiple inversions.

Materials and Methods

A total of 45 species, representing all sections of *Anemone* and seven other genera, were surveyed (Hoot et al. 1994). *Ranunculus*, *Caltha*, *Adonis*, *Clematis*, *Hepatica*, *Pulsatilla*, and *Knowltonia* were included as potential outgroups based on a phylogeny derived from conventional data (Hoot 1991) and past classifications (Gregory 1941; Tamura 1967). Leaf material was sent or collected from various locations or grown from seed or tubers in the garden of coauthor S.B.H. (Voucher specimens are deposited at MICH.)

Total cellular DNA was isolated from fresh leaves according to the miniprep method of Doyle and Doyle (1987) and further purified on a CsCl gradient. Samples were digested with ten restriction enzymes: *Bam*HI, *Bcl*I, *Bgl*II, *Cl*aI, *Dra*I, *Eco*RI, *Eco*RV, *Hind*III, *Sac*I and *Xba*I. The resulting fragments were separated on agarose gels, transferred to filters, and hybridized using the methods described in Palmer et al. (1988b). A series of filter hybridizations was performed using as probes 40 clones that cover virtually the entire tobacco chloroplast genome (Fig. 1; Olmstead and Palmer 1992). Two additional tobacco clones, tobacco coordinate units (TCUs; Shinozaki et al. 1986; Knox et al. 1993) 80652–82469 and TCU 82469–82752, were used to locate more precisely the endpoints of the extended inverted repeat found in most taxa. Four additional subclones that span probe 4 (Fig. 1) were used to better define several inversion endpoints.

To better locate the endpoints of the multiple inversions and their arrangement, two additional filters were created for four representative taxa: *Adonis aestivalis*, *Anemone canadensis*, *Anemone multifida*, and *Clematis hexapetala*. Single digests using the above 10 enzymes and nine double digests using *Bam*HI as the second enzyme were performed for all taxa except *Adonis*. *Adonis* DNA was digested singly with *Hind*III, *Bgl*II, *Eco*RI, and *Bam*HI and doubly with *Bam*HI and each of the first three enzymes. These filters were probed with various tobacco clones and with four synthetic oligonucleotides whose sequences correspond to regions internal to tRNA genes from tobacco: *trnS*-GGA (39 bases), *trnT*-UGU (41 bases), *trnL*-UAA (35 bases), and *trnF*-GAA (40 bases). The oligonucleotides were labeled with ³²P using the tailing reaction described by Collins and Hunsaker (1985) with some modifications. Each 10- μ l reaction was carried out using 20 ng oligonucleotide DNA, 20 units of terminal deoxynucleotidyl transferase (purchased from IBI), and 70.0 μ Ci of α -labeled (³²P)dATP, in a buffer of 0.14 M K-cacodylate (pH 7.2), 30 mM Tris base, 1.0 mM CoCl₂ and 0.1 mM DTT. The probes were purified on a 10-ml Sephadex G-25 column. The filters were hybridized overnight at 42°C in the same hybridization mixture used above except for the addition of formamide (15% of total volume). The hybridized filters were washed four times in 1 L of 0.5 \times SSC and 0.1% SDS at 42°C.

Polymerase chain reaction (PCR) amplification to determine the presence or absence of intron 2 of *rps12* was performed according to the conditions described in Kuhnel et al. (1990), except that the synthesis step was carried out at 55°C rather than 50°C. One amplifying primer is located at the 3' end of exon 2 of *rps12*, has the same orientation as *rps12*, and starts at position 100671 of the tobacco chloroplast genome sequence (Shinozaki et al. 1986). The sequence of this 22-mer is CGAGGA(A/G)CCCTAGATGCTGTCG. The second amplifying primer is located at the 5' end of *rps7* (the gene immediately downstream of *rps12*), has the opposite orientation as *rps12*, and starts at tobacco position 99926. The sequence of this 19-mer is CCA(A/C)CATGTAACTAATCG.

Phylogenetic analyses of restriction site data (excluding rearrangements) were conducted using the heuristic search option of PAUP (Swofford 1990), implemented on a MacIntosh II computer.

Analyses were performed using 50 replications of the random addition sequence. Because the relationship of *Caltha*, *Ranunculus*, and *Adonis* was not resolved by the restriction site data, only character states found in agreement for all three taxa were scored to create an "outgroup" for a further phylogenetic analysis. Further details on the selection of outgroups, the mapping of restriction sites, and the phylogenetic analyses of restriction site mutations may be found in Hoot et al. (1994).

Results

Length Mutations

Complete maps were constructed for all 10 enzymes employed. The data matrix for the chloroplast restriction site data is available upon request from S. Hoot. The 10 restriction enzymes mapped recognized approximately 510 sites in each cpDNA. Most of the tobacco probes hybridized strongly to the DNAs of *Anemone* and other genera. This suggests that the chloroplast gene content is for the most part conserved.

