Molecular Evolution of Chloroplast DNA Sequences[1]

Stephanie E. Curtis
North Carolina State University

Michael T. Clegg
University of Georgia

Comparative data on the evolution of chloroplast genes are reviewed. The chloroplast genome has maintained a similar structural organization over most plant taxa so far examined. Comparisons of nucleotide sequence divergence among chloroplast genes reveals marked similarity across the plant kingdom and beyond to the cyanobacteria (blue-green algae). Estimates of rates of nucleotide substitution indicate a synonymous rate of 1.1×10^{-9} substitutions per site per year. Noncoding regions also appear to be constrained in their evolution, although addition/deletion events are common. There have also been evolutionary changes in the distribution of introns in chloroplast encoded genes. Relative to mammalian mitochondrial DNA, the chloroplast genome evolves at a conservative rate.

Introduction

The chloroplast genome has become a major focus for studies of plant molecular evolution. The small size of the chloroplast genome together with rapid progress in the molecular characterization of chloroplast encoded genes (reviewed by Whitfeld and Bottomley [1983]) have combined to facilitate evolutionary investigations. In this article, we provide a brief account of research on the molecular evolution of chloroplast encoded genes. The comparative analysis of complete DNA sequences will be our primary concern, although some results from the restriction analysis of chloroplast DNA will also be discussed.

Discussion

Organization of the Chloroplast Genome

To facilitate our later discussion and to define terms, it is helpful to first consider the structural organization of the chloroplast genome (fig. 1). A major feature of this circular DNA molecule (averaging about 150 kbp in angiosperms) is a large duplicated region of reverse orientation that is separated by two regions of single copy DNA (referred to as the small and large single copy regions). The

1. Key words: chloroplast, DNA, evolution, nucleotide substitution.

Address for correspondence and reprints: Professor Stephanie E. Curtis, Department of Genetics, North Carolina State University, Raleigh, North Carolina 27650.

Mol. Biol. Evol. 1(4):291–301. 1984.

0737-4038/84/0104-0003$02.00

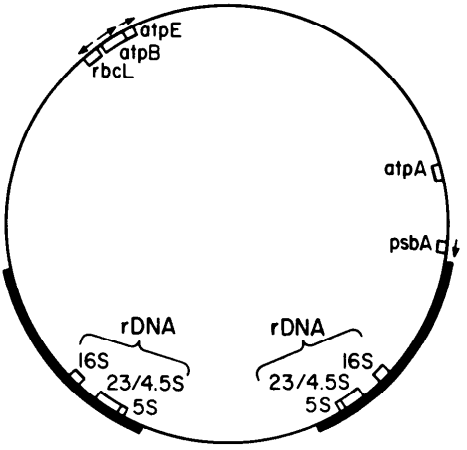

FIG. 1.—Structural organization of the typical higher plant chloroplast genome. Included are map positions of the genes coding for the large subunit of ribulose-1, 5-bisphosphate carboxylase (*rbcL*), the α, β, and ϵ subunits of ATP synthase (*atpA, atpB, atpE*), the 32 kilodalton thylakoid membrane protein (*psbA*), and the 23S, 16S, 5S, and 4.5S ribosomal RNA genes. The duplicated region of reverse orientation is shown by a heavy line. Arrows indicate directions of transcription.

structural organization depicted in figure 1 is characteristic of most higher plants, although pea (*Pisum sativa*) and broad bean (*Vicia faba*) lack one of the duplicate regions (Palmer and Thompson 1981). The inverted repeats contain genes coding for the 23S, 16S, 5S, and 4.5 S ribosomal RNAs (rRNA), several tRNAs (Whitfeld and Bottomley 1983), and ribosomal proteins (G. Zurawski, personal communication). A number of genes have been mapped onto the large single copy region, including the genes coding for the large subunit of ribulose-1, 5-bisphosphate carboxylase (*rbcL*), the α, β, ϵ, and DCCD-binding proteolipid subunits of ATP synthase (*atpA, atpB, atpE,* and *atpH,* respectively), the 32-kilodalton thylakoid membrane protein (*psbA*), numerous tRNAs (Alt et al. 1983*a;* Deno et al. 1983; Whitfeld and Bottomley 1983), two ribosomal proteins (Subramanian et al. 1983; Sugita and Sugiura 1983), the p700 chlorophyll A apoprotein (Westhoff et al. 1983) and three polypeptides of the cytochrome b6-f complex (Alt et al. 1983*b;* Willey et al. 1983).

