Curr Genet (1996) 29: 572–581

ORIGINAL PAPER

Hironori Katayama · Yasunari Ogihara

Phylogenetic affinities of the grasses to other monocots as revealed by molecular analysis of chloroplast DNA

Received: 14 July / 7 November 1995

Abstract The distribution of structural alterations of the chloroplast genome found in grass chloroplast (cp) DNA in comparison with that of tobacco was systematically surveyed in the cpDNAs of monocots. Southern hybridization and/or PCR analyses for the detection of (1) three inversions in the large single-copy region, (2) loss of an intron in the *rpoC1* gene, (3) an extra-sequence insertion in the *rpoC2* gene, (4) the deletion of ORF2280, (5) rearrangements of the *accD* (ORF512) gene, and (6) non-reciprocal translocation of the *rpl23* gene, were carried out on cpDNAs isolated from 58 species, 22 families, and 11 orders, which covered almost all families of monocots. These structural alterations of cpDNA mostly occurred at the family level. However, only part of the Restionaceae possessed the inversion that characterizes the lineage of grass differentiation. The order of mutational events made it possible to reconstruct grass phylogeny in monocots. Since no variations in structural alterations of the cpDNA were found among the Poaceae, grass plants were inferred to have originated from an ancestor harboring these structural alterations of the chloroplast genome. These phylogenetic relationships were supported by the sequence data of *rbcL*.

Key words Rearrangements of the chloroplast genome · *rbcL* sequence · Phylogenetic relationships · Origin of grasses

Introduction

Although the basic genomic structure of cpDNAs is conserved in most higher plants, some structural alterations of cpDNAs have been found in comparison with the structure of cpDNAs in related plants. From comparative studies of the entire cpDNA sequences of rice (Poaceae) (Hiratsuka et al. 1989) with those of tobacco (Solanaceae) (Shinozaki et al. 1986), six kinds of structural alterations of the chloroplast genome have been identified in grass cpDNAs. These alterations are as follows: (1) three inversions (Howe et al. 1988; Hiratsuka et al. 1989) in the large single-copy region, (2) loss of an intron in the *rpoC1* gene, (3) insertion of an additional sequence in the *rpoC2* gene (Igloi et al. 1990; Shimada et al. 1990), (4) deletion of ORFs in the inverted repeat region (Downie and Palmer 1992), (5) rearrangements of the *accD* gene (Ogihara et al. 1992), and (6) a non-reciprocal translocation of the *rpl23* gene to a region downstream from *rbcL* (Bowman et al. 1988; Ogihara et al. 1988). With the limited exception of structural changes, the gene order of tobacco cpDNA is regarded as the standard among angiosperms, since most of them possess a tobacco-type cpDNA (Palmer et al. 1988). Structural alterations of the chloroplast genome thus seem to have occurred only in certain plant lineages. Accordingly, tracing of the mutational events enable us to reconstruct plant phylogeny (Downie and Palmer 1992).

Previously, we (Katayama and Ogihara 1993) reported that structural alterations of cpDNA are restricted to grasses. But, because only a limited number of monocot and grass species were used, it is still unclear why only grass cpDNAs possess these structural alterations, and in which step of differentiation these mutational events occurred. In the present study, the structural rearrangements of the chloroplast genome found in grasses was systematically surveyed in monocots. The cpDNAs of 58 species, 22 families (covering almost all monocot families), and 11 orders were examined.

The recent accumulation of DNA sequence data for chloroplast genes such as *rbcL* (the gene for the large subunit of ribulose 1,5-bisphosphate carboxylase) makes it possible to deduce the phylogenetic relationships among divergent plants and provides a powerful tool for plant taxonomy (Clegg 1993). The DNA sequencing of *rbcL* was also carried out to infer the genetic relatedness of

H. Katayama · Y. Ogihara (✉)
Kihara Institute for Biological Research,
Yokohama City University, Maioka-cho 641-12,
Yokohama 244, Japan

Communicated by B. B. Sears

rearranged chloroplast genomes in monocots (Duvall et al. 1993). Additionally we conducted DNA sequencing around the intervening sequences of the *rpoC1* and *rpoC2* genes in plants related to the Poaceae, to illustrate the mutational events of intron loss/gain in monocot cpDNAs.

Materials and methods

Plant materials. Monocotyledonous plants of 58 species (including 30 species of the Poaceae), 22 families, and 11 orders (see Table 1), which cover almost all families of monocots (Dahlgren et al. 1985), were used for the present investigation. Two dicot plants, tobacco and water lily, were examined as controls.

DNA extraction. Total DNAs were extracted from green leaves according to the cetyltrimethylammonium bromide (CTAB) method described by Murray and Thompson (1980). Extracted DNAs were employed for Southern hybridization and/or polymerase chain reaction (PCR) analyses.

DNA probes and Southern hybridization. Specific probes for ORF2280, *accD*, and the *rpl23* genes were constructed by the polymerase chain reaction (PCR) method (ORF2280) or by digestion of the clones with the appropriate restriction enzymes (*accD* and *rpl23*), as described earlier (Ogihara et al. 1992). For PCR, oligonucleotide primers were synthesized according to the published sequences (ORF2280, TATTGATATGAAGATTGCCG and TGAAACCTTGGCATATATCT; Shinozaki et al. 1986) so as to produce specific probes (Katayama and Ogihara 1993).

PCR analysis for the detection of structural alterations. PCR was applied for the detection of three inversions and to check for the loss and gain of the intervening sequence in *rpoC1* and *rpoC2*, respectively. Oligonucleotide primers complementary to the conserved sequences (mostly tRNA genes) flanking the end points of regions involved in each of the three inversions (Shinozaki et al. 1986; Hiratsuka et al. 1989) and exon sequences of *rpoC1* and *rpoC2* near junctions to the insertional mutations (Igloi et al. 1990; Shimada et al. 1990) were synthesized. The oligonucleotides used for the detection of the three inversions were the same as those reported by Doyle et al. (1992). For the amplification of target sequences, PCR was carried out under the conditions previously reported (Katayama and Ogihara 1993). PCR products containing junctions of the *rpoC1* exon-intron and the *rpoC2* exon-insertional sequence were cloned into the pT7 vector (Marchuk et al. 1990), and sequencing was done by the dideoxy method (Sanger et al. 1977).