Because of the substantial degree of restriction-site variation and the numerous inversions, length mutations were difficult to evaluate in some regions of the genome of the *Anemone* complex. Restriction-fragment size variation seemed common in the large single-copy region near the inversion endpoints. This variation took the form of a range of fragment lengths with no distinctive cutoff points between sizes. Two of these are found in the regions TCU 6149–6652 and TCU 48602–49800 (TCU—tobacco coordinate units, Shinozaki et al. 1986). The range of size variation was 50–500 bp. The region TCU 111920–113110 in the small single-copy region next to the end of the inverted repeat contained a series of length polymorphisms. Differences in size varied from 500 to 1,500 bp. Like the previous length variations, these were phylogenetically uninformative.

One phylogenetically informative length mutation was found shared by *Anemone canadensis*, *A. dichotoma*, and *A. richardsonii*. These species all displayed a decrease in length of 500–600 bp in a region within tobacco probe 32 that mapped to the approximate position of *rps12* within the inverted repeat. Because the size of this apparent deletion corresponds well to the known size of intron 2 of *rps12* [intron 1 is trans-spliced (Fig. 1) and thus unlikely to be involved in this length mutation], we performed PCR and DNA sequencing experiments to investigate the possibility that this deletion represented the loss of the entire intron 2 in the *A. canadensis* lineage. As shown in Fig. 2, PCR using primers closely flanking this intron produced fragments of the expected size for the sequenced *rps12* genes from tobacco (lane 2; intron size = 536 bp; expected PCR size = 746 bp; Shinozaki et al. 1986), soybean (lane 3; 531 bp; 741 bp; von Allmen and Stutz 1987), and maize (lane 4; 540 bp; 750 bp; Giese et al.

