Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 769–773, February 1987
Evolution

Unusual structure of geranium chloroplast DNA: A triple-sized inverted repeat, extensive gene duplications, multiple inversions, and two repeat families

(genome evolution/rearrangement/*Pelargonium hortorum*)

JEFFREY D. PALMER*, JACQUELINE M. NUGENT, AND LAURA A. HERBON†

Department of Biology, University of Michigan, Ann Arbor, MI 48109

Communicated by Warren H. Wagner, Jr., October 13, 1986

ABSTRACT **Physical and gene mapping studies reveal that chloroplast DNA from geranium (*Pelargonium hortorum*) has sustained a number of extensive duplications and inversions, resulting in a genome arrangement radically unlike that of other plants. At 217 kilobases in size, the circular chromosome is about 50% larger than the typical land plant chloroplast genome and is by far the largest described to date, to our knowledge. Most of this extra size can be accounted for by a 76-kilobase inverted duplication, three times larger than the normal chloroplast DNA inverted repeat. This tripling has occurred primarily by spreading of the inverted repeat into regions that are single copy in all other chloroplast genomes. Consequently, 10 protein genes that are present only once in all other land plants are duplicated in geranium. At least six inversions, occurring in both the inverted repeat and large single-copy region, must be postulated to account for all of the gene order differences that distinguish the geranium genome from other chloroplast genomes. We report the existence in geranium of two families of short dispersed repeats and hypothesize that recombination between repeats may be the major cause of inversions in geranium chloroplast DNA.**

The chloroplast genome is highly conserved in size and arrangement among the vast majority of land plants (1–10). Most angiosperm chloroplast DNAs (cpDNAs) are between 135 and 160 kilobase pairs (kb) in size, contain a 21- to 28-kb inverted repeat, and feature a nearly invariant arrangement of genes around the circular chromosome. cpDNAs from representative angiosperms, ferns, and gymnosperms share a common size and gene order (5). These shared features suggest a consensus genome arrangement for the common ancestor of vascular plants, which existed some 400 million years ago. Furthermore, complete sequence analysis reveals only a single major cpDNA rearrangement between plants representing the two basic lineages of land plants, vascular and nonvascular (6). Only among a single group of leguminous cpDNAs, which have deleted one entire segment of the inverted repeat, does one find an accelerated frequency of even one class of structural mutations, namely inversions (2, 11–14).

In this study we show that cpDNA from geranium possesses a unique combination of structural alterations relative to all other land plants. The geranium genome is unusually large, is highly rearranged by inversion, has duplications of many typically single-copy genes, and contains two families of short dispersed repeats. We discuss the possibility that these repeats play a direct role in generating inversions.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

cpDNA was isolated from geranium (*Pelargonium hortorum* cv. Irene) by a sucrose gradient procedure (15). Methods for agarose gel electrophoresis, bidirectional filter transfers, hybridizations, cpDNA cloning in plasmid vectors, and isolation of plasmid DNA were as described (15). All filters were washed in 2× SSC (300 M NaCl/30 mM trisodium citrate) and 0.5% $NaDodSO_4$ at 65°C prior to autoradiography.

RESULTS

Physical Structure of the Geranium Chloroplast Genome. The fragments produced by digestion of geranium cpDNA with the four restriction enzymes chosen for mapping are displayed in Fig. 1. Summation of restriction fragment sizes yields a genome size estimate of 217 kb, far greater than the average land plant cpDNA of 150 kb (16) and the largest previously known genome of 180 kb (*Spirodela*; ref. 17). Summations of double bright bands (Fig. 1) range from 54.4 kb for *Pst* I to 72.3 kb for *Sac* I, suggesting an inverted repeat far larger than any previously reported in land plants (10–28 kb; ref. 16).

To construct a physical map of the fragments shown in Fig. 1, recombinant plasmids were made containing *Pst* I fragments covering 85% of the geranium chloroplast genome. Each plasmid was hybridized to a filter containing single digests with *Pst* I, *Pvu* II, *Xho* I, and *Sac* I and double digests with *Pst* I–*Pvu* II, *Pst* I–*Xho* I, and *Pst* I–*Sac* I. Additional mapping information to cover the regions of uncloned *Pst* I fragments was gained by analysis of the double digestion products of all pairwise combinations of the four enzymes and by consideration of the gene mapping results described below. These mapping data reveal that the geranium chloroplast genome exists as a 217-kb circular molecule organized into four segments (Fig. 2). A pair of duplicated segments, of minimum size 76 kb and arranged as an inverted repeat, separates the remainder of the chromosome into single-copy regions of maximum sizes 7 and 58 kb.