DNA sequencing of the rbcL gene. A part of the *rbcL* gene (1232 nucleotides) from five monocots, *Allium cepa*, *Tradescantia ohiensis*×*T. poludosa*, *Juncus effusus*, *Cyperus alternifolius*, and *Restio tetraphyllus* was amplified by PCR under the conditions previously described (Katayama and Ogihara 1993). The resultant PCR products were cloned into the pT7 vector (Marchuk et al. 1990), and sequenced by the dideoxy method (Sanger et al. 1977). Sequence data of *rbcL* from the other 17 monocots and two dicots were cited from GenBank (Chase et al. 1993). The phylogenetic tree was constructed by the parsimony method (Swofford 1991).

Results

We have systematically surveyed the distribution of the following six structural alterations of the chloroplast genome found in grasses in the monocots.

Three inversions

The chloroplast genomes of grasses possess three inversions relative to the cpDNA gene arrangement found in most flowering plants such as tobacco (Howe et al. 1988; Hiratsuka et al. 1989). Amplification primers for PCR were designed to detect each of these inversions. PCR products responsible for the first inversion were found in the cpDNAs of the Restionaceae, Joinvilleaceae, and Poaceae, indicating that these three taxa harbor this inversion. It should be emphasized, however, that only one species of Restionaceae, out of four species examined, possessed the first inversion, thus showing polymorphism for the inversion within the family (Fig. 1). The primer combination for detection of the second inversion yielded PCR products only with the cpDNAs of the Joinvilleaceae and Poaceae, but not with those of the Restionaceae, including the species with the first inversion. The PCR to detect the third inversion gave a positive signal with the cpDNAs of the Poaceae. These results (Table 1) clearly show that some Restionaceae species (one species in this case, but a different one from that of Doyle et al. 1992) possess only the first inversion, while the Joinvilleaceae contains the first and second inversion, and the Poaceae harbors all three inversions.

Loss of the intron in the *rpoC1* gene

The *rpoC1* gene of tobacco, coding for the β' subunit of RNA polymerase, contains a 738-bp intron. This intron, however, is deleted in the cpDNA of rice and maize (Igloi et al. 1990; Shimada et al. 1990). The occurrence of this intron deletion was examined in cpDNAs from 58 species of monocots (Table 1). The existence of the intron was directly confirmed by the analysis of PCR products. DNA fragments containing the intron were shifted to higher molecular weights than those without the intron, as shown in Fig. 2 (only the data for *R. tetraphyllus* are presented in this figure). Nucleotide sequence analysis at the exon-intron junctions of the *rpoC1* gene in *R. tetraphyllus* (Restionaceae) and *Joinvellea plicata* (Joinvilleaceae) verified the site of the deletion. Sequence data clearly show that the cpDNA of the Joinvilleaceae contains the group-II intron (Rieger et al. 1991) at the same position as that of tobacco, and this intron was completely deleted from the cpDNAs of the Restionaceae and grasses (Fig. 3).

Insertional sequence into the *rpoC2* gene

The *rpoC2* gene, coding for the β'' subunit of RNA polymerase of grass cpDNA, contains an extra insertional sequence (Igloi et al. 1990; Shimada et al. 1990). The PCR analysis (Fig. 4) revealed that the insertional sequence was only found in grass cpDNAs; whereas the cpDNAs of the Restionaceae and Joinvilleaceae, both belonging to the Poales, did not harbor the insertional sequence (Fig. 4).

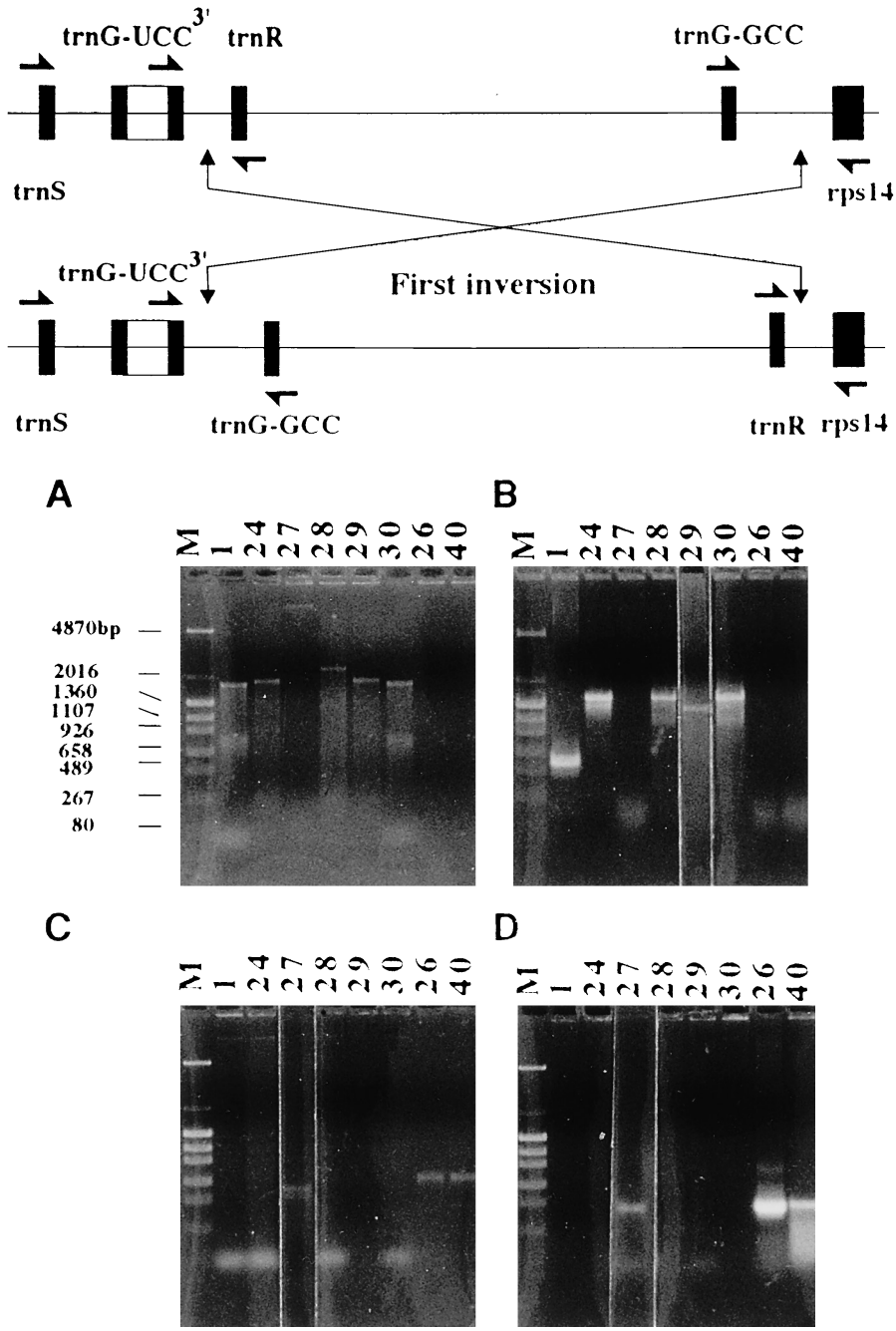


Fig. 1 Physical map around the first inversion found in grass cpDNA, and representative PCR amplifications with primers diagnostic for the first inverstion. *Arrows* over the gene stand for the primers used: (**A**) primer combinations of trnS with trnR, (**B**) trnG-GCC and rps14, (**C**) trnG-UCC$^{3'}$ and trnG-GCC, (**D**) trnR and rps14. *Numbers* on each panel indicate the code numbers of the plants analyzed (see Table 1)