While figure 1 refers to the chloroplast genome of angiosperms, some algae (e.g., *Chlamydomonas rheinhardtii* [Rochaix 1981]), a fern (*Osmunda cinnamomea* [Palmer and Stein 1982]), a liverwort (*Marchantia polymorpha* [Ohyama et al. 1983]), and the cyanelle of *Cyanophora paradoxa* (Bohnert and Loffelhardt 1982) exhibit a similar organization. On the other hand, the alga *Euglena gracilis,* which lacks the inverted repeat structure, has instead an rRNA gene cluster with one to 5½ tandem repeating units (Koller and Delius 1982; Whitfeld and Bottomley 1983). Finally, the map location of *rbcL* and *psbA* shown in figure 1 is typical of most higher plants so far examined except legumes, where a 50-kb inversion places *rbcL* within a few kb of *psbA* (Palmer and Thompson 1982; Spielmann et al. 1983). In addition, an inversion of about 20 kb, which is apparently unique to monocots, places *atpA* closer to *rbcL* in wheat (*Triticum aestivum* [Howe et al. 1983]) and maize (*Zea mays* [Palmer and Thompson 1982]) than depicted by figure 1.

Analyses of Restriction Fragment Variation

The analysis of restriction fragment patterns of chloroplast DNA is most useful for population samples or for the study of evolutionary relationships among closely related species. Restriction endonucleases were first employed to study chloroplast DNA evolution by Atchison et al. (1976). A variety of studies have accumulated since 1976 that indicate relatively low rates of nucleotide substitution, either between individuals within species or between related species (Timothy et al. 1979; Kung et al. 1982; Bowman et al. 1983; Palmer and Zamir 1982; Tsunewaki and Ogihara 1983).

When examined in greater detail, the distribution of variant sites appears to be nonrandom with respect to duplicated and single copy regions of the molecule. In particular, the inverted repeat region evolves at a slower rate than single copy regions (Bowman et al. 1983; Palmer et al. 1983*a*, 1983*b*; Clegg et al. 1984). Data from *Nicotiana* and *Brassica* species suggest a clustering of variable sites in the small single copy region and in the large single copy region near the junction of the inverted repeats (Kung et al. 1982; Palmer et al. 1983*a*); however, these impressions are still tentative because the number of events observed is too small to enable us to establish statistical significance.

The proportion of nucleotide substitutions per nucleotide position (p) can be estimated from the proportion of fragment differences among sample digests (Nei and Li 1979). Estimates of p for the small single copy region among five grass species belonging to the genus *Pennesitum* average around 0.0034 (Clegg et al. 1984). Similar estimates for the whole chloroplast genome from seven *Brassica* species also average about 0.003 (Palmer et al. 1983*a*). However, these estimates should be interpreted with caution because the estimation model assumes that all restriction site changes are due to nucleotide substitution. Detailed analyses of cpDNA restriction fragment patterns reveal instead that small additions, deletions, and rearrangements are common (Gordon et al. 1982; Bowman et al. 1983; Palmer et al. 1983*a*). To achieve an unambiguous classification of the different kinds of mutational events, it is necessary to compare complete DNA sequences.

Comparisons of Complete DNA Sequences

In many respects, the analysis of complete DNA sequences is most informative about evolutionary processes, especially when taxonomically diverse comparisons are available. A fundamental issue in plant evolution is the origin of the chloroplast which has many prokaryotic characteristics (discussed by Gray and Doolittle [1982]). The endosymbiotic theory suggests that the cyanobacteria (blue-green algae), primitive prokaryotes which carry out a plantlike oxygen-evolving photosynthesis, are related to the evolutionary progenitors of plant chloroplasts. Comparisons of homologous chloroplast and cyanobacterial genes may provide definitive evidence on the origin of the chloroplast genome. We have thus included sequence and organizational data that have recently become available for cyanobacterial genes which have homologues in chloroplast DNA. Our primary focus will concern protein coding genes, and in particular the *rbcL* gene, where sequence data are available from a wide range of plant taxa, including cyanobacteria. In addition, more limited data for several other chloroplast encoded genes will be discussed. We will not consider nucleotide sequence comparisons among chloroplast rRNA genes which are reviewed by Whitfeld and Bottomley (1983).

Ribulose-1, 5-bisphosphate carboxylase/oxygenase (RBC) is a key enzyme in the photosynthetic carbon metabolism of both prokaryotic and eukaryotic autotrophs. RBC is composed of eight large (LSU) and eight small (SSU) subunits of approximately 50,000 and 15,000 daltons, respectively, in both plants and cyanobacteria (Miziorko and Lorimer 1983). The LSU contains the active site for both enzymatic activities and is highly conserved antigenically from cyanobacteria to higher plants (Dorner et al. 1958; Curtis and Haselkorn 1983). The gene encoding the LSU (*rbcL*) is present in one copy per chloroplast DNA molecule. The SSU, whose functions are not understood, is poorly conserved and nuclear encoded in a small multigene family (Berry-Lowe et al. 1982; Coruzzi et al. 1983).