Gene Duplication and Inversion in Geranium. The enlarged size of the inverted repeat in geranium (three times that in most plants), and the correspondingly smaller single-copy regions, suggested that many chloroplast genes that are normally present once per genome might be duplicated in geranium. To test this hypothesis, hybridization experiments were performed between filter-bound geranium cpDNA fragments and cloned fragments containing 28 chloroplast protein

Abbreviation: cpDNA, chloroplast DNA.

*To whom reprint requests should be addressed.
†Present address: Department of Landscape Architecture and Regional Planning, University of Massachusetts, Amherst, MA 01002.

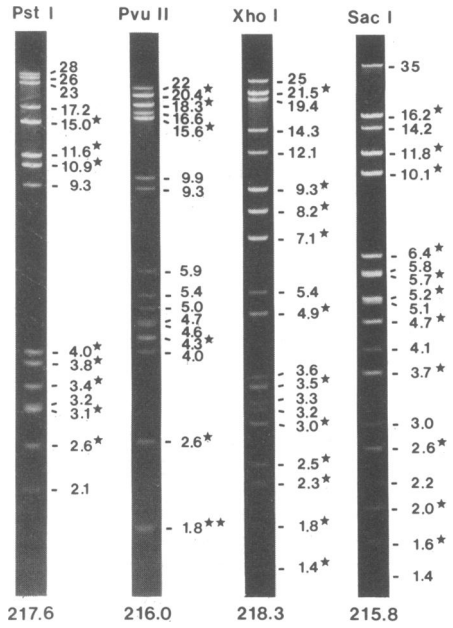

FIG. 1. Separation of geranium cpDNA restriction fragments on a 0.7% agarose gel. Fragment sizes are given in kb. Doublet bands are marked with one star and a triplet band with two stars. Additional fragments were observed on higher percentage gels in digests with *Pvu* II (0.8 kb), *Xho* I (1.0 kb), and *Sac* I (0.4 kb, and doublet fragments of 1.0, 0.7, and 0.6 kb). Numbers at the bottom indicate summations of restriction fragment sizes, taking into account fragment stoichiometries.

and rRNA genes from spinach, tobacco, pea, and mung bean (Table 1). These hybridizations allowed us to locate all 28

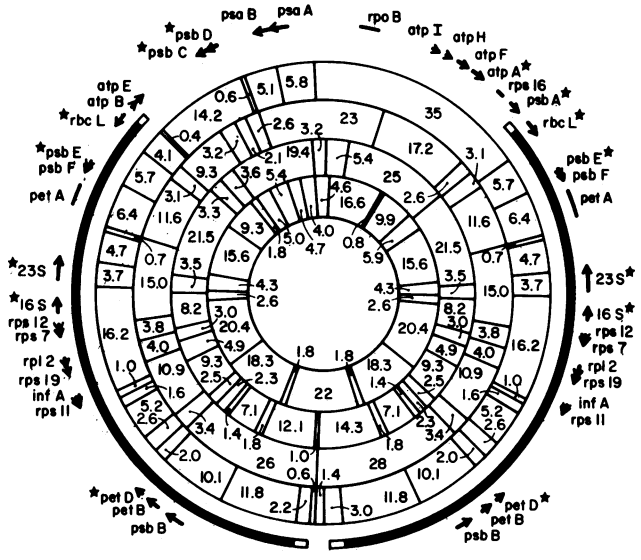

FIG. 2. Physical and gene maps of the geranium chloroplast chromosome. *Sac* I sites are shown on the outermost complete circle, *Pst* I sites on the second-outermost circle, *Xho* I sites on the second-innermost circle, and *Pvu* II sites on the innermost circle. Fragment sizes are given in kb. The two long, filled lines represent the minimum mapped extent of the inverted repeat and the open extensions of these lines its maximum possible extent. The single-copy regions are shown in only one of two possible relative orientations (reviewed in refs. 8 and 9). Stars indicate genes whose orientations were determined directly; all other gene orientations shown were inferred (see text). The tobacco *rps12* gene is split into two portions separated by 30 kb of DNA and numerous genes (18, 19). We probed for only the 3′ portion of this gene (Table 1), and therefore only this 3′ segment is represented in this and succeeding figures.

