American Fern Journal 90(2):51–59 (2000)

The Chloroplast Genome Structure of the Vascular Plant *Isoëtes* is Similar to That of the Liverwort *Marchantia*

R. JOEL DUFF
Department of Biology, University of Akron, Akron, OH 44315-3908
EDWARD E. SCHILLING
Department of Botany, University of Tennessee, Knoxville, TN 37996-1100

ABSTRACT.—Restriction site mapping was used to characterize the chloroplast genome of the lycophyte, *Isoëtes melanopoda*. The *Isoëtes* chloroplast genome is approximately 139–145 kb in size with an inverted repeat of 12–13 kb. The gene content and consensus gene order are similar to that of *Marchantia*. A distinctive feature of the *Isoëtes* genome is an increased size of the small single copy (SSC) region possibly due to the insertion of a piece of DNA (3–8 kb) of unknown composition. The inferred insertion, along with a slightly larger inverted repeat, are responsible for the apparent size difference in the total genome relative to *Marchantia*. Patterns of restriction fragments were also consistent with the presence of a small inversion (2–3 kb) in the large single copy (LSC) region.

Structural differences in chloroplast genomes have proven to be powerful characters in understanding relationships among land plants because of their rarity and corresponding apparent low levels of homoplasy (Palmer, 1985a, b, 1987, 1991; Palmer and Stein, 1986; Palmer et al., 1988; Manhart and Palmer, 1990; Raubeson and Jansen, 1992; Downie and Palmer, 1992; Doyle, 1992; Lew and Manhart, 1993; Raubeson and Stein, 1995). The chloroplast genome exhibits a remarkably consistent structure and gene order and content among a wide range of members of the plant kingdom (reviewed in: Stein et al., 1986; Palmer et al., 1988; Olmstead and Palmer, 1994). Among photosynthetic plants this circular molecule is usually 120–160 kilobases (kb) in length and has a capacity to code for approximately 120 genes.

Our understanding of the chloroplast genome comes primarily from angiosperms, but little has been reported about the size, overall structure, or variation of the chloroplast genomes of lower vascular plants (Lycophyta, Psilotophyta, and Sphenopsida) or most of the bryophytes. Among the lower vascular plants only the presence or absence of a 30 kb inversion (Raubeson and Jansen, 1992) has been demonstrated. Among the pteridophytes Stein et al. (1992) and Conant, et al. (1994) have shown that multiple structural rearrangements exist among the diverse groups of ferns and that these structural features may be useful phylogenetic markers at the familial and higher levels. More information about the genomes of the primitive land plants will be needed to attain a comprehensive understanding of the chloroplast genome.

In this study we provide the first complete restriction site maps of the chloroplast genome of a lycophyte, the quillwort *Isoëtes melanopoda*. We find that the *Isoëtes* chloroplast genome shares significant features with those of two

non-vascular plants, the bryophytes *Marchantia* (liverwort) and *Physcomitrella* (moss). The *Isoëtes* chloroplast genome has the following features: (1) its inverted repeat (IR) is several kb larger than the IR of either of the bryophytes but is significantly smaller than the IR of most ferns and seed plants, (2) the small single copy region (SSC) has a large (3–8 kb) region of extra genetic material of unknown origin, and (3) the large single copy region (LSC) appears to have several small inversions (2–4 kb).