The *rbcL* genes from a number of chloroplast DNAs (Whitfeld and Bottomley 1983) and two cyanobacterial taxa (Curtis and Haselkorn 1983; Reichelt and Delaney 1983; Shinozaki et al. 1983) have been isolated and sequenced. The cyanobacterial genes were isolated using heterologous *rbcL* probes from various chloroplast DNAs. Antigenic relatedness between plant and cyanobacterial proteins and the ability to use chloroplast genes to isolate analogous cyanobacterial genes were early indications that *rbcL* is highly conserved from bacteria to plants. The amino acid sequences derived from translation of *rbcL* genes are compared in figure 2. It is immediately obvious from the comparison that the LSU protein sequences have been highly constrained throughout evolution. The plant sequences are 86%–94% identical, and the cyanobacterial/plant identities are only slightly less, at 78%–85%. Interestingly, the two cyanobacterial proteins are no more similar (84%) than are some plant and cyanobacterial sequences (Curtis and Haselkorn 1983). Included in the highly conserved regions of LSU are those containing residues that have been shown to participate in the active site of the enzyme or the binding of the activator CO_2 molecule (Miziorko and Lorimer 1983). The greatest variation in LSU occurs at the carboxy and amino termini where there is variability in the number of amino acids.

Comparisons of cyanobacterial and chloroplast *rbcL* nucleotide sequence divergence reflect the high levels of similarity seen at the amino acid sequence level. Estimates of the proportion of nucleotide substitutions for first and second codon positions are 0.20 and 0.11, respectively (Shinozaki et al. 1983). Third position sites are, however, effectively randomized in this comparison (Shinozaki et al. 1983).

Although the coding region of *rbcL* is conserved between cyanobacteria and plants, there is little similarity in flanking sequences. With the exception of putative ribosome binding sites (Curtis and Haselkorn 1983; Reichelt and Delaney 1983), there is no obvious similarity between *rbcL* flanking regions of cyanobacteria and higher plants. This is true even of the region to which the initiation of *rbcL* transcription in *Anabaena* has been mapped (Nierzwicki-Bauer et al. 1984). This region does not contain sequences similar to the highly conserved *rbcL* promoter region identified in higher plants (Zurawski et al. 1984).

Moreover, the linkage relationships of *rbcL* and other gene sequences have generally not been conserved between cyanobacteria and plants. In contrast to the organization in chloroplast DNAs, the gene for the subunit of the ATP synthase (*atpB*) in *Anabaena* does not map adjacent to *rbcL* but is at least 10 kbp away (Curtis and Haselkorn 1983). In addition, in cyanobacteria the gene for the small subunit of ribulose-1, 5-bisphosphate carboxylase (*rbcS*) is linked to, and transcribed together with, *rbcL* (Shinozaki and Sugiura 1983; Nierzwicki-Bauer et al.