Table 1. Chloroplast gene probes

Gene*	Fragment		Ref.
	Size, kb	Endonuclease(s)	
3′ 23S rRNA	3.5	*Pst* I–*Sac* I	2
5′ 23S rRNA	3.5	*Sac* I	2
3′ 16S rRNA	1.3	*Sac* I	2
5′ 16S rRNA	2.7	*Sac* I	2
3′ *rps12*	1.0	*Pst* I–*Bam*HI	19
rps7	1.7	*Pst* I–*Sal* I	19
rpl2	0.772	*Xho* I–*Sal* I	20
rps19	0.7	*Sal* I–*Pst* I	20
infA	0.670	*Sal* I	21
rps11	0.635	*Sal* I–*Xba* I	21
3′ *petD*	0.416	*Bam*HI–*Xba* I	22
5′ *petD*	0.296	*Bam*HI	22
petB	2.4	*Sal* I–*Bam*HI	22
3′ *psbB*	1.597	*Bam*HI–*Sal* I	23
5′ *psbB*	0.338	*Bam*HI	23
5′ *psbE*	0.65	*Eco*RI	24
3′ *psbE–psbF*	0.50	*Eco*RI	24
3′ *petA*	1.1	*Bam*HI	25
5′ *petA*	0.9	*Hin*dIII–*Bam*HI	25
3′ *rbcl*	0.8	*Bam*HI–*Hin*dIII	26
rbcL	1.167	*Hin*dIII–*Pst* I	26
5′ *rbcL*	0.660	*Pst* I–*Xba* I	26
5′ *atpB*	1.171	*Xba* I–*Pst* I	27
atpE	0.420	*Eco*RI–*Xba* I	28
psaA	2.4	*Bam*HI	29
psaB	1.6	*Bam*HI	29
3′ *psbC*	0.367	*Bam*HI–*Pst* I	29
5′ *psbC*–3′ *psbD*	1.150	*Pst* I	30
5′ *psbD*	0.707	*Pst* I–*Bam*HI	30
rpoB	1.063	*Bam*HI	31
atpI–5′ atpH	2.2	*Hin*dIII–*Pst* I	32
3′ *atpH*	0.8	*Pst* I–*Bam*HI	32
atpF–5′ atpA	1.5	*Sal* I–*Hin*dIII	33
3′ *atpA*	0.9	*Hin*dIII–*Sal* I	33
rps16	1.9	*Hpa* II–*Nde* I	34
5′ *psbA*	0.532	*Eco*RI–*Pst* I	35
3′ *psbA*	1.2	*Pst* I–*Eco*RI	35

*Genes are listed in order of their position and orientation in spinach (see Figs. 3 and 5). See ref. 36 for details on source species, coding region coordinates, and methods of construction of gene probes.

tested genes in the geranium genome (Fig. 2). For nine of these genes, the differential hybridization of 5′ and 3′ probes also allowed us to assign the gene's orientation in the genome (Fig. 2). We indirectly inferred the orientation of most other genes on the basis of their known orientations in spinach, tobacco, and pea, their clustering in these species as part of polycistronic transcription units (7–9, 19–30, 32, 33), and their highly similar clustering in geranium.

Both rRNA genes and 13 of the 26 mapped protein genes are located within the inverted repeat in geranium and thus are fully duplicated (Fig. 2). In contrast, only three of the protein genes (*rpl2, rps7, rps12*) are duplicated in the ancestral angiosperm chloroplast genome, here approximated by spinach and tobacco (Fig. 3; refs. 1, 2, 4, 5, and 8). We therefore infer that the greatly enlarged inverted repeat in geranium is the result of spreading of a spinach-sized repeat into both the large and small single-copy regions, leading to duplication of the cluster of 10 genes extending from *rps19* through *rbcL* (Fig. 3).