Materials and Methods

Leaves of *Isoëtes melanopoda* Gay & Durieu (Isoetaceae) were collected from a single population in Alabama and a voucher specimen (*Duff 9201*) was deposited at the University of Tennessee Herbarium. Total DNA was isolated using the procedure of Doyle and Doyle (1987). Single digests of sixteen restriction endonucleases (*Bam*HI, *Ban*I, *Ban*II, *Dra*I, *Eco*RI, *Eco*RV, *Hae*II, *Hind*III, *Nci*I, *Nco*I, *Pst*I, *Pvu*II, *Sac*I, *Sal*I, *Stu*I, and *Xho*I) and selected double digests (*Ban*I/*Ban*II, *Eco*RI/*Eco*RV, *Pst*I/*Sal*I, *Sac*I/*Pvu*II, and *Stu*I/*Xho*I) were made and the fragments separated on 0.9% agarose gels run out 15 cms and transferred by dry blotting to Amersham (Hybond N+) nylon membranes. Cloned cpDNA fragments from lettuce (Jansen and Palmer, 1987) and tobacco (Olmstead and Palmer, 1992; Shinozaki et al., 1986) were used as probes for physical mapping and gene localization. Membrane-bound DNAs were hybridized to ^{32}P-dCTP labeled probes using random primer oligolabeling (Feinberg and Vogelstein, 1983, 1984; Gibco BRL labeling kit) for 24 hr at 55° C. Hybridization buffers and conditions were used as described by the manufacturer (Gibco BRL) except that hybridizations were done at 55° C. Tobacco DNA cut with *Bam*HI provided a control lane to verify the identity of probes. Fragment sizes were estimated by comparison with fragments from *Hind*III digestion of Lambda Phage. Mapping followed the general strategies outlined in Palmer (1982, 1986) and Jansen and Palmer (1987).

Results

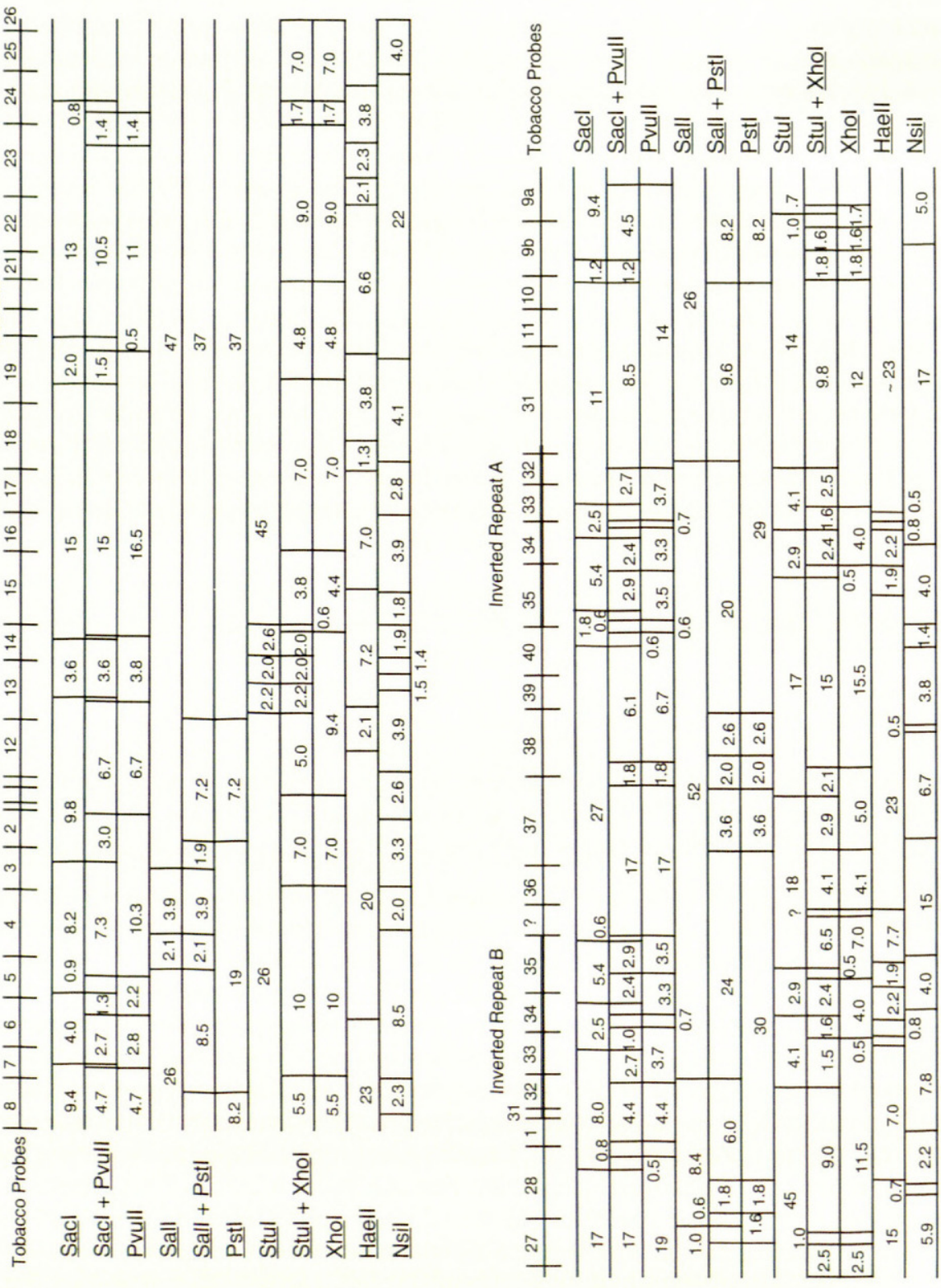
Hybridization to both lettuce and tobacco probes generally gave good results although fine scale mapping of several regions of the *Isoëtes* cpDNA was difficult because there was only very limited hybridization to several tobacco probes. As a result it was possible to generate complete maps only for restriction enzymes that cut the DNA into large fragments spanning these regions. The problematic regions are those covered by tobacco probes 3, 20a, 21, 29, and 30, each of which contain at least some gene sequences not found in the genome of *Marchantia* (Ohyama et al. 1986). Stein et al. (1992) in studies of *Adiantum* also reported a lack of hybridization with tobacco probe 3, which contains the gene *rps*16 that is not found in *Marchantia*. Probes 20b, 21, 29, and 30 account for several open reading frames (ORFs) that are of comparable sizes in *Marchantia*.