DNA sequencing at the border of the exon-insertional sequence confirmed the PCR analysis (Fig. 5). This result indicates that the common ancestor of grass plants must have conferred the extra sequence. No sequence homologous to the extra sequence could be detected in the DNA data bases of EMBL and GenBank (1995).

Deletion of ORF2280 in IR regions

The ORF2280, which is located in the IR region of tobacco cpDNA, is so highly rearranged in grass cpDNA (Downie et al. 1994) that the sequence gives no Southern hybridization signals under the present conditions. Since the translation product of ORF2280 is detected in the chloroplasts of spinach, tobacco, and *Oenothera* (Glick and Sears 1993), the sequence is evidently active in those cpDNAs. The probe of ORF2280 was constructed to be specific for the internal region of the gene, a region which has no homology with the rice counterpart. Southern hybridization patterns with this probe showed that this sequence was not detected in the cpDNAs of the Cyperaceae, Juncaceae, Centrolepidaceae, Restionaceae, Joinvilleaceae, and Poaceae (Fig. 6A and Table 1).

Table 1 Distribution of structural alterations in the cpDNAs of monocots

Code	Species[a]	Order[a]	Family[a]	Structural alterations in cpDNAs[b]					
				(1)	(2)	(3)	(4)	(5)	(6)
1	*Nicotiana tabacum*	Solanales	Solanaceae	+	–	+	+	–	–
2	*Nymphaea* spp.	Ranunculales	Nymphaeaceae	+	–	+	+	–	–
3	*Sagittaria trifolia*	Alismatales	Alismataceae	+	–	+	+	–	–
4	*Potamogeton distinctus*	Najadales	Potamogetonaceae	+	–	+	+	–	–
5	*Rhapis flabelliformis*	Arecales	Arecaceae	+	–	+	+	–	–
6	*Colocasia antiquorum*	Arales	Araceae	+	–	+	+	–	–
7	*Sparganium stloniferum*	Typhales	Sparganiaceae	+	–	+	+	–	–
8	*Lilium elegans*	Liliales	Liliaceae	+	–	+	+	–	–
9	*Allium cepa*	Liliales	Liliaceae	+	–	+	+	–	–
10	*Iris laevigata*	Liliales	Iridaceae	+	–	+	+	–	–
11	*Bletilla striata*	Liliales	Orchidaceae	+	–	+	+	–	–
12	*Ananas comosus*	Bromeliales	Bromeliaceae	+	–	+	+	–	–
13	*Zingiber mioga*	Zingiberales	Zingiberacea	+	–	+	+	–	–
14	*Thalia dealbata*	Zingiberales	Marantaceae	+	–	+	+	–	–
15	*Canna generalis*	Zingiberales	Cannaceae	+	–	+	+	–	–
16	*Tradescantia ohiensis* *T. poludosa*	Commelinales	Commelinaceae	+	–	+	+	–	–
17	*Eriocaulon nudicuspe*	Commelinales	Eriocaulaceae	+	–	+	–	–	–
18	*Juncus effusus*	Cyperales	Juncaceae	+	–	–	–	–[c]	–
19	*Cyperus alternifolius*	Cyperales	Cyperaceae	+	–	–	–	–[c]	–
20	*Eleocharis kuroguwai*	Cyperales	Cyperaceae	+	–	–	–	–[c]	–
21	*Scirpus lacustris*	Cyperales	Cyperaceae	+	–	–	–	–[c]	–
22	*Carex ciliato-marginata*	Cyperales	Cyperaceae	+	–	–	–	–[c]	–
23	*Flagellaria* spp.	Poales	Flagellariaceae	+	–	+	–	–	–
24	*Centrolepis* sp.	Poales	Centrolepidaceae	–	–	–	–	–	–
25	*Anarthria prolifera*	Poales	Anarthriaceaea	–	–	+	–	–	–
26	*Joinvillea plicata*	Poales	Joinvilleaceae	+	–	–	–	–	I,II
27	*Restio tetraphyllus*	Poales	Restionaceae	–	–	–	–	–	I
28	*Elegia cuspidata*	Poales	Restionaceaea	–	–	–	–	–	–
29	*Desmocladus castaneus*	Poales	Restionaceaea	–	–	–	–	–	–
30	*Lepyrodia riparia*	Poales	Restionaceaea	–	–	–	–	–	I
31	*Semiarundinaria fastuosa*	Poales	Poaceae	–	+	–	–	+	I,II,III
32	*Sasa palmata*	Poales	Poaceae	–	+	–	–	+	I,II,III
33	*Phragmites australis*	Poales	Poaceae	–	+	–	–	+	I,II,III
34	*Zizania latifolia*	Poales	Poaceae	–	+	–	–	+	I,II,III
35	*Triticum aestivum*	Poales	Poaceae	–	+	–	–	+	I,II,III
36	*Hordeum vulgare*	Poales	Poaceae	–	+	–	–	+	I,II,III
37	*Secale cereale*	Poales	Poaceae	–	+	–	–	+	I,II,III
38	*Avena sativa*	Poales	Poaceae	–	+	–	–	+	I,II,III
39	*Zea mays*	Poales	Poaceae	–	+	–	–	+	I,II,III
40	*Oryza sativa*	Poales	Poaceae	–	+	–	–	+	I,II,III
41	*Agrostis storonifera*	Poales	Poaceae	–	+	–	–	+	I,II,III
42	*Agrostis tenuis*	Poales	Poaceae	–	+	–	–	+	I,II,III
43	*Phleum beltolonii*	Poales	Poaceae	–	+	–	–	+	I,II,III
44	*Phleum pratense*	Poales	Poaceae	–	+	–	–	+	I,II,III
45	*Dactylis glomerata*	Poales	Poaceae	–	+	–	–	+	I,II,III
46	*Festuca arundinaceae*	Poales	Poaceae	–	+	–	–	+	I,II,III
47	*Festuca ovina*	Poales	Poaceae	–	+	–	–	+	I,II,III
48	*Festuca rubura*	Poales	Poaceae	–	+	–	–	–	I,II,III
49	*Lolium perenne*	Poales	Poaceae	–	+	–	–	–	I,II,III
50	*Poa pratensis*	Poales	Poaceae	–	+	–	–	+	I,II,III
51	*Buchloe dactyloides*	Poales	Poaceae	–	+	–	–	+	I,II,III
52	*Cynodon dactylon*	Poales	Poaceae	–	+	–	–	+	I,II,III
53	*Eragrostis curvula*	Poales	Poaceae	–	+	–	–	+	I,II,III
54	*Zoysia japonica*	Poales	Poaceae	–	+	–	–	+	I,II,III
55	*Zoysia tenuifolia*	Poales	Poaceae	–	+	–	–	+	I,II,III
56	*Eremochloa ophiuroides*	Poales	Poaceae	–	+	–	–	+	I,II,III
57	*Axonopus affinis*	Poales	Poaceae	–	+	–	–	+	I,II,III
58	*Paspalum notatum*	Poales	Poaceae	–	+	–	–	+	I,II,III
59	*Pennisetum cladestinum*	Poales	Poaceae	–	+	–	–	+	I,II,III
60	*Stenotaphrum secundatum*	Poales	Poaceae	–	+	–	–	+	I,II,III

