


## Hypothesis

# Are Cyanobacteria an Ancestor of Chloroplasts or Just One of the Gene Donors for Plants and Algae?

Naoki Sato 

Graduate School of Arts and Sciences, University of Tokyo, Meguro-ku, Tokyo 153-8902, Japan;  
naokisat@bio.c.u-tokyo.ac.jp

**Abstract:** Chloroplasts of plants and algae are currently believed to originate from a cyanobacterial endosymbiont, mainly based on the shared proteins involved in the oxygenic photosynthesis and gene expression system. The phylogenetic relationship between the chloroplast and cyanobacterial genomes was important evidence for the notion that chloroplasts originated from cyanobacterial endosymbiosis. However, studies in the post-genomic era revealed that various substances (glycolipids, peptidoglycan, etc.) shared by cyanobacteria and chloroplasts are synthesized by different pathways or phylogenetically unrelated enzymes. Membranes and genomes are essential components of a cell (or an organelle), but the origins of these turned out to be different. Besides, phylogenetic trees of chloroplast-encoded genes suggest an alternative possibility that chloroplast genes could be acquired from at least three different lineages of cyanobacteria. We have to seriously examine that the chloroplast genome might be chimeric due to various independent gene flows from cyanobacteria. Chloroplast formation could be more complex than a single event of cyanobacterial endosymbiosis. I present the “host-directed chloroplast formation” hypothesis, in which the eukaryotic host cell that had acquired glycolipid synthesis genes as an adaptation to phosphate limitation facilitated chloroplast formation by providing glycolipid-based membranes (pre-adaptation). The origins of the membranes and the genome could be different, and the origin of the genome could be complex.

**Keywords:** chloroplast origin; cyanobacterial endosymbiosis; glycolipids; host-directed chloroplast formation; peptidoglycan; *Paulinella chromatophore*; phylogenetic analysis



**Citation:** Sato, N. Are Cyanobacteria an Ancestor of Chloroplasts or Just One of the Gene Donors for Plants and Algae? *Genes* **2021**, *12*, 823.  
<https://doi.org/10.3390/genes12060823>

Academic Editors: Denis Baurain and Luc Cornet

Received: 26 March 2021  
Accepted: 25 May 2021  
Published: 27 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Cyanobacteria are a distinct group of bacteria that perform oxygenic photosynthesis. They have specialized internal membranes called thylakoid membranes in which photosynthetic systems reside. Galactolipids and sulfolipids that are major constituents of the thylakoid membranes have been believed to be necessary for the photosynthetic activity. Cyanobacteria and chloroplasts of plants and algae share thylakoid membranes with galactolipids and sulfolipids, various components of photosystems, and photosynthetic pigments. These traits of substance, structure, and function as a whole were taken as good evidence for the cyanobacterial origin of chloroplasts. The endosymbiotic hypothesis of chloroplast origin was first proposed by C. Mereschkowsky in 1905 [1] who obtained a hint from a short note of A. Schimper in 1883 [2]. Since then, it remained a well-known but unsupported hypothesis. Various authors in the 1960s revitalized the hypothesis based on the discovery of chloroplast DNA, and the similarities of chloroplasts and cyanobacteria became evident thanks to the developments in cell biology and biochemistry of photosynthetic organisms in the 1960s and 1970s. It was the phylogenetic analysis in the 1980s and later that provided convincing evidence for the cyanobacterial origin of chloroplasts. Historical views on the endosymbiotic origin of chloroplasts are explained in some review articles [3–10].

All textbooks of biology explain the origin of chloroplasts by cyanobacterial endosymbiosis that occurred 1 or 2 billion years ago. The origin of mitochondria is also explained

by the endosymbiosis of a proteobacterium. These origin stories are currently considered as well-established facts that form the basis of the formation of eukaryotic cells. Many phenomena of eukaryotic cells are currently understood in terms of the endosymbiotic origin of organelles. Stiller [11] criticized this situation: “While many research articles and reviews refer to a single primary origin of all plastids as an established fact, it remains a working hypothesis that gained popularity under a set of assumptions that turn out to have little empirical support in nature (p. 467)”.

There are still two major questions about the endosymbiotic theory or hypothesis. First, what is exactly the decisive evidence that the chloroplasts and mitochondria originated from cyanobacterial and proteobacterial endosymbionts, respectively, and how and when this theory was established? It is well known that, in the late 1960s and 1970s, Lynn Margulis claimed her original version of the origin of eukaryotic cells, which involved the endosymbiotic origin of chloroplasts, mitochondria, and mitotic apparatus or flagella [12,13]. Nevertheless, it was the molecular phylogeneticists that presented convincing evidence for the close relationship between chloroplasts and cyanobacteria, and between mitochondria and  $\alpha$ -proteobacteria. The accumulation of supporting data continued until the early 2000s when the genomic sequences of major model organisms became available. Under these situations, not only the organellar genes but also nuclear-encoded organelle proteins were used for the phylogenetic analysis that finally confirmed the close relationships, chloroplasts/cyanobacteria, and mitochondria/ $\alpha$ -proteobacteria. Nevertheless, this is just a story of organellar DNA. DNA origin is not enough to affirm the two endosymbiotic events as illustrated in many books and review articles. Membranes and genomes are essential components of a cell or an organelle such as chloroplasts and mitochondria. Can we deduce an endosymbiotic image of the origin of an organelle directly from its DNA origin?

Another question, related to this point, arises from more recent studies on the extensive phylogenetic studies on many enzymes involved in the characteristic, common traits of chloroplasts and cyanobacteria. Most of the enzymes involved in the synthesis of thylakoid membrane lipids in plants and algae do not originate from cyanobacteria. How can we explain this? In mitochondria, systematic phylogenetic analysis of the full complement of mitochondrial proteins also raised a comparable question: namely, how can we explain that 90% of mitochondrial enzymes [14,15] existed in the eukaryotic cell long before the formation of mitochondria? Keeling [16] pointed out the “cognitive inertia” of endosymbiotic thinking, and suggested removing the historically biased view: “It is interesting, therefore, to consider what kinds of conclusions we might draw if we learned what we know today in a different order, or learned it all at once (p. 10)”. It is worth trying to re-examine all available evidence for the endosymbiotic theory and to re-organize the history of the formation of chloroplasts and mitochondria. Therefore, this is a discussion or hypothesis article with a limited length on the historical and biological questions on the endosymbiotic origins of organelles, but I focus mainly on the relationship between cyanobacteria and chloroplasts.

Because many readers in biology might not be familiar with this kind of half-philosophy, half-biology article, I first present the overall plan of the article. There are three major, mutually related topics: (1) History section (Section 2) deals with the ambiguous meaning of “endosymbiosis”. Endosymbiosis notion includes physical integration of cyanobacterial cell into a eukaryotic cell, but this cannot be proven by cytological experiments. Only DNA origin can be identified by phylogenetic analysis. Then, I discuss the visually identifiable components (membranes) and genome in the following two sections. (2) The lipid and peptidoglycan section (Section 3) deals with the problem that the structural elements of chloroplasts did not originate from cyanobacteria. (3) Chloroplast gene section (Section 4) deals with putative, multiple origins of chloroplast genome within cyanobacteria. Indeed, this is mostly a re-interpretation of previous results. Section 5 also discusses the origins of some other chloroplast enzymes. After these three parts, Section 6 is devoted to discussing the *Paulinella* model of chloroplast formation, and I present a role of pre-adaptation that

determined the fate of chloroplasts. This is not a simple rejection of the endosymbiotic hypothesis, but a broadening of the hypothesis. I also have to mention that the topic of the present article is limited to the formation of “primary” chloroplast. The secondary endosymbiosis must be a different story.

## 2. Historical and Theoretical Problems on the Origin of Chloroplasts

### 2.1. Historical Demarcation

The relationship between cyanobacteria and chloroplasts has been repeatedly argued since Mereschkowsky [1] (see also an English translation [3]). I already presented a detailed historical analysis of the endosymbiotic theories [8,9]. Readers are referred to these as well as [4–6,17] for the history of endosymbiotic theories. In the early 20th century, the similarity of cyanobacteria and chloroplasts was restricted to the similarity in size, the presence of photosynthetic pigments, and the activity of carbon fixation with oxygen evolution (“photosynthetic oxygen production” has been conventionally called “oxygen evolution”), all of which reflected the status of knowledge at the time. The history of endosymbiotic arguments may be divided into three major phases: In the first phase (1905–1961), the similarity of cyanobacteria and chloroplasts in terms of photosynthetic properties as defined above was recognized by many researchers in plant and algal studies, but the majority of top researchers, such as Pascher [18] and Buchner [19], clearly rejected endosymbiotic origin of any eukaryotic organelle (chloroplasts or mitochondria). However, Pascher [18] distinguished cyanelles from chloroplasts and described that the cyanelles of *Cyanophora paradoxa* and *Paulinella chromatophora* were authentic cyanobacterial cells living in a eukaryotic cell. Cyanelles were evidence of the cyanobacterial origin of chloroplasts for some researchers but were distinct from chloroplasts for many others. Lederberg [20] suggested later that the genetic system supposed to reside in chloroplasts was a kind of plasmid, an exogenous genetic element. Symbiogenesis was also supported by some speculative biologists or philosophers [21,22].

In the second phase (1962–1977), the presence of organellar DNA in chloroplasts and mitochondria was taken as evidence for the bacterial origin of these organelles. In parallel, cytological and biochemical characterization of chloroplasts and mitochondria enabled precise comparison of these organelles with cyanobacteria and other aerobic bacteria. The prokaryotic nature of the organellar ribosomes (70S sedimentation-based size, partial rRNA sequence, sensitivity to antibiotics) was an important, novel fact that supported the similarity of organelles and bacteria. Moreover, physical characterization of photosynthetic mechanisms, biochemical identification of photosystem components, and chemical analysis of lipid components of thylakoid membranes, all these data were additional evidence for the similarity of chloroplasts and cyanobacteria. In 1966, Echlin presented a comprehensive table comparing cyanobacteria and chloroplasts [23]. Hybridization of chloroplast rRNA and cyanobacterial DNA was also innovative molecular evidence that favored the similarity of cyanobacteria and chloroplasts. Nevertheless, all these similarities could be used in both ways, namely, for and against the endosymbiotic origin of chloroplasts. In 1967, Goksøyr published a short paper in *Nature*, illustrating the sequential endosymbiotic origins of mitochondria and chloroplasts [24] (note that this preceded the 1967 paper of Lynn Margulis (Lynn Sagan then) [12], and the Serial Endosymbiosis Theory (SET) of Taylor published in 1974 [25]). The autogenous origin of chloroplasts could also explain the data [26–28], although it was difficult to explain the origin of both chloroplasts and mitochondria within a single scheme. Prudent opinions remained, while changes began to emerge [29–32].