Complete restriction site maps for eight single restriction digests and three

double digests of *Isoëtes melanopoda* were successfully generated (Fig. 1). Partial maps for seven additional enzymes were obtained for a great majority of the genome including the inverted repeat and small single copy region, and were utilized in determining portions of the genome structure. Completed restriction site maps gave total chloroplast genome size estimates for *Isoëtes* of 139–145 kb with an average of 141 kb. The estimates for the minimum size of the inverted repeat varied from 11.8 (*Eco*RI + *Eco*RV) to 13.2 kb (*Nsi*I). The size of the large single copy region (LSC) was approximately 85 kb and for the small single copy region (SSC) was 24–29 kb. Consistent with the results of Raubeson and Jansen (1992), the *Isoetes* genome exhibits the presumed ancestral gene order exhibited in bryophytes and lycophtes. Evidence of this genome architecture came from the fact that two non-adjacent pairs of tobacco probes, 31–11 and 2–12, respectively, consistently hybridized to overlapping fragments (Fig. 1). A large insertion in the SSC, relative to both *Marchantia* and tobacco, was inferred from mapping studies and appears to be a feature unique to *Isoëtes*. An additional feature of the *Isoëtes* genome was the apparent presence of a small inversion (1.5–3.0 kb) found in the large single copy region (LSC) as well as several other smaller inversions which may be postulated from individual restriction site maps but cannot be characterized more completely due to the resolution of the current data set.

INVERTED REPEAT.—The size of the inverted repeat in *Isoëtes* (11.8–13.2 kb) is larger than that of *Marchantia* (10 kb) and *Physcomitrella* (9.4 kb). Based on hybridization data, this appears to be due to the inclusion in the IR of *Isoëtes* of regions homologous to those of tobacco fragment 32 (*rps'*12, *rps*7) and a very small portion of probe 31, which are found in the IR of tobacco but are restricted to the large single copy region adjacent to IR_A in *Marchantia* (Ohyama et al., 1986). Just as in *Marchantia*, regions of cpDNA homologous to tobacco probes 28, 29, 30, 36 and 1, all of which are part of the IR of tobacco, were mapped to the large single copy region of the *Isoëtes* chloroplast genome.