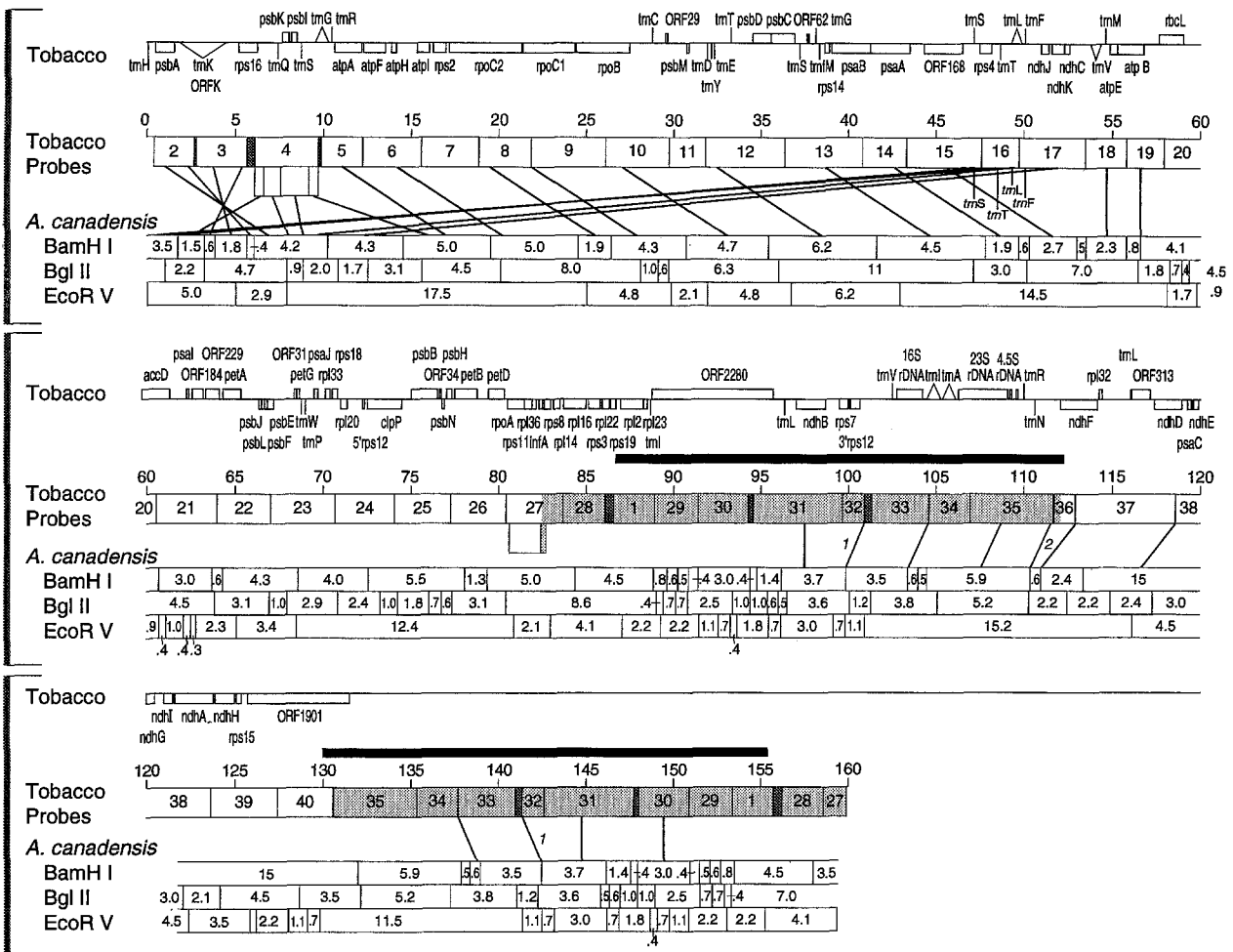


Fig. 1. Comparison of chloroplast genome organization in *Anemone canadensis* and tobacco. Genes above the tobacco genome map (Shinozaki et al. 1986) are transcribed from left to right; those below the line are transcribed from right to left. The dark bars represent the 25.3-kb inverted repeat of tobacco. The 40 tobacco probes used are numbered as in Olmstead and Palmer (1992). Numbers projected above probes are TCUs \times 1,000. Darkly shaded blocks are gaps between tobacco probes. Lightly shaded regions indicate the location of the 29.4-kb inverted repeat in *A. canadensis* and other *Anemone*, *Hepatica*, *Pulsatilla*, *Knowltonia*, and *Clematis* species included in this study. Additional tobacco clones and oligonucleotides used to locate inverted repeat and inversion endpoints are shown projecting below

1987). A single product of this same approximate size was obtained for a large number of species in *Anemone* and related genera, including the four members of *Anemone*, *Knowltonia*, and *Pulsatilla* shown in lanes 8–11 of Fig. 2. However, three species, *A. canadensis*, *A. dichotoma*, and *A. richardsonii*, consistently produced a single major PCR product of about 210–220 bp (Fig. 2, lanes 5–7). This size matches well with that expected (210 bp) for an *rps12* gene that lacks intron 2. We then sequenced directly the gel-isolated PCR product from *A. canadensis* and found that intron 2 was indeed cleanly missing from its *rps12* gene (data not shown). We therefore conclude that intron 2 was lost from the *rps12* gene in a common ancestor of these

the probe map. (See tobacco probes 4, 15–17, and 27.) The lines between the *Anemone* and the tobacco genomes locate similar regions as indicated by the probes. Lines were only drawn when similar regions of the two genomes were not aligned. The two length mutations found in *A. canadensis* are numbered in *italic* near their locations. The location and size (kb) of the various restriction site fragments are shown for three enzymes: *Bam*HI, *Bgl*II, and *Eco*RV. The tobacco map is modified from Shinozaki et al. (1986) and Palmer (1991). Part of this modification includes the duplication of probe 28 and part of probe 27 to show tobacco's hypothetical alignment with the expanded IR in *Anemone*.

three *Anemone* species. Introns are quite stable components of land-plant chloroplast genomes; only a few other cases of intron loss have been described (Downie et al. 1991; Palmer 1991; Downie and Palmer 1992). This is the first such case known to have occurred at the intragenetic level.

Extension of the Inverted Repeat

Mapping experiments using 40 clones spanning the entire tobacco chloroplast genome showed that the inverted repeat in *Caltha*, *Ranunculus*, and *Adonis* is indistinguishable in size from that of tobacco (Shinozaki 1986). An extension of the inverted repeat to a length

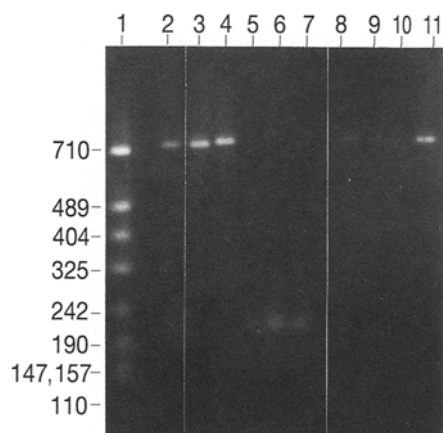


Fig. 2. Absence of *rps12* intron 2 from the *A. canadensis* clade. PCR amplification was performed as described in Materials and Methods, yielding products of ~750 bp in taxa containing intron 2 (lanes 2–4 and 8–11) and of ~210 bp in taxa completely lacking this intron (lanes 5–7). Lane 1, *Hpa*II digest of plasmid pBluescript SK+ (Stratagene); lane 2, tobacco (*Nicotiana tabacum*); lane 3, soybean (*Glycine max*); lane 4, maize (*Zea mays*); lane 5, *Anemone richardsonii*; lane 6, *A. canadensis*; lane 7, *A. dichotoma*; lane 8, *A. vitifolia*; lane 9, *A. caffra*; lane 10, *Knowltonia vesicatoria*; lane 11, *Pulsatilla vulgaris*.

