J Mol Evol (1997) 45:688–695

JOURNAL OF MOLECULAR EVOLUTION

Analysis of the Cluster of Ribosomal Protein Genes in the Plastid Genome of a Unicellular Red Alga *Cyanidioschyzon merolae:* Translocation of the *str* Cluster as an Early Event in the Rhodophyte-Chromophyte Lineage of Plastid Evolution

Niji Ohta,[1] Naoki Sato,[2] Hisayoshi Nozaki,[3] Tsuneyoshi Kuroiwa[3]

[1] School of Human Science, Waseda University, 2-579-15 Mikajima, Saitama, 359 Japan
[2] Department of Biochemistry and Molecular Biology, Faculty of Science, Saitama University, 255 Shimo-Ohkubo Urawa, Saitama, 338 Japan
[3] Department of Biological Sciences, Graduate School of Sciences, University of Tokyo, 7-3-1 Hongo, Tokyo, 113 Japan

Received: 2 June 1997 / Accepted: 15 July 1997

Abstract. The nucleotide sequence of a cluster of ribosomal protein genes in the plastid genome of a unicellular red alga, *Cyanidioschyzon merolae*, which has been supposed to be the most primitive alga, was determined. The phylogenetic tree inferred from the amino acid sequence of ribosomal proteins of two rhodophytes, a chromophyte, a glaucophyte, two chlorophytes (land plants), a cyanobacterium, and three eubacteria suggested a close relationship between the cyanobacterium *Synechocystis* PCC6803 and the plastids of various species in the kingdom Plantae, which is consistent with the hypothesis of the endosymbiotic origin of plastids. In this tree, the two species of rhodophytes were grouped with the chromophyte, and the glaucophyte was grouped with the chlorophytes. Analysis of the organization of the genes encoding the ribosomal proteins suggested that the translocation of the *str* cluster occurred early in the lineage of rhodophytes and chromophytes after these groups had been separated from chlorophytes and glaucophytes.

Key words: Cyanidiophyceae — *Cyanidioschyzon merolae* — Plastid genome — Protoflorideophyceae — Ribosomal protein — Translocation of *str* cluster

Introduction

Plastids are thought to be descendants of cyanobacterial endosymbionts according to the endosymbiosis theory (Margulis 1970; Taylor 1974). Plastids of rhodophytes might be a likely link between the cyanobacteria and the plastids of land plants, but our knowledge on the plastid genome of rhodophytes is limited. Studies on the relationship of cyanobacteria and various plastids that are based on the nucleotide sequence of 16S rRNA are still controversial (Nelissen et al. 1995). Information on phylogenetic relationship that is inferred from sequence data other than the nucleotide sequence of 16S rRNA may also be helpful, because the phylogenetic trees inferred from 16S rRNA are often affected by the G+C content and some other factors (Helmchen et al. 1995).

Until now, complete nucleotide sequence of the genome has been determined in a cyanobacterium (*Synechocystis* PCC6803; Kaneko et al. 1996), plastids of a glaucophyte (*Cyanophora paradoxa;* Stirewalt et al. 1995), a multicellular rhodophyte (*Porphyra purpurea,* Reith and Munholland 1995), a chromophyte (*Odontella sinensis,* Kowallik et al. 1995), and chlorophytes (plants, for example, *Marchantia polymorpha,* Ohyama et al. 1986; *Nicotiana tabacum,* Shinozaki et al. 1986). These complete nucleotide sequences made it possible to use long conserved stretches for the analysis of the phylogenetic relationships of various photosynthetic organisms. Sequence data of algae which are likely to be the most

The nucleotide sequence reported in this paper has been submitted to EMBL/Gen bank/DDBJ with accession number D89930
Correspondence to: N. Ohta; e-mail: niji@human.waseda.ac.jp

primitive would be useful in analyzing the relationship between cyanobacteria and plastids. *Cyanidioschyzon merolae* is a unicellular rhodophyte that belongs to Protoflorideophyceae or Cyanidiophyceae (Seckbach 1992), which is supposed to be the most primitive group of rhodophytes, and is closely related to *Cyanidium caldarium* (Ohta et al. 1997). The primitiveness of this group has been supported by various observations such as mode of cell division, localization of plastid nuclei (we prefer to use this word rather than ''nucleoids''), and the small size of the nuclear genome (Ohta et al. 1993; Kuroiwa et al. 1994). Molecular phylogenetic analysis inferred from *rbcL* (Fujiwara et al. 1994; Delwiche and Palmer 1996) and 16S rRNA (Nelissen et al. 1995) also supported the primitiveness of Protoflorideophyceae. We attempted to use the ribosomal protein gene cluster to compare various plastids, cyanobacteria, and noncyanobacterial bacteria, because ribosomal proteins are present in a wide variety of prokaryotes and plastids.

The gene cluster encoding the ribosomal proteins has been extensively studied in *Escherichia coli.* In *E. coli,* ribosomal protein genes are arranged as ribosomal gene clusters. In these clusters, multiple genes are cotranscribed as ribosomal protein operon (Cerretti et al. 1983). Similar operons are present in the plastid genomes that have been sequenced (Sugiura 1992) and the cyanobacterial genome (Kaneko et al. 1996). In these genomes, more than half of the ribosomal proteins belong to one of the three operons, *S10* (Zurawski and Zurawski 1985), *spc* (Cerretti et al. 1983), and α (Bedwell et al. 1985). These operons are, in turn, parts of a large ribosomal gene cluster. The *str* operon is located 16.9 kbp away from this large cluster in *E. coli* (Post and Nomura 1980).