The gene mapping experiments (Fig. 4) were particularly useful in locating the two boundaries between the large single-copy region and inverted repeat of geranium. A 1167-base-pair (bp) fragment (*rbcL* in Fig. 4) internal to the pea *rbcL* gene and a 660-bp fragment (5′ *rbcL*), containing the

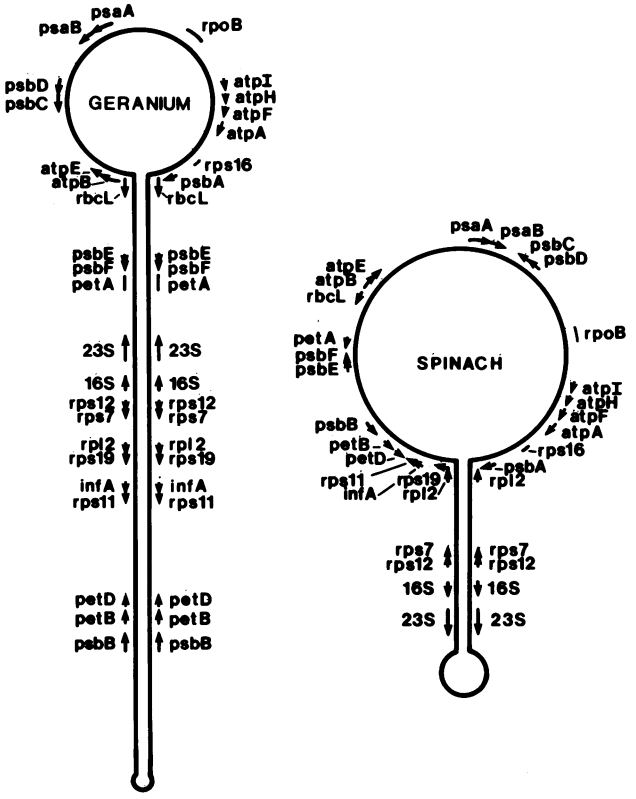

FIG. 3. Comparison of gene order and inverted repeat size in geranium and spinach cpDNAs. Spinach data are from refs. 20–24, 28, 29, and 33. *rps12*, *rps7*, *rps16*, and *rpoB* are not mapped in spinach; their positions are assigned on the basis of mapping data (19, 31, 34) for the colinear (1) genome of tobacco.

first 173 bp of the pea *rbcL* coding region and 487 bp of 5′ noncoding sequences, hybridize equally strongly to the large single-copy ends of the inverted repeat (Figs. 2 and 4). This strongly suggests that the entire *rbcL* coding region is duplicated in geranium. In contrast, a 1171-bp fragment (5′ *atpB*), which abuts the 660-bp 5′ *rbcL* pea fragment and

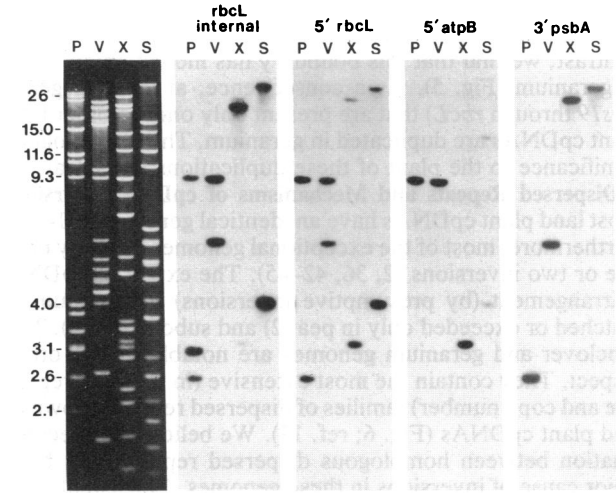

FIG. 4. Localization of the large single-copy ends of the geranium inverted repeat by gene mapping. Geranium cpDNA fragments produced by digestion with *Pst* I (P), *Pvu* II (V), *Xho* I (X), and *Sac* I (S) were separated on a 0.7% agarose gel (ethidium bromide staining is shown on the left), transferred to a filter, and hybridized with the indicated gene probes (Table 1) from pea (four autoradiograms on the right).

contains the first 989 bp of *atpB* and 182 bp of 5′ flanking sequences, hybridizes exclusively to single-copy sequences located just outside the geranium inverted repeat (Figs. 2 and 4). These results suggest that the left inverted repeat segment ends within the short [about 700 bp in most plants (26, 27)] spacer that separates the divergently oriented *rbcL* and *atpB* genes (Fig. 2). Similar gene hybridizations (Fig. 4) reveal that the boundary between the right inverted repeat segment and the large single-copy region is located within the 2-kb interval between *rbcL* and *psbA* (Fig. 2).