[a] After Dahlgren et al. 1985

[b] These structural alterations are as follows: (1) existence of an *rpoC1* intron, (2) an extra sequence in *rpoC2*, (3) existence of a full ORF2280, (4) existence of a full *accD*, (5) translocation of *rpl23*, and (6) inversions in the large single-copy region (I: first inversion, II: second inverstion, III: third inversion)

[c] Deletion of the *rpl23* gene from the cpDNA

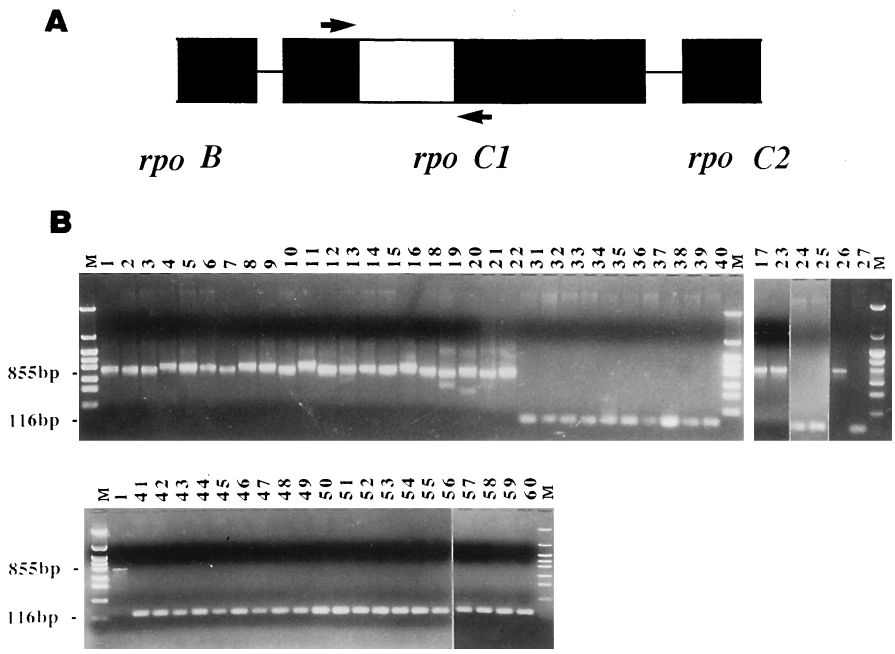


Fig. 2 The gene structure of *rpoC1* and the position of synthesized primers for PCR are schematically presented in **A**, and agarose-gel electrophoresis patterns of PCR products with cpDNA from 55 species are shown in **B**. The PCR product of the tobacco *rpoC1* gene is 855 bp, and that of rice is 116 bp in this system

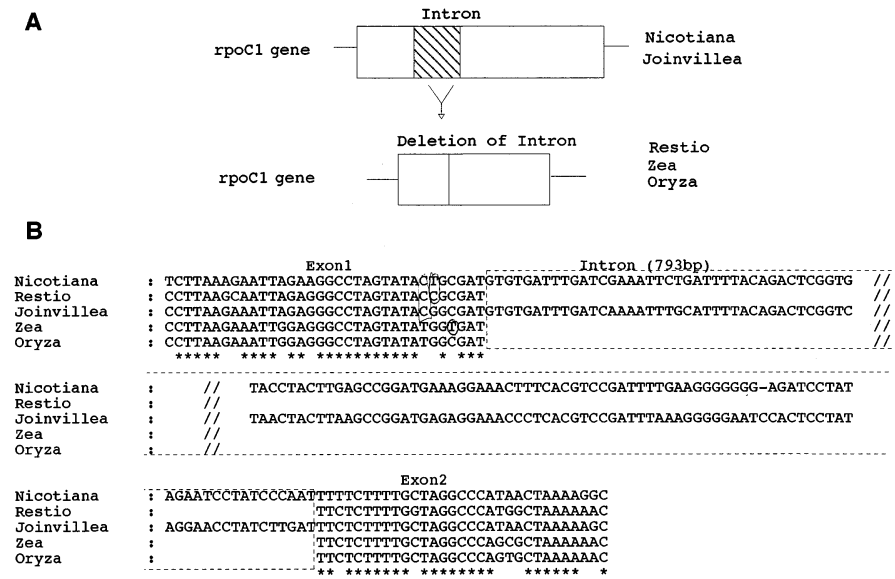


Fig. 3 Structure (**A**) and DNA sequencing comparison (**B**) around the exon-intron junctions of the *rpoC1* gene from five cpDNAs, tobacco, *Restio tetraphyllus* (Restionaceae), *Joinvillea plicata* (Joinvilleaceae), *Zea mays* (Poaceae), and *Oryza sativa* (Poaceae)

Rearrangements of the *accD* gene (ORF512)

The fifth case involves rearrangements of the *accD* gene (ORF512), which is located downstream from *rbcL* in tobacco cpDNA (Ogihara et al. 1992, Morton and Clegg 1993). Approximately 90% of the tobacco *accD* gene is deleted in grass cpDNA, so that no Southern hybridization signal in grass cpDNA is detected with this probe. With the exception of the Commelinaceae, Southern hybridization gave no positive signals with the cpDNAs of the superorder Commeliniflorae under the present conditions (Fig. 6B and Table 1). The data suggest that rearrangements of the *accD* gene occurred in the progenitor of *Eriocaulon*, and that the members of the Poales and Cyperales were descended from this plant group.