In the present study, we cloned and sequenced the gene cluster that encodes ribosomal proteins in a unicellular red alga, *Cyanidioschyzon merolae,* which has been supposed to be the most primitive alga. We compared the order of the genes in this cluster to that in the homologous gene cluster of other species. We also constructed a phylogenetic tree inferred from ribosomal proteins. We found that the *str* cluster has been translocated in the lineage of rhodophytes and chromophytes, suggesting that this translocation occurred early after the separation of this lineage from the chlorophyte and glaucophyte lineage.

Materials and Methods

Cells of *Cyanidioschyzon merolae* were grown in the medium of Allen (1959) as previously described (Suzuki et al. 1992). Plastid DNA was isolated according to the methods described by Suzuki et al. (1992). The DNA was partially digested with restriction endonuclease *Sau*3AI and the resultant fragments were cloned in lambda DASH II (Stratagene, La Jolla, CA, USA). Subcloning into pBluescript II SK+ (Stratagene, La Jolla, CA, USA) was performed using *E. coli* XL-1 blue as the host bacterium. A part of the ribosomal protein gene cluster was obtained first by random sequencing of cloned fragments, and then the whole gene cluster was cloned and sequenced by genome walking. Similarity search of the putative open reading frames against the SwissProt and CyanoBase databases, respectively, was performed with the BLAST program (Karlin and Altschul 1990) at the Genome Net WWW Server through the Internet. The deduced amino acid sequences from *rpl2, rps19, rpl22, rps3, rpl16, rpl14,* and *rps8* were combined as single sequences, and then they were aligned with Clustal W program version 1.6 for the PowerMac (Thompson et al. 1994). Phylogenetic trees based on the combined amino acid sequences were constructed by the maximum-likelihood method (Felsenstein 1981) with the PUZZLE program for the PowerMac (Strimmer and von Haeseler 1996) using the JTT model of sequence evolution (Jones et al. 1992) and 1,000 puzzling steps. The following database entries of amino acid sequences were used in the calculation. *Bacilus subtilis,* U43929 (*rpl2, rps19, rpl22, rps3, rpl16*) and L47971 (*rpl14, rps8*); *Cyanophora paradoxa,* U30821; *Escherichia coli,* U18997; *Haemophilus influenzae,* U32761 (*rpl2, rps19, rpl22, rps3, rpl16*) and U32762 (*rpl14, rps8*); *Marchantia polymorpha,* X04465; *Nicotiana tabacum,* Z00044; *Odontella sinensis,* Z67753; *Porphyra purpurea,* U38804; *Synechocystis* PCC6803, D90905.

Results and Discussion

Gene Order

Nucleotide sequence of the ribosomal protein gene cluster of the plastid genome of *C. merolae* was determined. We identified the following genes in the following order: *rpl3, rpl4, rpl23, rpl2, rps19, rpl22, rps3, rpl16, rpl29, rps17, rpl14, rpl24, rpl5, rps8, rpl6, rpl18, rps5, secY* (gene for the preprotein translocater subunit), *rpl36, rps13, rps11, rpoA* (gene for the RNA polymerase alpha subunit), *rpl13, rps9, rpl31, rps12, rps7, tufA* (gene for the elongation factor Tu), and *rps10.* The organization of the gene cluster for the ribosomal proteins of the plastid genome of *C. merolae* was compared to that of the gene cluster of *Porphyra purpurea* (Reith and Munholland 1995), *Odontella sinensis* (Kowallik et al. 1995), *Cyanophora paradoxa* (Stirewalt et al. 1995), *Marchantia polymorpha* (Ohyama et al. 1986), and *Nicotiana tabacum* (Shinozaki et al. 1986). We also compared these with the structure of the ribosomal gene clusters of mitochondrial genomes of *Acanthamoeba castellanii* (Burger et al. 1995) and *M. polymorpha* (Oda et al. 1992) and the genomes of *Synechocystis* PCC6803 (Kaneko et al. 1996), *Escherichia coli* (*E. coli* database collection; http://susi.bio.uni-giessen.de/usr/local/www/html/ecdc.html), and *Haemophilus influenzae* (*H. influenzae* database collection; http://susi.bio.uni-giessen.de/usr/local/www/html/hidc.htm). The results are shown in Fig. 1. Although the ribosomal protein operons in *E. coli* are called *S10, spc,* α, and *str,* respectively, we use the names, *L2, spc,* α, and *str,* respectively. We use ''*L2* cluster'' instead of ''*S10* cluster'' since the *rps10* gene is not always present in the gene clusters that correspond to the *S10* operon in *E. coli.* We have chosen ''*L2*'' as the name of this gene cluster because *rpl2* is located most upstream among the genes that are conserved in the gene cluster of all of the species analyzed in the present study.