Comparison of the geranium and spinach genomes (Fig. 3) reveals not only differences in gene copy number but also many changes in gene location and orientation. For example, the geranium rRNA genes are transcribed toward the large single-copy region, in contrast to those in spinach and all other land plants. Fig. 5 summarizes these differences schematically in the form of an evolutionary model that postulates the fewest number of rearrangements necessary to derive the geranium genome from a spinach-like genome. These rearrangements are of three types: (*i*) expansion of the inverted repeat into both single-copy regions and the resulting duplication of a minimum of 10 protein genes; (*ii*) insertion of extra sequences within the inverted repeat (primarily between the small single-copy region and gene cluster 5, and between clusters 3–4 and 5), as spreading alone cannot account for the extra size of the geranium repeat relative to a spinach-type genome (Fig. 5); and (*iii*) no fewer than six inversions, occurring both in the large single-copy region and in the inverted repeat. None of these inversions appear to disrupt known transcriptional linkages. The step-wise portrayal of these rearrangements in Fig. 5 is intended only to clarify the nature of the individual mutations. We emphasize that we have no knowledge of whether these changes occurred in a concerted fashion, during one great eruption of genomic rearrangement, or whether they occurred individually in some unknown temporal order.

Dispersed Repeated Sequences in Geranium cpDNA. Certain cloned *Pst* I fragments of the geranium chloroplast genome hybridize only to themselves (e.g., fragments of 2.1, 3.8, and 4.0 kb; Fig. 6), while others hybridize to one or more additional *Pst* I fragments. For example, the cloned 26-kb *Pst* I fragment hybridizes to itself and to a *Pst* I fragment of 10.9 kb (Fig. 6). The reciprocal hybridization, of the cloned 10.9-kb fragment to the 26-kb fragment, confirms the existence of a region of homology between these two regions of the inverted repeat. A second family of short dispersed repeats is illustrated by the hybridization of a cloned 2.6-kb *Pst* I fragment to no fewer than three other *Pst* I fragments (Fig. 6). Detailed mapping experiments reveal that the first repeat family contains at least eight repeat elements, four of which are clustered within a 15-kb region in each of the large inverted repeat segments (Fig. 5). The nine members of the second repeat family have a more complex chromosomal distribution: three are clustered in a 15-kb region at the large single-copy end of each segment of the inverted repeat, and the other three are in a 15-kb interval within the large single-copy region.

DISCUSSION

Evolution of geranium cpDNA is exceptional in three ways relative to cpDNA evolution in other land plants. First, the inverted repeat has spread through adjacent single-copy sequences and tripled in size to 76 kb, compared to the typical angiosperm inverted repeat of 21–28 kb (16). Consequently, over half the chloroplast genome is duplicated in geranium. Second, gene order is highly scrambled in geranium as the result of at least six inversions, whereas most land plant cpDNAs have the same gene order (1–10). Third, two short sequences are repeated and dispersed to a number of chro-