The third phase was marked by the development of phylogenetic analysis, which was the breakthrough that changed the status of the endosymbiotic theory. An epoch-making paper of Schwartz and Dayhoff appeared in 1978, reporting a reliable phylogenetic relationship between cyanobacteria and chloroplasts [33]. Phylogenetic analysis was indeed the only rational way of identifying the direct lineage from cyanobacteria to chloroplasts. The methodology of phylogenetic inference became more and more accurate and reliable.

Innovations in both computer hardware and computational technology, and dramatic speedup in DNA sequencing and assembly helped the progress of phylogenetic analysis. From the 1980s onwards, chloroplasts became regarded as descendants of cyanobacteria. Nevertheless, it is difficult to distinguish when the conceptual change occurred. The change seemed to proceed gradually and slowly during the 1990s and 2000s. Phylogenetic data accumulated, and all data tended to be interpreted in the light of the endosymbiotic theory, namely, the cyanobacterial origin of chloroplasts. Then, no other way of thinking became mentioned in the literature.

## 2.2. Genocentrism and the Changes in Thinking at the Beginning of the 21st Century

I have to point out curious changes in the thinking of researchers from the 1980s to the early 2000s regarding the endosymbiotic origin of organelles, especially the cyanobacterial origin of chloroplasts. The review article of Gray and Doolittle [34] affirmed the cyanobacterial origin of chloroplasts, based on various phylogenetic trees. Among them, the most important was the phylogenetic tree of rRNA, because the ribosome is the core of the gene expression system in any organism. Wallace [35] also discussed the origin of organelle genomes and concluded:

“All of these data are consistent with the hypothesis that the chlDNA genes diverged from the blue-green algae well after the divergence of the eubacterial and eucaryotic nuclear lines. Therefore, chlDNAs must have originated by a symbiotic event between members of the eubacteria and eucaryotic nuclear lineages. Similarly, sequence data have indicated that the mtDNA rRNA genes are more homologous to bacterial rRNA genes than to rRNA genes of eukaryotic nuclei. They too probably also had an endosymbiotic origin.” ([35] p. 227)

The origin of organelle DNA was considered convincing evidence for the endosymbiotic origin of organelles because no other way of acquiring exogenous DNA was known in the 1980s. The notion of horizontal gene transfer appeared later, but the phylogenetic data supporting endosymbiosis were no longer examined in the light of horizontal gene transfer. Delwiche et al. [36] reported the results of robust phylogenetic analyses of the *tufA* gene, and concluded: “The finding from *tufA* that all plastids are derived from cyanobacteria is compatible with analyses of 16S rRNA and *atpB* sequences” (p. 124). Moreira and Philippe [37] wrote in 2001: “a wealth of data strongly suggests that all known photosynthetic plastids are monophyletic and that they have a cyanobacterial origin” (p. 771) by citing Delwiche [38], who, nevertheless, noted cautiously in the legend for the illustration of endosymbiotic origin of plastids: “One hypothesis for endosymbiotic events in the evolution of plastids. . . . This figure assumes a single origin of plastids, but this remains uncertain, as does the precise relationship among the primary plastid lineages” (p. S168). Looking back at the literature, the beginning of the 21st century was a turning point, in which the cyanobacterial origin of chloroplasts became accepted as almost established, although the debate on the monophyly of primary plastids continued [39,40]. Certainly, the genome sequencing of the cyanobacterium *Synechocystis* sp. PCC 6803 [41] and other cyanobacteria opened a way of directly comparing the genomes of cyanobacteria and chloroplasts, and the accumulation of sequence data of both cyanobacteria and chloroplasts facilitated the comparison of the two entities. McFadden [42] listed various lines of convincing evidence, and wrote also in 2001: “the endosymbiotic origin of plastids from cyanobacterial ancestors thus seems confirmed.” (p. 953)

The similarity of RNA polymerase in cyanobacteria and chloroplasts was also known in the 1990s. As described in biochemistry textbooks, bacterial RNA polymerases consist of  $\alpha$ ,  $\beta$ ,  $\beta'$ , and  $\sigma$  subunits. However, in the cyanobacterial counterpart, the  $\beta'$  subunit is split into two parts, called  $\beta'$  and  $\beta''$ . This split is found at the level of the gene, as well as protein. This strange situation is also found in RNA polymerases encoded by the chloroplast genome (see the references cited in [36] for the genes that served as evidence for the cyanobacterial origin of plastids). Essential components of gene expression in the chloroplast were found almost cyanobacterial. The chloroplast genome was also

found to be similar to the cyanobacterial genomes, except for the large difference in sizes. Almost all enzymes and RNA that are encoded by the chloroplast genome had clearly identifiable cyanobacterial counterparts. The similarity of photosynthetic proteins was evident: the components encoded by the chloroplast genome had orthologs in cyanobacteria. However, in many cases, the photosynthesis-related components were only found in cyanobacteria and chloroplasts, and it was not possible to identify the root; namely, it was not possible to find where the evolution started. Another problem was unidentified DNA polymerase (see Section 5.1) in chloroplasts. Nevertheless, the majority of opinions favored the cyanobacterial origin of chloroplasts.

In retrospect, this was certainly motivated by the importance of genome and gene expression systems in biology. The cyanobacterial nature of the genetic apparatus of chloroplasts including DNA, RNA polymerase, and ribosomal components was taken as evidence that the chloroplast was a descendant of cyanobacteria. I call this “genocentrism” (gene-centered view), which was certainly the decisive impetus that changed the view on the origin of chloroplasts. I have to point out, however, that the genocentrism has been hidden behind the illustrations clearly showing the membranes to explain the endosymbiotic origin of organelles in many review articles [5,6,43–45] and textbooks (see for example [46]). In contrast, only a vague relationship was known between the cyanobacterial membranes and the chloroplast membranes. The presence of galactolipids and sulfolipids, as well as various photosynthetic components, was known in both cyanobacteria and chloroplasts, which was the evidence for the similarity of thylakoid membranes. The origin of the chloroplast envelope remained ambiguous: it could originate from the cyanobacterial plasma membrane, modified endoplasmic reticulum, phagosome, or peri-algal vacuole. Any origin was possible.

### 2.3. Spread of Endosymbiotic Gene Transfers (EGT)

Many components of the chloroplast ribosome were found to be encoded by the nuclear genome. Phylogenetic analysis suggested that these nuclear genes originated from chloroplast genes. This was explained by the endosymbiotic gene transfer (EGT), namely, chloroplast genes were somehow transferred to the nucleus [47]. This was supported by artificial EGT experiments [48]. A fairly efficient gene flow from the chloroplast genome (and the mitochondrial genome) to the nucleus was supposed to work during the long history of plant evolution (for a review of the time, see [49]).

However, as the cases of EGT became abundant, many other nuclear genes encoding chloroplast proteins have been simply considered to result from EGT, such as the peptidoglycan synthesis enzymes that were identified in the nuclear genome of the moss *Physcomitrella patens*. Various other chloroplast enzymes were also identified in plant nuclear genomes. In many cases, phylogenetic analysis was not performed. In other cases, only cyanobacterial sequences and plant sequences were used in phylogenetic analysis. Bacterial sequences were rarely used to identify the origin of the nuclear genes.

### 2.4. Currently Accepted Evidence for the Endosymbiotic, Cyanobacterial Origin of Chloroplasts

Currently, several different lines of evidence for the cyanobacterial origin of chloroplasts exist (Table 1. See also [9,50]). Photosynthesis with oxygen evolution (Item 1) is the best trait that suggests a direct relationship between cyanobacteria and chloroplasts, which also have structural and biochemical similarities. Structurally, both have internal, flattened membranes called “thylakoid” (meaning a sac) membranes, which are the site of light-driven reactions of photosynthesis. The chloroplast is surrounded by a set of limiting membranes called envelope membranes, whereas the cyanobacterial cell is enclosed by inner and outer membranes. Note that *Gloeobacter* lacks thylakoid membranes, and photosynthetic reaction centers reside in the inner cell membrane. Additional and related shared traits in cyanobacteria and chloroplasts include the presence of photosynthetic pigments such as chlorophyll *a* and the presence of galactolipids in the membranes. The presence of DNA and its associated genetic system (Item 2) is evidence for the autonomy



of chloroplasts, and therefore, evidence for the exogenous origin of chloroplasts, implying endosymbiotic origin. Reproduction by binary fission (Item 3) is another similarity between chloroplasts and cyanobacteria. Another classical evidence is the presence of peptidoglycan (Item 4) in some chloroplasts, which will be explained in detail later (Section 3.1). The presence of cyanelles, which were believed to be living cyanobacterial cells within a photosynthetic eukaryote (see Section 2.1), is no longer evidence for the visual cyanobacterial endosymbiosis that can explain chloroplasts origin. The cyanelle of *Cyanophora* is now recognized as a chloroplast, rather than a cyanobacterium. The cyanelle of *Paulinella* is not an independent organism, which will be discussed later.