SMALL SINGLE COPY REGION.—The single copy region was estimated to be 24–29 kb in size. The variation in this estimate was due to lack of hybridization and presence of genetic material in *Isoëtes* not represented in the tobacco probes resulting in difficulties in resolving the boundaries of the region. The lack of hybridization to several probes combined with the size of spanning fragments made precise estimations of the size of the SSC and the entire genome difficult. Figure 1 shows that the amount of DNA in the area adjacent to the edge of IR_B was more than could be accounted for from the sizes of the tobacco probes used and than its expected content compared with the *Marchantia* genome (Ohyama et al., 1986). For example, a 17 kb chloroplast DNA fragment, the result of digestion by *Pvu*II, only hybridized to probes 36, 37, and very weakly to 35 which account for a maximum of 10 kb of DNA in tobacco. The best estimate of total size of the SSC came from the map of *Sac*I. This enzyme yielded only two fragments that span the SSC; a 23–27 kb fragment and a 1.8 kb fragment each of which hybridized to probe 35. For the same enzyme two 5.4 kb fragments hybridized strongly to probe 35 and very


Tobacco Probes
SacI
SacI + PvuII
PvuII
SalI
SalI + PstI
PstI
StuI
StuI + XhoI
XhoI
HaeII
NsiI
Inverted Repeat A
Inverted Repeat B

weakly to probe 34 indicating that the 5.4 kb fragment accounts for the majority of the 4.7 kb tobacco probe 35. The approximately 27 kb and 1.8 kb fragments that weakly hybridize to probe 35 must lie primarily within the SSC. This gives an upper estimate of the SSC of approximately 29 kb. Considering maps of more than 10 enzymes completed for the SSC, an estimate of 24–29 kb can be made. Compared with *Marchantia* (20 kb SSC), *Physcomitrella* (21 kb SSC, Calie and Hughes, 1986), and *Adiantum* (20–22 kb SSC, Hasebe and Iwatsuki, 1990), this would place the amount of extra DNA in the *Isoëtes* SSC at 3–8 kb.

ABSENCE OF 30 KB INVERSION.—Based on the presence of fragments from *Bam*HI and *Hind*III digests that show overlap to tobacco probes 31 and 10, and 12 and 2, Raubeson and Jansen (1992) report the absence of a 30 kb inversion found in ferns and seed plants. Our data supported the absence of this inversion in these higher plants. Figure 1 shows that genetic material homologous to tobacco probes 29, 30 and a portion of probe 1 were detected between fragments homologous to probes 2 and 12. Tobacco probe 1 hybridized to fragments in two areas: between probes 2 and 29 and between probes 28 and 32 at the edge of IR_B. This is what would be expected were the genome of *Isoëtes* to have the same arrangement as *Marchantia*. In *Marchantia* the transfer RNA, H(GUG), found in probe 1 of tobacco, is present next to probe 2 but the remaining portion of probe 1; *trn*I(CAU), *rpl*23, *rpl*2, can be found in the LSC at the edge of IR_B in *Marchantia*. The presence of very weak hybridization to probe 1 adjacent to probe 2 suggested that the arrangement of genes in *Isoëtes,* in the region of this 30 kb inversion in ferns and seed plants, is identical to that in *Marchantia*.

ADDITIONAL INVERSIONS.—A small (1–3 kb) inversion was detected in the LSC but its precise size was difficult to determine. The restriction enzymes used produced DNA fragments that were too large to allow localization of the endpoints of the inversion. The presence of the inversion was supported by the detection of multiple fragments that hybridized identically to both tobacco fragments 14 and 15. Additional, smaller inversions may be hypothesized from several individual restriction maps. These inversions could not be accurately characterized because they are located in a portion of the genome for which there was poor hybridization signal. This region corresponds to the position of a large open reading frame (ORF2136 = tobacco fragments 29,30). Only