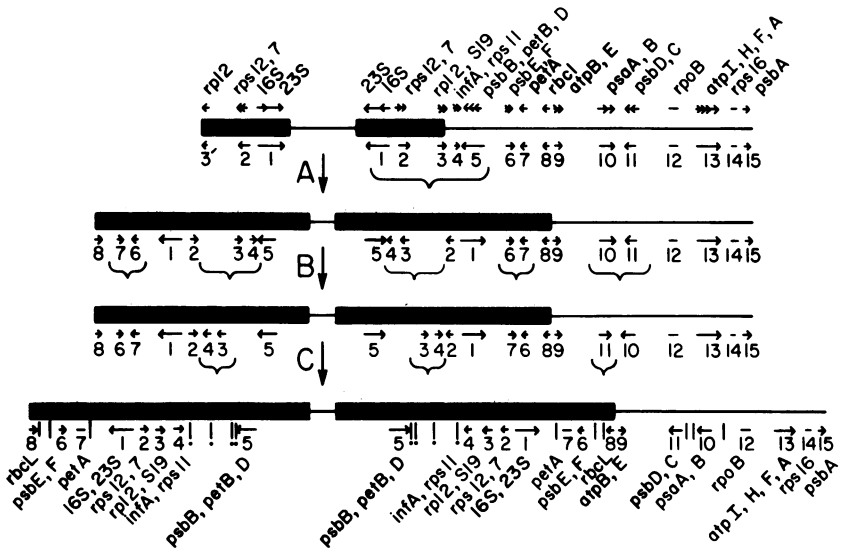

FIG. 5. Model for the evolution of geranium cpDNA (bottom map) from a spinach-like ancestral chloroplast genome (top map). Step A postulates spreading at both ends of the spinach inverted repeat and inversion of gene clusters 1–5. Step B postulates three inversions—of clusters 3 and 4 and flanking sequences, of clusters 6 and 7, and of clusters 10 and 11. Step C postulates two inversions—of clusters 3 and 4 and of cluster 11—and the addition of sequences to two regions of the inverted repeat. To simplify the diagram each arrow below each of the four maps represents a cluster of from one to four genes thought to be cotranscribed in spinach, tobacco, and pea. The vertical lines below the geranium map indicate the positions of elements belonging to the two families of dispersed repeats. The two families are distinguished by the presence or absence of dots below the lines.

mosomal locations, whereas most chloroplast genomes lack any detectable dispersed repeats (8, 9).

The extent of both duplication and inversion of cpDNA sequences observed in geranium is approached only among species in the green algal genus *Chlamydomonas* (8, 37, 38). However, the *Chlamydomonas* alterations involve very anciently diverged species, and probably did not take place as rapidly or as recently as the geranium changes (37).

Inverted Repeat Evolution. The expansion of the inverted repeat in geranium results in an overall genome size of 217 kb, far larger than the largest previously known land plant cpDNA (180 kb) (17). Including those legume cpDNAs that lack any inverted repeat (2, 11–14), the total range of land plant chloroplast genome sizes is almost 100 kb (120–217 kb; refs. 7–10 and 16) and the range of inverted repeat sizes is 76 kb (0–76 kb). Thus, much of the known size variation in cpDNAs results from expansion or contraction of the repeat, unaccompanied by any change in sequence complexity. Factoring out the repeat size variation, we find a range of sequence complexities of only 40 kb (110–150 kb) among land plant cpDNAs. In sharp contrast, angiosperm mitochondrial DNAs vary over 10-fold in absolute size and in sequence complexity (9), while plant nuclear genomes vary almost 1000-fold in size (39). These comparisons suggest, therefore, the existence of relatively strong constraints on chloroplast genome size. The nature of such constraints remains to be determined.

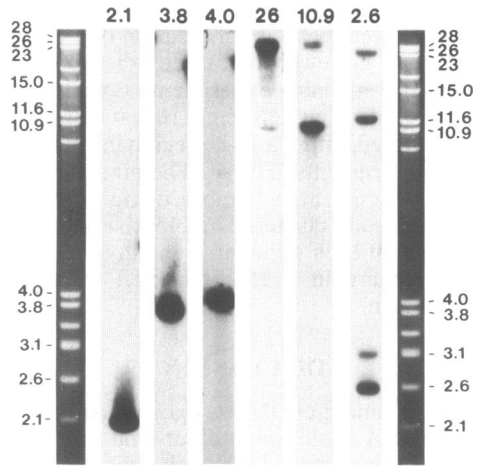

FIG. 6. Short dispersed repeats in the geranium chloroplast genome. Filters containing geranium *Pst* I fragments that had been separated on a 0.7% agarose gel were hybridized with each of six plasmid clones containing geranium *Pst* I fragments (sizes in kb of the *Pst* I inserts are given above each filter strip). Faint bands in the 4.0-kb hybridization lane represent partial digestion products that are 7.9 and 14.9 kb in size (cf. Fig. 2).

Expansion of the geranium inverted repeat is notable in that it obliterates a repeat–large single-copy junction that is rather fixed in most angiosperms. Sequence studies have shown that this junction occurs within or close to the *rps19* gene in five diverse dicots and monocots (20, 40, 41). In contrast, we find that this boundary has moved about 30 kb in geranium (Fig. 5). As a consequence, at least 10 genes (*rps19* through *rbcL*) that are present only once in other land plant cpDNAs are duplicated in geranium. The physiological significance to the plant of these duplications is unclear.

