2

Use of Chloroplast DNA Rearrangements in Reconstructing Plant Phylogeny

Stephen R. Downie and *Jeffrey D. Palmer*

Reconstructing phylogenies among genera and at higher taxonomic levels always has been fraught with difficulties. Conventional plant classifications employ a diverse array of approaches (phytochemical, anatomic, morphologic, etc.) and often offer a synthesis of these data sets. Many of these traditional characters are susceptible to convergent evolution by natural selection; the ensuing homoplasy largely precludes robust phylogenies. Only recently have we been able to examine the genetic material itself to investigate phylogenetic relationships. Chloroplast DNA (cpDNA) variation has proven to be immensely valuable in reconstructing phylogenies at the species level, and the application of cpDNA comparisons at higher taxonomic levels is now being pursued actively.

One approach to extracting phylogenetic information from cpDNA is by analyzing the distribution of major structural rearrangements. Because of their infrequent occurrence, rearrangements usually can provide strong evidence of monophyly. In this chapter, we demonstrate the utility and significance of cpDNA rearrangements in reconstructing plant phylogeny. After an introduction to the salient features of the chloroplast chromosome, we briefly review the approaches used to detect and analyze rearrangements and discuss our current survey to detect and circumscribe cpDNA rearrangements among angiosperms. We will then examine the different classes of rearrangements and for each provide examples of their use in phylogenetic reconstruction. This chapter will deal exclusively with land plants. Algal genomes may have different structural dynamics than land plant genomes and are too poorly characterized to warrant discussion of their phylogenetic utility at the present time.

This research was supported by NSF Grant BSR-8996262 to JDP and a Natural Sciences and Engineering Research Council (NSERC) of Canada Postdoctoral Fellowship to SRD. We thank Richard G. Olmstead and Claude W. dePamphilis for critical reading of the manuscript and the many scientists who have generously provided us with DNAs to include in our survey of angiosperm cpDNA rearrangements.

The Chloroplast Chromosome

The chloroplast genomes of photosynthetic land plants are circular DNA molecules ranging in size from 120 to 217 kilobase pairs (kb) (Table 2.1). A list of land plants for which complete restriction endonuclease cleavage site maps of plastid genomes have become available since 1985 is presented in Table 2.1. This tabulation updates the previous compilation by Crouse et al. (1985). The chloroplast genome contains, with few exceptions, two duplicate regions in reverse orientation, known as the inverted repeat (IR). These repeated regions separate the remainder of the molecule into large single-copy (LSC) and small single-copy (SSC) regions (Fig. 2.1). The expansion or contraction of the IR into, or out of, adjacent single-copy regions, and changes in sequence complexity due to insertions or deletions of unique sequences are largely responsible for variation in size of the molecule. Several reviews of cpDNA structure, function, and evolution have been published recently (Whitfeld and Bottomley, 1983; Palmer, 1985a, 1985b, 1991; Zurawski and Clegg, 1987; Sugiura, 1989).

Recent studies of chloroplast genome evolution have revealed a high degree of conservation in size, structure, gene content, and linear order of genes among major lineages of land plants (Palmer, 1985b, 1991; Palmer and Stein, 1986). This conservative mode of cpDNA evolution suggests that any change in structure, arrangement, or content of the chloroplast genome may have significant phylogenetic implications.

Mutations in cpDNA are of two kinds: nucleotide substitutions (point mutations) and rearrangements. The detection of nucleotide substitutions through restriction site analysis or direct comparisons of homologous sequences currently is used widely in phylogenetic reconstruction (e.g., Palmer et al., 1988a, and other chapters in this volume). Major rearrangements of the chloroplast molecule include inversions, the insertion or deletion of genes and introns, and loss of one copy of the IR. Minor rearrangements consist of small insertions and deletions (1–1,000 bp). These events are much more common than major rearrangements and occur principally in noncoding intergenic spacer regions and introns (Palmer, 1985b). Because small length mutations have a tendency to cluster in "hot spot" regions (Kung et al., 1982; Palmer et al., 1988a), the assignment of exact homology for each mutation may be difficult. Owing to their homoplasious nature, they are difficult to use for systematic purposes, and are sometimes not included in phylogenetic analyses (Sytsma and Gottlieb, 1986; Palmer et al., 1988a), although in studies of closely related taxa they may provide useful information (Doebley et al., 1987a; Soltis et al., 1989, 1990).

Detection and Analysis of Rearrangements

Chloroplast DNA rearrangements most often are revealed using a heterologous filter hybridization approach in which cloned restriction fragments from one chloro-

*Table 2.1. Land plant species for which plastid genome size (in kb) and complete restriction endonuclease maps are available.**

Taxon	Size	Reference
Angiosperms[a]		
Asteridae		
Apocynaceae		
Vinca minor	150	Palmer (unpublished)
Asteraceae		
16 tribes, 267 genera	148–151	Jansen et al. (this volume)
16 tribes, 57 genera, 57 spp.	148–151	Jansen et al. (1990)
Heliantheae (6 genera, 33 spp.)	151	Schilling and Jansen (1989)
Mutisieae (13 genera, 13 spp.)	151	Jansen and Palmer (1988)
Madiinae (5 genera, 26 spp.)	151	Baldwin et al. (1990)
Barnadesia caryophylla	151	Jansen and Palmer (1987a)
Carthamus tinctorius	151	Ma and Smith (1985)
Helianthus (7 spp.)	152	Heyraud et al. (1987)
Lactuca sativa	151	Jansen and Palmer (1987a)
Dipsacaceae		
Scabiosa sp.	154	Palmer (unpublished)
Orobanchaceae		
Epifagus virginiana[b]	71	dePamphilis and Palmer (1989)
Plantaginaceae		
Plantago sp.	144	Palmer (unpublished)
Solanaceae[c]		
Capsicum annuum	143	Gounaris et al. (1986)
Solanum tuberosum	155	Heinhorst et al. (1988)
Caryophyllidae		
Caryophyllaceae		
Cerastium arvense	147	Palmer (unpublished)
Chenopodiaceae		
Beta (11 spp.)	148	Kishima et al. (1987)
Phytolaccaceae		
Phytolacca heterotepela	153	Palmer (unpublished)
Polygonaceae		
Rumex sp.	158	Palmer (unpublished)
Dilleniidae		
Actinidiaceae		
Actinidia deliciosa	160	Hudson and Gardner (1988)
Begoniaceae		
Begonia sp.	170	Palmer (unpublished)
Clusiaceae		
Hypericum sp.	140	Palmer (unpublished)
Cucurbitaceae		
Cucumis (21 spp.)	150	Perl-Treves and Galun (1985)
Primulaceae		
Anagallis arvensis	148	Palmer (unpublished)
Primula sp.	149	Palmer (unpublished)
Salicaceae		
Populus (10 spp.), *Salix exigua*	155	Smith and Sytsma (1990)
Hamamelidae		
Urticaceae		
Pilea microphylla	150	Palmer (unpublished)

(*continued*)

Table 2.1.
(Continued)