←

FIG. 1. Linearized restriction site maps of the chloroplast genome of *Isoëtes melanopoda*. Enzymes used in single or double digestions to map the genome are indicated at each end of the map. The top line represents the tobacco probe set (Olmstead and Palmer, 1992) in positions relative to their hybridization with *Isoëtes* cpDNA. Large solid bars indicate relative position and boundaries of the inverted repeat in *Isoëtes*. Asterisks denote tobacco probes for which very week hybridization was observed to *Isoëtes* fragments. Fragments less than 0.5 kb were not consistently observed and therefore are postulated based on data from double digests. Fragment sizes larger than 18 kb could not be accurately calculated and so reflect estimations based on additive sizes of fragments detected by double digests.

fragments spanning the entire region were utilized in the mapping data presented here. Raubeson and Jansen (1992) also report the probable occurrence of up to two small inversions in this region as well, but did not characterize them further.

Discussion

Although *Isoëtes* is a highly distinctive genus, whose evolutionary lineage is separated from the majority of vascular plants by up to 300 million years and from the non-vascular plants for as long or longer, it exhibits a chloroplast genome remarkably similar to the liverwort *Marchantia*. Most vascular plants appear to have significantly larger chloroplast genomes than the bryophytes because the inverted repeat is larger. The *Isoëtes* cpDNA genome is also larger than the bryophyte genome in size but the difference is the result of an increased size of the SSC relative to other characterized cpDNAs and to a lesser extent an increase in the size of the IR. The overall size, including the IR, and structure of the chloroplast genome of *Selaginella* (Duff, unpublished data) also appears to be very similar to that of *Marchantia*.

One possible explanation for the remarkable similarity in size and structure of the genomes of *Marchantia, Physcomitrella,* and *Isoëtes* is that the absence of the 30 kb inversion confers some structural integrity to these molecules. The 30 kb inversion that took place in the ancestor of the ferns and seed plants appears to be correlated with a relaxation of physical constraints on the size of the inverted repeat region. Hence the IR subsequently underwent rapidly expand to incorporate genetic material from the LSC in the vast majority of ferns and seed plants (Palmer and Thompson 1982; Jansen and Palmer 1987; Howe et al. 1988). Stein et al. (1992) has proposed that the region around the *psb*A gene is especially prone to recombination. In *Marchantia, Physcomitrella,* and *Isoëtes* this gene is found far from the IR, unlike those plants lacking the inversion, and thus one would expect its potential effects on the structural integrity of the genome not to affect the IR. These taxa are by all accounts highly divergent and yet the sizes of their respective inverted repeats are not significantly different nor have they been demonstrated to have undergone inversion events to the extent observed in plants which contain the 30 kb inversion relative to *Marchantia*. The implication, although a correlation only at this point, is that the similarity in the chloroplast genomes of *Marchantia, Physcomitrella,* and *Isoëtes* is directly related to their gene order.

The variation in the estimated size of SSC (24–29 kb) in *Isoëtes* was unexpected and exemplifies the difficulty in resolving the boundaries of the SSC with the inverted repeats and the lack of probe homologies in this region. For example, it has been shown that the *chl*L gene has been lost from angiosperm cpDNAs although it is present in *Isoëtes* as well as all bryophytes, lycophytes, ferns except *Psilotum,* and gymnosperms except *Welwitschia* (Burke et al., 1993). This gene is typically located in the SSC adjacent to the IR_A though the restriction map generated for *Isoëtes* does not clearly define this region due to the present of several large fragments spanning this region. Furthermore, the