**Table 1.** Evidence for the cyanobacterial origin of chloroplasts.

Item	Evidence	Explanation
1	Similarity of photosynthesis	(a) Photosynthesis with oxygen evolution (b) Presence of chlorophyll <i>a</i> and carotenoids (c) Similarity in the structure consisting of two limiting membranes and internal thylakoid membranes containing galactolipids
2	Similarity of the genetic system	Prokaryotic RNA polymerase and ribosome
3	Similarity in reproduction	Binary fission
4	Presence of peptidoglycan	Glaucophytes, some green algae, charophytes, bryophytes, pteridophytes, some gymnosperms
5	Phylogenetic relationship	(a) rRNA genes (b) Ribosomal proteins and house-keeping enzymes (c) Photosynthetic components (orthologs found only in cyanobacteria and chloroplasts)
6	Synteny of gene clusters	(a) Ribosomal proteins (b) ATPase subunits (c) RNA polymerase subunits

Next, phylogenetic analysis of chloroplast-encoded genes (Item 5; see Section 4) showed their cyanobacterial origin. Nearly all protein-coding genes and RNA genes encoded by the chloroplast genomes of algae and plants originated from cyanobacteria. Rare exceptions will be explained in Section 4.

Another line of evidence is the synteny of related genes in the chloroplast genomes and the cyanobacterial genomes (Item 6) [51,52]. The order of gene arrangement in the ribosomal protein cluster is highly conserved in bacteria, and the same order is found in the chloroplast genome if missing genes are allowed. The missing ribosomal proteins are supposed to be in the nuclear genome as a result of EGT. This idea was often verified by subsequent genomic analyses. Synteny is also found in the genes coding for ATPase subunits and RNA polymerase subunits, although there are many supposedly missing genes. Synteny is considered superior to phylogenetic trees because the latter could be affected by calculation methods and data sets, but there is no doubt about the order of genes.

### 2.5. Archaeplastida and the Cyanobacterial Origin of Chloroplasts

Many researchers argue that the monophyletic origin of the three lineages, namely, green algae/plants, red algae, and glaucophytes, is an important piece of evidence for the cyanobacterial origin of chloroplasts. Phylogenetic analysis of chloroplast genes invariably shows (with some exceptions as already mentioned) that the chloroplast genes diverge from the cyanobacterial clade as a single clade [53,54]. This is taken as evidence for the cyanobacterial origin of chloroplasts. Besides, extensive phylogenetic analysis of the 143 nuclear genes supports the monophyly of the three lineages [55]. Additionally, we showed that the chimeric origins of carbon fixation cycle (also called Calvin-Benson cycle) enzymes

are common in the red algae and the green plants [56], which was subsequently confirmed in glaucophytes [57]. Phylogenetic analysis showed that four enzymes in the Calvin-Benson cycle originated from cyanobacteria, but others are of eukaryotic origin, except rubisco, whose origin is different in the green and red lineage. These chimeric origins are conserved in the three lineages. This chimerism must have been established early in the chloroplast formation before the diversification of the three lineages. In other words, this is another evidence for the monophyly of the three lineages.

In contrast, Stiller [58] argued that the similarity of green and red algae (both chloroplasts and nuclear genomes) was a result of convergent evolution with a large influx of plastid DNA into nuclear genomes, and rejected a rhodophyte-green plant sister relationship. Using slowly evolving nuclear genes, Nozaki et al. [59] suggested that the three lineages are not really monophyletic, because some non-photosynthetic clades are also included in “Plantae” that they define. These two anti-monophyletic views were not identical, but presented serious concerns about “primary endosymbiosis”. Nevertheless, Nozaki et al. [59] still try to interpret their results “under the assumption of the single plastid primary endosymbiosis”. The recent identification of non-photosynthetic predators sister to red algae named *Rhodolphis* could also break the monophyly of Archaeplastida [60]. However, the majority of researchers seem to affirm that not only the chloroplasts but also the entire cells of the three lineages are monophyletic.

The three lineages, green algae/plants, red algae, and glaucophytes, are classified in a single phylum called Archaeplastida [61]. Curiously, the definition of Archaeplastida involved the notion of cyanobacterial endosymbiosis: “Photosynthetic plastid with chlorophyll *a* from an ancestral primary endosymbiosis with a cyanobacterium; plastid secondarily lost or reduced in some; usually with cellulose cell wall; flat cristae; starch storage product.” ([61], p. 420. My emphasis). In the most recent revision [62], this was slightly changed: “Photosynthetic plastid with chlorophyll type-*a* from an ancestral primary endosymbiosis with a cyanobacterium; plastid with two membranes without periplastid endoplasmic reticulum; plastid reduced in some; usually with cell wall or other extracellular covering; flat mitochondrial cristae; starch storage product (p. 35)”. That is all that the authors gave as the criteria of classification of Archaeplastida.

We should distinguish two problems here: monophyly of the three lineages of algae/plants on the one hand, and the cyanobacterial origin of their chloroplasts on the other. The common origin of the three lineages of Archaeplastida (with or without non-photosynthetic clades) does not imply that their chloroplasts are cyanobacterial descendants.

### 2.6. Recent Discussions on the Mitochondrial Origin

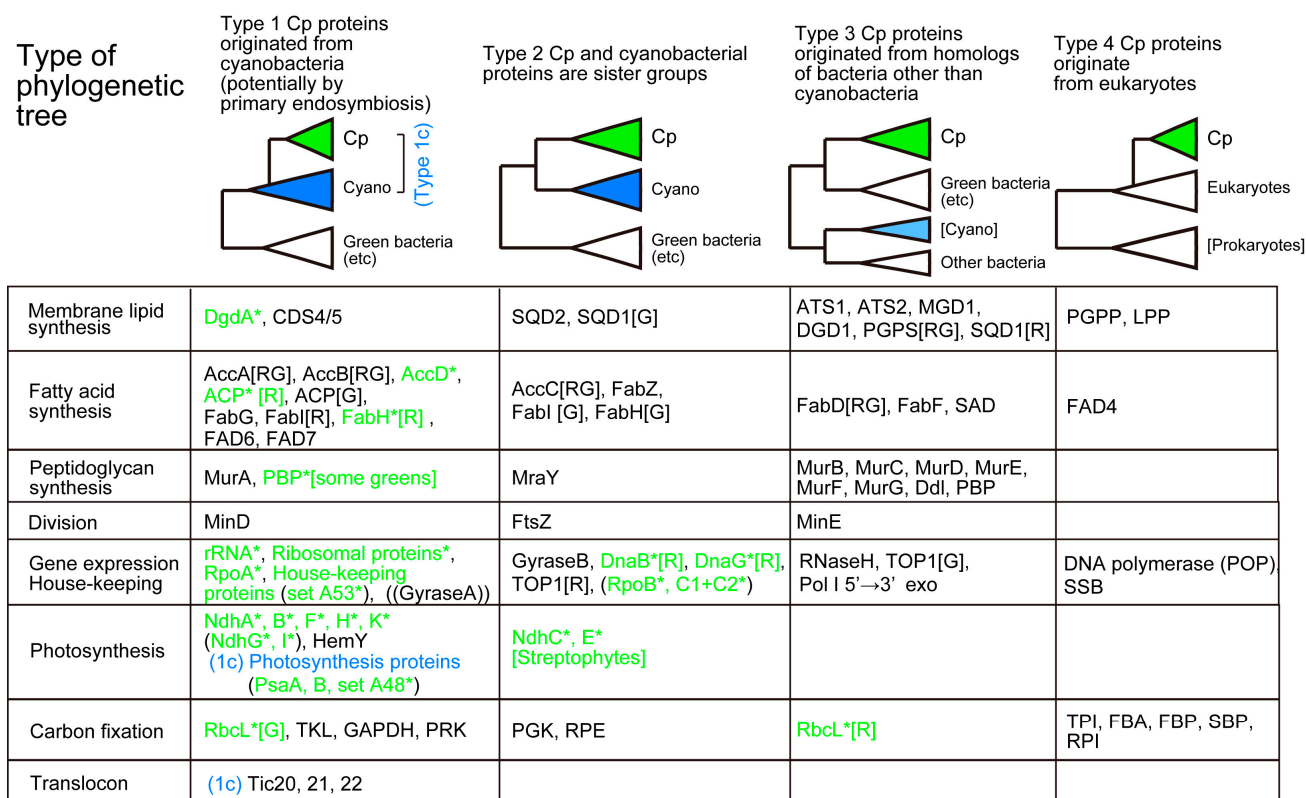
At this point, I briefly summarize the situation in mitochondria. The postgenomic era opened a new way of understanding. About 1000 proteins were identified in the mitochondrial proteome. Large-scale phylogenetic analysis suggested that about 90% of them did not originate from  $\alpha$ -proteobacteria, which is supposed to be the origin of the mitochondrial ribosome and mitochondrial DNA. Gray [14,15] proposed the “pre-endosymbiont hypothesis”. Pitis and Gabaldón [63] and Gabaldón [64] proposed the “pre-mitochondrial symbioses” hypothesis. Harish and Kurland [65] clearly stated: “mitochondria are not captive bacteria”. This could be a beginning of a new way of thinking of mitochondrial origin. We can imagine various scenarios to explain the situation: (1) Many genes for the future mitochondrial proteins had been acquired by the eukaryotic host before the formation of mitochondria by the acquisition of a proteobacterial endosymbiont. The old proteins changed their localization and function upon the formation of mitochondria. (2) A different form of mitochondrion was initially present in the eukaryotic cell, and then the mitochondrial genome was replaced by the one from a proteobacterium. At the same time, some more proteobacterial enzymes were introduced. Various other scenarios are also probable. Mitochondrial origin might not be a simple event of endosymbiosis, although many researchers still believe that an endosymbiosis of a proteobacterium was essential in

establishing the mitochondria. We now focus on the situation in chloroplasts, which also have similar problems to re-examine.