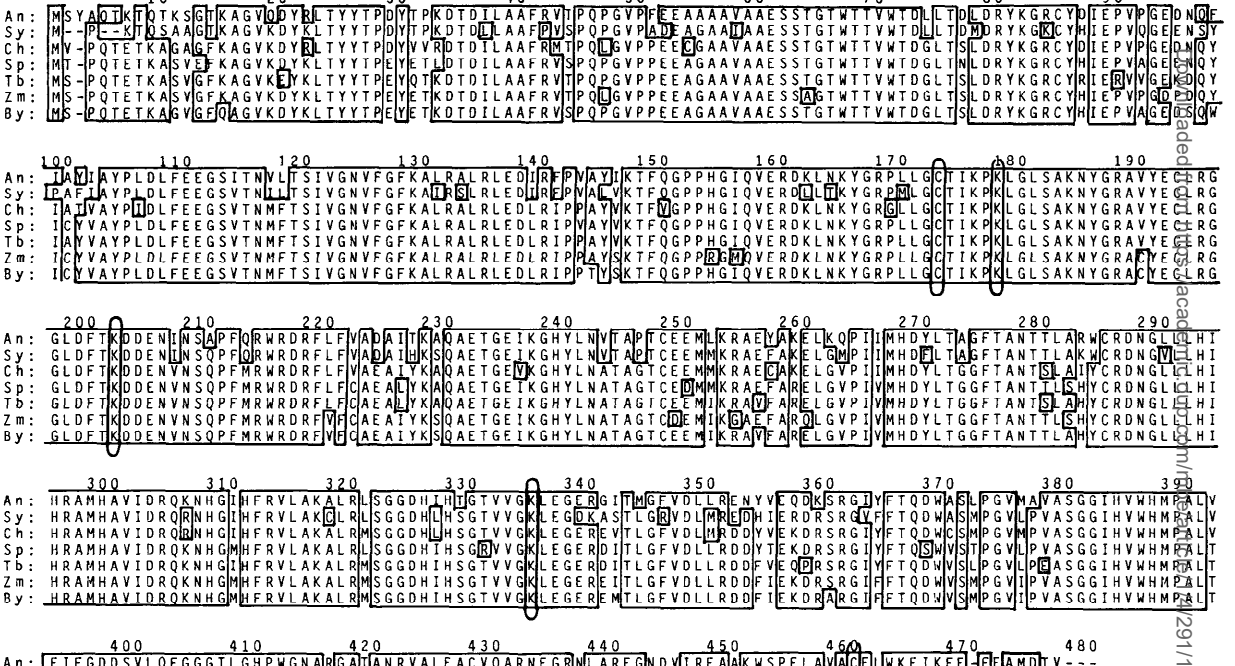

FIG. 2.—Alignment of amino acid sequences derived from translation of *rbcL* genes from *Anabaena* (An [Curtis and Haselkorn 1983]), *Synechococcus* (Sy [Reichelt and Delaney 1983; Shinozaki et al. 1983]), *Chlamydomonas* (Ch [Dron et al. 1982]), spinach (Sp [Zurawski et al. 1981]), tobacco (Tb [Shinozaki and Sugiura 1982]), maize (Zm [McIntosh et al. 1980; Poulsen 1981]), and barley (By [Zurawski et al. 1984*b*]). Sequence data are only available for 427 amino acids of the barley *rbcL* data. Residue numbering refers to the *Anabaena* sequence. Boxes surround positions at which at least five of the seven proteins have identical amino acids, except after residue 428, where at least four of the six are boxed. Circled residues have been suggested as participants in the catalytic site or activator CO_2 binding (Miziorko and Lorimer 1983).

1984). In most plants, *rbcS* is part of a nuclear multigene family, although *rbcL* and *rbcS* have been shown to be linked in the cyanelle DNA of the primitive eukaryote *Cyanophora paradoxa* (Heinhorst and Shively 1983).

A second cyanobacterial/plant comparison is now possible using sequence data for *psbA*, the gene encoding a 32-kilodalton thylakoid membrane protein. The *psbA* genes from spinach (*Spinacia oleracea*) and *Nicotiana debneyii* chloroplasts have been sequenced, and their translation gives proteins of identical amino acid sequence (Zurawski et al. 1982*a*). In contrast to the organization in chloroplast DNA, the cyanobacterium *Anabaena* contains at least three copies of *psbA*, two of which have been sequenced (Curtis and Haselkorn 1984). The two *psbA* genes of *Anabaena* are 90% identical with one another at the amino acid level and 84% identical with the *psbA* genes from spinach and *Nicotiana*. Comparisons of *psbA* flanking sequences between *Anabaena* and higher plants are also consistent with the *rbcL* data in that little sequence similarity is observed.

Taken together, the *rbcL* and *psbA* comparisons between cyanobacteria and higher plants strongly support the inference that these genes derived from a common ancestor. Estimates of divergence times between cyanobacterial and chloroplast genes, while still tentative, also suggest that divergence occurred after the evolution of the eukaryotic cell (Shinozaki et al. 1983). These data may therefore be taken as support for the endosymbiotic origin of the chloroplast genome. Further support for the endosymbiotic theory can be found in comparisons of 16S rDNA sequences from *Escherichia coli*, cyanobacteria, and chloroplast DNA (Tomioka and Sugiura 1983).

To obtain a more detailed view of molecular evolutionary processes, it is necessary to compare more recently diverged sequences. Zurawski et al. (1984) have compared the distribution of nucleotide substitutions in *rbcL* between barley (*Hordeum vulgare*) and maize. Their analysis includes the coding region for *rbcL*, the 5′ leader sequences (~320 bp) of the *rbcL* message, the nontranscribed region (~160 bp) separating the 5′ leaders of *rbcL* and *atpB* and the 5′ leader of *atpB* (~300 b). There is a significant bias in favor of transitions over transversions when nucleotide substitutions are classified into transition or transversion events. In coding regions, a twofold excess of transitions is observed, while the excess is only about 1.6-fold for noncoding regions in the barley-maize data (Zurawski et al. 1984). The formula of Kimura (1981) was used to estimate the proportion of nucleotide substitutions per site for first, second, and third codon positions and for synonymous substitutions in the *rbcL* coding region, as 0.036, 0.019, 0.135, and 0.116, respectively. Interestingly, the noncoding region appears to be accepting nucleotide substitutions at approximately half (0.076) the third position rate. Zurawski et al. (1984) show that the reduced rate can, in part, be accounted for by selective constraints associated with promoter and ribosome binding functions in the noncoding region.