Taxon	Size	Reference
Liliidae		
Alliaceae		
Allium cepa	145	Chase and Palmer (1989)
Asparagaceae		
Asparagus sprengeri	149	Chase and Palmer (1989)
Amaryllidaceae		
Narcissus X hybridus	157	Chase and Palmer (1989)
Narcissus pseudonarcissus	161	Hansmann (1987)
Dioscoreaceae		
Dioscorea bulbifera	152	Terauchi et al. (1989)
Orchidaceae		
Oncidium (9 spp.), *Psychopsis sanderae*, *Rossioglossum schlieperianum*, *Trichocentrum* (2 spp.)	143	Chase and Palmer (1989)
Commelinidae		
Poaceae		
Aegilops (10 spp.)	135	Ogihara and Tsunewaki (1988)
Avena (5 spp.)	135	Murai and Tsunewaki (1987)
Oryza sativa	135	Hiratsuka et al. (1989)
Secale (5 spp.)	136	Murai et al. (1989)
Sorghum bicolor	138	Dang and Pring (1986)
Triticum (10 spp.)	135	Ogihara and Tsunewaki (1988)
Zea (4 spp.), *Tripsacum dactyloides*	135	Doebley et al. (1987b)
Magnoliidae		
Aristolochiaceae		
Aristolochia durior	158	Palmer (unpublished)
Papaveraceae		
Eschscholzia californica	158	Palmer (unpublished)
Ranunculaceae		
Aquilegia sp.	151	Palmer (unpublished)
Delphinium sp.	149	Palmer (unpublished)
Ranunculus californica	157	Palmer (unpublished)
Rosidae		
Aceraceae		
Acer pseudoplatanus	142	Ngernprasirtsiri and Kobayashi (1990)
Apiaceae		
Coriandrum sativum	148	Palmer (unpublished)
Balsaminaceae		
Impatiens sp.	156	Palmer (unpublished)
Crassulaceae		
Sedum oreganum	146	Sundberg et al. (1990)
Fabaceae		
Lupinus polyphyllus	147	Palmer et al. (1987b)
Medicago sativa	126	Palmer et al. (1987b)
Trifolium subterraneum	142	Milligan et al. (1989)
Wisteria floribunda	130	Palmer et al. (1987b)
Geraniaceae		
Pelargonium X hortorum	217	Palmer et al. (1987a)

(*continued*)

Table 2.1.
(Continued)

Taxon	Size	Reference
Hippocastanaceae		
Aesculus californica	153	Palmer (unpublished)
Linaceae		
Linum grandiflorum	154	Palmer (unpublished)
Linum (8 spp.)	160, 173	Coates and Cullis (1987)
Onagraceae		
Clarkia (14 spp.)	170	Sytsma et al. (1990)
Fuchsia sp.	151	Palmer (unpublished)
Fuchsia (6 spp.)	150	Sytsma et al. (in press)
Oxalidaceae		
Oxalis oregana	152	Palmer (unpublished)
Rutaceae		
Citrus (7 spp.), *Poncirus trifoliata*, *Microcitrus* sp.	166	Green et al. (1986)
Saxifragaceae		
10 genera, 40 spp.	151	Soltis et al. (1990; unpublished)
Gymnosperms		
Ginkgoaceae		
Ginkgo biloba	158	Palmer and Stein (1986)
Pinaceae		
Pinus monticola	120	White (1990)
Pinus radiata	120	Strauss et al. (1988)
Pseudotsuga menziesii	120	Strauss et al. (1988)
Pteridophytes		
Adiantum capillus-veneris	153	Hasebe and Iwatsuki (1990)
Osmunda (3 spp.)	144	Stein et al. (1986)
Bryophytes		
Marchantia polymorpha	121	Ohyama et al. (1986)
Physcomitrella patens	122	Calie and Hughes (1987)

[*]Information not previously compiled by Crouse et al. (1985), who listed mapped genomes for 16 families, 29 genera, and 32 species of land plants, is presented. Unless otherwise stated (see text and other tables presented herein) all maps are colinear with that of *Nicotiana tabacum*.

[a]Restriction mapping data are currently being analyzed for 99 families (211 spp.) of angiosperms including 36 families from the Asteridae (Downie and Palmer, unpublished data). To date, 40 families (71 spp.) have been found whose cpDNA genomes are colinear with that of *Nicotiana tabacum*. Those taxa whose cpDNA genomes possess rearrangements are listed in Tables 2.2 and 2.3.

[b]*Epifagus virginiana* is nonphotosynthetic.

[c]A comprehensive comparative restriction mapping analysis of the Solanaceae (55 genera, 132 species) is currently underway (Olmstead and Palmer, unpublished data).

plast genome are hybridized to filter-bound fragments of a second. Defined segments of cloned restriction fragments obtained from mapped reference genomes, such as that of *Nicotiana tabacum*, are used as hybridization probes. Heterologous probes can be used effectively across widely divergent lineages of angiosperms because most cpDNAs are highly conserved in sequence and arrangement.

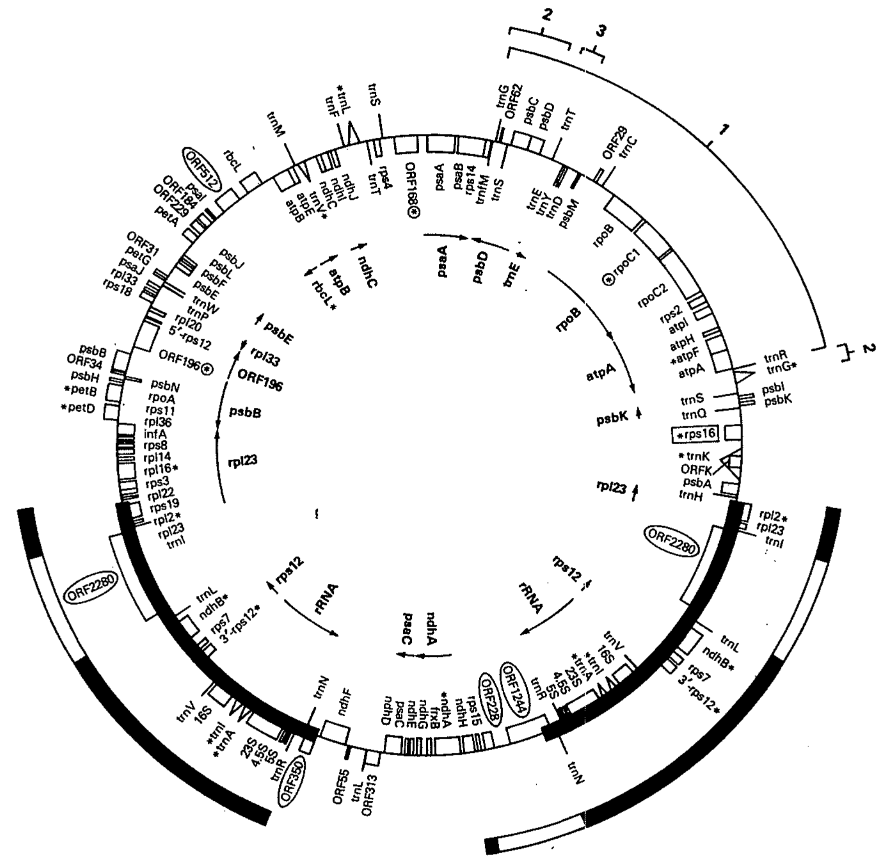

Figure 2.1. Physical and gene map of the 156-kb *Nicotiana tabacum* chloroplast genome, showing selected rearrangements found in *Oryza sativa* and *Marchantia polymorpha*. Genes transcribed clockwise are shown on the inside of the circle; those transcribed counterclockwise are on the outside. Arrows on the inside of the circle indicate sets of genes thought to constitute operons. The operon names are indicated in boldface. Circled gene names indicate genes present in *N. tabacum* and *M. polymorpha* but absent from *O. sativa*. The boxed gene name indicates a gene present in *N. tabacum* and *O. sativa* but absent in *M. polymorpha*. Asterisks denote genes that have the same intron(s) in all three sequenced genomes; circled asterisks denote genes that are split in the three genomes with two exceptions (ORF196 and *rpo*C1 are not split in *O. sativa*, see Table 2.3). The thickened parts of the circle represent the 25.3-kb IR of *N. tabacum*. The thick lines outside of the *N. tabacum* IR represent the extent of the *O. sativa* IR (20.8 kb), with regions deleted from *O. sativa* left open. The numbered brackets outside the circle indicate three overlapping inversions in *O. sativa* relative to *N. tabacum*. Gene nomenclature follows Hallick and Bottomley (1983), Hallick (1989), and Palmer (1991). *Nicotiana tabacum* cpDNA map is based on Shinozaki et al. (1986); *O. sativa* and *M. polymorpha* data are based on Hiratsuka et al. (1989) and Ohyama et al. (1986), respectively.