### 2.7. Phylogenetic Analysis of Chloroplast Enzymes

We consider here chloroplast enzymes that are encoded by both nuclear and chloroplast genomes and examine their phylogenetic relationships to the cyanobacterial homologs. Phylogenetic trees of the enzymes described in this section have been published as Supplementary Materials of a previous paper [50]. Figure 1 presents four types of phylogenetic relationships between the cyanobacterial clade and the chloroplast clade. Type 1 tree indicates that the chloroplast gene originated from cyanobacteria, which is consistent with the idea of chloroplast origin by cyanobacterial endosymbiosis. When homologs are only found in cyanobacteria and chloroplasts, we cannot determine the position of the root, and hence, the direction of evolution in a strict sense. In some special cases such as PsaA and PsaB, we can root the phylogenetic tree, but in most cases, the root position is undetermined. These cases are classified as Type 1c. In this case, *Gloeobacter* or basal *Synechococcus* are often taken as the root. In the Type 2 tree, the chloroplast clade is sister to the cyanobacterial clade. This means that the chloroplast originated from an ancestor of all extant cyanobacteria, which seems unbelievable. However, we often obtain this type of tree. This could be an artifact of phylogenetic reconstruction, but the tree form is rather robust in most cases, even if taxon samplings and phylogenetic methods are varied. We will have to find an explanation for this type of tree. Type 3 tree indicates that the chloroplast enzymes originated from bacteria other than cyanobacteria. Type 4 tree shows that eukaryotic enzymes are re-targeted to the chloroplast. Enzymes found only in cyanobacteria are classified as Type 5.





(5) Cyanobacterial proteins having no homologs in Cp: MgdA (except the one in *C. merolae* having no activity), MgdE

**Figure 1.** Four major types of phylogenetic relationships between cyanobacterial and chloroplast proteins. Chloroplast-encoded proteins (or RNA) are marked with an asterisk and colored. [R] and [G] indicate red and green lineages, respectively. This is an extended, revised version of the original in [9,50].

### 3. Origin of Structural Elements of Chloroplasts

#### 3.1. Peptidoglycan Synthesis Enzymes

Peptidoglycan is a meshwork structure that exists between the outer and inner membranes of Gram-negative bacteria including cyanobacteria. In Gram-positive bacteria, a thick layer of peptidoglycan covers the whole cell. The presence of a wall-like layer in chloroplasts or similar photosynthetic structures was first suggested in *P. chromatophora* and *C. paradoxa* in the early 20th century. The chromatophore of *Paulinella* is now considered a structure that is phylogenetically distinct from the chloroplasts of plants and algae, but in the 1920s, it was identified as a living cyanobacterial cell within the eukaryotic cell. The organelles surrounded by the wall-like layer in *Paulinella* and *Cyanophora* were subsequently called cyanelles, meaning cyanobacterial cells living in the cell as an organelle [18] and were distinguished from the common chloroplasts. The identity of peptidoglycan in *Cyanophora* cyanelles was established by chemical analysis [66]. Peptidoglycan has long been considered important evidence for the endosymbiotic origin of not only cyanelles but also chloroplasts (see a microbiology textbook [67]. see also Item 4 of Table 1).

The presence of peptidoglycan in green plants was suggested by the genome analysis of the moss *P. patens* [68]. The complete set of enzymes involved in the synthesis of peptidoglycan was identified in the nuclear genome of *P. patens*. The presence of peptidoglycan was suggested by treatment with ampicillin, an inhibitor of peptidoglycan synthesis: namely, chloroplast division was inhibited by ampicillin in the moss. Although the role of peptidoglycan in chloroplast division is still not understood, this was taken as evidence suggesting that peptidoglycan exists in the moss chloroplast. Genes for the complete set of peptidoglycan synthesis enzymes were also found in other organisms,

such as green algae (*Micromonas*), charophytes (*Klebsormidium*), liverworts (*Marchantia*), pteridophytes (*Selaginella*), and even in some gymnosperms [69]. *Arabidopsis thaliana* also possesses homologs of some of the peptidoglycan synthesis enzymes. Curiously, AtMurE was shown to be involved in chloroplast development, rather than biosynthesis of a glycan-like substance [70].

In contrast, all enzymes of peptidoglycan synthesis except MurF are encoded by the chromatophore genome of *P. chromatophora* and *Paulinella micropora*. MurF was found to be encoded by a nuclear gene, which originated from proteobacteria [9,50,71–74].

Peptidoglycan has not been observed by electron microscopy as a clearly defined layer in plant chloroplasts. Nevertheless, a well-designed experiment using the ‘Click chemistry’ was a breakthrough that identified peptidoglycan as a layer surrounding the moss chloroplast [75]. Subsequently, peptidoglycan materials were located between the outer and inner envelope membranes of moss chloroplasts by quantitative densitometry of transmission electron micrographs [76].

Then, a question is asked: “Does the presence of peptidoglycan supports the endosymbiotic origin of chloroplasts?” Recent phylogenetic analysis showed clearly that the peptidoglycan synthesis enzymes in plants and algae are monophyletic, but the majority of them originated from bacteria other than cyanobacteria [77,78]. Note that this conclusion was obtained not only by our study [77] but also by an independent study [78]. Rare examples of cyanobacteria-related enzymes include MurA and MraY (Figure 1). A class of penicillin-binding proteins (PBP) encoded by some green algal chloroplast genomes also diverged from cyanobacteria.

These results suggest that the peptidoglycan synthesis system was introduced into green algae and glaucophytes from various bacteria. However, important questions remain. First, when (or at which stage of evolution) was the peptidoglycan synthesis system introduced? The timing could be either in the common origin of Archaeplastida, or in the common ancestor of green algae and glaucophytes (if red algae diverged first), or separately in green algae and glaucophytes. In the first case, red algae must have lost the peptidoglycan synthesis system. Some *Cyanophora* enzymes are not closely related to green algal homologs and could have different origins. We also have to consider that many green algae such as *Chlamydomonas* do not have any peptidoglycan synthesis gene. We have another question: How were the genes introduced into algal cells? If each gene was introduced one by one from different bacteria, then the introduced genes could not function in peptidoglycan synthesis until all necessary genes are provided. There is no evolutionary advantage of keeping individual genes during the gene acquiring process. A possible scenario might be like this: a complete system of peptidoglycan synthesis was present at first in an organism, and each gene was replaced by horizontal gene transfer. After gene exchanges, a new complete, functional set of genes was introduced into an ancestor of algae (either of the three possibilities as above). Continued studies are necessary on this topic.

### 3.2. Lipid Biosynthesis Enzymes

Chloroplast is the major site of fatty acid synthesis in plant cells. A large proportion of fatty acids produced in the chloroplast are exported to the cytosol for the lipid synthesis in the endoplasmic reticulum (ER), while the remaining fatty acids are used for the chloroplast lipid synthesis. The proportion of the two metabolic flows depends on organisms and tissues. All the enzymes of lipid biosynthesis in the ER of plants and algae are eukaryotic enzymes that have closely related orthologs in other eukaryotes (animals and protists). To discuss the cyanobacterial origin of chloroplasts, we have to examine fatty acid synthesis and lipid synthesis within the chloroplasts.

#### 3.2.1. Fatty Acid Synthesis

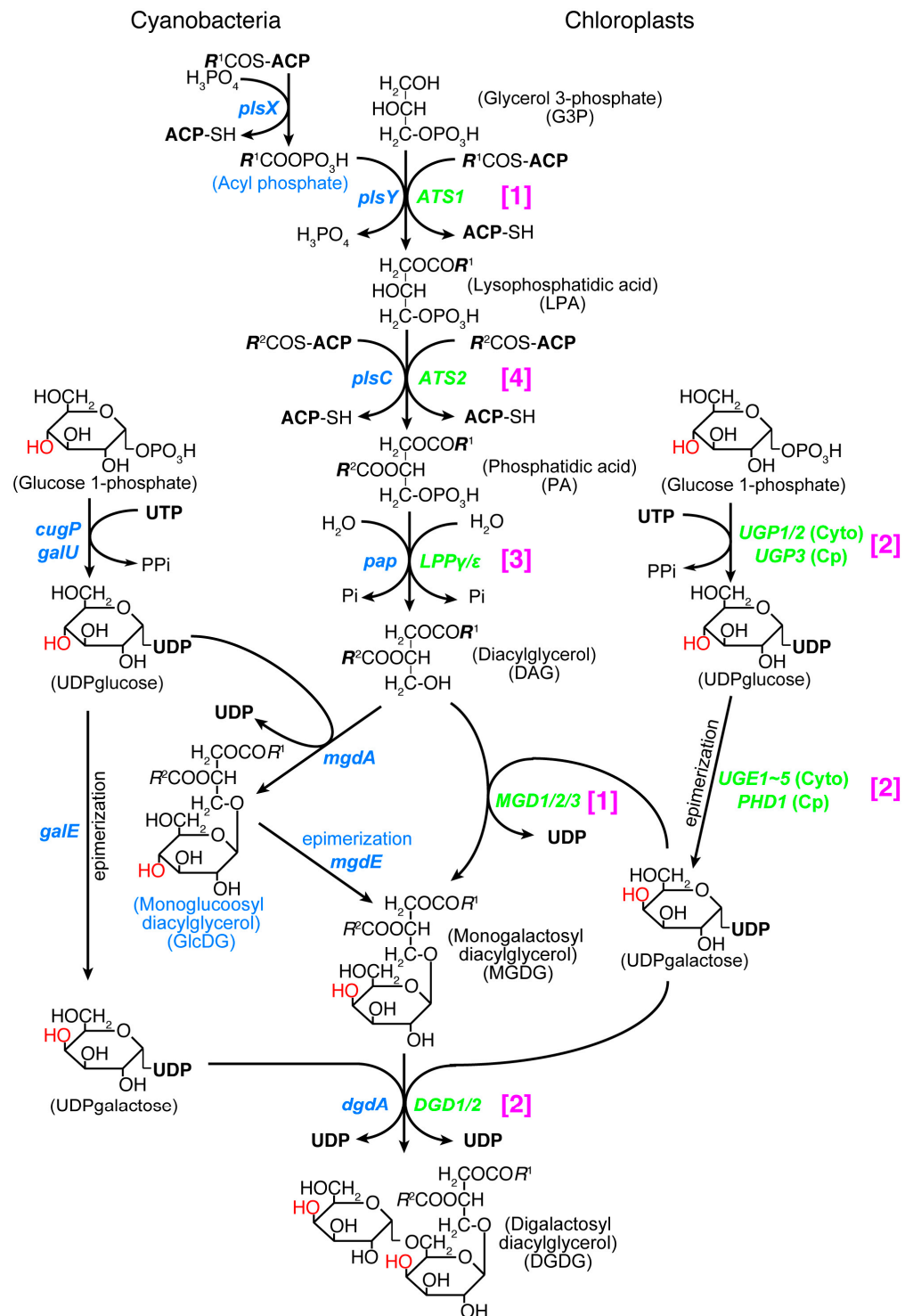
Two types of fatty acid synthases (FAS) are known: Type II FAS is found in bacteria including cyanobacteria and consists of four separate subunits catalyzing the main