Comparisons of homologous sequence data for noncoding regions of the chloroplast genome also show that short addition/deletion events are common. Moreover, these events are frequently associated with short repeats (Takaiwa and Sugiura 1982; Zurawski et al. 1984). Short addition/deletion events occur approximately one-quarter as often as nucleotide differences in the noncoding region between *rbcL* and *atpB* in the barley-maize comparison (Zurawski et al. 1984).

Recently, Zurawski and Clegg (1984) have estimated the proportion of nucleotide substitutions for *atpB* and *atpE* coding regions from the barley-maize

contrast. An unexpected finding was that the estimates of rates of nucleotide substitutions were statistically homogeneous for *rbcL, atpB,* and *atpE*. Specifically, first, second, and third codon positions gave the same estimates for all three genes, although each codon position has a different rate of substitution. This was surprising because the comparison of derived amino sequences between *E. coli* and spinach or maize *atpB* and *atpE* showed approximately 67% and 26% homology, respectively (Krebbers et al. 1982; Zurawski et al. 1982*b*). The homogeneous rates for the barley-maize comparison suggest a remarkable degree of stochastic regularity on this shorter time scale (50–65 million years since divergence from a common ancestor [Stebbins 1981]). A similar result has been observed for *atpH* where derived amino acid sequences are identical between wheat (Howe et al. 1982) and spinach (Alt et al. 1983*a*) but exhibit only 35% amino acid sequence identity to the homologous *E. coli* gene.

The analysis of complete nucleotide sequence data supports the conclusion from the restriction fragment analyses that the chloroplast genome is conservative in evolution. First, complete sequence data spanning 1,874 nucleotides among two cultivated barleys and 686 nucleotides between *Hordeum vulgare* (cultivated barley) and *H. spontaneum* (the wild progenitor of cultivated barley) show no nucleotide substitutions (Zurawski and Clegg 1984). These sequence runs derive from coding and noncoding regions among three genes. Second, the estimated synonymous rate of nucleotide substitution (approximately 1.1 substitutions per nucleotide per 10^9 years, assuming 50 million years since a common ancestor) between barley and maize for *rbcL, atpB,* and *atpE* (Zurawski and Clegg 1984; Zurawski et al. 1984) is approximately 100-fold less than similar estimates for primate mitochondrial DNA (Brown et al. 1982).

Evolution of Introns

Many chloroplast tRNA genes are interrupted by intervening sequences (Whitfeld and Bottomley 1983). In some instances, long open reading frames have been detected (e.g., maize $tRNA^{Ile}$ and $tRNA^{Ala}$ [Koch et al. 1981]). However, the position and translational sense of these open reading frames are not preserved when homologous sequences are compared among different taxa (Takaiwa and Sugiura 1982). Some statistical evidence for a conserved region in an intron has also been obtained in comparing barley to tobacco $tRNA^{Val}$ (Deno et al. 1983; Zurawski and Clegg 1984). The possibility that functional constraints may influence the evolution of tRNA introns is tantalizing but as yet unproved.

Even more intriguing is the complete loss (or addition) of introns over a broad evolutionary scale. For instance, *rbcL* has nine introns in *Euglena* (Stiegler et al. 1982; R. Hallick, personal communication) but none in higher plants or cyanobacteria. The 23S ribosomal RNA genes of *Chlamydomonas* are unique among 23S rRNA genes studied to date in having a large intron (Allet and Rochaix 1979). In addition, the *psbA* gene contains four introns in *Euglena* (R. Hallick, personal communication) but has none in cyanobacteria or higher plants. Finally, the gene for a large ribosomal subunit protein has an intron in all higher plants so far examined except spinach (G. Zurawski, personal communication). The evolutionary processes which govern the addition or loss of intervening sequences are unknown.

Acknowledgments

We thank an anonymous reviewer for many helpful comments. This work was supported in part by National Science Foundation grant DEB-8118414 to M.T.C. and National Institutes of Health grant GM32 766-01 to S.E.C.

LITERATURE CITED

ALLET, B., and J. D. ROCHAIX. 1979. Structure analysis at the ends of the intervening DNA sequences in the chloroplast 23S ribosomal genes of *C. reinhardii*. Cell **18**:55–60.

ALT, J., W. SEBALD, J. G. MOSER, R. SCHEDEL, P. WESTHOFF, and R. G. HERRMANN. 1983*a*. Localization and nucleotide sequence of the gene for the ATP synthase proteolipid subunit on the spinach plastid chromosome. Current Genet. **7**:129–138.

ALT, J., P. WESTHOFF, B. B. SEARS, N. NELSON, E. HURT, G. HAUSKA, and R. G. HERRMANN. 1983*b*. Genes and transcripts for the polypeptides of the cytochrome b6/f complex from spinach thylakoid membranes. EMBO J. **2**:979–986.