Currently, we are investigating chloroplast genome structure in angiosperms with emphasis on the subclass Asteridae. Our survey for major structural rearrangements encompasses 88 species, representing 36 families of Asteridae (see Cronquist, 1981), and an additional 123 species comprising representatives from all subclasses of monocotyledons and dicotyledons. From an initial set of large, cloned restriction fragments (Sugiura et al., 1986), we have subcloned small fragments specific for many of the genes and introns found in the cpDNA of *N. tabacum*. To date we have used over 120 hybridization probes, ranging in size from 0.2 to 3.5 kb, which together comprise the entire *N. tabacum* chloroplast genome. These probes are smaller than the 5-to-15-kb probes commonly used in restriction site mapping studies. The use of small probes permits the detection of small rearrangement events that often are undetected when larger fragments are used as probes. Our methods are adapted primarily to survey, as rapidly and as economically as possible, large numbers of taxa for rearrangements. Once a mutation is detected and characterized at the molecular level it can, depending upon the availability of material, be circumscribed in related taxa.

Our survey for rearrangements employs a modification of the approach outlined in Palmer et al. (1988a) to accommodate both increased numbers of taxa and probes. To minimize the number of hybridization cycles, each cpDNA sample is run in triplicate. To accomplish this, triple-size restriction enzyme digests are prepared initially, then one-third of each digest is loaded on each gel. Double-sided blotting yields six identical filters, which greatly reduces the amount of time necessary to complete the required number of hybridizations. To lessen the size and cost of such an undertaking, care is taken so that the bromophenol blue dye marker in each of the gels migrates no more than 6 cm. In this way, four 20-cm-wide filters can be placed on a single sheet of standard 20 × 25-cm x-ray film. Resolution is sufficient to detect gene and intron losses, inversions, and changes in size of the IR, but often it is difficult to map inversions precisely and to measure fragment sizes accurately. When this becomes necessary, follow-up gels can be run in which the dye marker migration is 12–20 cm.

Rearrangements can be detected by arranging the autoradiograms according to the order in the chloroplast genome of the hybridization probes and by observing both fragment number and size as one "walks" along the chloroplast chromosome from one hybridization probe to the next. Any anomaly in the number of fragments detected, their size, or the intensity of hybridization may be indicative of a mutation. The construction of restriction site maps aids in diagnosing certain types of rearrangements. The detection of specific rearrangements is described in the appropriate sections below.

DNA sequence analysis is necessary when filter hybridization data are ambiguous; it can be used to corroborate the results of the hybridization experiments. In some instances, extreme base sequence divergence or shuffling of sequences by rearrangement may preclude a significant level of cross-hybridization. The putative absence of genes or introns, or portions thereof, can be confirmed by sequencing the region of the suspected absence. DNA sequence analysis may also be necessary to ascertain homologous mutations via correct sequence alignment and can be used to determine the presence and extent of length mutations within coding sequences.

The use of the polymerase chain reaction (PCR) in systematic studies is rapidly gaining in popularity (Arnheim et al., 1990). Although comparative DNA sequence data are often routinely obtained using PCR, the use of PCR in surveying for structural rearrangements has been limited. Once a rearrangement has been found by filter hybridizations, PCR is well suited for rapid screens of large numbers of DNAs. The use of PCR to diagnose inversions and the presence or absence of genes and introns is described in the appropriate sections below.

The Phylogenetic Utility of Chloroplast DNA Rearrangements

Compared to the large amount of phylogenetic data now available from the analysis of restriction site variation, relatively little effort has been made to survey and identify major rearrangements among land-plant cpDNAs, even though several identified rearrangements have been quite useful in phylogenetic reconstruction. To illustrate the phylogenetic utility of cpDNA rearrangements and the value of searches to find them, examples of each of the major classes of structural rearrangements are given below. Other rearrangements that will be discussed include the expansion or contraction of the IR, and the occurrence of small length mutations within evolutionarily constrained regions. We elaborate on how we can exploit this variation as phylogenetic information and offer insight into how these structural rearrangements are detected.

Inversions

An inversion occurs when any segment of the chloroplast chromosome has been rotated 180° relative to the regions on either side. The gene order exemplified by *N. tabacum* (Fig. 2.1) is similar to the ancestral vascular plant gene order, because it is found, with few exceptions, in all other examined angiosperms, ferns, and *Ginkgo biloba* (Palmer, 1985a, 1985b; Palmer and Stein, 1986; Palmer et al., 1988a). In most altered genomes, the order of genes can be derived from the ancestral form by one to a few inversions. For example, three inversions characterize the chloroplast genome of the monocot *Oryza sativa* relative to *N. tabacum* (Hiratsuka et al., 1989; Sugiura, 1989). Similar gene arrangements have been documented in *Triticum aestivum* (Quigley and Weil, 1985; Howe et al., 1988) and *Zea mays* (Palmer and Thompson, 1982), suggesting that these rearrangements predate the divergence of the grasses from other monocots. All other monocots examined exhibit the consensus gene order found in *N. tabacum* (de Heij et al., 1982; Palmer et al., 1988a; Chase and Palmer, 1989; Downie and

Palmer, unpublished data). The chloroplast genomes of the liverwort *Marchantia polymorpha* and the moss *Physcomitrella patens* differ from that of *N. tabacum* by only one 30-kb inversion (Calie and Hughes, 1987; Ohyama et al., 1988), despite over 400 million years of evolutionary divergence (Stewart, 1983). The polarity of this 30-kb inversion is unknown. A summary of known inversions in land plants is presented in Table 2.2. Relatively few taxa, most notably in the conifers, Geraniaceae, Fabaceae, Campanulaceae, and Lobeliaceae, contain multiple rearrangements.

An example of the utility of a cpDNA inversion in phylogenetic reconstruction has been documented recently. Chloroplast DNAs from subtribe Barnadesiinae (tribe Mutisieae) of the Asteraceae have the typical gene order found in most land plants, whereas all other Asteraceae share a derived 22-kb inversion (Jansen and Palmer, 1987a, 1987b). This finding and congruent results obtained through a phylogenetic analysis of restriction site mutations (Jansen and Palmer, 1988) and nucleotide substitutions (Jansen et al., Chapter 11, this volume) demonstrate that the Mutisieae is not monophyletic (because its three other subtribes possess the inversion) and that the Barnadesiinae should be considered the sister group to the remainder of the family (see also Jansen et al., Chapter 11, this volume). Consequently, it is now possible to root unambiguously phylogenetic trees using Barnadesiinae as the outgroup in cladistic analyses of the family.