cycle of chain elongation. Type I FAS is a large molecule containing the four enzymatic activities (plus acyl carrier protein (ACP) and other related activities) within one or two multifunctional polypeptide(s), which is found in the cytosol of fungi and animals. Fungal and animal FAS are also different in the arrangement of enzymatic active centers within the polypeptides. Because plant and algal chloroplasts contain Type II FAS, the fatty acid synthesis activity in the chloroplasts has been compared with the cyanobacterial activity. Type II FAS is also present in the mitochondria of eukaryotes including animals, plants, and other organisms, but only limited production of fatty acids is ascribed to mitochondria. Note that all enzymes involved in fatty acid synthesis (except some algal enzymes encoded by the chloroplast genome, such as red algal ACP and FabH) are encoded by the nuclear genome. Under these circumstances, fatty acid synthesis in the chloroplast was simply believed to originate from cyanobacteria. Recent phylogenetic analysis showed indeed cyanobacterial origin of many components of fatty acid synthesis, but not all (Figure 1). The condensing enzyme called FabF that catalyzes the chain elongation reaction in the chloroplast originates from green sulfur bacteria, whereas the mitochondrial homolog originates from proteobacteria [9,50]. Chloroplast FabD, which transfers malonyl CoA to ACP, is also closely related to green sulfur bacteria and green non-sulfur bacteria. The green bacteria (both sulfur and non-sulfur) are photosynthetic bacteria that do not produce oxygen but utilize light energy for carbon fixation. Curiously, in many phylogenetic analyses, green bacteria are often found to be the most related group of chloroplasts.

In the chloroplasts, the initial products of FAS are palmitic and stearic acids. In the chloroplasts of green plants and algae, stearic acid is desaturated (introduction of a double bond is called desaturation) to oleic acid by stearoyl ACP desaturase (SAD). Therefore, the primary products of fatty acid synthesis in these chloroplasts are palmitic and oleic acids. SAD is not present in red algae and glaucophytes, nor in cyanobacteria. Nevertheless, diatoms have homologs of SAD. Desaturation in cyanobacteria, including oleic acid synthesis, uses acyl lipids as the substrate rather than acyl ACP or acyl CoA.

### 3.2.2. Chloroplast Lipid Biosynthesis

Both chloroplasts and cyanobacteria contain two galactolipids, monogalactosyl diacylglycerol (MGDG) and digalactosyl diacylglycerol (DGDG), as well as a sulfolipid, sulfoquinovosyl diacylglycerol (SQDG), and a phospholipid, phosphatidylglycerol (PG). A trace amount of phosphatidylcholine is present in the chloroplast. Tri- and tetragalactosyl diacylglycerols are also found in land plants depending on growth conditions. The set of MGDG, DGDG, and SQDG, or 'glycolipid trio', is indeed a common trait of oxygen-evolving photosynthetic organisms. A rare exception is *Gloeobacter* which lacks SQDG. The lack of SQDG and thylakoid membrane in this group of cyanobacteria could be related to each other, but little is known about this question. Photosynthetic bacteria without oxygen evolution also contain some of these lipids, but not the complete trio. Therefore, the glycolipid trio seemed to be good evidence for the cyanobacterial origin of chloroplasts. This prompted lipid biologists to call the lipid synthesis pathway in the chloroplast "prokaryotic pathway" [79]. This name had a connotation that the lipid biosynthetic pathway of chloroplast originated from cyanobacteria. However, this turned out to be false. The entire pathways in cyanobacteria and chloroplasts are unrelated in terms of biochemistry and phylogeny (Figure 2).



**Figure 2.** Glycolipid biosynthesis pathways in cyanobacteria and chloroplasts. The distinction between galactose and glucose is shown by the 4'-OH group in red. Cyanobacterial genes are shown in cyan, whereas eukaryotic or non-cyanobacterial genes are shown in green. Note that the biochemical reactions are different in the two steps marked [1] in magenta. In other steps, biochemically identical reactions are catalyzed by structurally unrelated (isofunctional non-homologous) enzymes marked [2]. Eukaryotic paralogs and phylogenetically distant orthologs are marked [3] and [4], respectively.

Essential differences in the galactolipid biosynthetic pathways in cyanobacteria and chloroplasts were found by labeling experiments already in the 1980s ([80,81] and the refer-

ences therein). Namely, a glucolipid, monoglucosyl diacylglycerol (GlcDG), acts as an intermediate in producing MGDG in cyanobacteria. All genes encoding the galactolipid synthesis enzymes were identified in chloroplasts by the end of the 20th century [82,83], whereas the cyanobacterial genes for the synthesis of galactolipids were identified later [84–87]. The conversion of GlcDG to MGDG is an isomerization of sugar moiety called epimerization [88,89]. The enzymes involved in the synthesis of UDP-glucose and UDP-galactose, UDP-glucose pyrophosphorylase and UDP-glucose 4-epimerase, respectively, are both different in cyanobacteria and chloroplasts: UDP-glucose pyrophosphorylase is encoded by the *galU* or *cugP* genes in cyanobacteria [90], whereas it is encoded by the *UGP1/2* gene (cytoplasmic enzyme) and *UGP3* gene (chloroplast enzyme) in plants and algae (see Table 4 in [91]). These are nonhomologous, isofunctional genes. UDP-glucose 4-epimerase is encoded by the *galE* gene in cyanobacteria, while it is encoded by the *UGE1–5* genes (cytoplasmic enzyme) and *PHD1* gene (chloroplast enzyme) in plants and algae. These three kinds of genes are also nonhomologous, isofunctional genes.

Acyltransferases that produce lysophosphatidic acid and phosphatidic acid were also identified in the chloroplast in the early days, but cyanobacterial acyltransferases were identified later. *Escherichia coli* acyltransferases encoded by the *plsB* and *plsC* genes were found in the 1970s, but the acyltransferase for the first acylation was identified quite recently in other bacteria including cyanobacteria. The bacterial system uses acyl phosphate as the acyl donor of the first acylation step and two genes named *plsX* and *plsY* were identified [92]. Identification and phylogenetic analysis [93] and functional characterization [94] of the cyanobacterial genes were performed recently. The cyanobacterial pathway of galactolipid synthesis is, therefore, biochemically different from the corresponding pathway in chloroplasts in at least two points (see Figure 2). The second galactosylation step is catalyzed by two different types of galactosyltransferases in cyanobacteria (*DgdA*) and chloroplasts (*DGD1*). These two enzymes belong to the CAZy GT4 family, but their domain structures are different [95]. Note that the chloroplast genome of some red algae in Cyanidiales encodes a *dgdA* gene [96].

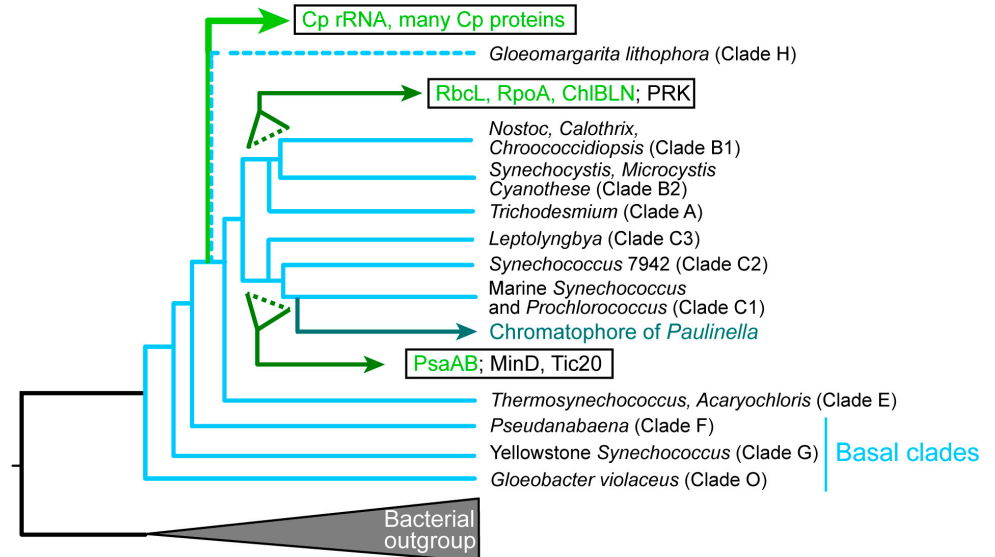
In addition, phylogenetic analysis of all the enzymes involved in the synthesis of galactolipids revealed that none of the chloroplast enzymes originated from cyanobacteria [9,50]. In terms of biochemical reaction, the second acylation and the dephosphorylation are common in chloroplasts and cyanobacteria. Nevertheless, phylogenetic analysis showed that the chloroplast enzymes and cyanobacterial enzymes belong to distant clades. Figure 2 shows that all enzymes in the galactolipid synthesis pathway are biochemically or phylogenetically different in cyanobacteria and chloroplasts (see Figure 1 for the summary of phylogenetic analysis). They are classified into three biochemical categories and a phylogenetic category (this is shown in magenta brackets in Figure 2). (1) Different reactions in cyanobacteria and chloroplasts: *PlsX* + *PlsY* vs. *ATS1*, *MgdA* + *MgdE* vs. *MGD1/2/3*. (2) Structurally different (non-homologous, isofunctional) enzymes: *CugP*/*GalU* vs. *UGP1/2/3*, *GalE* vs. *UGE1–5*/*PHD1*, *DgdA* vs. *DGD1/2*. (3) Chloroplast enzymes are paralogs of ER enzymes: *Pap* vs. *LPPγ/ε*. Therefore, the galactolipid biosynthesis system is unrelated to cyanobacteria and chloroplasts, and this is mostly supported by evidence that does not rely on phylogenetic analysis. (4) Only the contrast *PlsC* vs. *ATS2* was based on phylogenetic analysis that robustly supported their distant relationship.