of approximately 29.4 kb was found in *Anemone*, *Clematis*, *Hepatica*, *Knowltonia*, and *Pulsatilla* (Table 1, Fig. 1). The endpoints of this extension were more precisely located by probing with two additional clones, as illustrated by the hybridizations to *Bgl*II digests of the three *Anemone* DNAs shown in Fig. 3. Tobacco probe 28 (TCU 83807–85632), which is entirely single copy in tobacco, hybridizes equally strongly to two *Bgl*II fragments of 8.6 and 7.0 kb (Fig. 3, lanes 7–9). A 283-bp tobacco fragment, TCU 82469–82752, hybridizes to these same two *Bgl*II fragments, but more strongly to the larger of the two (Fig. 3, lanes 4–6). That two different probe fragments both hybridize to these two different *Bgl*II fragments implies the existence of a shared repeat element on these two *Bgl*II fragments; that the 284-bp fragment hybridizes differentially to these two fragments implies that the repeat ends within this probe region. This inference is supported by the observation that tobacco fragment TCU 80652–82469, which is adjacent to the 283-bp fragment, hybridizes exclusively to the larger *Bgl*II fragment (Fig. 3, lanes 1–3). Thus, assuming that it ends near the middle of the 283-bp fragment, the *Anemone* inverted repeat is extended by approximately 4.1 kb relative to the repeat in tobacco (which ends at TCU 86684) and most other angiosperms (Palmer 1991). Assuming conservation of gene content between *Anemone* and tobacco, this implies duplicate copies in *Anemone* of six ribosomal protein genes, *rps8*, *rpl14*, *rpl16*, *rps3*, *rpl22*, and *rps19*, that are single copy in *Caltha*, *Ranunculus*, and *Adonis* (Fig. 1), as well as in tobacco and most other angiosperms (Palmer 1991).

Inversions

Excluding differences relating to the expansion of the inverted repeat, four different gene orders were found among the chloroplast genomes of the taxa in this study. *Caltha* and *Ranunculus* have the same chloroplast gene order as tobacco and most other vascular plants; this corresponds to the primitive gene order found within seed plants (Palmer 1991; Raubeson and Jansen 1992). Each of the other three gene orders can be derived from this ancestral, tobaccolike gene order by postulating between one and six inversions, whose endpoints, sizes, and distribution are summarized in Table 1. The first gene order is that found in *Adonis aestivalis*, which has a single inversion of 41.9 kb between endpoints 2 and 7. This same inversion was found for another species, *Adonis annua*, in a restriction-site study of the Ranunculaceae (Johansson and Jansen 1993). But a second inversion of 24 kb (endpoints at TCU 48.6 and 73.0) found in *A. annua* was not detected in our study. The second gene order, found in virtually all species within the *Anemone* complex and *Clematis*, results from four inferred inversions (Table 1). The third gene order, found in both *Clematis* species examined and in *Anemone canadensis*, *A. dichotoma*, and *A. richardsonii*, results from the same four inversions mentioned above plus two additional inversions.

Figures 1 and 4 illustrate the location of the inversions for all species except *Adonis* in relation to the tobacco genome. We cannot logically infer any temporal order among the four inversions found in the *Anemone* complex and *Clematis* (the second gene order mentioned above). There are 24 possible orders, only one of which is illustrated in Fig. 4. Based on gene order and phylogenetic evidence, inversions 5 and 6 occurred after inversions 1–4, but temporal order cannot be distinguished for inversions 5 and 6.

No matter which temporal order is hypothesized for the various inversions, the same endpoints are used, but in different combinations. Table 1 gives only the endpoint combinations that match the inversion order illustrated in Fig. 4. The locations of the seven inversion endpoints (numbered as in Fig. 4) with respect to tobacco genes and their intergenic spacer lengths are as follows: Endpoint 1 maps in the 454-bp spacer between *trnH* and *psbA*; endpoint 2 maps in the 861-bp region between *rps16* (5' exon) and *trnQ*; endpoint 3 probably occurs in the 612-bp spacer surrounded by *psbI* and *trnS*, but could be within or just before *psbI*; endpoint 4 occurs in the 780-bp spacer bordered by *trnS* and the *trnG*; endpoints 5 is located in the 331-bp spacer between *trnS* and *rps4*; endpoint 6 maps between *trnT* and *trnL*; and endpoint 7 is found in the 1088-bp spacer between *ndhC* and *trnV*, apparently very close to *trnV*. Three of the inversion endpoints (2–4) occupy the approximately 3.3 kb region flanked by *rps16* and *trnG*.

Table 1. Major structural rearrangements in the chloroplast genomes of *Anemone* and related genera

Mutation	Endpoints ^a		Size	Taxa ^b	
Extension of IR into LSC region	82.5–112.0 (IR-B) 130–0.4 (IR-A)		4.1 kb (total IR size = 29.4 kb)	A,C,H,K,P	
Endpoint					
		1	2		
Inversions ^c	1	0.4	48.6	48.2	A,C,H,K,P
	2	8.6	47.4	38.8	A,C,H,K,P
	3	8.6/47.4	0.4/48.6	47.0	A,C,H,K,P
	4	0.4/47.4	6.7	6.3	A,C,H,K,P
	5	8.6/47.4	53.6	43.8	Ac,Ad,Ar,C
	6	8.6/53.6	9.2	5.6	Ac,Ad,Ar,C
	7	6.7	48.6	41.9	<i>Adonis</i>

^a Endpoint coordinates of rearrangement events are presented in tobacco coordinate units (TCUs) and match with the inversion order illustrated in Fig. 4

^b A = *Anemone* (all taxa examined), C = *Clematis* (2 species), H = *Hepatica*, K = *Knowltonia*, P = *Pulsatilla* (4 species), Ac = *Anemone canadensis*, Ad = *A. dichotoma*, Ar = *A. richardsonii*