Inversions can be detected in the following ways. The hybridization of two nonadjacent restriction fragments from a genome lacking an inversion to the same two fragments in another genome indicates that an inversion has occurred (see Figs. 2 and 3 in Jansen and Palmer, 1987a). Conversely, two adjacent fragments in an uninverted genome that have become separated by an inversion will hybridize to different fragments in the inverted genome. The hybridization of small probes to those fragments containing the presumed inversion endpoints can provide a more precise localization of both the endpoints and size of the inverted segment. Once restriction maps are constructed, gene mapping and sequencing studies can confirm the difference in gene order and the direction of transcription (via the differential hybridization of 5′ and 3′ gene probes). Additional taxa can be surveyed for the inversion by performing filter hybridization using cloned restriction fragments that contain the inversion endpoints. The PCR technique can also be used to detect inversions. Primers synthesized for two conserved sequences closely flanking a known inversion endpoint will yield a small PCR product after amplification when applied to a species possessing the same inversion. However, if an inversion is not present, the size of the PCR product will be much larger, if a product is produced at all.

Gene/Intron Loss

The complete sequences of the chloroplast genomes of *N. tabacum* (Shinozaki et al., 1986), *M. polymorpha* (Ohyama et al., 1986), and *O. sativa* (Hiratsuka et al., 1989) provide invaluable information on chloroplast gene content and organization. Comparisons of homologous sequences between *N. tabacum* and the liverwort *M. polymorpha* (Wolfe and Sharp, 1988) and between *N. tabacum* and *O. sativa* (Sugiura, 1989) reveal a high degree of conservatism in gene content. Of the 20 distinct introns previously demonstrated or tentatively identified in *N. tabacum* cpDNA (Shinozaki et al., 1986), 18 are present also in *M. polymorpha* (Ohyama et al., 1986) and 17 in *O. sativa* (Hiratsuka et al., 1989). Among these three sequenced chloroplast genomes, there are no known examples of gene or intron gains. Moreover, our observations and those of many other laboratories indicate that genes and introns have been gained rarely, if at all, during land-plant evolution. Consequently, our discussion will deal only with the loss of these sequences.

Any disruption in gene integrity will result in a loss of function. In some instances, gene losses are viewed more appropriately as gene transfers because some genes lost from the chloroplast genome have been found in the nucleus (Baldauf and Palmer, 1990; Gantt and Palmer, unpublished data). Nucleotide substitutions and length mutations occur readily in intron sequences (Ritland and Clegg, 1987; Zurawski and Clegg, 1987) and are of little systematic value. However, the loss of entire intron sequences is a relatively rare event and therefore phylogenetically informative. Palmer (1991) reviews the evolutionary processes and mechanisms responsible for the loss/gain of genes and introns in cpDNA.

The stability in gene/intron content among land-plant cpDNAs can make their absence valuable as a systematic marker at a number of taxonomic levels. For example, the genes *rpoA*, *rpl22*, and *tufA* are absent from the chloroplast genomes of *Pelargonium*, the Fabaceae, and land plants, respectively (Table 2.3). The intron in the gene *rpl2* is absent from all members examined of the Caryophyllales (Zurawski et al., 1984; Downie et al., 1991). Furthermore, filter hybridization surveys of more than 300 chloroplast genomes show that this intron is absent also from members of the Convolvulaceae, *Cuscuta*, *Menyanthaceae*, two genera within the Geraniaceae (*Sarcocaulon* and *Monsonia*), Saxifragaceae *s.s.*, and *Drosera filiformis* (Table 2.3; Downie et al., 1991). This intron loss can be considered to have occurred independently in at least six different lineages of dicots.

The presence or absence of a particular gene or intron may be assayed by hybridization using a probe specific to that gene or intron. Subsequent sequencing of the region in question can confirm its presence or absence, its fragmentation, or its change in position. Currently, in our laboratory, we are using the PCR technique to survey for the presence or absence of introns in some tRNA genes (Kuhsel et al., 1990). Primers are synthesized for conserved sequences flanking the region of interest and the intervening sequence is amplified by PCR. Comparing the size of the resultant PCR product to a sequence of known length on an agarose or polyacrylamide gel can indicate the presence or absence of a specific gene or intron (Bruzdzinski and Gelehrter, 1989). In general, deletions or inser-

Table 2.2. Summary of known inversions in land plant cpDNA (relative to vascular plant consensus gene order, as exemplified by Nicotiana tabacum*).*

Taxon	Size (or number)[a]	Reference
Angiosperms		
Asteraceae		
(all subtribes except Barnadesiinae)	22	Jansen and Palmer (1987a)
Lactuca sativa	4	Downie, Knox, Jansen, and Palmer (unpublished)
Cactaceae		
Pereskia sacharosa	(1)	Wallace (unpublished)
Campanulaceae (4 spp.)	(several)	Downie and Palmer (unpublished)
Chenopodiaceae		
Atriplex (60 spp.)	(1)	Palmer (unpublished)
Fabaceae		
Fabaceae	50	Palmer and Thompson (1982)
Phaseolinae	78	Palmer et. al. (1988b) Bruneau et al. (1990)
Robinieae (except *Sesbania*)	30	Lavin (unpublished)
Trifolium subterraneum	(ca. 8)	Milligan et al. (1989)
Vicia faba	(2 or 3)	Palmer et al. (1987b)
Pisum (4 spp.)	(ca. 8)	Palmer et al. (1988b)
Pisum humile	4	Palmer et al. (1985)
Medicago lupulina	11	Johnson and Palmer (unpublished)
Medicago (3 spp.)	62	Johnson and Palmer (unpublished)
Medicago arabica	(1)	Johnson and Palmer (unpublished)
Medicago tornata	(1)	Johnson and Palmer (unpublished)
Geraniaceae		
Erodium chamaedryoides	(1 or 2)	Calie and Palmer (unpublished)
Geranium grandiflorum	(several)	Calie and Palmer (unpublished)
Pelargonium X hortorum	(ca. 6)	Palmer et al. (1987a)
Sarcocaulon vanderietiae	(several)	Calie and Palmer (unpublished)
Lobeliaceae		
Lobelia (27 spp.), *Sclerotheca jayorum*	(2)	Knox, Downie, and Palmer (unpublished)
Lobelia erinus, L. fervens	(5)	Knox, Downie, and Palmer (unpublished)
Lobelia cardinalis, L. holstii, Monopsis lutea	(3)	Knox, Downie, and Palmer (unpublished)
Oleaceae		
Jasminum (2 spp.)	(2)	Downie and Palmer (unpublished)
Onagraceae		
Oenothera spp.	50	Herrmann et al. (1983) Sytsma (unpublished)
Orobanchaceae		
Conopholis americana	(1)	Downie and Palmer (unpublished)
Poaceae		
Triticum, Oryza, Zea	(3)[b]	Howe et al. (1988) Quigley and Weil (1985) Palmer and Thompson (1982) Hiratsuka et al. (1989)

(continued)

Table 2.2.
(Continued)

Taxon	Size (or number)[a]	Reference
Ranunculaceae		
Adonis aestivalis	(1)	Hoot and Palmer (unpublished)
Anemone (40 spp.), *Hepatica, Knowltonia, Pulsatilla*	(4)	Hoot and Palmer (unpublished)
Anemone (3 spp.), *Clematis* (2 spp.)	(6)	Hoot and Palmer (unpublished)
Scrophulariaceae		
Striga asiatica	(3)	Downie and Palmer (unpublished)
Gymnosperms		
Conifers	(several)	Strauss et al. (1988) Raubeson and Jansen (unpublished)
Pseudotsuga menziesii	45	Strauss et al. (1988)
Pteridophytes		
Adiantum capillus-veneris	(2)	Hasebe and Iwatsuki (1990)
Bryophytes		
Marchantia polymorpha	30	Ohyama et al. (1988)
Physcomitrella patens	30	Calie and Hughes (1987)

[a]Estimated sizes in kb provided. If more than one inversion is present, or if inversion is not well characterized, the number of postulated inversions is presented in parentheses.