Dispersed Repeats and Mechanisms of cpDNA Inversion. Most land plant cpDNAs have an identical gene order (1–10). Furthermore, most of the exceptional genomes differ by only one or two inversions (2, 36, 42–45). The extent of cpDNA rearrangement (by presumptive inversions) in geranium is matched or exceeded only in pea (2) and subclover (13). The subclover and geranium genomes are notable in one other respect. They contain the most extensive (in terms of repeat size and copy number) families of dispersed repeats known in land plant cpDNAs (Fig. 6; ref. 13). We believe that recombination between homologous dispersed repeats may be a major cause of inversions in these genomes. Consistent with this hypothesis, most of the geranium repeats are located close to inversion endpoints (Fig. 5). This hypothesis can be tested more directly by studying the endpoints of recent inversions found among close relatives of geranium and subclover. We predict that short repeats will be found in an inverted orientation relative to one another at the ends of such inversions. We also point out that short repeats have

been found at the ends of two inversions in wheat cpDNA (43, 44).

If repeat-mediated recombination is the major mechanism of inversion in cpDNA, then the primary event in the destabilization of the geranium, subclover, and wheat genomes is probably their "invasion" by such repeats—i.e., the amplification and subsequent dispersal of repeat elements. How these repeats arise, and conversely, why they and their associated inversions are absent from most cpDNAs, are open questions. The general absence of repeats could reflect constraints on chloroplast genome size, as postulated in the preceding section.

We have shown that the geranium chloroplast genome possesses a truly remarkable structure relative to the highly constrained genomes of most land plants. Moreover, the geranium genome is more than just a curiosity, for it may provide the context in which to study mechanisms of rapid cpDNA rearrangement and inversion, the origin and consequences of dispersed repeats in cpDNA, and the effects of duplication and inversion on the expression and function of chloroplast genes.

Note Added in Proof. A description (46) of small dispersed repeats and their role in restructuring the wheat chloroplast genome appeared while this paper was in press.

We thank M. Sugiura and K. Ko for providing several of the gene clones used in this study, and M. Chase, R. Jansen, C. Makaroff, B. Milligan, and M. Zolan for critical reading of the manuscript. This research was supported by grants to J.D.P. from the U.S. Department of Agriculture (85-CRCR-1-1778) and the National Science Foundation (DCB-84-51660).