^c The numbering of inversions 1–4 and inversions 5 and 6 do not reflect temporal order

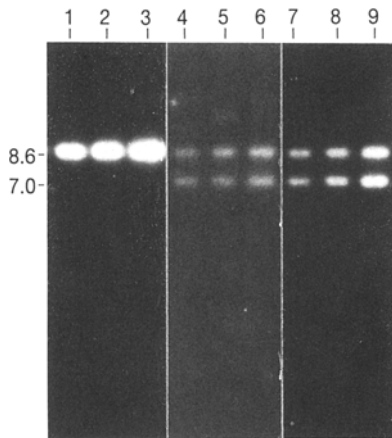


Fig. 3. Mapping the large single-copy boundary of the extended repeat in *Anemone*. Autoradiograms resulting from *Bgl*III restriction enzyme digest and subsequent hybridization: Probe TCU 80652–82469 (lanes 1–3) hybridized to only a single fragment; probes TCU 82469–82752 (lanes 4–6) and TCU 83807–85632 (lanes 7–9) each hybridize to the same two repeated fragments. Lanes 1, 4, and 7—*Anemone multifida*; lanes 2, 5, and 8—*A. canadensis*; lanes 3, 6, and 9—*A. dichotoma*.

Discussion

Phylogeny and Rearrangements

The phylogenetic information available from the rearrangements found in the present study is of limited but, in most cases, convincing value. At the generic level, the extension of the inverted repeat into the large single-copy region and the first four inversions (Table 1) reinforce the monophyletic status of *Clematis* and the *Anemone* complex, clearly separating these genera from *Adonis*, *Caltha*, and *Ranunculus*. This is corroborated by cpDNA restriction-site variation (Fig. 5).

All anemones and the genera *Clematis*, *Hepatica*, *Pulsatilla*, and *Knowltonia* share the first four inversions (Table 1). For this reason, these inversions are not useful taxonomic markers at the sectional level. These four inversions must have occurred early in the history of the *Anemone* complex, before diversification had occurred.

Two additional inversions separate *Clematis* (two species), *Anemone canadensis*, *A. dichotoma*, and *A. richardsonii* from all other anemones (Table 1). The relationship suggested by these inversions is not congruent with the cladograms resulting from chloroplast DNA restriction sites (Fig. 5; Hoot et al. 1994). In the restriction site trees (rooted or unrooted), the *A. canadensis* clade is separated from the *Clematis* species by a number of characters and by three branches of moderate (70%) to high (98%) bootstrap support (Fig. 5). To force *Clematis* and the *Anemone canadensis* clade together as a monophyletic group costs an additional 16 steps in tree length.

In addition, the morphology of *Clematis* is very different from *Anemone*. The only features shared by *Clematis* and members of the *A. canadensis* clade are opposite leaves (*A. dichotoma*) and styles completely covered by hairs (*A. canadensis*). Neither of these are convincing characters, given their sporadic occurrence in the clade and their homoplastic recurrence in distantly related anemones (Hoot et al. 1994). *Clematis* also differs from *Anemone* in base chromosome number, an important systematic character in the *Anemone* complex (Hoot et al. 1994).

In light of the above, these appear to be the first two documented cases of parallel or homoplastic inversions in cpDNA—i.e., inversions 5 and 6 both appear to have occurred twice, once in the *Clematis* lineage and once in the *A. canadensis* lineage. Considering the robust topology of Fig. 5, a scenario of parallel inversions is