[b]These three inversions are 28 kb, 10 kb, and approximately 1 kb (see Fig. 1).

tions of moderate size (less than 2 kb) can be detected in this way. Furthermore, amplified intron or gene-sequence products can be readily isolated and subsequently cloned and/or sequenced.

A summary of genes and introns known (by DNA sequencing) or suspected (by filter hybridization) to be lost from the cpDNAs of various land plants is presented in Table 2.3. Our results suggest that the unidentified open reading frames (ORFs) are most amenable to loss. In several situations it appears that similar losses have occurred in parallel. The ribosomal protein genes and introns are also sometimes lost, whereas no photosynthetic gene is known to have been lost from any cpDNA of a photosynthetic land plant.

Loss of the Inverted Repeat

One of the most intriguing rearrangements is the loss of one copy of the IR. Although the presence of the IR may confer a certain stability upon the cpDNA molecule, making it less prone to rearrangement (Palmer and Thompson, 1982; Strauss et al., 1988), the absence of one copy suggests that it is not fundamental to genome function. With few exceptions, all angiosperm cpDNAs possess a large IR, usually approximately 20 to 30 kb in size, that encodes a duplicate set of ribosomal RNA genes (Fig. 2.1). The deletion of one entire segment of this duplicated sequence is a significant mutation, which, when considered in a phylogenetic context, can define monophyletic groups. This rare deletion has

Table 2.3. Summary of known cpDNA gene and intron losses in land plants.

Gene/Intron[a]	Taxon[b]	Reference
many*[c]	*Epifagus virginiana*	dePamphilis and Palmer (1989)
	Conopholis americana	Downie and Palmer (unpublished)
*tuf*A*	land plants	Baldauf and Palmer (1990)
*rpo*A*	*Pelargonium* (40 spp.)	Calie and Palmer (unpublished)
*rpl*20	*Sarcocaulon* (2 spp.)	Downie and Palmer (unpublished)
*rpl*22*	Fabaceae (3 subfamilies)	Palmer and Doyle (unpublished) Spielmann et al. (1988)
*rps*7	*Podophyllum peltatum*	Downie and Palmer (unpublished)
*rps*16	Fabaceae (5/9), *Linum grandiflorum*, *Malpighia coccigera*, *Passiflora* sp., *Polygala lindheimeri*, *Populus deltoides*, *Salix amygdaloides*, *Securidaca diversifolia*, *Turnera ulmifolia*, *Viola* (2 spp.)	Downie and Palmer (unpublished)
*clp*P (orf196)	*Geranium* (2 spp.), *Jasminum* (2 spp.), *Linum grandiflorum*, *Lobelia holstii*, *Lonicera subsessilis*, *Monopsis lutea*, *Monsonia* (2 spp.), *Oenothera missouriensis*, *Sarcocaulon* (2 spp.)	Downie and Palmer (unpublished)
*ndh*F	*Hebestigma cubense*	Downie and Palmer (unpublished)
*zfp*A* (orf512)	*Oryza sativa*	Hiratsuka et al. (1989)
	Bambusa sp., *Zea mays*, Campanulaceae (3/4), Geraniaceae (4/8), Lobeliaceae (4/5), Oleaceae (2/3)	Downie and Palmer (unpublished)
orf184*	*Pisum sativum*	Sasaki et al. (1989)
	Fabaceae (11/16)	Downie and Wolfe (unpublished)
orf228*	*Oryza sativa*	Hiratsuka et al. (1989)
orf1244*	*Oryza sativa*	Hiratsuka et al. (1989)
	Bambusa sp., *Zea mays*, Campanulaceae (3/4), Convolvulaceae (3/4), Lobeliaceae (4/5), *Cuscuta* sp., *Linum grandiflorum*, *Pisum sativum*	Downie and Palmer (unpublished)
orf2280*	*Oryza sativa*	Hiratsuka et al. (1989)
	Bambusa sp., *Zea mays*, Campanulaceae (2/2), Geraniaceae (3/6)	Downie and Palmer (unpublished)

(continued)

Table 2.3. (Continued)

Gene/Intron[a]	Taxon[b]	Reference
*rpo*C1 intron*	*Oryza sativa*	Hiratsuka et al. (1989)
*rpl*2 intron*	Caryophyllales	Zurawski et al. (1984) Downie et al. (1991)
	Convolvulaceae (4/5), Menyanthaceae (4/5), Saxifragaceae (24/50), *Cuscuta* sp., *Drosera filiformis*, *Monsonia* (2 spp.), *Sarcocaulon* (2 spp.)	Downie et al. (1991)
*rpl*16 intron*	Geraniaceae (5/44), *Limonium gmelinii*	Downie et al. (unpublished)
*trn*I intron	*Campanula garganica*	Downie and Palmer (unpublished)
orf196 introns 1 and 2*	*Oryza sativa*	Hiratsuka et al. (1989)
	Zea mays, Bambusa sp.	Downie and Palmer (unpublished)

[a]Gene/intron absence postulated only on the basis of filter hybridizations. Asterisks denote those genes/introns whose absence has been confirmed by DNA sequencing in at least one of the taxa.

[b]Numbers in parentheses indicate # of genera/ # of species exhibiting loss.

[c]*Epifagus virginiana* and *Conopholis americana* are nonphotosynthetic, parasitic plants. Most photosynthetic genes, NADH dehydrogenase genes, and ORFs are missing.

now been found in four independent lineages of vascular plants (Table 2.4). Particularly notable from a phylogenetic standpoint is the absence of one copy of this repeat from six tribes and the putatively allied genus *Wisteria* within the subfamily Papilionoideae (Lavin et al., 1990; Doyle et al., Chapter 10, this volume), and from the conifers, including Taxaceae (Lidholm et al., 1988; Strauss et al., 1988; Raubeson and Jansen, unpublished data). The strategy used to detect the presence or absence of the IR involves hybridization assays using small probes homologous with the conserved ends of the IR and single-copy regions (Palmer et al., 1988a; Lavin et al., 1990).

Expansion/Contraction of the Inverted Repeat

The expansion or contraction of the IR into, or out of, the two single-copy regions significantly influences the variability in size of the chloroplast genome. At one extreme is the 217-kb *Pelargonium X hortorum* cpDNA possessing a greatly enlarged IR of 76 kb, almost three times the size found in most angiosperms (Palmer et al., 1987a). Consequently, many protein genes that are present only once in most other plants are duplicated in *P. X hortorum*. At the other extreme in angiosperms with the IR is *Coriandrum sativum* with an IR less than half the normal size (Palmer, 1985b; Downie, unpublished data). In *C. sativum*, the gene *rpl*2, which normally is located near the terminus of the IR (Fig. 2.1), is a single-copy gene some 10 kb away from the end of the repeat. In *Oryza sativa* the IR segments have expanded into single-copy regions; however, a series

Table 2.4. Land plant cpDNAs possessing only one inverted repeat segment.

Taxon	Reference
Conifers	Strauss et al. (1988) Lidholm et al. (1988)
Fabaceae	
Papilionoideae (6 tribes and *Wisteria*)	Lavin et al. (1990) Palmer et al. (1987b)
Geraniaceae	
Erodium and *Sarcocaulon*	Calie and Palmer (unpublished)
Orobanchaceae	
Conopholis americana	Downie and Palmer (unpublished)

of deletions within the repeat makes the IR smaller than that found in *N. tabacum* (Hiratsuka et al., 1989). Variation in IR size is common but has not yet been used in phylogenetic analyses. Any length mutation that occurs within the IR undoubtedly will affect its size, thus making homologous size variants difficult to assess.