Translocation of the *rpl23* Gene

In grass cpDNA, the *rpl23* gene, which was originally located in the IR region, has been non-reciprocally translocated to a region downstream from *rbcL* (Bowman et al. 1988; Ogiahara et al. 1988). Consequently, by choosing appropriate restriction enzymes three hybridization signals

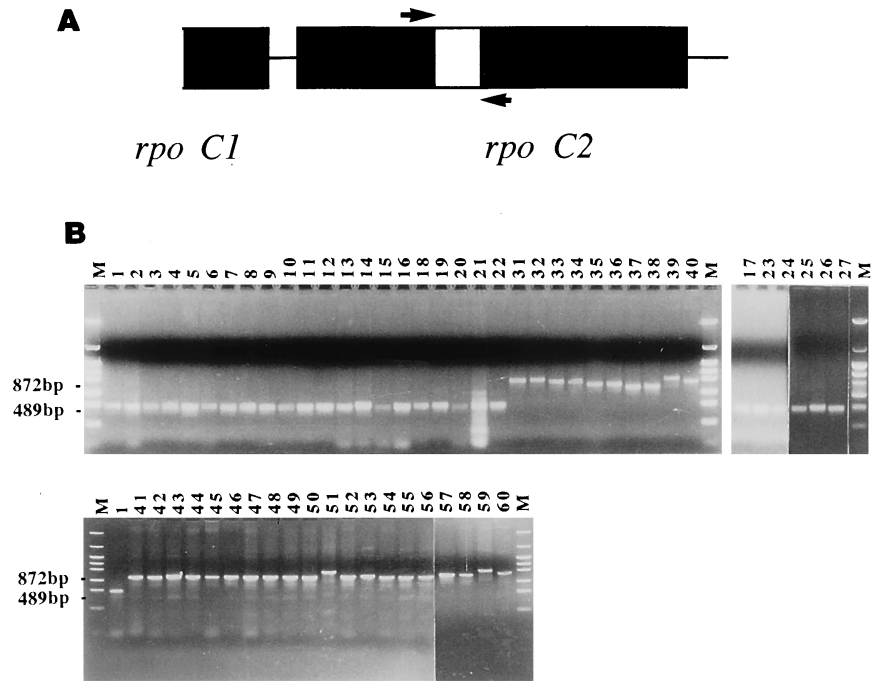


Fig. 4 The gene structure of *rpoC2* and the position of synthesized primers for PCR are schematically presented in **A**, and agarose-gel electrophoresis patterns of PCR products with cpDNA from 55 species are shown in **B**. The PCR product of the tobacco *rpoC2* gene is 489 bp, and that of rice is 872 bp in this system

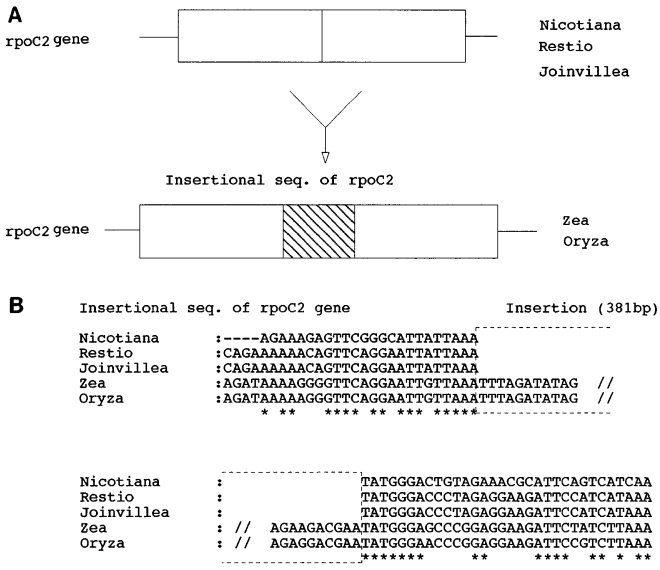


Fig. 5 Structure (**A**) and DNA sequencing comparison (**B**) around the exon-insertion junctions of the *rpoC2* gene from five cpDNAs, *Nicotiana* (Solanaceae), *Restio tetraphyllus* (Restionaceae), *Joinvillea plicata* (Joinvilleaceae), *Zea mays* (Poaceae), and *Oryza sativa* (Poaceae)

can be detected in grass cpDNAs. On the other hand, only one or two bands that correspond to the authentic *rpl23* gene were detectable in the other cpDNAs. Southern hybridization data show that three hybridization signals were evident in the cpDNAs of grasses, whereas the cpDNAs of the other plants gave only one or two signals (a typical hybridization pattern obtained by digestion with *Bam*HI is presented in Fig. 6C). It is striking that in this case no hybridization signal was detected in the cpDNAs of the Cyperaceae or Juncaceae (Fig. 6C). These results show that intragenomic translocation of the *rpl23* gene occurred only in grass cpDNAs and that *rpl23* was rearranged in the cpDNA of the Cyperales.

Based on the detected mutational events of cpDNAs, the evolutionary course of grass plants in monocots was reconstructed as shown in Fig. 7.