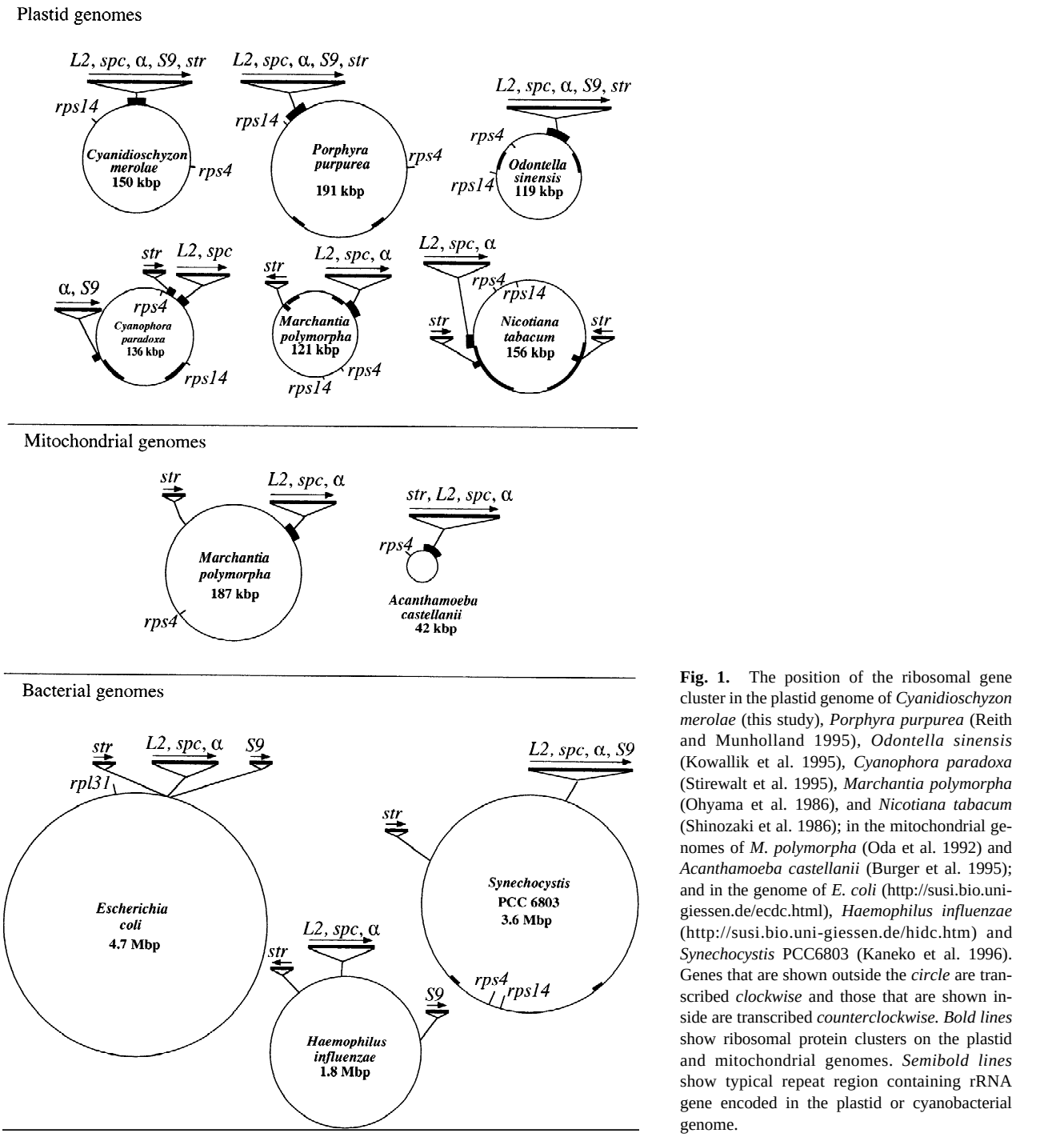

Fig. 1. The position of the ribosomal gene cluster in the plastid genome of *Cyanidioschyzon merolae* (this study), *Porphyra purpurea* (Reith and Munholland 1995), *Odontella sinensis* (Kowallik et al. 1995), *Cyanophora paradoxa* (Stirewalt et al. 1995), *Marchantia polymorpha* (Ohyama et al. 1986), and *Nicotiana tabacum* (Shinozaki et al. 1986); in the mitochondrial genomes of *M. polymorpha* (Oda et al. 1992) and *Acanthamoeba castellanii* (Burger et al. 1995); and in the genome of *E. coli* (http://susi.bio.uni-giessen.de/ecdc.html), *Haemophilus influenzae* (http://susi.bio.uni-giessen.de/hidc.htm) and *Synechocystis* PCC6803 (Kaneko et al. 1996). Genes that are shown outside the *circle* are transcribed *clockwise* and those that are shown inside are transcribed *counterclockwise. Bold lines* show ribosomal protein clusters on the plastid and mitochondrial genomes. *Semibold lines* show typical repeat region containing rRNA gene encoded in the plastid or cyanobacterial genome.

Since the gene cluster that is composed of three genes, *rpl13, rps9,* and *rpl31,* has not been described in *E. coli,* we define this gene cluster as *S9.* We use the term ''cluster'' instead of ''operon'' to designate the group of genes in the algal plastomes since the exact transcription units are not established in these organisms, and since, at least in one case, *secY* in a rhodophyte *C. caldarium* is reported to be transcribed monocistronically (Vogel et al. 1996).

In the plastid genomes of *C. merolae, P. purpurea,* and *O. sinensis,* the ribosomal protein genes investigated in this study were present in a single cluster. In *Synechocystis,* the *L2, spc,* α, and *S9* clusters made a large single cluster, whereas the *str* cluster was located apart. In *E. coli,* the three operons named *L2 (S10), spc,* and α made a large single cluster but the *S9* cluster is located downstream of the large cluster. In the plastid genome of *C. paradoxa,* the clusters *L2* and *spc* were grouped in a large cluster, and the clusters α and *S9* were also grouped as a separate large cluster. In the plastid genome of *M. polymorpha* and *N. tabacum,* the *S9* cluster was not present in the plastid genome. Though the clusters *L2, spc,*

