Chloroplast DNA variation in the Geraniaceae - a preliminary report

Robert A. Price, Patrick J. Calie*, Stephen R. Downie, John M. Logsdon, Jr., and Jeffrey D. Palmer

Department of Biology, Indiana University, Bloomington, Indiana 47405 U. S. A.
*Department of Microbiology, University of Michigan Medical School, Ann Arbor, Michigan 48109 U. S. A.

Abstract

Restriction enzyme analyses of chloroplast DNA from selected representatives of the five genera of Geraniaceae s. s. reveal that the family is perhaps the most variable group of angiosperms in chloroplast genome size and structure. Among the photosynthetic flowering plants the largest and smallest chloroplast genomes known to date are found in Pelargonium x hortorum and Erodium chamaedryoides, respectively. Unusual events seen in the family include substantial increase in the size of the large, ribosomal RNA-encoding inverted repeat in Pelargonium and Geranium, loss of one copy of the inverted repeat in Erodium and Sarcocaulon, multiple inversions, loss of highly conserved introns, and loss or inactivation of several genes. The Geraniaceae are supported as a monophyletic group by the shared loss of the rpl16 intron as well as by sequence data from the rpoA gene. RpoA sequence data and the further loss of the rpl2 intron support a sister-group relationship between Monsonia and Sarcocaulon. An extensive restriction site mapping study of the chloroplast genome in the family is now being conducted to further explore structural variability and to provide a detailed phylogenetic framework. Greatest emphasis will be placed on relationships within Pelargonium, where approximately 60 species representing each of the currently or previously recognized sections are being compared. RbcL sequence comparisons will also be conducted to assess the position of the family relative to other families of the subclass Rosidae.

Introduction

The chloroplast genome has been found to be highly conserved in size and structure in a wide variety of photosynthetic flowering plants and exhibits a range in rates of point mutation or structural rearrangement that makes it particularly valuable in phylogenetic comparisons

both among and within families (Palmer et al., 1988; Downie and Palmer, 1991). Comparisons of restriction site polymorphisms in chloroplast DNA (cp DNA) are now being widely used in evolutionary studies within families or genera of angiosperms, while sequence comparisons of the relatively conservative rbcL gene are being extensively used to examine relationships among families of flowering plants and gymnosperms. Both restriction mapping and gene sequencing will be applied in our studies of the Geraniaceae in order to examine relationships within and among genera and also to establish which other families are most closely related to the group.

Preliminary restriction site mapping studies of cp DNA have been conducted in selected taxa of four of the five genera of Geraniaceae s. s. (Erodium chamaedryoides, Geranium grandiflorum, Pelargonium x hortorum, and Sarcocaulon vanderietiae). Limited information on gene or intron losses is also available for a broader sample of species of these genera and for two species of Monsonia. Restriction enzyme analyses have revealed that the family spans the range from the smallest to the largest chloroplast genomes of any photosynthetic flowering plants studied to date and vary greatly in the degree of rearrangement of the genome relative to the standard pattern exemplified by spinach and Nicotiana tabacum (Figure 1).

Genome size and structural rearrangements

Flowering plant cp DNAs are circular molecules and usually have a size of approximately 140 to 160 kilobase pairs (kb), including a ribosomal RNA-encoding inverted repeat region typically of 20 to 30 kb (e.g., spinach in Fig. 1; Palmer, 1985; Downie and Palmer, 1991). In contrast, the chloroplast genome of Erodium chamaedryoides has lost one copy of the inverted repeat and is reduced to a size of approximately 115 kb (Price et al., unpublished data), while that of Pelargonium x hortorum has the largest known inverted repeat of any species of flowering plant (ca. 76 kb) and has an overall size of about 217 kb (Fig. 1; Palmer et al., 1987a). The chloroplast genome of Geranium grandiflorum also has an expanded inverted repeat and a total size of roughly 200 kb and preliminary examination of restriction fragments from purified chloroplast DNA indicates that a similarly large genome is also found in representatives of several sections of the genus Pelargonium. The chloroplast genome of Sarcocaulon vanderietiae is more similar to that of E. chamaedryoides in having lost all or almost all of the inverted repeat and has a size in the range of 120-130 kb.