Length Mutation in Constrained Regions

Although small length mutations occur predominantly in noncoding DNA, they are also occasionally found within genes and other evolutionarily constrained portions of the genome. In order not to disrupt the reading frame in protein-coding genes, only insertions or deletions of just 3 bp, or in multiples of three, are permitted (e.g., Blasko et al., 1988). Small insertions or deletions within coding regions that are conserved evolutionarily may be considered as phylogenetic characters independent of nucleotide substitutions (Meyer et al., 1986; Morden and Golden, 1989).

Use of Rearrangements in Phylogenetic Reconstruction

The paucity of major structural rearrangements found to date within the chloroplast genomes of vascular plants suggests that they occur rarely during cpDNA evolution. However, once a rearrangement is found, characterized, and its distribution circumscribed in related taxa, its presence can make a profound phylogenetic statement. These unique characters are prominent and powerful systematic markers that offer the following advantages to phylogenetic reconstruction: (1) assessing the homology of the rearrangement usually is straightforward; (2) the polarity of each structural mutation is ascertained readily when it is compared to an outgroup; and (3) once a particular rearrangement is identified, it can be surveyed rapidly in other taxa through simple filter hybridization or PCR assays.

Although shared structural mutations can provide strong evidence of common ancestry, it is apparent now that similar rearrangements can occur independently, such as the loss of the *rpl2* intron. However, because the intron is absent in otherwise distantly related clades and present in their immediate, respective outgroups, the assessment of homology can be made confidently, and the loss can be considered to have occurred independently in at least six different lineages of dicots.

As in any systematic endeavor, the selection of representative specimens is critical to the outcome of the analysis. Because of the conservative nature of cpDNA evolution, specifically as it relates to its generally invariant order and content of genes, relatively few species are necessary to represent most taxa at the generic level and above. However, as previous studies attest, rearrangements can identify major dichotomies within clades at any level, so unless the group in question is well represented, some rearrangements may go undetected. Incorporating additional specimens into the analysis also is necessary when doubts arise concerning monophyly.

Extensively rearranged genomes are encountered rarely in land plants, and have so far been well characterized only in *Pisum* (Palmer et al., 1988b), *Trifolium subterraneum* (Palmer et al., 1987b; Milligan et al., 1989), *Pelargonium X hortorum* (Palmer et al., 1987a), and conifers (Strauss et al., 1988). The processes that contributed to the formation of these rearranged chloroplast genomes are not clear (Palmer, 1991) but may have involved either some major alteration of the IR (its loss or manyfold expansion) (Palmer 1985a, 1985b; Palmer et al., 1987a; Strauss et al., 1988) or the occurrence of dispersed, recombinogenic repeat elements (Howe, 1985; Palmer et al., 1987a; Blasko et al., 1988; Bowman et al., 1988; Milligan et al., 1989). Extensive genome rearrangement makes it almost impossible to align restriction fragment maps and largely precludes phylogenetic analyses of comparative restriction site variation. Furthermore, determining the nature and polarity for each mutation, providing they can be delimited, would be an arduous task. Fortunately, most altered genomes can be explained by a few discrete inversions (Palmer, 1985b; Downie and Palmer, unpublished data).

The analysis of major structural rearrangements is a complementary approach to comparative sequencing for studying the higher-level relationships among angiosperms. As phylogenetic distance increases among taxa, comparative restriction site mapping is plagued by excessive homoplasy and length mutation. Since many genes are conserved more than the genome as a whole, the direct comparison of homologous coding sequences (such as *rbc*L) is more appropriate for studying higher levels of plant phylogeny. Accordingly, in collaboration with others from our laboratory, we are in the process of sequencing *rbc*L from representatives of the Asteridae and outgroups from the Rosidae to acquire complementary information to that obtained from the rearrangement study described herein. Sequencing provides a large number of phylogenetically informative characters, whereas fewer cpDNA rearrangements are expected simply due to the rarity of these events. However, once found, these mutations should be considered to be more powerful characters than individual nucleotide substitutions, as data sets constituting the latter are inevitably afflicted with certain levels of homoplasy. Moreover, comparative sequence data may not resolve relatively ancient and compressed evolutionary radiations, whereas each rearrangement has the potential to resolve with confidence a particular branching point in a phylogeny (Palmer et al., 1988a). The integration of rearrangement data with other cpDNA-derived data (such as restriction site mutations and nucleotide substitutions) in phylogenetic analyses is an issue that has not yet been seriously explored.

Conclusions

The distribution of major structural rearrangements has the potential to illuminate the deeper branches of plant evolution and in doing so to define monophyletic groups. We have initiated a large-scale survey to detect and circumscribe major structural rearrangements in the chloroplast genomes of angiosperms, with special reference to the subclass Asteridae. Even though rearrangements alone are unlikely to provide a comprehensive framework of plant relationships, simply because of the small number of phylogenetically informative characters they represent, when used in conjunction with other molecular and traditional approaches, they have the power to help resolve many questions of plant phylogeny.

References

Arnheim, N., White, T., and Rainey, W.E. (1990) Application of PCR: organismal and population biology. *Bioscience* **40,** 174–182.

Baldauf, S., and Palmer, J.D. (1990) Evolutionary transfer of the chloroplast *tuf*A gene to the nucleus. *Nature* **344,** 262–265.

Baldwin, B.G., Kyhos, D.W., and Dvořák, J. (1990) Chloroplast DNA evolution and adaptive radiation in the Hawaiian silversword alliance (Asteraceae-Madiinae). *Ann. Missouri Bot. Gard.* **77,** 96–109.

Blasko, K., Kaplan, S.A., Higgins, K.G., Wolfson, R., and Sears, B.B. (1988) Variation in copy number of a 24-base pair tandem repeat in the chloroplast DNA of *Oenothera hookeri* strain Johansen. *Curr. Genet.* **14,** 287–292.

Bowman, C.M., Barker, R.F., and Dyer, T.A. (1988) In wheat ctDNA, segments of ribosomal protein genes are dispersed repeats, probably conserved by nonreciprocal recombination. *Curr. Genet.* **14,** 127–136.

Bruzdzinski, C.J., and Gelehrter, T.D. (1989) Determination of exon-intron structure: a novel application of the polymerase chain reaction technique. *DNA* **8,** 691–696.

Bruneau, A., Doyle, J.J., and Palmer, J.D. (1990) A chloroplast DNA inversion as a subtribal character in the Phaseoleae (Leguminosae). *Syst. Bot.* **15,** 378–386.

Calie, P.J., and Hughes, K.W. (1987) The consensus land plant chloroplast gene order is present, with two alterations, in the moss *Physcomitrella patens*. *Mol. Gen. Genet.* **208,** 335–341.

Chase, M.W., and Palmer, J.D. (1989) Chloroplast DNA systematics of lilioid monocots: resources, feasibility, and an example from the Orchidaceae. *Amer. J. Bot.* **76,** 1720–1730.

Coates, D., and Cullis, C.A. (1987) Chloroplast DNA variability among *Linum* species. *Amer. J. Bot.* **74,** 260–268.

Cronquist, A. (1981) *An Integrated System of Classification of Flowering Plants*, Columbia University Press, New York.

Crouse, E.J., Schmitt, J.M., and Bohnert, H. (1985) Chloroplast and cyanobacterial genomes, genes and RNAs: a compilation. *Plant Mol. Biol. Reporter* **3,** 43–89.