1. Fluhr, R. & Edelman, M. (1981) *Nucleic Acids Res.* **9,** 6841–6853.
2. Palmer, J. D. & Thompson, W. F. (1982) *Cell* **29,** 537–550.
3. Palmer, J. D., Singh, G. P. & Pillay, D. T. N. (1983) *Mol. Gen. Genet.* **190,** 13–19.
4. De Heij, H. T., Lustig, H., Moeskops, D. J. M., Bovenberg, W. A., Bisanz, C. & Groot, G. S. P. (1983) *Curr. Genet.* **7,** 1–6.
5. Palmer, J. D. & Stein, D. B. (1986) *Curr. Genet.* **10,** 823–833.
6. Gray, J. (1986) *Nature (London)* **322,** 501–502.
7. Whitfeld, P. R. & Bottomley, W. (1983) *Annu. Rev. Plant Physiol.* **34,** 279–310.
8. Palmer, J. D. (1985) *Annu. Rev. Genet.* **19,** 325–354.
9. Palmer, J. D. (1985) in *Monographs in Evolutionary Biology: Molecular Evolutionary Genetics*, ed. MacIntyre, R. J. (Plenum, New York), pp. 131–240.
10. Gillham, N. W., Boynton, J. E. & Harris, E. H. (1985) in *DNA Evolution: Natural Selection and Genome Size*, ed. Cavalier-Smith, T. (Wiley, New York), pp. 299–351.
11. Mubumbila, M., Crouse, E. J. & Weil, J. H. (1984) *Curr. Genet.* **8,** 379–385.
12. Palmer, J. D., Jorgensen, R. A. & Thompson, W. F. (1985) *Genetics* **109,** 195–213.
13. Palmer, J. D., Osorio, B., Aldrich, K. J. & Thompson, W. F. (1987) *Curr. Genet.* **11,** in press.
14. Michalowski, C., Breunig, K. & Bohnert, H. J. (1987) *Curr. Genet.* **11,** in press.
15. Palmer, J. D. (1986) *Methods Enzymol.* **118,** 167–186.
16. Crouse, E. J., Schmitt, J. M. & Bohnert, H. J. (1985) *Plant Mol. Biol. Rep.* **3,** 43–89.
17. Van Ee, J. H., Vos, Y. J. & Planta, R. J. (1980) *Gene* **12,** 191–200.
18. Torazawa, K., Hayashida, N., Obokata, J., Shinozaki, K. & Sugiura, M. (1986) *Nucleic Acids Res.* **14,** 3143.
19. Fromm, H., Edelman, M., Koller, B., Goloubinoff, P. & Galun, E. (1986) *Nucleic Acids Res.* **14,** 883–898.
20. Zurawski, G., Bottomley, W. & Whitfeld, P. R. (1984) *Nucleic Acids Res.* **12,** 6547–6558.
21. Sijben-Muller, G., Hallick, R. B., Alt, J., Westhoff, P. & Herrmann, R. G. (1986) *Nucleic Acids Res.* **14,** 1029–1044.
22. Heinemeyer, W., Alt, J. & Herrmann, R. G. (1984) *Curr. Genet.* **8,** 543–549.
23. Morris, J. & Herrmann, R. G. (1984) *Nucleic Acids Res.* **12,** 2837–2850.
24. Herrmann, R. G., Alt, J., Schiller, B., Widger, W. R. & Cramer, W. A. (1984) *FEBS Lett.* **176,** 239–244.
25. Willey, D. L., Auffret, A. D. & Gray, J. C. (1984) *Cell* **36,** 555–562.
26. Zurawski, G., Whitfeld, P. R. & Bottomley, W. (1986) *Nucleic Acids Res.* **14,** 3975.
27. Zurawski, G., Bottomley, W. & Whitfeld, P. R. (1986) *Nucleic Acids Res.* **14,** 3974.
28. Zurawski, G., Bottomley, W. & Whitfeld, P. R. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 6260–6264.
29. Alt, J., Morris, J., Westhoff, P. & Herrmann, R. G. (1984) *Curr. Genet.* **8,** 597–606.
30. Rasmussen, O. F., Bookjans, G., Stummann, B. M. & Henningsen, K. W. (1984) *Plant Mol. Biol.* **3,** 191–199.
31. Ohme, M., Tanaka, M., Chunwongse, J., Shinozaki, K. & Sugiura, M. (1986) *FEBS Lett.* **200,** 87–90.
32. Cozens, A. L., Walker, J. E., Phillips, A. L., Huttly, A. K. & Gray, J. C. (1986) *EMBO J.* **5,** 217–222.
33. Hennig, J. & Herrmann, R. G. (1986) *Mol. Gen. Genet.* **203,** 117–128.
34. Shinozaki, K., Deno, H., Sugita, M., Kuramitsu, S. & Sugiura, M. (1986) *Mol. Gen. Genet.* **202,** 1–5.
35. Oishi, K. K., Shapiro, D. R. & Tewari, K. K. (1984) *Mol. Cell. Biol.* **4,** 2556–2563.
36. Jansen, R. K. & Palmer, J. D. (1987) *Curr. Genet.* **11,** in press.
37. Lemieux, B. & Lemieux, C. (1985) *Curr. Genet.* **10,** 213–219.
38. Lemieux, C., Turmel, M., Lee, R. W. & Bellemare, G. (1985) *Plant Mol. Biol.* **5,** 77–84.
39. Bennett, M. D. & Smith, J. B. (1976) *Philos. Trans. R. Soc. London Ser. B* **274,** 227–274.
40. Spielmann, A. & Stutz, E. (1983) *Nucleic Acids Res.* **11,** 7157–7167.
41. Posno, M., Torenvliet, D. J., Lustig, H., van Noort, M. & Groot, G. S. P. (1985) *Curr. Genet.* **9,** 211–219.
42. Herrmann, R. G., Westhoff, P., Alt, J., Winter, P., Tittgen, J., Bisanz, C., Sears, B. B., Nelson, N., Hurt, E., Hauska, G., Viebrock, A. & Sebald, W. (1983) in *Structure and Function of Plant Genomes*, eds. Cifferi, O. & Dure, L., III (Plenum, New York), pp. 143–153.
43. Howe, C. J. (1985) *Curr. Genet.* **10,** 139–145.
44. Quigley, F. & Weil, J. H. (1985) *Curr. Genet.* **9,** 495–503.
45. Hirai, A., Ishibashi, T., Morikami, A., Iwatsuki, N., Shinozaki, K. & Sugiura, M. (1985) *Theor. Appl. Genet.* **70,** 117–122.
46. Bowman, C. M. & Dyer, T. A. (1986) *Curr. Genet.* **10,** 931–941.