Biochemically, the biosynthetic pathways of SQDG and PG are both similar in cyanobacteria and chloroplasts, but the phylogenetic relationship of the enzymes is rather complex [50]. In SQDG synthesis, the UDP-sulfoquinovose is synthesized from UDP-glucose by *SQD1*/*SqdB* (chloroplasts/cyanobacteria, respectively [97]), and then sulfoquinovose is transferred to diacylglycerol (DAG) by *SQD2*/*SqdX* (chloroplasts/cyanobacteria, respectively [98]). *SqdC* replaces *SqdX* in some cyanobacteria. All enzymes of *SQD1*/*SqdB* are homologous, but divided into two major clades, one containing the cyanobacterial clades A-B1-B2 (see Figure 3A for the distinction between the major groups of cyanobacteria), green non-sulfur bacteria, and green algae/plants as well as glaucophytes, and another

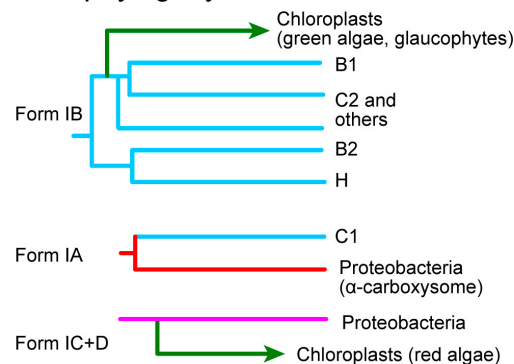


containing the cyanobacterial clades C1-C2,  $\alpha$ -proteobacteria, and red algae (Figure 1). Chloroplast SQD2 is a single clade, sister to the cyanobacterial SqdX.

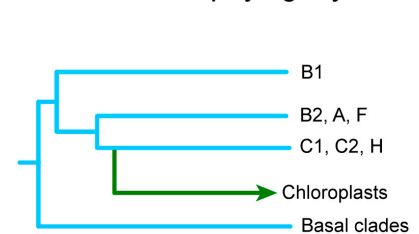
### A. Canonical tree



### B. RbcL phylogeny



### C. PsaA-PsaB phylogeny



**Figure 3.** Schematic phylogenetic trees showing different origins of chloroplast enzymes (or RNA) within the cyanobacterial diversity. **(A)** Canonical phylogenetic tree of cyanobacteria and chloroplasts. Various clades of cyanobacteria are shown according to [53,54]. Chloroplast-encoded proteins and RNA are shown in green. Nuclear-encoded chloroplast proteins are shown in black. The clade names are taken from [53]. According to [54], *G. lithophora* is sister to the main chloroplast clade represented by chloroplast rRNA. For simplicity, Clade H is used for *G. lithophora*. The positions of branching of RbcL and PsaA/B are shown with dotted lines because they do not match exactly with the canonical tree. **(B)** Different origins of chloroplast and cyanobacterial RbcL. **(C)** Probable origin of PsaA and PsaB. As explained in the text and Supplementary Materials, the phylogeny of PsaA and PsaB is very difficult. This diagram, which is consistent with previous results, is still a plausible hypothesis among many other alternatives. Different line colors are used to show different lineages such as cyan for cyanobacteria and green for chloroplast. Colored protein names are used to show chloroplast-encoded proteins.

In PG synthesis, there are two different enzymes of CDP diacylglycerol synthase (CDS): the eukaryotic CDS is found in ER in plants (CDS1 and 2 in *Arabidopsis*), algae, animals, and fungi, whereas the prokaryotic CDS is found in the chloroplast (CDS4 and 5) and bacteria. Chloroplast CDS is likely to originate from cyanobacteria. This is indeed the rare

case (with Cyanidiales DgdA), in which a direct phylogenetic relationship is demonstrated in cyanobacterial and chloroplast orthologs (Type 1 in Figure 1).

In summary, chloroplasts and cyanobacteria are similar in the compositions of lipids that make up thylakoid membranes, but they have essentially unrelated pathways of lipid synthesis. This could be convergent evolution, rather than vertical inheritance, that sustains the functioning of membrane-embedded components of photosynthetic machinery (photosystems, electron transfer chain, and ATP synthase). Many of the enzymes acting within the chloroplasts are eukaryotic enzymes or enzymes that are distantly related to cyanobacterial counterparts. However, many components of the chloroplast fatty acid synthesis system originate from cyanobacteria. This suggests a close relationship between photosynthesis and fatty acid synthesis, which was discussed in the context of the “fatty acid hypothesis” that explains the origin of eukaryotes [9].

#### 4. Diverse Origins of Chloroplast-Encoded Genes

The chloroplast genome has been believed as a reduced cyanobacterial genome that was introduced into the hypothetical algal ancestor. Phylogenetic analysis of the rRNAs was the representative support for the cyanobacterial origin of the chloroplast. However, as pointed out in Section 2.2, this was strongly affected by the genocentric view of the time. Phylogenetic analysis of chloroplast genome has been traditionally performed with concatenated amino acid sequences of conserved genes [53,54] because chloroplast-encoded proteins (e.g., ribosomal proteins) are generally small. In practice, we obtain divergent phylogenetic trees for individual chloroplast ribosomal proteins, but a fairly reliable phylogenetic tree of concatenated ribosomal sequences that resembles the tree of rRNA [9,50]. In the phylogenetic trees of the chloroplast rRNA as well as the concatenated chloroplast-encoded proteins, the chloroplast clade branches from the basal clade of cyanobacteria after the clade E, but before the split of the clade A-B1-B2 and the clade C1-C2-C3 (Figure 3A) [9,34,35,53,54]. Recent work suggested that *Gloeomargarita lithophora* (assigned “clade H” in the present article) is the closest cyanobacterium to the chloroplasts [54]. This is the canonical phylogenetic tree of chloroplasts and cyanobacteria that supports the current notion of the cyanobacterial origin of chloroplasts (Figure 3A). It should be noted, however, that the unity of the chloroplast genome should not be the first assumption to make.

Each of the large proteins, such as the large subunit of rubisco, the RNA polymerase subunits, and the two large subunits of Photosystem I reaction center P700, has sufficient phylogenetic signals for constructing a reliable, individual phylogenetic tree. Examination of these individual trees suggested that the relationship between the cyanobacteria and the chloroplasts might not be straightforward as indicated by the canonical tree. The carbon-fixation enzyme, rubisco, was the first to present evidence for the multiple origins of the chloroplast genome. Let us examine the RbcL phylogeny.

The classification of the large subunit of rubisco, RbcL, is quite complicated. Tabita [99] presented a classification of bacterial and chloroplast RbcL based on biochemical and phylogenetic data. He showed that the red algal rubisco belongs to Form ID, whereas the rubisco of the green plants and algae, as well as many cyanobacteria, belong to Form IB. Form ID was closely related to Form IC in various proteobacteria (see Figure S1A). Delwiche and Palmer [100] proposed that the chloroplast *rbcL* (and *rbcS*, too) gene, which had been originally introduced by a cyanobacterial endosymbiont, was replaced by an  $\alpha$ -proteobacterial homolog in the red lineage. This was a novel idea that the chloroplast genome can be a target of horizontal gene transfer. Non-cyanobacterial origin of chloroplast-encoded genes was also found in the phylloquinone/menaquinone biosynthesis genes (*menA*, *B*, *C*, *D*, *E*, *F*, *G*, and *H*) [101], which was interpreted as multiple horizontal gene transfer events in the ancestral endosymbiont and cyanobacteria.

Further phylogenetic analysis of the RbcL proteins revealed that bacterial RbcL is highly diversified and that the RbcL of some cyanobacteria, later classified as clade C1 [53], is sister to a clade including various proteobacteria containing  $\alpha$ -carboxysomes (Figure S1.

See also Supplementary Material 4 of [50]). The name “ $\alpha$ -cyanobacteria” was used to call these cyanobacteria including *Prochlorococcus* and marine *Synechococcus*. Many other cyanobacteria contain  $\beta$ -carboxysomes and are called  $\beta$ -cyanobacteria, but these were a mixture of various clades as revealed in [53,54]. The two types of carboxysomes are biochemically different and correspond to different genomic structures [102,103].

Another complication of the RbcL phylogeny was found in the point of divergence of chloroplasts within the cyanobacterial clade. Note that the C1 clade has Form IA rubisco, which is not related to the chloroplast rubisco. Figure S1B–D,E (with different taxon sampling and different methods) show that the chloroplast clade of green plants/algae diverged from the base of the B1–C2 clade rather than the B2–H clade, although the branching pattern within the cyanobacteria was different from the canonical tree (Figure 3B).

Phylogenetic analysis of various chloroplast-encoded genes (see also Supplementary Material 4 of [50]) revealed that the *rpoA* gene (encoding the  $\alpha$  subunit of prokaryotic RNA polymerase) as well as the *chlB*, *L*, *N* genes (encoding the light-independent protochlorophyllide oxidoreductase) originated from the base of the cyanobacterial clade A–B1–B2, whereas the *psaA* and *psaB* genes (encoding the P700 reaction center proteins of Photosystem I) originated from the base of the cyanobacterial clade C1–C2 (Figure 3C; see also [104] for *chlB*, *L*, *N*, and [105] for *psaA* and *psaB*). Phylogenetic analysis was performed to confirm these previous results on RpoA, PsaA, and PsaB. The result of RpoA was clearly confirmed (Figure S2), whereas the analysis of PsaA and PsaB turned out to be very difficult (see descriptions in Supplementary Materials). Various analyses using different taxon sampling and different phylogenetic methods gave inconsistent results. Finally, the previous result is identified as the most plausible hypothesis (Figure 3C). This is really a rare case of difficult phylogenetic analysis, and this tentative conclusion should be re-examined in further studies.

In these phylogenetic trees, only the *rpoA* tree showed a sister relationship between *G. lithophora* and chloroplasts. However, in this case, the position of *G. lithophora* was different from the position found in the canonical phylogenetic tree [54]. Therefore, the phylogenetic relationship of cyanobacterial species is variable depending on genes.