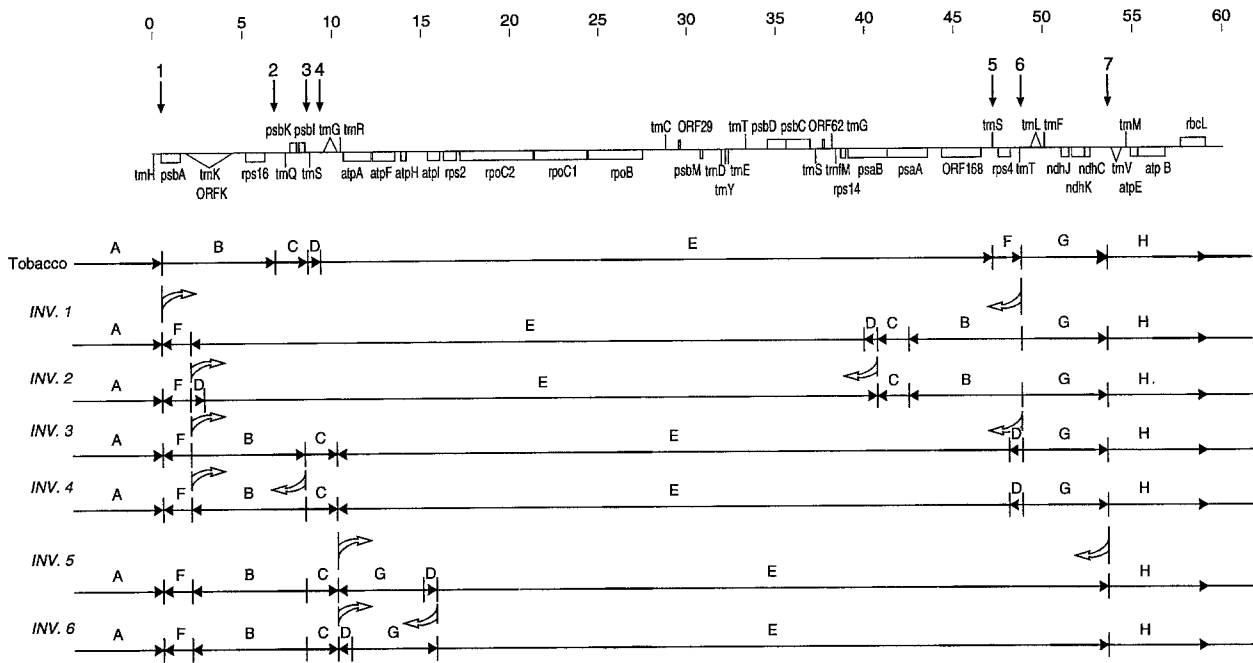


Fig. 4. Maps of a portion of the large single-copy region for tobacco and *Anemone* species. The gene map for tobacco is modified from Shinozaki et al. (1986). The scale is TCUs \times 1,000. The locations of the inversion endpoints are indicated by the numbered vertical arrows. The tobacco genome has been broken up into 12 lettered fragment

blocks to illustrate the changes in order. Actual inverted regions are indicated by the large curved arrows. The orientation of the fragment blocks is indicated by arrows. Only one of 24 possible orders is shown for inversions 1–4. Likewise, our data do not distinguish between the two possible orders of inversions 5 and 6.

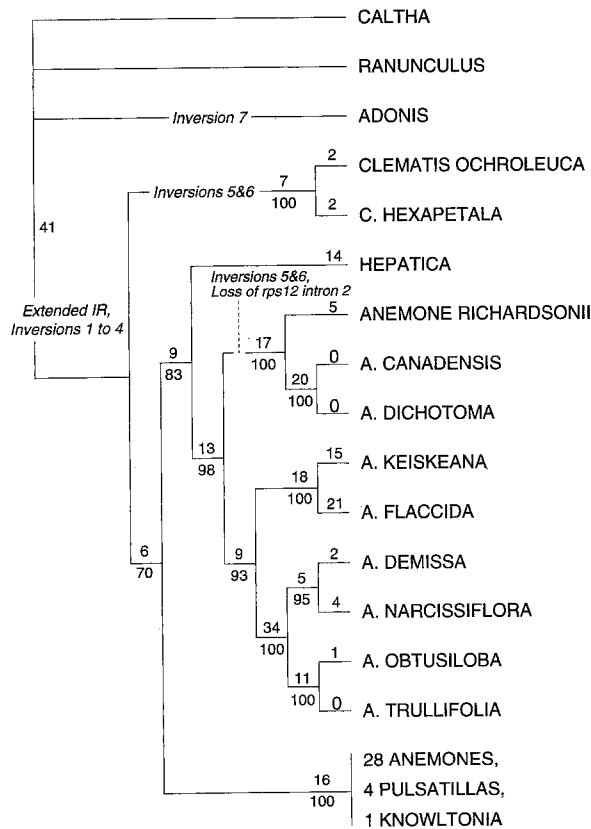


Fig. 5. Chloroplast DNA restriction site tree. C.I. (excluding non-informative data) = 0.69, R.I. = 0.91. Numerals above lines indicate the number of restriction sites supporting each branch. Numbers below lines indicate the bootstrap values for 100 replications.

to be preferred on parsimony grounds over one in which inversions 5 and 6 occurred earlier in a common ancestor of *Clematis* and the entire *Anemone* complex (i.e., contemporaneously with inversions 1–4), followed by no fewer than three independent reversals (reversions). According to this parallel inversion model, the total number of inversion events for *Adonis*, *Clematis*, and the *Anemone* complex is nine. More detailed fine structural mapping and sequencing of the pertinent regions of the genome for relevant taxa may demonstrate that the parallel inversions do not share common endpoints, confirming the lack of homology. These results indicate the importance of precise mapping and/or sequencing of endpoint regions to assess homology in phylogenetic studies utilizing inversion data. Homoplasy in inversions has also recently been found for single inversions in Cactaceae and *Chenopodium* (Downie and Palmer 1994) and in *Oenothera* and legumes (Palmer and Thompson 1982; K. Sytsma, pers. comm.).

The large number of inversions within the *Anemone* complex and *Clematis* suggest that the occurrence of the various inversions may be causally linked. For example, in considering the two homoplastic inversions, it could be that once the fifth inversion occurred, the probability of the sixth was greatly increased. The fifth inversion may have positioned repeated sequences in the genome in such a way as to facilitate the sixth inversion. Therefore, the two inversions may not constitute two totally independent events. This same explanation is also

applicable when considering inversions 1–4 within the *Anemone* complex and *Clematis* (see Inversion Characteristics below).