Dang, L.A., and Pring, D.R. (1986) A physical map of the sorghum chloroplast genome. *Plant Mol. Biol.* **6,** 119–123.

de Heij, H.T., Lustig, H., Moeskops, D.M., Bovenberg, W.A., Bisanz, C., and Groot, G.S.P. (1983) Chloroplast DNAs of *Spinacia, Petunia,* and *Spirodela* have a similar gene organization. *Curr. Genet.* **7,** 1–6.

dePamphilis, C.W., and Palmer, J.D. (1989) Evolution and function of plastid DNA: a review with special reference to nonphotosynthetic plants. In: *Physiology, Biochemistry, and Genetics of Nongreen Plastids* (eds. C.D. Boyer, J.C. Shannon, and R.C. Hardison), American Society of Plant Physiologists, pp. 182–202.

Doebley, J.F., Ma, D.P., and Renfroe, W.T. (1987a) Insertion/deletion mutations in the *Zea* chloroplast genome. *Curr. Genet.* **11,** 617–624.

Doebley, J., Renfroe, W., and Blanton, A. (1987b) Restriction site variation in the *Zea* chloroplast genome. *Genetics* **117,** 139–147.

Downie, S.R., Olmstead, R.G., Zurawski, G., Soltis, D.E., Soltis, P.S., Watson, J.C., and Palmer, J.D. (1991) Six independent losses of the chloroplast DNA *rpl*2 intron in dicotyledons: molecular and phylogenetic implications. *Evolution,* in press.

Gounaris, I., Michalowski, C.B., Bohnert, H.J., and Price, C.A. (1986) Restriction and gene maps of plastid DNA from *Capsicum annuum*. *Curr. Genet.* **11,** 7–16.

Green, R.M., Vardi, A., and Galun, E. (1986) The plastome of *Citrus*. Physical map, variation among *Citrus* cultivars and species and comparison with related genera. *Theor. Appl. Genet.* **72,** 170–177.

Hallick, R.B. (1989) Proposals for the naming of chloroplast genes. II. Update to the nomenclature of genes for thylakoid membrane polypeptides. *Plant Mol. Biol. Reporter* **7,** 266–275.

Hallick, R.B., and Bottomley, W. (1983) Proposals for the naming of chloroplast genes. *Plant Mol. Biol. Reporter* **1,** 38–43.

Hansmann, P. (1987) Daffodil chromoplast DNA: comparison with chloroplast DNA, physical map, and gene localization. *Z. Naturforsch.* **42c,** 118–122.

Hasebe, M., and Iwatsuki, K. (1990) Chloroplast DNA from *Adiantum capillus-veneris* L., a fern species (Adiantaceae); clone bank, physical map and unusual gene localization in comparison with angiosperm chloroplast DNA. *Curr. Genet.* **17,** 359–364.

Heinhorst, S., Gannon, G.C., Galun, E., Kenschaft, L., and Weissbach, A. (1988) Clone bank and physical and genetic map of potato chloroplast DNA. *Theor. Appl. Genet.* **75,** 244–251.

Herrmann, R.G., Westhoff, P., Alt, J., Winter, P., Tittgen, J., Bisanz, C., Sears, B.B., Nelson, N., Hurt, E., Hauska, G., Viebrock, A., and Sebald, W. (1983) Identification and characterization of genes for polypeptides of thylakoid membrane. In: *Structure and Function of Plant Genomes* (eds. O. Ciferri and L. Dure), Plenum Press, New York, pp. 143–153.

Heyraud, F., Serror, P., Kuntz, M., Steinmetz, A., and Heizmann, P. (1987) Physical map and gene localization on sunflower (*Helianthus annuus*) chloroplast DNA: evidence for an inversion of a 23.5-kbp segment in the large single copy region. *Plant Mol. Biol.* **9,** 485–496.

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., and 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.

Howe, C.J. (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, C.J., Barker, R.F., Bowman, C.M., and Dyer, M. (1988) Common features of three inversions in wheat chloroplast DNA. *Curr. Genet.* **13,** 343–349.

Hudson, K.R., and Gardner, R.C. (1988) Organisation of the chloroplast genome of kiwifruit (*Actinidia deliciosa*). *Curr. Genet.* **13,** 339–342.

Jansen, R.K., and Palmer, J.D. (1987a) Chloroplast DNA from lettuce and *Barnadesia* (Asteraceae): structure, gene localization, and characterization of a large inversion. *Curr. Genet.* **11,** 553–564.

Jansen, R.K., and Palmer, J.D. (1987b) A chloroplast DNA inversion marks an ancient evolutionary split in the sunflower family (Asteraceae). *Proc. Natl. Acad. Sci. USA* **84,** 5818–5822.

Jansen, R.K., and Palmer, J.D. (1988) Phylogenetic implications of chloroplast DNA restriction site variation in the Mutisieae (Asteraceae). *Amer. J. Bot.* **75,** 753–766.

Jansen, R.K., Holsinger, K.E., Michaels, H.J., and Palmer, J.D. (1990) Phylogenetic analysis of chloroplast DNA restriction site data at higher taxonomic levels: an example from the Asteraceae. *Evolution* **44,** 2089–2105.

Kishima, Y., Mikami, T., Hirai, A., Sugiura, M., and Kinoshita, T. (1987) *Beta* chloroplast genomes: analysis of fraction I protein and chloroplast DNA variation. *Theor. Appl. Genet.* **73,** 330–336.

Kuhsel, M.G., Strickland, R., and Palmer, J.D. (1990) An ancient Group I intron shared by eubacteria and chloroplasts. Science **250,** 1570–1573.

Kung, S.D., Zhu, Y.S., and Shen, G.F. (1982) *Nicotiana* chloroplast genome III. Chloroplast DNA evolution. *Theor. Appl. Genet.* **61,** 73–79.

Lavin, M., Doyle, J.J., and Palmer, J.D. (1990) Evolutionary significance of the loss

of chloroplast-DNA inverted repeat in the Leguminosae subfamily Papilionoideae. *Evolution* **44,** 390–402.

Lidholm, J., Szmidt, A.E., Hällgren, J.E., and Gustafsson P. (1988) The chloroplast genomes of conifers lack one of the rRNA-encoding inverted repeats. *Mol. Gen. Genet.* **212,** 6–10.

Ma, C., and Smith, M.A. (1985) Construction and mapping of safflower chloroplast DNA recombinants and location of selected gene markers. *Theor. Appl. Genet.* **70,** 620–627.

Meyer, T.E., Cusanovich, M.A., and Kamen, M.D. (1986) Evidence against use of bacterial amino acid sequence data for construction of all-inclusive phylogenetic trees. *Proc. Natl. Acad. Sci. USA* **83,** 217–220.

Milligan, B.G., Hampton, J.N., and Palmer, J.D. (1989) Dispersed repeats and structural reorganization in subclover chloroplast DNA. *Mol. Biol. Evol.* **6,** 355–368.

Morden, C.W., and Golden, S.S. (1989) *psbA* genes indicate common ancestry of prochlorophytes and chloroplasts. *Nature* **337,** 382–385.

Murai, K., and Tsunewaki, K. (1987) Chloroplast genome evolution in the genus *Avena. Genetics* **116,** 613–621.

Murai, K., Naiyu, X., and Tsunewaki, K. (1989) Studies on the origin of crop species by restriction endonuclease analysis of organellar DNA. III. Chloroplast DNA variation and interspecific relationships in the genus *Secale. Jap. J. Genet.* **64,** 35–47.

Ngernprasirtsiri, J., and Kobayashi, H. (1990) Application of an efficient strategy with a phage λ vector for constructing a physical map of the amyloplast genome of sycamore (*Acer pseudoplatanus*). *Arch. Biochem. Biophys.* **276,** 172–179.