ATCHISON, B. A., P. R. WHITFELD, and W. BOTTOMLEY. 1976. Comparison of chloroplast DNAs by specific fragmentation with *EcoRI* endonuclease. Mol. Gen. Genet. **148**:263–269.

BERRY-LOWE, S. L., T. D. MCKNIGHT, D. M. SHAH, and R. B. MEAGHER. 1982. The nucleotide sequence, expression, and evolution of one member of a multigene family encoding the small subunit of ribulose-1, 5-bisphosphate carboxylase in soybean. J. Mol. Appl. Genet. **1**:483–498.

BOHNERT, H. J., and W. LOFFELHARDT. 1982. Cyanelle DNA from *Cyanophora paradoxa* exists in two forms due to intramolecular recombination. FEBS Lett. **150**:403–406.

BOWMAN, C. M., G. BONNARD, and T. A. DYER. 1983. Chloroplast DNA variation between species of *Triticum* and *Aegilops*. Location of the variation of the chloroplast genome and its relevance to the inheritance and classification of the cytoplasm. Theoret. Appl. Genet. **65**:247–262.

BROWN, W. M., E. M. PRAGER, and A. C. WILSON. 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. J. Mol. Evol. **18**:225–239.

CLEGG, M. T., J. R. Y. RAWSON, and K. THOMAS. 1984. Chloroplast DNA variation in pearl millet and related species. Genetics **106**:449–461.

CORUZZI, G., R. BROGLIE, A. CASHMORE, and N.-H. CHUA. 1983. Nucleotide sequences of the two pea cDNA clones encoding the small subunit of ribulose-1, 5-bisphosphate carboxylase and major chloroplyll a/b-binding thylakoid polypeptide. J. Biol. Chem. **258**:1399–1402.

CURTIS, S. E., and R. HASELKORN. 1983. Isolation and sequences of the gene for the large subunit of ribulose-1, 5-bisphospate carboxylase from the cyanobacterium *Anabaena* 7120. Proc. Nat. Acad. Sci. USA **80**:1835–1839.

———. 1984. Structure and expression of two members of the 32 kilodalton thylakoid membrane protein gene family in the cyanobacterium *Anabaena* 7120. Plant Mol. Biol. (accepted).

DENO, H., K. SHINOZAKI, and M. SUGIURA. 1983. Nucleotide sequence of tobacco chloroplast gene for the subunit of proton-translocating ATPase. Nucleic Acids Res. **11**:2185–2191.

DORNER, R. W., A. KAHN, and S. G. WILDMAN. 1958. Proteins of green leaves. VIII. The distribution of fraction I protein in the plant kingdom as detected by precipitin and ultracentrifugal analyses. Biochem. Biophys. Acta **29**:240–245.

DRON, M., M. RAHIRE, and J.-D. ROCHAIX. 1982. Sequence of the chloroplast DNA region of *Chlamydomonas reinhardii* containing the gene of the large subunit of ribulose bisphospate carboxylase and parts of its flanking genes. J. Mol. Biol. **162**:775–793.

GORDON, K. H. J., E. J. CROUSE, H. J. BOHNERT, and R. G. HERRMANN. 1982. Physical mapping of differences in chloroplast DNA of the five wild-type plastomes in *Oenothera* subsection *euoenothera*. Theoret. Appl. Genet. **61:**373–384.

GRAY, M. W., and W. F. DOOLITTLE. 1982. Has the endosymbiont hypothesis been proven? Microbiol. Rev. **46:**1–42.

HEINHORST, S., and J. M. SHIVELY. 1983. Encoding of both subunits of ribulose-1, 5-bisphosphate carboxylase by organelle genome of *Cyanophora paradoxa*. Nature **304:**373–374.

HOWE, C. J., C. M. BOWMAN, T. A. DYER, and J. C. GRAY. 1983. The genes for the alpha and proton-translocating subunits of wheat chloroplast ATP synthase are close together on the same strand of chloroplast DNA. Mol. Gen. Genet. **190:**51–55.

HOWE, C. J., A. D. SUFFRET, A. DOHERTY, C. M. BOWMAN, T. A. DYER, and J. C. GRAY. 1982. Location and nucleotide sequence of the gene for the proton-translocating subunit of wheat chloroplast ATP synthase. Proc. Nat. Acad. Sci. USA **79:**6903–6907.

KIMURA, M. 1981. Estimation of evolutionary distances between homologous nucleotide sequences. Proc. Nat. Acad. Sci. USA **78:**454–458.