Accepting that inversions 5 and 6 are homoplastic, they are important independent phylogenetic markers uniting the clade consisting of *Anemone canadensis*, *A. dichotoma*, and *A. richardsonii*. This clade is strongly united by the loss of *rps12* intron 2, 17 restriction-site mutations, and a bootstrap value of 100% (Fig. 5). *Anemone canadensis* and *A. dichotoma* have long been recognized as sister species with disjunct distributions between North America and Asia. In contrast, *A. richardsonii* differs markedly in morphology from the other two anemones and had been placed in a different section by previous workers (Hoot et al. 1994).

Inversion Characteristics

Most seed-plant chloroplast genomes retain an ancestral gene order such as that found in tobacco (Palmer et al. 1988b; Palmer 1991; Downie and Palmer 1992; Raubeson and Jansen 1992); the large number of inversions found in *Anemone* and closely related genera is unusual. Other taxa reported with large numbers of inversions are the conifers (Strauss et al. 1988), grasses (Quigley and Weil 1985; Howe et al. 1988; Hiratsuka et al. 1989), legumes (Palmer et al. 1988a; Milligan et al. 1989), geraniums (Palmer et al. 1987a), Lobeliaceae (Knox et al. 1993), and Campanulaceae (M.B. Cosner, pers. comm.). Because of their rarity, these rearrangements often helped in elucidating taxonomic relationships.

Many of the same *Anemone* inversion endpoints have been reused in multiple inversions. These nine inversions appear to use only seven different endpoints out of the 18 endpoints that are numerically possible (Table 1). Including the homoplastic inversions, no endpoint is found in just one inversion. Endpoints 3 and 5 seem to have been used in six and five different inversions, respectively (including the homoplastic inversions).

Many of the inversions found in *Anemone* have endpoints in regions similar to those found in distantly related taxa. A number of inversion endpoints from a variety of taxa (*Lobelia*, pea, mung bean, *Oenothera*, grasses, Douglas fir, and Asteraceae) occur near *trnS*, *trnG*, and *trnR* (Knox et al. 1993). *Adonis* and all members of the *Anemone* clade have endpoints in this same region (endpoints at TCU 8.6 and 9.2). Another inversion “hot spot” occurs in the region bounded by *trnS* and *trnM*. In addition to three *Anemone* endpoints (TCU 47.4, 48.6, and 53.6), broadbean (Palmer et al. 1987b), *Trachelium* (M.B. Cosner, unpubl. data), *Chenopodium* and Cactaceae (Downie and Palmer 1994), and lobelias (Knox et al. 1993) have endpoints in this region. Lobelias and anemones have endpoints in four of the same spacer regions (located by *Anemone* endpoints 4–7, Fig. 5).

The large number of common inversion endpoints among diverse plant lineages and the two likely homoplastic inversions in the *Anemone* complex suggest a common causal mechanism. It has been proposed that multiple inversions in chloroplast genomes are probably caused by recombination between indirect repeats bordering the inverted regions (Palmer 1991). Repeated sequences, which are generally absent or rare in cpDNA, have been reported in a number of taxa with multiple inversions: Conifers (Strauss et al. 1988), grasses (Quigley and Weil 1985; Howe et al. 1988; Moon et al. 1988; Hiratsuka et al. 1989), legumes (Palmer et al. 1988a; Milligan et al. 1989; Bruneau et al. 1990), and geraniums (Palmer et al. 1987a).

In addition, it has been noted that inversion endpoints are often associated with short repeated sequences near or within tRNA genes (Palmer 1991). All *Anemone* endpoints are next to at least one tRNA gene; endpoint 4 is bordered by two tRNA genes. However, the role of short repeated elements and their association with tRNA genes in chloroplast genome inversions is unlikely to be elucidated by further study of the genome structure and base sequences found in *Anemone* and related genera. There have been no additional inversions since the divergence of *Anemone* and *Clematis*. During this period, both genera have undergone great morphological differentiation and speciation as well as divergence at the cpDNA sequence level (Fig. 5; Hoot et al. 1994). This suggests that the original cause of the inversions, possibly indirect repeats, may no longer be present.

All of the inversion endpoints found in this study appear to map to spacer regions between genes that are transcribed either singly or as parts of operonlike gene clusters. With the exception of two rearrangements in legumes (Palmer et al. 1988a; Milligan et al. 1989), all mapped cpDNA inversions have their endpoints in spacer regions between transcriptional units, thus avoiding disruption of gene functions.

Acknowledgments. We thank Anton A. Reznicek, William R. Anderson, and Warren H. Wagner for critical readings of the manuscript and helpful suggestions, Kathy Scott for assistance with the *rps12* intron analysis, and numerous individuals, the Edinburgh Botanical Garden, and other botanical gardens for providing leaf material. Thanks also to the Horace H. Rackham School of Graduate Studies for facilitating this research with a Rackham Predoctoral Fellowship awarded to S.B.H. in 1989–1990. This work was supported in part by NSF grant BSR-8717600 to J.D.P.