tobacco probes used in constructing the map lack this gene region and thus it is possible that several small fragments generated by *Nsi*I, *Sac*I, and *Pvu*II could have escaped detection in this region between probes 40 and 35. Even accounting for these sources of uncertainty the total size of the SSC appears to be increased over that of all other previously characterized land plant chloroplast genomes. Furthermore this increase in DNA content is most apparent in the fragments that hybridized to tobacco probe 35, found in the IR of both tobacco and *Isoëtes,* and tobacco probe 36. The latter represents genetic material found in the IR of tobacco but whose homologous content can be found in the SSC of *Marchantia* adjacent to IR_B. This extra DNA had insufficient sequence similarity with any portion of the tobacco or lettuce genomes to produce hybridization signals and its presence was only determined by the presence of fragments that spanned the entire region (see *Stu*I, *Sal*I, *Pst*I in Fig. 1). This DNA found in the *Isoëtes* SSC may be the result of either an increased size of spacer regions between the genes or may be due to an insertion of foreign DNA. Very few definitive cases of extra, non-homologous DNA found in the chloroplast genome are known. Conant et al. (1994) report the presence of about 0.9 kb of DNA in the IR of *Cyathea furfuracea* (Cyatheaceae) for which there is no counterpart in the *Adiantum* genome and thus cannot be explained by duplication or inversion events. A more dramatic example can be seen in the unusual choloroplast genome of *Trachelium* (Campanulaceae) in which as many as seven insertions of foreign DNA have been postulated (Cosner et al. 1997). In any case, more detailed analyses will need to be undertaken to confirm the existence and to determine the nature of any additional sequences in the *Isoëtes* chloroplast genome.

In summary, we have established that the genome of *Isoëtes* has the overall structure of the bryophytes, *Marchantia* and *Physcomitrella.* The only significant difference between the bryophyte genome and that of *Isoëtes* is the possible increase in DNA content of the SSC region of the *Isoëtes* genome and increased IR content. Further work to characterize the insertion in *Isoëtes* and to map the genomes of *Lycopodium* sens. lat., *Equisetum,* and *Psilotum* will most likely result in a better understanding of the evolution of the chloroplast genome and possibly provide clues regarding the relationships of the lower vascular plants.

ACKNOWLEDGMENTS

This paper represents a portion of a Ph.D. dissertation submitted to the University of Tennessee. We thank Diana Stein and Linda Raubeson for their advice and help on the mapping of the genome and the comments of one anonymous reviewer. This research was supported by National Science Foundation Doctoral Dissertation Improvement Grant DEB-9321767, and grants from the A. J. Sharp Fund, Department of Botany, University of Tennessee.

LITERATURE CITED

BURKE, D. H., L. RAUBESON, M. ALBERTI, J. HEARTST, E. JORDAN, S. KIRCH, A. VALISNKI, D. CONANT, and D. STEIN. 1993. The *chl*L (*frx*C) gene: phylogenetic distribution in vascular plants and

DNA sequence from *Polystichum acrostichoides* (Pteridophyta) and *Synechococcus* sp. 7002 (Cynanobacteria). Pl. Syst. Evol. 187:89–102.

CALIE, P. J., and K. W. HUGHES. 1987. The consensus land plant chloroplast gene order is present, with two alterations, in the moss *Physcomitrella patens*. Molecular Gen. Genet. 208:335–341.

CONANT, D. S., D. B. STEIN, A. VALINSKI, and P. SUDARSANAM. 1994. Phylogenetic implications of chloroplast DNA variation in the Cyatheaceae. I.. Syst. Bot. 19:60–72.

COSNER, M. E., R. K. JANSEN, J. D. PALMER, and S. R. DOWNIE. 1997. The highly rearranged chloroplast genome of *Trachelium caeruleum* (Campanulaceae): multiple inversions, inverted repeat expansion and contraction, transposition, insertions/deletions, and several repeat families. Curr. Genet. 21:419–429.

DOWNIE, S. R., and J. D. PALMER. 1992. Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. Pp. 14–35. In: Soltis P. S., D. E. Soltis, and J. J. Doyle (eds.) *Molecular systematics of plants*. Chapman and Hall, New York.

DOYLE, J. J. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. Syst. Bot. 17:144–163.

DOYLE, J. J., and J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19:11–15.

FEINBERG, A. P., and VOGELSTEIN. 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Analyt. Biochem. 132:6–13.

——— 1984. Addendum: "A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity." Analyt. Biochem. 137:266–267.