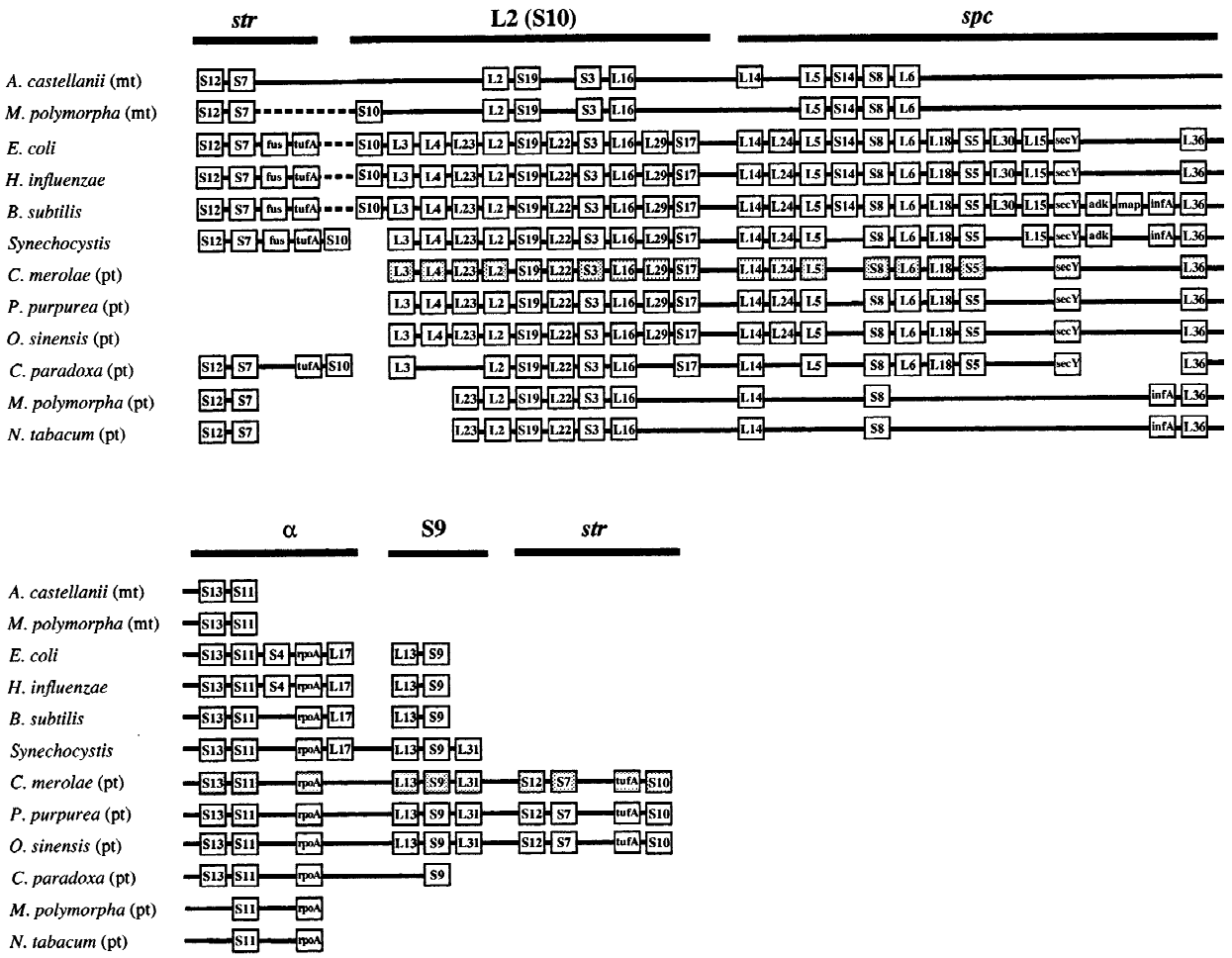

Fig. 2. Comparison of the organization of the ribosomal protein gene clusters, *L2, spc,* α, *S9,* and *str*. The organizations in the genomes of *Escherichia coli, Haemophilus influenzae, Bacillus subtilis,* and *Synechocystis* PCC6803, the plastid genomes of *Cyanidioschyzon merolae, Porphyra purpurea, Odontella sinensis, Cyanophora paradoxa, Marchantia polymorpha,* and *Nicotiana tabacum,* as well as the mitochondrial genomes of *Acanthamoeba castellanii* and *Marchantia polymorpha* are compared. The genes that are contiguous are linked by *bold lines.*

and α were grouped as a single large cluster, the *str* cluster was located at a distant locus. In the mitochondrial genome of *A. castellanii,* the *str, L2, spc,* and α clusters made a large single cluster, while in the mitochondrial genome of *M. polymorpha,* the *L2, spc,* and α clusters made a single cluster but the *str* cluster was located at a distant locus. The *S9* cluster was not present in both of the mitochondrial genomes.

Gene organization of the ribosomal gene clusters from various organisms is compared in detail in Fig. 2. In *E. coli, H. influenzae,* and *B. subtilis,* the first gene of the *L2* cluster was *rps10,* whereas the first gene of the corresponding cluster was *rpl3* in *Synechocystis* and in the plastid genomes of *C. merolae, P. purpurea, O. sinensis,* and *C. paradoxa.* In *Synechocystis* and *C. paradoxa, rps10* was present at the end of the *str* cluster, which was located upstream of the *L2* cluster. In *C. merolae, P. purpurea,* and *O. sinensis, rps10* was present at the end of the *str* cluster, which was located downstream of the large ribosomal gene cluster. The first gene in the *L2* cluster was *rpl23* in the plastid genomes of *M. polymorpha* and *N. tabacum.* The *rps14* gene, which was a member of the *spc* cluster in *E. coli, H. influenzae,* and *B. subtilis* and the mitochondrial genome of *A. castellanii* and *M. polymorpha* was located away from the ribosomal gene cluster in the cyanobacterium and the plastids (Figs. 1 and 2).