KOCH, W., K. EDWARDS, and H. KOSSELL. 1981. Sequencing of the 16S-23S spacer in a ribosomal RNA operon of *Zea mays* chloroplast DNA reveals two split tRNA genes. Cell **25:**203–213.

KOLLER, B., and H. DELIUS. 1982. A chloroplast DNA of *Euglena gracilis* with five complete rRNA operons and two extra 16S rRNA genes. Mol. Gen. Genet. **188:**305–308.

KREBBERS, E. T., I. M. LARRINUA, L. MCINTOSH, and L. BOGORAD. 1982. The maize chloroplast genes for the β and ϵ subunits of the photosynthetic coupling factor CF_1 are fused. Nucleic Acids Res. **10:**4985–5002.

KUNG, S. D., Y. S. ZHU, and G. F. SHEN. 1982. *Nicotiana* chloroplast genome. III. Chloroplast DNA evolution. Theoret. Appl. Genet. **61:**73–79.

MCINTOSH, L., C. POULSEN, and L. BOGORAD. 1980. Chloroplast gene sequence for the large subunit of ribulose bisphospate carboxylase of maize. Nature **288:**556–560.

MIZIORKO, H. M., and G. H. LORIMER. 1983. Ribulose-1, 5-bisphosphate carboxylase-oxygenase. Annu. Rev. Biochem. **52:**507–535.

NEI, M., and W.-H. LI. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Nat. Acad. Sci. USA **76:**5269–5273.

NIERZWICKI-BAUER, S., S. E. CURTIS, and R. HASELKORN. 1984. The genes for the large and small subunits of ribulose-1, 5-bisphosphate carboxylase are linked and transcribed together in the cyanobacterium *Anabaena* 7120 (in preparation).

OHYAMA, K., Y. YAMANO, H. FUKUZAWA, T. KOMANO, H. YAMAGISHI, S. FUJIMOTO, and M. SUGIURA. 1983. Physical mappings of chloroplast DNA from liverwort *Marchantia polymorpha* L. cell suspension cultures. Mol. Gen. Genetics **189:**1–9.

PALMER, J. D., C. R. SHIELDS, D. B. COHEN, and T. J. ORTON. 1983*a*. Chloroplast DNA evolution and the origin of amphidiploid *Brassica* species. Theoret. Appl. Genet. **65:**181–189.

PALMER, J. D., G. P. SINGH, and D. T. N. PILLAY. 1983*b*. Structure and sequence evolution of three legume chloroplast DNAs. Mol. Gen. Genet. **190:**13–19.

PALMER, J. D., and D. B. STEIN. 1982. Chloroplast DNA from the fern *Osmunda cinnamomea:* physical organization, gene location and comparison to angiosperm chloroplast DNA. Current Genet. **5:**165–170.

PALMER, J. D., and W. F. THOMPSON. 1981. Rearrangements in the chloroplast genomes of mung bean and pea. Proc. Nat. Acad. Sci. USA **78:**5537–5544.

———. 1982. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. Cell **29:**537–550.

PALMER, J. D., and D. ZAMIR. 1982. Chloroplast DNA evolution and phylogenetic relationships in *Lycopersicon*. Proc. Nat. Acad. Sci. USA **79:**5006–5010.

POULSEN, C. 1981. Comments on the structure and function of the large subunit of the enzyme ribulose bisphosphate carboxylase-oxygenase. Carlsberg Res. Commun. **46:**259–273.

REICHELT, B., and S. DELANEY. 1983. The nucleotide sequence for the large subunit of ribulose-1, 5-bisphospate carboxylase from a unicellular cyanobacterium, *Cynechococcus* PCC6301. DNA **2:**121–130.

ROCHAIX, J. D. 1981. Organization, function and expression of the chloroplast DNA of *Chlamydomonas reinhardii*. J. Mol. Biol. **126:**596–617.

SHINOZAKI, K., and M. SUGIURA. 1982. The nucleotide sequence of the tobacco chloroplast gene for the large subunit of ribulose-1, 5-bisphosphate carboxylase/oxygenase. Gene **20:**91–102.

———. 1983. The gene for the small subunit of ribulose-1, 5-bisphosphate carboxylase/oxygenase is located close to the gene for the large subunit in the cyanobacterium *Anacytis nidulans* 6301. Nucleic Acids Res. **11:**6957–6964.

SHINOZAKI, K., C. YAMADA, N. TAKAHATA, and M. SUGIURA. 1983. Molecular cloning and sequence analysis of the cyanobacterial gene for the large subunit of ribulose-1, 5-bisphospate carboxylase/oxygenase. Proc. Nat. Acad. Sci. USA **80:**4050–4054.