These findings as well as previous results [9,50] suggest that the chloroplast genome is not just a reduced form of a single ancestral cyanobacterial genome. Rather, it could be a chimera resulting from multiple gene transfer events from different clades of cyanobacteria. Nevertheless, we still have to consider that the phylogenetic origins of *rbcL* and *psaA/B* as well as some genes that are shown in Figure 3 could result from artifacts of phylogenetic reconstruction. As described above, the origin of chloroplast *psaA* and *psaB* will have to be studied further. Nevertheless, if we never considered a possibility that the chloroplast genome could be a chimera of different cyanobacterial genomes, this is a good occasion to verify or falsify this hypothesis. This was already challenged by a separate origin of the red algal *rbcL* gene. There could be more examples than those that I described above.

Against the data shown above, various criticisms can be raised. An alternative interpretation could be an ancient gene duplication followed by differential gene loss as in the case of *psbA* [106]. Indeed, *psaA* and *psaB* show a trace of gene duplication in cyanobacteria. Another alternative is an artifact due to phylogenetic inference. Indeed, the cyanobacterial parts of the trees of *rbcL*, *psaA*, *psaB*, and others as described above are somewhat different compared to the canonical tree. Gene trees of cyanobacteria will have to be rigorously re-examined before we can conclude a single origin of chloroplast genome within the cyanobacteria based on the concatenated tree, which averages different evolutionary histories of individual genes.

## 5. Other Chloroplast Enzymes

### 5.1. Enzymes Related to DNA Replication

Enzymes involved in the replication of the chloroplast genome are now identified as Plastid/protist organellar DNA polymerase (POP), a single polypeptide of putative viral origin. POP functions in both mitochondria and chloroplasts in all photosynthetic eukary-

otes [107]. POP is also the mitochondrial replicase in most non-photosynthetic protists such as *Tetrahymena* and amoebozoans. In animals and fungi, DNA polymerase  $\gamma$  was identified in the early days of molecular biology, and many biochemistry textbooks still describe that the replication of the mitochondrial genome is catalyzed by DNA polymerase  $\gamma$ . Nevertheless, this is not valid in most eukaryotes other than animals and fungi. According to the typical endosymbiotic scenario, POP is supposed to replace the proteobacterial DNA polymerase III in the initial eukaryotic cell that engulfed the proteobacterial endosymbiont. In the lineage leading to animals (Holozoa) and fungi, POP was replaced by DNA polymerase  $\gamma$ . POP became also used in the chloroplasts after the formation of the initial photosynthetic alga [108]. During the evolution of eukaryotes, various proteins involved in the organellar replication were also replaced by non-cyanobacterial components [109]. As shown in Figure 1, the origins of chloroplast replication machinery and the enzymes involved in chloroplast transcription and translation are different in chloroplasts and cyanobacteria.

I proposed ‘discontinuous evolution of plastid genomic machinery’ based on a comparison of transcription factors and RNA polymerases [110], emphasizing the vast loss of prokaryotic transcription factors in the chloroplasts of green lineage upon the primary endosymbiosis if chloroplasts originated from cyanobacterial endosymbiosis. This idea was based on the “primary endosymbiosis”, which I accepted as a premise at the time. In other words, discontinuity must be defined by setting a reference to something continuous, which was supposed to be cyanobacterial endosymbiont—chloroplast continuity. In the current state of knowledge, the discontinuity notion could be replaced by diverse origins of various components of the chloroplast nucleoid.

### 5.2. Division Machinery

Reproduction by binary fission is an important similarity of chloroplasts and cyanobacteria (Item 3 in Table 1). Various prokaryotic components of division machinery are known. FtsZ, which is a tubulin-like protein that forms a ring at the site of cell division in bacteria, is also involved in chloroplast division in both green plants and red algae (for reviews, see [111,112]). The division site is determined by the dynamic interaction of MinC, MinD, and MinE in bacteria, whereas the role of MinD and MinE in chloroplast division was identified in plants [111,113]. Various eukaryotic components are also known in chloroplast division, such as ARC5 (a family of dynamin), ARC3, and PDV1 and 2. Note that some components (MinD, MinE, ARC3, and PDV1/2) are not found in red algae. Nevertheless, the molecular architecture of the chloroplast division rings was elucidated in the model red alga, *Cyanidioschyzon merolae* [112]. Phylogenetic analysis of FtsZ proteins showed that the chloroplast clades are sister to the cyanobacterial clade (Type 2 tree in Figure 1. See also Supplementary Material 7 of [50]). This point was not discussed in the past by researchers (e.g., [111]), maybe because they believed that the chloroplast FtsZ originated from cyanobacteria. FtsZ is duplicated into FtsZ1 and Z2 in green algae and plants. FtsZ3 is also present in some plants.

Chloroplast MinD (only present in glaucophytes, green algae, and plants) seemed to originate from cyanobacteria, whereas chloroplast MinE is not directly related to cyanobacterial MinE (Figure 1). MinD diverged from the root of the cyanobacterial clade C1-C2, rather than the base of cyanobacteria (Figure 3A). MinE is a small protein (about 100 amino acid residues), and the curious phylogenetic origin could result from insufficient phylogenetic signals. These findings suggest that FtsZ, MinD, and MinE, which act in cell division in cyanobacteria, were introduced into chloroplasts via divergent processes, rather than by a single event of cyanobacterial endosymbiosis. This should also be studied further in the future.

### 5.3. Carbon Fixation and Amino Acid Synthesis

The origin of the enzymes of the Calvin-Benson cycle is diverse (Figure 1). This was first noted by Martin and Schnarrenberger [114] using cDNA data available at the time. We found that the same pattern of chimerism is conserved in *A. thaliana* and *C. merolae*

in the genome paper of *C. merolae* [56]. As described in Section 2.5, this was taken as evidence for the monophyly of red algae and plants. *C. paradoxa* also showed the same chimerism [57], supporting the monophyletic origin of the three lineages. Five enzymes in the chloroplast Calvin-Benson cycle are of the eukaryotic origin. Four enzymes originated from cyanobacteria, and two enzymes are sister to cyanobacteria (Type 2 tree). Nevertheless, the only enzyme encoded by the chloroplast genome, rubisco, must have been acquired by a process different from the main cyanobacterial gene transfers, either in the green lineage (from the B1-C2 clade. See Figure 3B) or in the red algae (from the C1 clade. See Figure 1).

Phylogenetic analysis of the amino acid synthesis enzymes in the chloroplast suggested that only 21 out of >100 Arabidopsis nuclear-encoded plastid enzymes involved in amino acid biosynthesis were unambiguously cyanobacteria-derived proteins [115]. Although the authors did not distinguish different types of phylogenetic relationships, my re-analysis using the Gclust database indicated that many of the enzymes described as cyanobacterial origin presented Type 2 tree (such as Fd-GOGAT, tryptophane synthase  $\alpha$  and  $\beta$  subunits, etc.). For some other enzymes, the origins of the enzymes in the red and green lineages are different. More analysis in depth should be necessary, although the conclusion stating the non-cyanobacterial origin of the majority of enzymes seems to be valid.

#### 5.4. Translocon Components

Translocon is the machinery that incorporates nuclear-encoded organellar proteins from the cytosol, where they are synthesized by the eukaryotic ribosomes (for a recent review, see [116]). In principle, translocon acts as the identifier of organelle that maintains the identity of a particular organelle, although some exceptional proteins use a different pathway to enter chloroplasts [117]. The structure of chloroplast translocon is being elucidated by extensive studies. An interesting similarity of chloroplast translocon and bacterial protein export system has been noted (see Figure 2 of [116]), although not all components are homologous in the two systems. An essential question remains as to how a similar mechanism can explain the reverse direction of protein transport. Translocon components are present in both the inner and outer envelope membranes of the chloroplast and are named Tic and Toc, respectively. Most of the translocon components are shared by green algae/plants, red algae, and glaucophytes, and this is taken as good evidence for the monophyly of the three lineages [56,118]. See also a detailed table on the distribution of components in plants and algae in [119]), and thus, the common origin of all chloroplasts. Curiously, Tic20, Tic21, Tic22, Toc75, and Tic236 of chloroplast translocon have cyanobacterial homologs. Toc75 and Tic236 have also bacterial homologs Omp85 (TamA) and TamB, respectively. Tic20, Tic21, and Tic22 are present essentially in cyanobacteria and chloroplasts of plants and algae, and therefore, belong to Type 1c in Figure 1. Because the reliable phylogenetic relationships of chloroplast and cyanobacterial homologs were not obtained for Tic21, Tic22, Tic236/TamB, and Toc75/Omp75, only a Tic20 tree is presented in Supplementary Materials (Figure S5). Tic20 diverged from the C1 clade (Figure 3A and Figure S5).