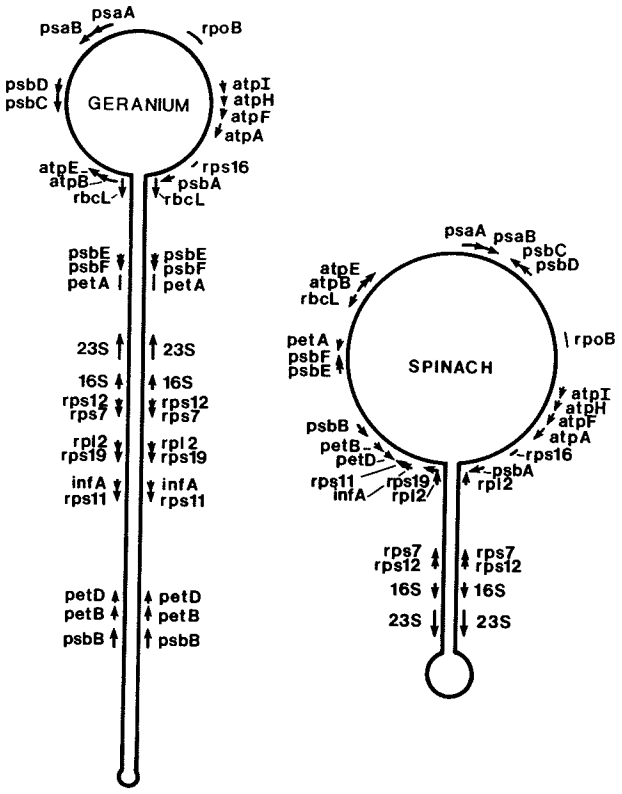

Figure 1. Comparison of gene order and inverted repeat size in the 217 kb and 150 kb chloroplast genomes of garden geranium (Pelargonium x hortorum) and spinach, respectively. Note the much expanded inverted repeat region (shown as a paired segment in the center portion of the circular genome) and correspondingly smaller large single copy region (circle at top) and small single copy region (circle at base). The gene order in Pelargonium is much rearranged relative to the standard arrangements in spinach and tobacco due to a series of six or more inversions. Reproduced with permission from Palmer et al., 1987a.

Both loss and substantial augmentation of the inverted repeat are very infrequent events among the land plants and thus the simplest hypothesis is that each occurred only once in the

evolutionary history of the family. Loss of the inverted repeat is otherwise known only from a major subset of the papilionoid legumes (Lavin et al., 1990), from all families of conifers (Strauss et al., 1988; Raubeson and Jansen, 1989), and from some parasitic members of the Scrophulariaceae and Orobanchaceae. Species in several families of flowering plants have enlarged chloroplast genomes of approximately 170-180 kb (see review in Downie and Palmer, 1991), but no other taxa studied to date approach the genome size seen in Geranium and Pelargonium. The putatively closely related outgroup taxon Oxalis oregana (Oxalidaceae) has a typical chloroplast genome of approximately 152 kb in size and thus our working hypothesis is that there are two major lineages in the Geraniaceae, one including Erodium, Sarcocaulon, and presumably, Monsonia, and the second including Geranium and Pelargonium. However, preliminary phylogenetic comparisons of rpoA gene sequence data have indicated that Monsonia emarginata and Sarcocaulon vanderietiae form a clade at least as closely related to Geranium grandiflorum and G. palmatum as they are to Erodium chamaedryoides, so more data are needed to provide a well supported phylogeny at the generic level.

Major rearrangements in the chloroplast chromosome are relatively infrequent events among the land plants, and inversions relative to the gene order in tobacco are useful evolutionary markers (Palmer, 1985, 1991; Downie and Palmer, 1991). Inversions appear to be much more frequent in the Geraniaceae than in most other families of flowering plants, a situation which may relate to destabilization following loss or alteration of the inverted repeat (Palmer et al., 1987b; Strauss et al., 1988). The genome most similar to that of tobacco among those studied in the Geraniaceae is that of Erodium chamaedryoides, which has one major inversion of about 20 kb in addition to the loss of the inverted repeat. Detailed restriction mapping of Pelargonium x hortorum has demonstrated the occurrence of at least six inversion events in the chloroplast chromosome relative to the standard tobacco pattern (Palmer et al., 1987a). These inversions are also unusual in occurring in the area of the inverted repeat as well as the large single copy region. Preliminary mapping of the chloroplast genomes of Geranium grandiflorum and Sarcocaulon vanderietiae also indicate the presence of at least two major inversions in both cases.

Intron and gene losses

The occurrence of some 20 introns (intervening sequences within gene coding regions) identified in the completely sequenced chloroplast genome of tobacco is a highly stable

feature in the flowering plants, so shared intron losses are very useful in defining monophyletic groups (Downie and Palmer, 1991). The five genera of the Geraniaceae s. s. share the loss of the intron in the ribosomal protein gene rpl16, while the intron is present in virtually all flowering plant families studied to date, including the putatively closely related Oxalis (Downie and Palmer, 1991). This strongly supports the family as a monophyletic group, in accord with evidence from the structure of the gynoecium and fruit. It will be of considerable interest to examine other genera with possible affinities to the Geraniaceae in this regard (e.g., the South American genera Hypseocharis, Balbisia, and Viviania), as well as to use sequence information from a gene such as rbcL. Within the Geraniaceae, the genera Monsonia and Sarcocaulon, which appear to be closely related on the basis of floral morphology (Venter, 1979), share the loss of the rpl2 intron, which is present in the remaining three genera of the family and all putatively closely related families (Downie et al., 1991).

Also of phylogenetic interest is the loss of the intact form of the RNA polymerase gene rpoA from the chloroplast, which apparently occurred early in the evolution of the genus Pelargonium. All Pelargonium species studied to date (including representatives of all of the sections) fail to yield hybridization signals on Southern blots probed with the intact rpoA gene from Geranium grandiflorum, which hybridizes strongly to representatives of all of the other genera. Several other genes have also apparently been lost from the chloroplast genome in Sarcocaulon or Geranium and are the subject of ongoing study.