References

- Bruneau A, Doyle JJ, Palmer JD (1990) A chloroplast DNA inversion as a subtribal character in the Phaseoleae (Leguminosae). *Syst Bot* 15:378–386
- Collins ML, Hunsaker WR (1985) Improved hybridization assays employing tailed oligonucleotide probes: a direct comparison with

- 5'-end-labeled oligonucleotide probes and nick-translated plasmid probes. *Anal Biochem* 151:211–224
- Downie SR, Palmer JD (1992) 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, p 14–35
- Downie SR, Palmer JD (1994) A chloroplast DNA phylogeny of the Caryophyllales based on structural and inverted repeat restriction site variation. *Syst Bot* 19(2)
- Downie SR, Olmstead RG, Zurawski G, Soltis DE, Soltis PS, Watson JC, Palmer JD (1991) Six independent losses of the chloroplast DNA *rp/2* intron in dicotyledons: molecular and phylogenetic implications. *Evolution* 45:1245–1259
- Doyle JJ, Davis GI, 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
- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh tissue. *Phytochem Bull* 19:11–15
- Giese K, Subramanian AR, Larrinua IG, Bogorad L (1987) Nucleotide sequence, promoter analysis, and linkage mapping of the unusually organized operon encoding ribosomal proteins S7 and S12 in maize chloroplast. *J Biol Chem* 262:15251–15255
- Gregory WC (1941) Phylogenetic and cytological studies in the Ranunculaceae Juss. *Trans Am Phil Soc, NS* 31:441–520
- 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 plastic DNA inversion during the evolution of the cereals. *Mol Gen Genet* 217:85–194
- Hoot SB (1991) Phylogeny of the Ranunculaceae based on epidermal microcharacters and macromorphology. *Syst Bot* 16:741–755
- Hoot SB, Reznicek AA, Palmer JD (1994) Phylogenetic relationships in *Anemone* (Ranunculaceae) based on morphology and chloroplast DNA. *Syst Bot* 19:172–203
- Howe CJ, Barker RF, Bowman CM, Dyer TA (1988) Common features of three inversions in wheat chloroplast DNA. *Curr Genet* 13:343–349
- Jansen RK, Palmer JD (1987) A chloroplast DNA inversion marks an ancient evolutionary split in the sunflower family (Asteraceae). *Proc Natl Acad Sci USA* 84:5818–5822
- Johansson JT, Jansen RK (1993) Chloroplast DNA variation and phylogeny of the Ranunculaceae. *Plant Syst Evol* 187:29–49
- Knox E, Downie SR, Palmer JD (1993) Chloroplast genome rearrangements and the evolution of giant lobelias from herbaceous ancestors. *Mol Biol Evol* 10:414–430
- Kuhnel MG, Strickland R, Palmer JD (1990) An ancient group I intron shared by eubacteria and chloroplasts. *Science* 250:1570–1573
- Milligan BG, Hampton JN, Palmer JD (1989) Dispersed repeats and structural reorganization in subclover chloroplast DNA. *Mol Biol Evol* 6:355–368
- Moon E, Kao T-H, Wu R (1988) Rice mitochondrial genome contains a rearranged chloroplast gene cluster. *Mol Gen Genet* 213:247–253
- Olmstead R, Palmer JD (1992) A chloroplast DNA phylogeny of the Solanaceae: subfamilial relationships and character evolution. *Ann Missouri Bot Gard* 79:346–360
- Palmer JD (1991) Plastid Chromosomes: Structure and Evolution. In: Bogorad L, Vasil IK (eds) *Cell culture and somatic cell genetics of plants 7A*. Academic, San Diego, p 5
- Palmer JD, Thompson WF (1982) Chloroplast DNA inversions are more frequent when a large inverted repeat is lost. *Cell* 29:537–550
- Palmer JD, Nugent JM, Herbon LA (1987a) 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 (1988a) Evolutionary significance of inversions in legume chloroplast DNAs. *Curr Genet* 14:65–74
- Palmer JD, Osorio B, Aldrich J, Thompson WF (1987b) Chloroplast DNA evolution among legumes: loss of a large inverted repeat occurred prior to other sequence rearrangements. *Curr Genet* 11:275–286
- Palmer JD, Jansen RK, Michaels HJ, Chase MW, Manhart JR (1988b) Chloroplast DNA variation and plant phylogeny. *Ann Missouri Bot Gard* 75:1180–1206
- 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) Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. *Science* 255:1697–1699
- 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 *Osunda*. *Curr Genet* 10:835–841
- Strauss SH, Palmer JD, Howe GT, Doerksen AH (1988) Chloroplast genomes of two conifers lack a large inverted repeat and are extensively rearranged. *Proc Natl Acad Sci USA* 85:3898–3902
- Swofford DL (1990) PAUP: Phylogenetic analysis using parsimony, version 3.0. Illinois Natural History Survey, Champaign, IL
- Tamura M (1967) Morphology, ecology and phylogeny of the Ranunculaceae VII. *Sci Repts Osaka Univ* 17:41–56
- von Allmen JM, Stutz E (1987) Complete sequence of 'divided' *rps12* (r-protein S12) and *rps7* (r-protein S7) gene in soybean chloroplast DNA. *Nucleic Acids Res* 15:2387