HASEBE, M., and K. IWATSUKI. 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.

HOWE, C. J., R. F. BARKER, C. M. BOWMAN, and T. A. DYER. 1988. Common features of three inversions in wheat. Curr. Genet. 13:343–349.

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

LEW, K. A, and J. R. MANHART. 1993. The *rps*12 gene in *Spirogyra maxima* (Chlorophyta) and its evolutionary significance. J. Phycol. 29:500–505.

MANHART, J. R., and J. D. PALMER. 1990. The gain of two chloroplast tRNA introns marks the green algal ancestors of land plants. Nature 345:268–270.

OHYAMA, K., H. FUKUZAWA, T. KOHCHI, H. SHIRAI, T. SANO, S. SANO, K. UMESONO, Y. SHIKI, M. TAKEUCHI, Z. CHANG, S. AOTA, H. INOKUCHI, and H. OZEKI. 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322:572–574.

OLMSTEAD, R. G., and J. D. PALMER. 1994. Chloroplast DNA systematics: a review of methods and data analysis. Amer. J. Bot. 81:1205–1224.

PALMER, J. D. 1982. Physical and gene mapping of chloroplast DNA from *Atriplex triangularis* and *Cucumis sativa*. Nucl. Acids Res. 10:1593–1605.

———. 1985a. Comparative organization of chloroplast genomes. Ann. Rev. Genet. 19:325–354.

———. 1985b. Chloroplast DNA and molecular phylogeny. Bioessays 2:263–267.

———. 1986. Isolation and structural analysis of chloroplast DNA. Meth. Enzymol. 118:167–186.

———. 1987. Chloroplast DNA evolution and biosystematic uses of chloroplast DNA variation. Amer. Naturalist 130:S6–S29.

———. 1991. Plastid chromosomes: structure and evolution. In: L. Bogorad and I. K. Vasil (eds.) *Cell culture and somatic cell genetics in plants, vol. 7, the molecular biology of plastids*. Academic Press, New York.

PALMER, J. D., and D. B. STEIN. 1986. Conservation of chloroplast genome structure among vascular plants. Curr. Genet. 10:823–833.

PALMER, J. D., and J. F. THOMPSON. 1982. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. Cell 29:537–550.

PALMER, J. D., R. K. JANSEN, H. J. MICHAELS, M. W. CHASE, and J. R. MANHART. 1988. Chloroplast DNA variation and plant phylogeny. Ann. Missouri Bot. Gard. 75:1180–1206.

RAUBESON, L. A., and R. K. JANSEN. 1992. Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. Science 255:1697–1699.

RAUBESON, L. A., and D. B. STEIN. 1995. Insights into fern evolution from mapping chloroplast genomes. Amer. Fern J. 85:193–204.

SHINOZAKI, K, M. OHME, M. TANAKA, T. WAKAGUSI, N. HAYASHIDA, T. MATSUBAYASHI, N. SAITA, J. CHUNWONGSE, IBOKATA, J. YAMAGUCHI-SHINOZAKI, C. OHTO, K. TORASAWA, B. Y. MENG, M. SUGITA, H. DENO, T. KAMOGASHIRA, K. YAMADA, J. KUSUDA, F. TAKAIWA, A. KATO, N. TOHDOH, H. SHIMADA, and M. SUGIURA. 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene order and expression. EMBO J. 5:2043–2049.

STEIN, D. B., J. D. PALMER, and W. F. THOMPSON. 1986. Structural evolution and flip-flop recombination of chloroplast DNA in the fern genus *Osmunda*. Curr. Genet. 10:835–841.

STEIN, D. B., D. S. CONANT, M. E. AHEARN, E. T. JORDAN, S. A. KIRCH, M. HASEBE, K. IWATSUKI, M. K. TAN, and J. A. THOMSON. 1992. Structural rearrangements of the chloroplast genome provide an important phylogenetic link in ferns. Proc. Natl. Acad. Sci. USA 89:1856–1860.