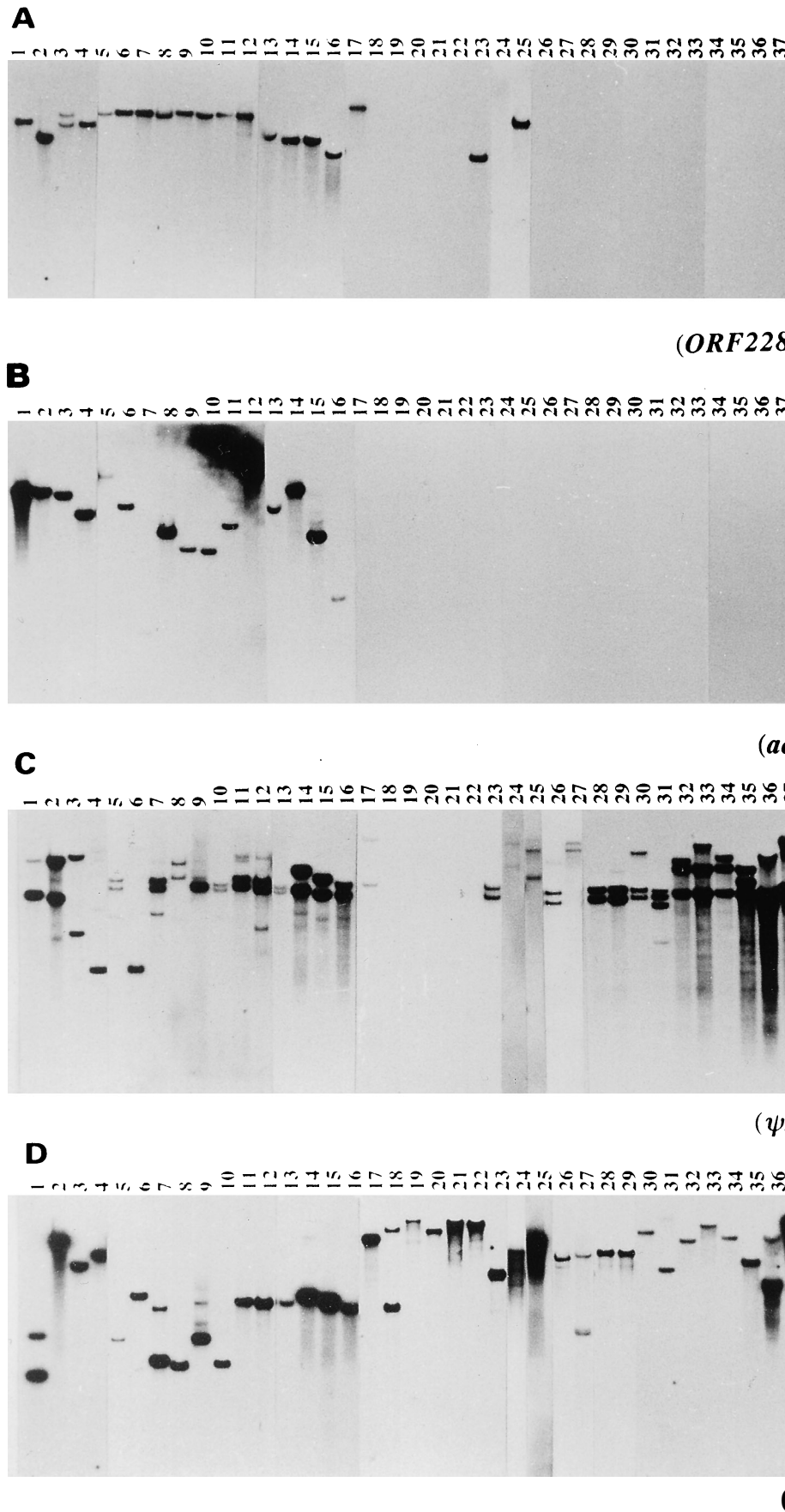
Phylogenetic tree inferred by *rbcL* sequencing

Part of the *rbcL* gene (1232 nucleotides) from five plant species was sequenced, and compared with the sequences of other monocots registered with GenBank (Chase et al. 1993). A phylogenetic tree among them was constructed according to the PAUP method (Swofford 1991), as shown in Fig. 8. This tree is mostly in good agreement with that inferred from morphological data (Dahlgren et al. 1985). The closest relative to grasses is the Joinvilleaceae, and the Restionaceae is the next closest. Nucleotide substitution rates were accelerated during the grass course of differentiation (Fig. 8 and see Gaut et al. 1992).

Discussion

In comparison with that of tobacco (Solanaceae), we have systematically surveyed the distribution of six structural rearrangements of the chloroplast genome found in grasses (Poaceae) throughout the cpDNAs from various plants including 58 species, 22 families, and 11 orders (Table 1).

Fig. 6A–D Southern hybridization analysis of cpDNAs in monocots probed with OF2280 (**A**), *accD* (**B**), *rpl23* (**C**), and *rbcL* (**D**) to check for structural alterations



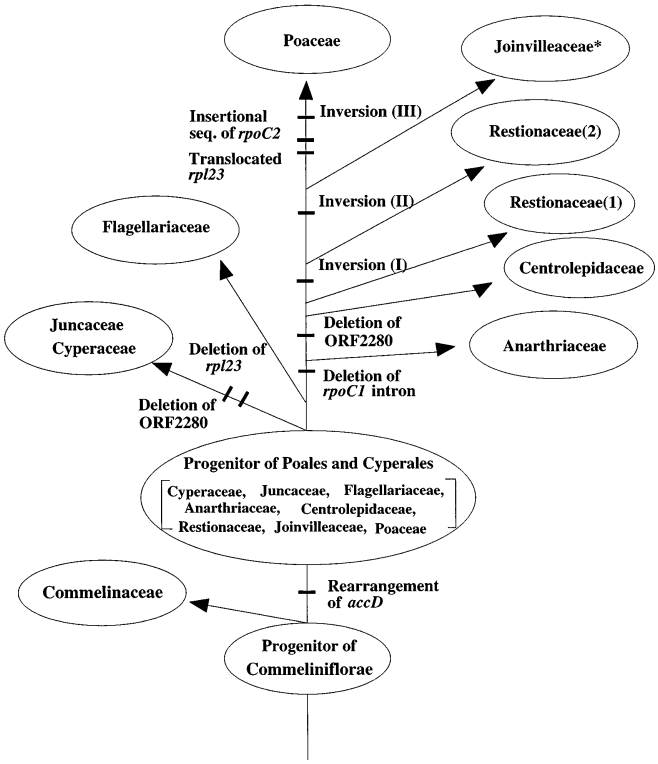


Fig. 7 Reconstruction of grass phylogeny as revealed by structural alterations of chloroplast DNA. *Bars* (–) indicate the occurrence and the type of rearrangement

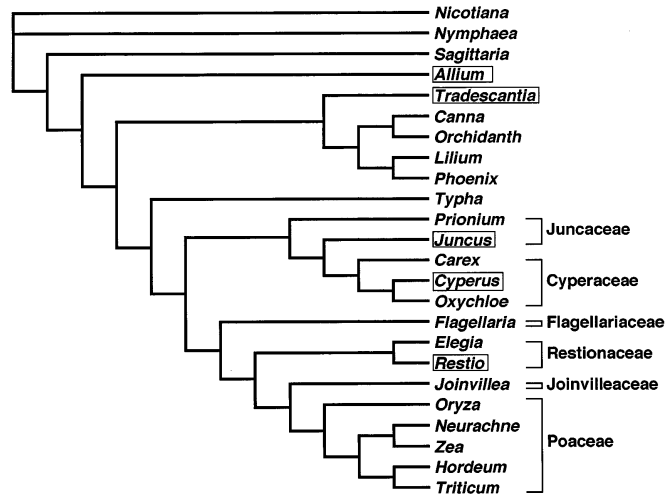


Fig. 8 A phylogenetic tree of the monocots constructed from data on the DNA sequence of the *rbcL* genes by the PAUP method. The sequenced plant species in the present study are *boxed*. The other sequences are cited from the EMBL data bank (Chase et al. 1993)

Additionally, part of the *rbcL* gene from five species was sequenced in order to infer the phylogenetic relationships among them.