Southern hybridization/restriction mapping studies

DNA samples have been isolated from over 60 species of Pelargonium, including members of each of the currently or previously recognized sections (Van der Walt and Vorster, 1988), in an effort to assess the delimitation and relationships of groups within the genus. Detailed restriction site analyses using 10 or more restriction endonucleases will be used in phylogenetic studies of the genus. Initial studies will involve mapping of the relatively unrearranged portion of the large single copy region. More extensive mapping will be conducted using representatives of most of the sections to assess the extent of differences in the structural arrangement of the chloroplast genome. Several members of each of the other genera of the family as well as putatively related taxa such as Oxalis and Averrhoa (Oxalidaceae), Limnanthes (Limnanthaceae), and Tropaeolum (Tropaeolaceae) will also be included in each of the surveys as outgroups.

Preliminary results from Southern blot hybridization appear to support expanding the delimitation of section Ciconium to include P. peltatum and much of section Eumorpha as suggested by Van der Walt and Vorster (1988). Two other species (P. aridum and P. barklyi) often placed in section Ligularia, but related to section Ciconium by their chromosome base number of x = 9 (Gibby and Westfold, 1986; Yu and Horn, 1988) also show stronger molecular affinities with section Ciconium than with other representatives of section Ligularia.

Sequence studies

Comparative sequence information from the rpoA gene is currently being obtained from representatives of each of the genera of the Geraniaceae and the outgroups Oxalis and Limnanthes. Preliminary sequences from over half of the length of the gene strongly support Erodium, Geranium, Monsonia, and Sarcocaulon as a monophyletic group (Pelargonium cannot be directly compared due to loss of the gene from the chloroplast) and also strongly support Monsonia and Sarcocaulon as sister taxa. Particular patterns of relationship among the other genera are only weakly supported by the current data set. Evolution of the gene within the family is also unusual in that the rate of change is much greater than that found in comparing other taxa, such as spinach and tobacco. Sequence comparisons of the more conservative gene rbcL will be used to examine the phylogenetic position of the Geraniaceae within the Rosidae, as well as to provide a check on the relative rate of sequence evolution within the Geraniaceae to see if it is generally elevated.

Acknowledgments

We are grateful to Ernest De Marie of Cornell University, who supplied us with living material of over 70 species of Pelargonium and other species of Geraniaceae along with useful discussions of the systematics of the genus, and also to Robin Parer, Harold Koopowitz, and John Trager for their help in obtaining other plant material for extraction. This research was supported by grants to J. D. P. from the NIH (GM-35087) and NSF (BSR-8717600).

Literature Cited

Downie, S. R., R. G. Olmstead, G. Zurawski, D. E. Soltis, P. S. Soltis, J. C. Watson, and J. D. Palmer. 1991. Loss of the chloroplast DNA rpl2 intron demarcates five lineages of dicotyledons: molecular and phylogenetic implications. submitted to Evolution.

Downie, S. R., and J. D. Palmer. 1991. Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. In P. Soltis, D. Soltis, and J. Doyle, eds. Plant molecular systematics. Chapman and Hall, New York.

Gibby, M., and J. Westfold. 1986. A cytological study of Pelargonium sect. Eumorpha (Geraniaceae). Pl. Syst. Evol. 153: 205-222.

Lavin, M., J. J. Doyle, and J. D. Palmer. 1990. Evolutionary significance of the loss of the chloroplast-DNA inverted repeat in the Leguminosae subfamily Papilionoideae. Evolution 44: 390-402.

Palmer, J. D. 1985. Comparative organization of chloroplast genomes. Ann. Rev. Genet. 19: 325-354.

Palmer, J. D. 1991. Plastid chromosomes: structure and evolution. In L. Bogorad and I. K. Vasil, eds., Cell culture and somatic cell genetics in plants: the molecular biology of plastids. Academic Press, Orlando.

Palmer, J. D., R. K. Jansen, H. J. Michaels, M. W. Chase, and J. R. Manhart. 1988. Chloroplast DNA variation and plant phylogeny. Annals Missouri Bot. Gard. 75: 1180-1206.

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

Raubeson, L. A., and R. K. Jansen. 1989. Molecular evidence on conifer phylogeny: structural variation in the chloroplast genome. (Abstr.) Amer. J. Bot. 76(6), Suppl.: 222.

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

Van der Walt, J. J. A., and P. J. Vorster. 1988. Pelargoniums of Southern Africa, vol. 3. Cape Town.

Venter, H. J. T. 1979. A monograph of Monsonia L. (Geraniaceae). Meded. Landbouwhog. Wageningen Nederland 79(9): 1-128.

Yu, S.-N., and W. A. H. Horn. 1988. Additional chromosome numbers in Pelargonium (Geraniaceae). Pl. Syst. Evol. 159: 165-171.