The *spc* cluster contained various genes that encode nonribosomal proteins in bacteria and plastids. The *secY* gene was present in this cluster in *E. coli, H. influenzae, B. subtilis, Synechocystis,* and the plastid genomes of *C. merolae, P. purpurea, O. sinensis,* and *C. paradoxa. B. subtilis* also contained *adk* (gene for the adenylate kinase), *map* (gene for the methionine aminopeptidase), and *infA* (gene for the initiation factor IF-1). *Synechocystis* contained *adk* and *infA,* while the plastid genomes of *M. polymorpha* and *N. tabacum* contained *infA.* The genes *rps14, rpl30, rps4, rpl17,* and *fus,* which were present in *E. coli* and *H. influenzae,* were not present in the corresponding clusters in the plastid genomes of *C. merolae, P. purpurea,* and *O. sinensis. Synechocystis* rather resembled *C. merolae* in this respect; it did not

contain the genes *rps14, rpl30,* and *rps4* in the ribosomal gene clusters.

The gene *rps4* was specific to *E. coli* and *H. influenzae.* These two bacteria as well as *B. subtilis* contained the genes *rpl30* and *rps14,* and the *rps10* gene preceded the *rpl3* gene. The *rpl31* gene was present at the end of the *S9* cluster in *Synechocystis* and the plastid genomes of *C. merolae, P. purpurea,* and *O. sinensis.* In these four genomes, the clusters *L2, spc,* α, and *S9* made a large cluster. The *str* cluster was also a part of this large cluster in *C. merolae, P. purpurea,* and *O. sinensis.* The genes *rpl3, rpl5, rpl6, rpl18, rps5, rps9, rps13, rps17, secY,* and *tufA* were not present in the plastid genomes of *M. polymorpha* and *N. tabacum.* The continuity of the α cluster and the *spc* cluster was not conserved in the plastid genome of *C. paradoxa.* We are interested in the fact that the *str* cluster is located at the end of the *S9* cluster in the plastid genomes of *C. merolae, P. purpurea,* and *O. sinensis,* whereas it is located upstream of the *L2* cluster in most other species. It is interesting to note that *rps10* is located at the 5′ end of the *L2* cluster (in bacteria) or at the 3′ end of the *str* cluster (in *Synechocystis* and plastids). In the mitochondrial genome of *A. castellanii,* the *str* and L2 clusters were contiguous. In *B. subtilis,* only a single open reading frame is located between the *str* cluster and the *L2* cluster. All of these results suggest that the plastid genomes of *C. merolae, P. purpurea,* and *O. sinensis* are especially similar, and they are distinct from the plastid genomes of chlorophytes (land plants) and *C. paradoxa.*

Phylogenetic Tree

The phylogenetic tree inferred from the amino acid sequences of the ribosomal proteins was constructed excluding the mitochondrial sequences (Fig. 3). We did not use mitochondrial ribosomal proteins, since the similarities of mitochondrial ribosomal genes to other sources are markedly low (Burger et al. 1995).

The deduced amino acid sequences from *rpl2, rps19, rpl22, rps3, rpl16, rpl14,* and *rps8* were combined as single sequences, and then they were aligned with Clustal W and a phylogenetic tree was constructed. We chose these ribosomal proteins to construct a phylogenetic tree because these ribosomal protein genes are common in the organisms analyzed in the present study, and they are all located within one large ribosomal gene cluster. Phylogenetic trees constructed by the maximum-likelihood method (Felsenstein 1981) are shown in Fig. 3. *E. coli, H. influenzae,* and *B. subtilis* were used as outgroups.

Synechocystis was a sister group of the plastid group, which suggests that an ancestral cyanobacterium was the origin of the plastid as proposed in previous studies (Bergsland and Haselkorn 1991; Morden et al. 1992; Helmchen et al. 1995). In the plastid group, two major

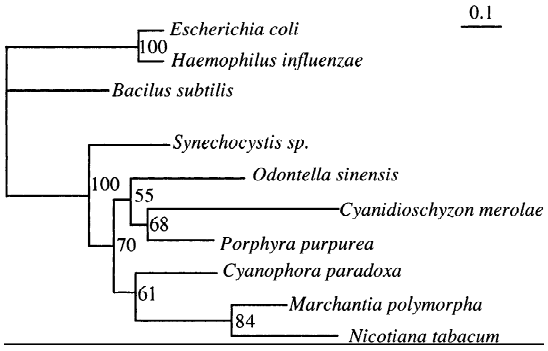

Fig. 3. Phylogenetic tree inferred from the amino acid sequences of ribosomal proteins. Amino acid sequences that are deduced from nucleotide sequences of *rpl2, rps19, rpl22, rps3, rpl16, rpl14,* and *rps8* were combined and used in the analysis. The tree was constructed by the maximum likelihood method (Felsenstein 1981) with the PUZZLE software (Strimmer and von Haeseler 1996). The *number at each node* shows the reliability value of the branch, which was calculated from 1,000 quartet puzzling steps. The *scale bar* represents 0.1 mutations per site. Branch lengths are drawn to scale.

groups can be distinguished. One contains chromophyte and rhodophyte plastids and the other contains glaucophyte and chlorophyte plastids. It is interesting that the glaucophyte plastid is grouped with chlorophytes (land plants). In the phylogenetic studies of 16S ribosomal RNA, glaucophyte plastids are grouped with rhodophyte-chromophyte plastids or chlorophyte plastids are grouped with rhodophyte-chromophyte plastids (Helmchen et al. 1995; Nelissen et al. 1995). In the present study, we used ribosomal protein genes to discuss phylogeny of these algae and plants. This tree (Fig. 3) was consistent with the result of comparison of gene organization (Figs. 1 and 2). On the contrary, the phylogenetic trees of 16S ribosomal RNA are not consistent with both the tree of ribosomal proteins and the organization of ribosomal genes.