Gene loss from the chloroplast genome in the lineage of grass differentiation

In a comparison of the grass cpDNA sequences with those of tobacco, losses of several genes, such as *accD* (Hiratsuka et al. 1989; Ogihara et al. 1992) and ORF2280 (Downie and Palmer 1992), were disclosed. In addition to these examples, we have shown that most, if not all, of the *rpl23* gene was deleted from the cpDNA of the Cyperaceae and Juncaceae (Fig. 6). Interestingly, the *rpl23* gene was found to have been intragenomically translocated into the hypervariable region of grass cpDNA (Ogihara et al. 1988) near the *rbcL* gene.

Sequences homologous with the *accD* gene (previously ORF512 in tobacco) were deleted in the cpDNAs of the Eriocaulaceae (Commelinales) and the progenitor of the Poales and Cyperales, as shown in Fig. 6. The *accD* gene encodes prokaryotic acetyl-coA carboxylase (Holt et al. 1993). Consequently, the loss of this gene introduces a sensitivity to the herbicides quizalofop and sethoxydim (Konishi and Sasaki 1994), and suggests an alteration of lipid metabolism in those plants.

The chloroplast genome of grasses (Poaceae) is smaller in size (approximately 20 kb) than that of other monocots, which possess chloroplast DNAs similar to the standard type having a molecular size of approximately 155 kb (Katayama et al. 1991; Palmer 1991). The reduction in size of grass cpDNAs is mainly due to deletions in the inverted repeats (Hiratsuka et al. 1989). Our Southern hybridization data showed that the sequence of ORF2280 was not detected in the cpDNAs of the Cyperaceae, Juncaceae, Restionaceae, Joinvilleaceae, and Poaceae, indicating the deletion of this gene from the cpDNA (Table 1). Since the translation product of ORF2280 is detectable in the soluble fraction of chloroplasts (Glick and Sears 1993), the deletion of the sequence suggests some unknown regulatory system in chloroplast function. As for ORF2280, multiple losses of the gene were observed to have occurred during dicot evolution (Downie et al. 1994). As in the case of dicot cpDNAs, a multiple occurrence of the loss of this gene was found in monocot cpDNAs; the gene was independently deleted in the cpDNAs of the Cyperales and in the progenitor of the Centrolepidaceae, Restionaceae, Joinvilleaceae, and Poaceae. Since we used the conserved region of ORF2280 as a probe, the deleted position(s) of ORF2280 might differ from those found in dicots (Downie et al. 1994).

Loss of the *rpoC1* intron in cpDNA

Intron loss was found in the *rpoC1* gene of grasses (Igloi et al. 1990; Shimada et al. 1990). The data so far available (Table 1) clearly show that intron loss of the *rpoC1* gene occurred in the cpDNAs of the Poaceae, Restionaceae, Arathriaceae, and Centrolepidaceae, but not in the Joinvilleaceae. The intervening sequence of *rpoC1* reveals the characters of a group-II intron (Rieger et al. 1991), and loss

of the intron may be mediated by processed mRNA. From the comparison of nucleotide sequences around the deletion points of the *rpoC1* genes of three related taxa, the Restionaceae, Joinvilleaceae and Poaceae, with the sequence of tobacco, the deletion of cpDNA in the Restionaceae appears to have occurred at the same position as in grass cpDNAs; and the cpDNA of the Joinvilleaceae harbors the complete intron (Fig. 3B). Although it is likely that multiple intron losses had occurred during the course of grass differentiation, more experiments are required to clarify the intron loss of the *rpoC1* gene, because DNA sequences of the intron and part of the exon revealed higher homology with tobacco than any other part of the chloroplast genome (Fig. 3B).

Reconstruction of grass phylogeny revealed by structural alterations of cpDNA

Structural alterations of chloroplast genomes in higher plants have been reported to have occurred at various taxonomic levels (e.g. Palmer 1991). Since the structure of chloroplast DNA is conserved among diverged plants, rearrangements of cpDNA are considered to be relatively uncommon. The conserved structure of the genome allows us to trace the lineage of the differentiation of plant species. But this rarity of genome mutations in the chloroplast DNA limits the utility of each mutation to providing only a broad picture of phylogenetic relationships. Combinations of mutations and alterations in chloroplast genome structure are needed to overcome this ambiguity, and they provide a powerful tool for reconstructing the phylogenetic relationships of grasses and their relatives.

The present results (Table 1 and Fig. 7) also indicate that the structural alterations of cpDNA in grasses are restricted to the Poaceae, Joinvilleaceae, Restionaceae, Centrolepidaceae, Anarthriaceae, Flagellariaceae, Juncaceae, Cyperaceae, and Eriocaulaceae, all of which belong to the superorder Commeliniflorae (Dahlgren et al. 1985). These structural analyses of cpDNA unite grasses with the Joinvilleaceae and Restionaceae. These groups were thus separated from the Cyperales and other members of the superorder Commeliniflorae. It is striking that two inversion types, namely the chloroplast genome types with the first inversion and that with no inversion, were found in the Restionaceae (Fig. 1 and Table 1). Thus, we consider that the progenitor of the Poaceae and Joinvilleaceae was derived from the inversion type of the Restionaceae. These phylogenetic relationships constructed on the basis of structural alterations of the cpDNA support the view that grasses are most closely related to a predominantly Southern Hemisphere group, and are consistent with cladistic studies (Dahlgren et al. 1985; Stebbins 1987).

Since variations in the rearrangements of chloroplast DNA structures within the Poaceae were not detected, grasses are inferred to have differentiated from a limited number of ancestral plants (monophyletic origin of grasses).

Grass phylogeny inferred by the sequence data of *rbcL*

Phylogenetic relationships of monocots inferred by use of the parsimony method on *rbcL* sequences were presented in Fig. 8. The phylogenetic relationships inferred from the *rbcL* sequence data were principally consistent with those constructed from structural rearrangements of the cpDNA: the Poaceae was grouped with the Joinvilleaceae and Restionaceae, and the Flagellariaceae was separated from the Joinvilleaceae, but grouped into the Poales.