Ogihara, Y., and Tsunewaki, K. (1988) Diversity and evolution of chloroplast DNA in *Triticum* and *Aegilops* as revealed by restriction fragment analysis. *Theor. Appl. Genet.* **76,** 321–332.

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

Ohyama, K., Kohchi, T., Sano, T., and Yamada, Y. (1988) Newly identified groups of genes in chloroplasts. *TIBS* **13,** 19–22.

Palmer, J.D. (1985a) Evolution of chloroplast and mitochondrial DNA in plants and algae. In: *Molecular Evolutionary Genetics* (ed. R.J. MacIntyre), Plenum Press, New York, pp. 131–240.

Palmer, J.D. (1985b) Comparative organization of chloroplast genomes. *Ann. Rev. Genet.* **19,** 325–354.

Palmer, J.D. (1991) Plastid chromosomes: structure and evolution. In: *Cell Culture and Somatic Cell Genetics in Plants,* Vol. 7, *The Molecular Biology of Plastids* (eds. L. Bogorad and I.K. Vasil), Academic Press, New York, pp. 5–53.

Palmer, J.D., and Stein, D.B. (1986) Conservation of chloroplast genome structure among vascular plants. *Curr. Genet.* **10,** 823–833.

Palmer, J.D., and Thompson, W.F. (1982) Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. *Cell* **29,** 537–550.

Palmer, J.D., Jorgensen, R.A., and Thompson, W.F. (1985) Chloroplast DNA variation and evolution in *Pisum*: patterns of change and phylogenetic analysis. *Genetics* **109,** 195–213.

Palmer, J.D., Nugent, J.M., and Herbon, L.A. (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, J.D., Osorio, B., Aldrich, J., and Thompson, W.F. (1987b) Chloroplast DNA evolution among legumes: loss of a large inverted repeat occurred prior to other sequence rearrangements. *Curr. Genet.* **11,** 275–286.

Palmer, J.D., Jansen, R.K., Michaels, H.J., Chase, M.W., and Manhart, J.R. (1988a) Chloroplast DNA variation and plant phylogeny. *Ann. Missouri Bot. Gard.* **75,** 1180–1206.

Palmer, J.D., Osorio, B., and Thompson, W.F. (1988b) Evolutionary significance of inversions in legume chloroplast DNAs. *Curr. Genet.* **14,** 65–74.

Perl-Treves, R., and Galun, E. (1985) The *Cucumis* plastome: physical map, intrageneric variation and phylogenetic relationships. *Theor. Appl. Genet.* **71,** 417–429.

Quigley, F., and Weil, J.H. (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.

Ritland, K., and Clegg, M.T. (1987) Evolutionary analysis of plant DNA sequences. *Am. Nat.* **130,** S74–S100.

Sasaki, Y., Nagano, Y., Morioka, S., Ishikawa, H., and Matsuno, R. (1989) A chloroplast gene encoding a protein with one zinc finger. *Nucleic Acids* Res. **17,** 6217–6227.

Schilling, E.E., and Jansen, R.K. (1989) Restriction fragment analysis of chloroplast DNA and the systematics of *Viguiera* and related genera (Asteraceae: Heliantheae). *Amer. J. Bot.* **76,** 1769–1778.

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., and Sugiura, M. (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO J.* **5,** 2043–2049.

Smith, R.L., and Sytsma, K.J. (1990) Evolution of *Populus nigra* L. (sect. *Aigeiros*): introgressive hybridization and the chloroplast contribution of *Populus alba* L. (sect. *Populus*). *Amer. J. Bot.* **77,** 1176–1187.

Soltis, D.E., Soltis, P.S., Ranker, T.A., and Ness, B.D. (1989) Chloroplast DNA variation in a wild plant, *Tolmiea menziesii*. *Genetics* **121,** 819–826.

Soltis, D.E., Soltis, P.S., and Bothel, K.D. (1990) Chloroplast DNA evidence for the origins of the monotypic *Bensoniella* and *Conimitella* (Saxifragaceae). *Syst. Bot.* **15,** 349–362.

Spielmann, A., Roux, E., Allmen, J. von, and Stutz, E. (1988) The soybean chloroplast genome: complete sequence of the *rps*19 gene, including flanking parts containing exon2 of *rpl2* (upstream), but not *rpl22* (downstream). *Nucleic Acids Res.* **16,** 1199.

Stein, D.B., Palmer, J.D., and Thompson, W.F. (1986) Structural evolution and flip-flop recombination of chloroplast DNA in the fern genus *Osmunda*. *Curr. Genet.* **10,** 835–841.

Stewart, W.N. (1983) *Paleobotany and the Evolution of Plants,* Cambridge University Press, Cambridge, UK.

Strauss, S.H., Palmer, J.D., Howe, G.T., and Doerksen, A.H. (1988) Chloroplast genomes of two conifers lack a large inverted repeat and are extensively rearranged. *Proc. Natl. Acad. Sci. USA* **85,** 3898–3902.

Sugiura, M. (1989) The chloroplast chromosomes in land plants. *Ann. Rev. Cell Biol.* **5,** 51–70.

Sugiura, M., Shinozaki, K., Zaita, N., Kusuda, M., and Kumano, M. (1986) Clone bank of the tobacco (*Nicotiana tabacum*) chloroplast genome as a set of overlapping restriction endonuclease fragments: mapping of eleven ribosomal protein genes. *Plant Sci.* **44,** 211–216.

Sundberg, S.D., Denton, M.F., and Rehner, S.A. (1990) Structural map of *Sedum oreganum* (Crassulaceae) chloroplast DNA. *Biochem. Syst. Ecol.* **18,** 409–411.

Sytsma, K.J., and Gottlieb, L.D. (1986) Chloroplast DNA evolution and phylogenetic relationships in *Clarkia* sect. *Peripetasma* (Onagraceae). *Evolution* **40,** 1248–1261.

Sytsma, K.J., Smith, J.F., and Gottlieb, L.D. (1990) Phylogenetics in *Clarkia* (Onagraceae): restriction site mapping of chloroplast DNA. *Syst. Bot.* **15,** 280–295.

Sytsma, K.J., Smith, J.F., and Berry, P.E. (1991) Biogeography and evolution of morphology, breeding systems, flavonoids, and chloroplast DNA in the four Old World species of *Fuchsia* (Onagraceae). *Syst. Bot.* **16,** 257–269.

Terauchi, R., Terachi, T., and Tsunewaki, K. (1989) Physical map of chloroplast DNA of aerial yam, *Dioscorea bulbifera* L. *Theor. Appl. Genet.* **78,** 1–10.

White, E.E. (1990) Chloroplast DNA in *Pinus monticola*. I. Physical map. *Theor. Appl. Genet.* **79,** 119–124.

Whitfeld, P.R., and Bottomley, W. (1983) Organization and structure of chloroplast genes. *Ann. Rev. Plant Physiol.* **34,** 279–310.

Wolfe, K.H., and Sharp, P.M. (1988) Identification of functional open reading frames in chloroplast genomes. *Gene* **66,** 215–222.

Zurawski, G., and Clegg, M.T. (1987) Evolution of higher-plant chloroplast DNA-encoded genes: implications for structure-function and phylogenetic studies. *Ann. Rev. Plant Physiol.* **38,** 391–418.

Zurawski, G., Bottomley, W., and Whitfeld, P.R. (1984) Junctions of the large single copy region of the inverted repeats in *Spinacia oleracea* and *Nicotiana debneyi* chloroplast DNA: sequence of the genes for $tRNA^{HIS}$ and the ribosomal proteins S19 and L2. *Nucleic Acids Res.* **12,** 6547–6558.