Another point that is different from previous studies is that the plastid of *C. merolae* is located within the group of rhodophytes (Fig. 3). In the tree inferred from *rbcL,* the plastid of *C. merolae* and with the plastid of *C. caldarium* belong to the first branch that separates from the rest of the plastid group (Ohta et al. 1997), and several previous analyses showed that *C. caldarium* was the first branch in the plastid group (Fujiwara et al. 1994; Delwiche and Palmer 1996; Nelissen et al. 1995).

Model

The results of phylogenetic analysis and comparisons of gene arrangement are most reasonably understood if we assume that the *str* and the *L2* clusters had been a part of a large gene cluster in the origin. A plausible gene organization in the origin might have been *rps12-rps7-fusA-tufA-rps10-rpl3-rpl4-rpl23-rpl2* . . . (etc.). This gene arrangement was broken in two different ways in different branches of evolution (Fig. 2): In bacteria, the linkage

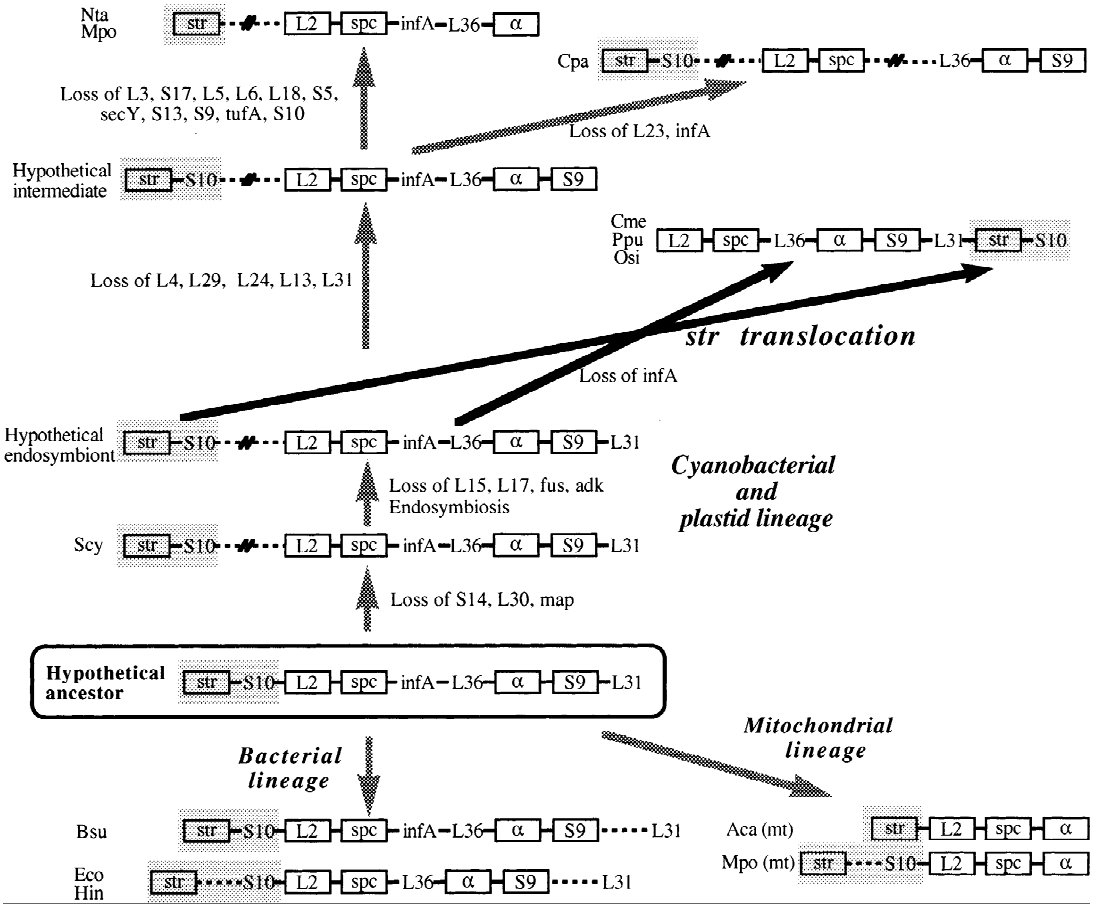

Fig. 4. A model of the evolution of the ribosomal protein gene cluster. A hypothetical gene arrangement was made from the gene organization of *Synechocystis* PCC6803 and *Bacillus subtilis.* Major gene clusters are shown by *boxes.* Some genes that are relevant in considering the reorganization of the gene cluster are also included. For simplicity, the names of the genes are indicated by the name of the protein, e.g., *L2* instead of *rpl2.* In the bacterial lineage, separation of the *rpl31* gene and the str cluster took place. We cannot give a comprehensive route for the diversification of the bacterial genomes because of lack of data. The mitochondrial lineage is a simple derivation of the hypothetical gene cluster by loss of several genes. In *Synechocystis,* a separation took place between the *rps10* gene and the *L2* cluster. Each of the ribosomal gene clusters of all the plastids can be explained as a derivative of the cyanobacterial gene cluster. After the endosymbiosis of a hypothetical cyanobacterial progenitor into a eukaryotic cell, two lineages emerged. Early in the rhodophyte-chromophyte lineage, the *str* cluster has been translocated to the end of the *S9* cluster. In the glaucophyte-chlorophyte cluster, various genes have been removed from the cluster into the nuclear genome. In the glaucophyte lineage, a separation took place between the *spc* cluster and the *rpl36* gene. In the chlorophyte lineage, a number of additional genes have been removed from the gene cluster. *Aca, Acanthamoeba castellanii; Bsu, Bacilus subtilis; Cme, Cyanidioschyzon merolae; Cpa, Cyanophora paradoxa; Eco, Escherichia coli; Hin, Haemophilus influenzae; Mpo, Marchantia polymorpha; Nta, Nicotiana tabacum; Osi, Odontella sinensis; Ppu, Porphyra purpurea; Scy, Synechocystis* PCC6803.