Acknowledgements We are grateful to Drs. C. Uematsu (Kisaichi Botanical Gardens, Osaka City University), T. Hoshino (Okayama Science University), B.G. Briggs (Royal Botanic Gardens Sydney), M. Fagg (Australian National Botanic Gardens), M. Hasebe (Botanical Garden, University of Tokyo), M. Ito (Chiba University), Mrs. T. Ohmura, and H. Yaneshita (Japan Turf Grass Co. Ltd.), who kindly supplied material for this study. This is contribution No. 160 from the Kihara Institute for Biological Research, Yokohama City University, Japan. The work was supported, in part, by Special Coordination Funds of the Science and Technology Agency of the Japanese Government.

References

Bowman CM, Barker RF, Dyer TA (1988) In wheat ctDNA, segments of ribosomal protein genes are dispersed repeats, probably conserved by nonreciprocal recombination. Curr Genet 14:127–136

Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Duvall MR, Price R, Hills HG, Qiu YL, Korn KA, Rettig JH, Conti E, Palmer JD, Nanhardt JR, Systema KJ, Michaels HJ, Kress WJ, Donoghue MJ, Clark WD, Hedren M, Gaut BS, Jansen RK, Kim KJ, Wimpee CF, Smith JF, Furnier GR, Straus SH, Xiang QY, Plunkett GM, Soltis PS, Eguiarte LE, Learn GH, Barrett SCH, Graham S, Albert VA (1993) Phylogenetics of seed plants: an analysis of nucleotide sequences from the plasmid gene *rbcL*. Ann Miss Bot Gard 80:528–550

Clegg MM (1993) Chloroplast gene sequences and the study of plant evolution. Proc Natl Acad Sci USA 90:363–367

Dahlgren RMT, Clifford HT, Yeo PF (1985) The families of the monocotyledons. Springer-Verlag, Berlin

Downie SR, Palmer JD (1992) Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. In: Soltis PS, Soltis DE, Doyle JJ (eds) Molecular systematics of plants. Chapman and Hall, New York, pp 14–35

Downie SR, Katz-Downie DS, Wolfe KH, Calie PJ, Palmer JD (1994) Structure and evolution of the largest chloroplast gene (ORF2280): internal plasticity and multiple gene loss during angiosperm evolution. Curr Genet 25:367–378

Doyle JJ, Davis JI, Soreng RJ, Garvin SD, Anderson MJ (1992) Chloroplast DNA inversions and the origin of the grass family (Poaceae). Proc Natl Acad Sci USA 89:7722–7726

Duvall MR, Learn Jr GH, Eguiarte LE, Clegg MT (1993) Phylogenetic analysis of *rbcL* sequences identifies *Acorus calamus* as the primal extant monocotyledon. Proc Natl Acad Sci USA 90:4641–4644

Gaut BS, Muse SV, Clark WD, Clegg MT (1992) Relative rates of nucleotide substitution at the *rbcL* locus on monocotyledonous plants. J Mol Evol 35:292–303

Glick RE, Sears BB (1993) Large unidentified open reading frame in plastid DNA (ORF2280) is expressed in chloroplasts. Plant Mol Biol 21:99–108

Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun CR, Meng BY, Li YQ, Kanno A, Nishizawa Y, Hirai A, Shinozaki K, Sugiura M (1989) The complete sequence of rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts

for a major plastid DNA inversion during the evolution of cereals. Mol Gen Genet 217:185–194
Holt JS, Powles SB, Holtum JAM (1993) Mechanisms and agronomic aspects of herbicide resistance. Annu Rev Plant Physiol Plant Mol Biol 44:203–229
Howe CJ, Barker RF, Bowman CM, Dyer TA (1988) Common features of three inversions in wheat chloroplast DNA. Curr Genet 13:343–349
Igloi GL, Meinke A, Dory I, Kossel H (1990) Nucleotide sequence of the maize chloroplast *rpoB/C1/C2* operon: comparison between the derived protein primary structures from various organisms with respect to functional domains. Mol Gen Genet 221:379–394
Katayama H, Ogihara Y (1993) Structural alterations of the chloroplast genome found in grasses are not common in monocots. Curr Genet 23:160–165
Katayama H, Sasakuma T, Ogihara Y (1991) Physical map of chloroplast DNA of the onion *Allium cepa* L., showing the location of photosynthesis-related genes. Jpn J Genet 66:421–431
Konishi T, Sasaki Y (1994) Compartmentalization of two forms of acetyl-CoA carboxylase in plants and the origin of their tolerance toward herbicides. Proc Natl Acad Sci USA 91:3598–3601
Marchuk D, Drumm M, Saulino A, Collins FS (1990) Construction of T-vector, a rapid and general system for direct cloning of unmodified PCR products. Nucleic Acids Res 19:1154
Morton BR, Clegg MT (1993) A chloroplast DNA mutational hot spot and gene conversion in a noncoding region near *rbcL* in the grass family (Poaceae). Curr Genet 24:357–365
Murray HG, Thompson WF (1980) Rapid isolation of high-molecular-weight plant DNA. Nucleic Acids Res 8:4321–4325
Ogihara Y, Terachi T, Sasakuma T (1988) Intramolecular recombination of chloroplast genome mediated by a short direct-repeat sequence in wheat species. Proc Natl Acad Sci USA 85: 8573–8577
Ogihara Y, Terachi T, Sasakuma T (1992) Structural analysis of length mutation in a hot-spot region of wheat chloroplast DNAs. Curr Genet 22:251–258
Palmer JD (1991) Plastid chromosome: structure and evolution. In: Bogorad L, Vasil IK (eds) The molecular biology of plastids. Academic Press, San Diego, pp 5–53
Palmer JD, Jansen Rk, Micaels HJ, Chase MW, Mnhart JR (1988) Chloroplast DNA variation and plant phylogeny. Ann Miss Bot Gard 75:1180–1206
Rieger R, Michaelis A, Green MM (1991) Glossary of genetics. Springer-Verlag, Berlin Heidelberg, New York, Tokyo
Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467
Shimada H, Fukuta M, Ishikawa M, Sugiura M (1990) Rice chloroplast RNA polymerase genes: the absence of an intron in *rpoC1* and the presence of an extra sequence in *rpoC2*. Mol Gen Genet 221:395–402
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 BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of tobacco chloroplast genome: its gene organization and expression. EMBO J 5:2043–2049
Stebbins GL (1987) Grass systematics and evolution: past, present and future. In: Soderstrom TR, Hilu KH, Campbell CS, Varkworth ME (eds) Grass systematics and evolution. Smithsonian Institution Press, Washington DC, pp 359–367
Swofford DL (1991) PAUP: Phylogenetic analysis using parsimony, Version 3.0r. Computer program distribution by the Illinois Natural History Survey, Champaign, Illinois