SPIELMANN, A., W. ORTIZ, and E. STUTZ. 1983. The soybean chloroplast genome: construction of a circular restriction site map and location of DNA regions encoding the genes for rRNAs, the large subunit of the ribulose-1, 5-bisphosphate carboxylase and the 32 KD protein of the photosystem II reaction center. Mol. Gen. Genet. **190:**5–12.

STEBBINS, G. L. 1981. Coevolution of grasses and herbivores. Ann. Missouri Bot. Garden. **68:**75–86.

STIEGLER, G. L., H. M. MATTHEWS, S. E. BINGHAM, and R. B. HALLICK. 1982. The gene for the large subunit of ribulose-1, 5-bisphospate carboxylase in *Euglena gracilis* chloroplast DNA: location, polarity, cloning and evidence for an intervening sequence. Nucleic Acids Res. **10:**3427–3444.

SUBRAMANIAN, A. R., A. STEINMETZ, and L. BOGORAD. 1983. Maize chloroplast DNA encodes a protein sequence homologous to the bacterial ribosomal protein 54. Nucleic Acids Res. **11:**5277–5286.

SUGITA, M., and M. SUGIURA. 1983. A putative gene of tobacco chloroplast coding for ribosomal protein similar to *E. coli* ribosomal protein S19. Nucleic Acids Res. **11:**1913–1918.

TAKAIWA, F., and M. SUGIURA. 1982. Nucleotide sequences of the 16S-23S spacer region in an rRNA gene cluster from tobacco chloroplast DNA. Nucleic Acids Res. **8:**2665–2676.

TIMOTHY, D. H., C. S. LEVINGS III, D. R. PRING, M. F. CONDE, and J. L. KERNICKE. 1979. Organelle DNA variation and systematic relationships in the genus *Zea:* teosinte. Proc. Nat. Acad. Sci. USA **76:**4220–4224.

TOMIOKA, N., and M. SUGIURA. 1983. The complete nucleotide sequence of a 16S ribosomal RNA gene from a blue-green alga, *Anacystis nidulans*. Mol. Gen. Genet. **191:**46–50.

TSUNEWAKI, K., and Y. OGIHARA. 1983. The molecular basis of genetic diversity among cytoplasms of *Triticum* and *Aegilops* species. II. On the origin of polyploid wheat cytoplasms as suggested by chloroplast DNA restriction fragment patterns. Genetics **104:**155–171.

WESTHOFF, P., J. ALT, N. NELSON, W. BOTTOMLEY, H. BUNEMANN, and R. G. HERRMANN. 1983. Genes and transcripts for the P_{700} chlorophyll *a* apoprotein and subunit 2 of the photosystem I reaction center complex from spinach thylakoid membranes. Plant Mol. Biol. **2:**95–107.

WHITFELD, P., and W. BOTTOMLEY. 1983. Organization and structure of chloroplast genes. Annu. Rev. Plant Physiol. **34:**279–326.

WILLEY, D. L., A. K. HUTTLEY, A. L. PHILLIPS, and J. C. GRAY. 1983. Localization of the genes for cytochrome f in pea chloroplast DNA. Mol. Gen. Genet. **189:**85–89.

WURTZ, E. A., and D. E. BUETOW. 1981. Intra-specific variation in the structural organization and redundancy of chloroplast ribosomal DNA cistrons in *Euglena gracilis*. Current Genet. **3**:181–187.

ZURAWSKI, G., H. J. BOHNERT, P. WHITFELD, and W. BOTTOMLEY. 1982*a*. Nucleotide sequence of the gene for the M_r 32,000 thylakoid membrane protein from *Spinacia oleracea* and *Nicotiana debneyi* predicts a totally conserved primary translation product of M_r 38,950. Proc. Nat. Acad. Sci. USA **79**:7699–7703.

ZURAWSKI, G., W. BOTTOMLEY, and P. R. WHITFELD. 1982*b*. Structures of the genes for the β and ϵ subunits of spinach chloroplast ATPase indicate a dicistronic mRNA and an overlapping translation stop/start signal. Proc. Nat. Acad. Sci. USA **79**:6260–6264.

ZURAWSKI, G., and M. T. CLEGG. 1984. The barley chloroplast DNA *atpBE, trnM2,* and *trnV1* loci. Nucleic Acids Res. (accepted).

ZURAWSKI, G., M. T. CLEGG, and A. H. D. BROWN. 1984.The nature of nucleotide sequence divergence between barley and maize chloroplast DNA. Genetics (accepted).

ZURAWSKI, G., B. PERROT, W. BOTTOMLEY, and P. R. WHITFELD. 1981. The structure of the gene for the large subunit of ribulose 1, 5-bisphosphate carboxylase from spinach chloroplast DNA. Nucleic Acids Res. **9**:3251–3269.

WALTER M. FITCH, reviewing editor

Received September 28, 1983; revision received November 22, 1983.