between *tufA* and *rps10* has been disrupted by insertion of unrelated genes. In *Synechocystis,* the linkage between *rps10* and *rpl3* has been broken by rearrangement, and this breakage is conserved in all the plastids that are descendants of a hypothetical cyanobacterial progenitor. In *C. merolae, P. purpurea,* and *O. sinensis,* the *str* cluster was again grouped with the main ribosomal gene cluster, but at the end of the *S9* cluster. This translocation seemed to take place only in the rhodophyte-chromophyte lineage. This translocation must have taken place very early during the evolution of the rhodophyte-chromophyte lineage.

We propose a model for the evolution of the ribosomal gene cluster (Fig. 4), which is essentially based on the phylogenetic tree inferred from ribosomal proteins (Fig. 3). This model was constructed so as to explain the organization of ribosomal genes in various prokaryotes, plastids, and mitochondria. The ribosomal gene cluster of a hypothetical ancestor is supposed to be composed of *str, L2, spc,* α, and *S9* clusters as well as *rps10, infA, rpl36,* and *rpl31* genes. In the bacterial lineage, *rpl31* has been separated from the cluster. In *E. coli* and *H. influenzae,* the *str* cluster has also been separated from the cluster. As a result, the *rps10* gene precedes the *L2-spc-*α*-S9* clusters in these bacteria. In *Synechocystis,* the separation has taken place between the *rps10* gene and the *L2* cluster. This type of genome is supposed to be the progenitor of plastids of all the algae and plants. After the endosymbiosis of the cyanobacterial progenitor into a eukaryotic cell, two branches emerged: the rhodophyte-chromophyte lineage and the chlorophyte-glaucophyte lineage. In the former branch, the *str* cluster has been

translocated at the end of the ribosomal gene cluster to give a continuous gene cluster. In the chlorophyte-glaucophyte branch, various ribosomal genes have been removed by transfer to the nuclear genome. After the glaucophyte branch had been separated from the chlorophyte branch, a rearrangement took place between the *spc* cluster and the *rpl36* gene. In the chlorophyte (this word is used in a broad sense, namely, a group of green algae and land plants) lineage, various other genes have been removed from the plastid genome and transferred to the cell nucleus. Although we do not have enough data on the mitochondria, the mitochondrial ribosomal gene cluster can also be considered to be derived from the hypothetical ancestor as illustrated on the lower right end of the figure.

These observations confirm that the cyanobacterium and the plastids are closely related. It is interesting that *Synechocystis*, which is a free-living prokaryote, is similar to the plastids of algae and plants rather than to the bacteria in the amino acid sequence of the ribosomal proteins, because free-living prokaryotes have 21 kinds of the small subunit of ribosomal proteins while in plastid 20 kinds of the 21 polypeptides in bacterial ribosomal small subunit have been detected, although some additional ribosomal proteins were reported in the plastids of several kinds of higher plants (Wada et al. 1993). This finding supports the monophyletic origin of plastids, namely, that all the plastids are descendants of a single cyanobacterial progenitor. Therefore, we propose that the translocation of the *str* cluster took place early in the rhodophyte-chromophyte lineage after the endosymbiosis. Alternatively, there were multiple events of endosymbiosis that gave rise to different lineages of plastids, namely rhodophyte-chromophyte lineage, glaucophyte lineage, and chlorophyte lineage, and the translocation of the *str* cluster occurred in the progenitor of the rhodophyte-chromophyte lineage. Nevertheless, the similarity of amino acid sequences of plastids and *Synechocystis* suggests that we have no sufficient reason to support the multiple endosymbiosis.

The present study clearly shows the early occurrence of the translocation of the *str* cluster in the rhodophyte-chromophyte lineage. This clear-cut separation of this lineage from other algae and plants is not evident in the analysis of comparison of amino acid sequences of the ribosomal proteins. We can conclude that the *str* translocation took place very early after the separation of the rhodophyte-chromophyte lineage. In this respect, it is interesting to note that, in the tree inferred from *rbcL*, the rhodophyte-chromophyte lineage has been clearly separated from the lineage of chlorophytes and glaucophytes as a result of the horizontal transfer of *rbcL* gene from β-proteobacteria to rhodophytes and chromophytes (Delwiche and Palmer 1996). The *str* translocation may have taken place at the same time when the horizontal transfer of *rbcL* occurred.

Acknowledgments. This work was supported in part by a Grant-in-Aid for Scientific Research to N. O. (No. 085404) and to N. S. (No. 08454254) from the Ministry of Education, Science, Sports and Culture of Japan, and a grant to N. O. from Funakoshi & Co. Ltd. Tokyo Japan.

References

Allen MB (1959) Studies with *Cyanidium caldarium*, an anomalously pigment chlorophyta. Arch Mikrobiol 32:270–277

Bedwell DM, Davis G, Gosink K (1985) Nucleotide sequence of the alpha ribosomal protein operon of *Escherichia coli*. Nucleic Acids Res 13:3891–3903

Bergsland KJ, Haselkorn (1991) Evolutionary relationships among eubacteria, cyanobacteria, and chloroplasts: evidence from the *rpoC1* gene of *Anabaena* sp. strain PCC7120. J Bacteriol 173:3446–3455

Burger G, Plante I, Lonergan KM, Gray MW (1995) The mitochondrial DNA of the Amoeboid Protozoon, *Acanthamoeba castellanii:* complete sequence, gene content and genome organization. J Mol Biol 245:522–537

Cerretti DP, Dean D, Davis GR, Bedwell DM, Nomura M (1983) The *spc* ribosomal protein operon of *Escherichia coli:* sequence and co-transcription of the ribosomal protein genes and a protein export gene. Nucleic Acids Res 11:2599–2616

Delwiche CF, Palmer JD (1996) Rampant horizontal transfer and duplication of RuBisCO genes in eubacteria and plastids. Mol Biol Evol 13:873–882

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17:368–376

Fujiwara S, Sawada M, Someya J, Minaka N, Kawachi M, Inoue I (1994) Molecular phylogenetic analysis of *rbcL* in the prymnesiphyta. J Phycol 30:863–871

Helmchen TA, Bhattacharya D, Melkonian M (1995) Analyses of ribosomal RNA sequence from glaucocystophyte cyanelles provide new insights into the evolutionary relationships of plastids. J Mol Evol 41:203–210

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275–282

Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, Nakamura Y, Miyajima N, Hirosawa M, Sugiura M, Sasamoto S, Kimura T, Hosouchi T, Matsuno A, Muraki A, Nakazaki N, Naruo K, Okumura S, Shimpo S, Takeuchi C, Wada T, Watanabe A, Yamada M, Yasuda M, Tabata S (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res 3:109–136

Karlin S, Altschul SF (1990) Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. Proc Natl Acad Sci USA 87:2264–68

Kowallik KV, Stobe B, Schaffran I, Kroth-Pancic P, Freier U (1995) The chloroplast genome of a chlorophyll a + c-containing alga, *Odontella sinensis.* Plant Mol Biol Rep 13:336–342

Kuroiwa T, Kuroiwa H, Mita T, Ohta N (1994) *Cyanidium caldarium* as a model cell for studying division of chloroplasts. In: Seckbach J (ed) Evolutionary pathways and enigmatic algae: *Cyanidium caldarium* (Rhodophyta) and related cells. Kluwer Academic, England, pp 239–253

Margulis L (1970) Origin of eukaryotic cells. Yale University Press, New Haven, USA

Morden CW, Delwiche CF, Kuhsel M, Palmer JD (1992) Gene phylogenies and the endosymbiotic origin of plastids. Biosystems 28: 75–90

Nelissen B, Van de Peer Y, De Wachter R (1995) An early origin of plastids within the cyanobacterial divergence is suggested by evolutionary trees based on complete 16S rRNA sequences. Mol Biol Evol 12:1166–1173

Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N,

Akasi K, Kanegae T, Ogura Y, Kohchi T, Ohyama K (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. J Mol Biol 223:1–7

Ohta N, Suzuki K, Kawano S, Kuroiwa T (1993) Direct evidence of mitochondrial nuclear division in the ultra-micro alga *Cyanidioschyzon merolae.* Cytologia 58:471–476

Ohta N, Sato N, Ueda K, Kuroiwa T (1997) Analysis of a plastid gene cluster reveals a close relationship between *Cyanidioschyzon* and *Cyanidium.* J Plant Res 110:235–245

Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota S, Inokuchi H, Ozeki H (1986) Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322:572–574

Post LE, Nomura M (1980) DNA sequence from the *str* operon of *Escherichia coli.* J Biol Chem 255:4660–4666

Reith M, Munholland J (1995) Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. Plant Mol Biol Rep 13: 333–335

Seckbach J (1992) The Cyanidiophyceae and the anomalous symbiosis of *Cyanidium caldarium.* In: Reiser W (ed) Algae and symbioses: plants, animals, fungi, viruses, interactions explored. Biopress, Bristol, England, pp 400–426

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Shinozaki KY, 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 the tobacco chloroplast genome: its gene organization and expression. EMBO J 5:2043–2049

Stirewalt VL, Michalowski CB, Loffelhardt W, Bohnert HJ, Bryant DA (1995) Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa.* Plant Mol Biol Rep 13:327–332

Strimmer K, von Haeseler A (1996) Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. Mol Biol Evol 13:964–969

Sugiura M (1992) The chloroplast genome. Plant Mol Biol 19:149–168

Suzuki K, Ohta N, Kuroiwa T (1992) Isolation of the cell-nuclear, mitochondrial, and chloroplast DNA from the ultra-small eukaryote *Cyanidioschyzon merolae.* Protoplasma 171:80–84

Taylor FJR (1974) Implications and extensions of the serial endosymbiosis theory of the origin of eukaryotes. Taxon 23:299–258

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Vogel H, Fischer S, Valentine K (1996) A model for the evolution of the plastid sec apparatus inferred from *secY* gene phylogeny. Plant Mol Biol 32:685–692

Wada A, Koyama K, Maki Y, Shimoi Y, Tanaka A, Tsuji H (1993) A 5 kDa protein (SCS23) from the 30 *s* subunit of the spinach chloroplast ribosome. FEBS Lett 319:115–118

Zurawski G, Zurawski SM (1985) Structure of the *Escherichia coli* S10 ribosomal protein operon. Nucleic Acids Res 13:4521–4526