## 6. Essential Differences in Chloroplasts from *Paulinella* Chromatophores

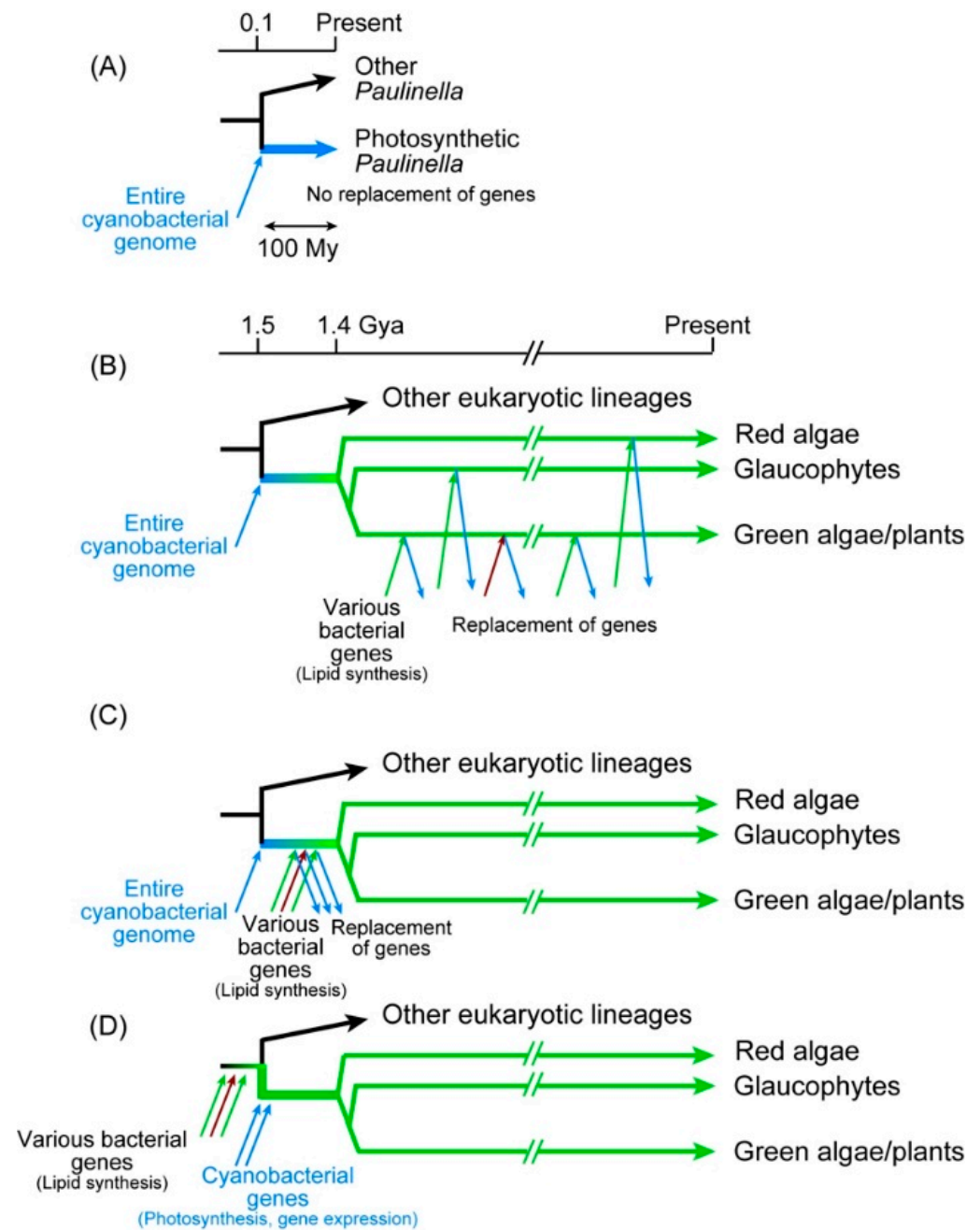
### 6.1. Chloroplasts and Chromatophores

*P. chromatophora* is a photosynthetic rhizopod having two photosynthetic organelles, called chromatophores or cyanelles [67,120,121]. As described in Section 3.1, the chromatophore genomes of *P. chromatophora* and *P. micropora*, as well as the nuclear genome of *P. chromatophora* were determined [71–73]. Phylogenetic analysis of the chromatophore genes revealed that the chromatophores diverged from the base of the cyanobacterial clade C1 (Figure 3A), and thus, believed to represent another case of primary endosymbiosis, or “plastid in the making” [71,122]. The relationship between chloroplasts and chromatophores has been discussed in the context of two distinct events of primary endosymbiosis [4,6,123]. The chromatophore was established about 90–140 Mya according



to the estimation using 18S rRNA [124]. Therefore, the emergence of chromatophores is quite new compared to chloroplasts. The common origin of chloroplasts is estimated to be established about 1.4–1.5 Gya [125,126], indicating that chloroplasts are about 15 times older than chromatophores. This ancestry difference is often taken as the reason for the greater trait differences between the chloroplast and the cyanobacteria than those between the chromatophores and the cyanobacteria. Nevertheless, we have to recognize that the chloroplast traits common in the three lineages, green algae/plants, red algae, and glaucophytes, must have been established before the divergence of these three lineages. The time range measured from the emergence of the common ancestor of the three lineages until their divergence must have been at most about 100 million years, which is roughly similar to the ancestry of photosynthetic *Paulinella* (Figure 4).

Then, a question is asked why the chromatophore of *Paulinella* keeps its large cyanobacterial genome encoding all the enzymes of cyanobacterial lipid biosynthesis pathways. We showed by genomic and biochemical analyses that *P. chromatophora* and *P. micropora* possess authentic cyanobacterial pathways of fatty acid and lipid biosynthesis in the chromatophore, in parallel with the eukaryotic lipid biosynthetic pathway in the ER and Type I fatty acid synthase in the cytosol [74]. All the enzymes of the cyanobacterial pathway of fatty acid and lipid synthesis are encoded by the chromatophore genome, and all of them diverged from the base of the cyanobacterial clade C1 [50,93]. Therefore, the chromatophore membranes are a legacy of the cyanobacterial endosymbiont, but the situation is different in the chloroplast.



**Figure 4.** Various scenarios on the origin of non-cyanobacterial proteins in the chloroplast. (A) The chromatophore of *Paulinella* is a reference, in which all components originate essentially from cyanobacteria. (B–D) Various possible versions of explanations for the non-cyanobacterial proteins in chloroplasts. (B) Cyanobacterial proteins introduced by the primary endosymbiosis were replaced by horizontal gene transfers during the long evolution of plants and algae. (C) Non-cyanobacterial proteins that are conserved in the three lineages must have been introduced before the divergence of the three lineages. (D) Non-cyanobacterial proteins might be introduced before chloroplast formation as a pre-adaptation. Note that the time ranges of chromatophore evolution and chloroplast diversification are about the same (100 million years). See text for details. The colors are used to indicate different lineages, such as cyan for cyanobacteria and green for chloroplast.

### 6.2. Possibly Different Histories of Chloroplasts and Chromatophores

Figure 4 shows possible explanations of the origin of chloroplast and chromatophore membrane lipids. The *Paulinella* system is shown in panel A. Photosynthetic *Paulinella*

emerged as a result of cyanobacterial endosymbiosis about 100 million years ago. The endosymbiont genome remains as a slightly reduced form (encoding 867 and 871 proteins in *P. chromatophora* and *P. micropora*, respectively) within the chromatophore, and encodes all the enzymes of the cyanobacterial lipid biosynthesis pathways. Glycolipids (MGDG, DGDG, and SQDG) are synthesized by the enzymes encoded by the chromatophore genome.

Figure 4B–D shows various possibilities of the origin of glycolipids in chloroplasts. Many researchers simply believe scenario B (Figure 4B), in which cyanobacterial endosymbiont was introduced into the proto-algal cell, implicating that the initial system of chloroplast lipid synthesis was completely cyanobacterial. Then, during the long time of chloroplast evolution (about 1400 million years), every gene for the glycolipid synthesis was replaced by a bacterial one via horizontal gene transfer. This is the opinion of many researchers of chloroplast evolution, who consider lipid synthesis as a minor metabolic activity of the chloroplast compared to photosynthesis. However, this scenario is unlikely, because all the enzymes of glycolipid synthesis (with limited exceptions) in green algae/plants, red algae, and glaucophytes are monophyletic. Note that the rare exceptions are, as described in Section 3.2.2, CDS and red algal DgdA, the latter of which will be discussed later. The majority of the glycolipid synthesis enzymes (Figure 2) of the three lineages are monophyletic and must have already existed in the common ancestor of the three lineages (or Archaeplastida). To explain this, we have to assume that all horizontal gene transfers that provided these enzymes occurred before the divergence of the three lineages (Figure 4C). The time window in which these numerous gene transfers occurred must be within about 100 million years (note that the values 1.5 and 1.4 Gya are obtained from the molecular clock-calibrated phylogenetic trees presented in [125,126] as rough estimates). Alternatively, the cyanobacterial endosymbiont had already acquired all the glycolipid synthesis genes before the endosymbiosis. However, this is also unlikely, because the cyanobacterium (before endosymbiosis) was assumed to have two complete sets of glycolipid synthesis systems. In bacterial genetics, redundant genes are easily lost during culture. If two parallel systems existed in the ancestral endosymbiont and only a single enzyme for each reaction was selected during the subsequent chloroplast evolution, the origins of enzymes would be a mixture of cyanobacterial and non-cyanobacterial, as found in the Calvin-Benson cycle enzymes (Section 5.3). This contrasts with the situation in *Paulinella*, which did not change the lipid biosynthesis system in a comparable time range. Horizontal gene transfers could be more frequent at the time of chloroplast formation than now. However, we might be able to extend the time range before the formation of chloroplast by “primary endosymbiosis” (Figure 4D). The common ancestor of proto-alga could have already possessed the enzymes for glycolipid synthesis.

### 6.3. Extrachloroplast Glycolipids Are Traces of Their Eukaryotic Origin

The acquisition of glycolipids before chloroplast formation might seem strange for most readers, because glycolipids are components of the thylakoid membranes, and are not necessary before the beginning of photosynthesis. Nevertheless, we know that glycolipids can replace phospholipids in a condition of phosphate limitation. In *Arabidopsis*, DGDG accumulates in phosphate deficiency to replace phospholipids in various cellular membranes other than chloroplast membranes [127–130]. Even in phosphate replete conditions, accumulation of DGDG in the plasma membrane was found in the *Arabidopsis* pollen tube [131]. PG deficiency is also partially complemented by the increase of SQDG in the chloroplast [132]. Adaptive accumulation of glycolipids is also known in bacteria [133].

In the common view of plant lipid biochemists, extrachloroplast localization of glycolipids is an emergency case of re-location of chloroplast lipids accelerated by stress conditions. In animals and fungi, phospholipids are synthesized in the ER and transported to various membranes. In plants, chloroplasts and other membranes are so different in lipid composition that lipid transfers from chloroplasts to other membranes have not been considered, although lipid flow from the ER to chloroplasts is well documented. In this context, glycolipids are specific to chloroplasts, and the re-localization of glycolipids in

other membranes is an exceptional phenomenon. However, the presence of glycolipids in extrachloroplast membranes in plants can be understood differently. The ability to synthesize glycolipids could be originally a eukaryotic property that is hidden behind the chloroplast glycolipids in the extant plants but becomes beneficial in particular situations such as phosphate limitation or specific cell types.

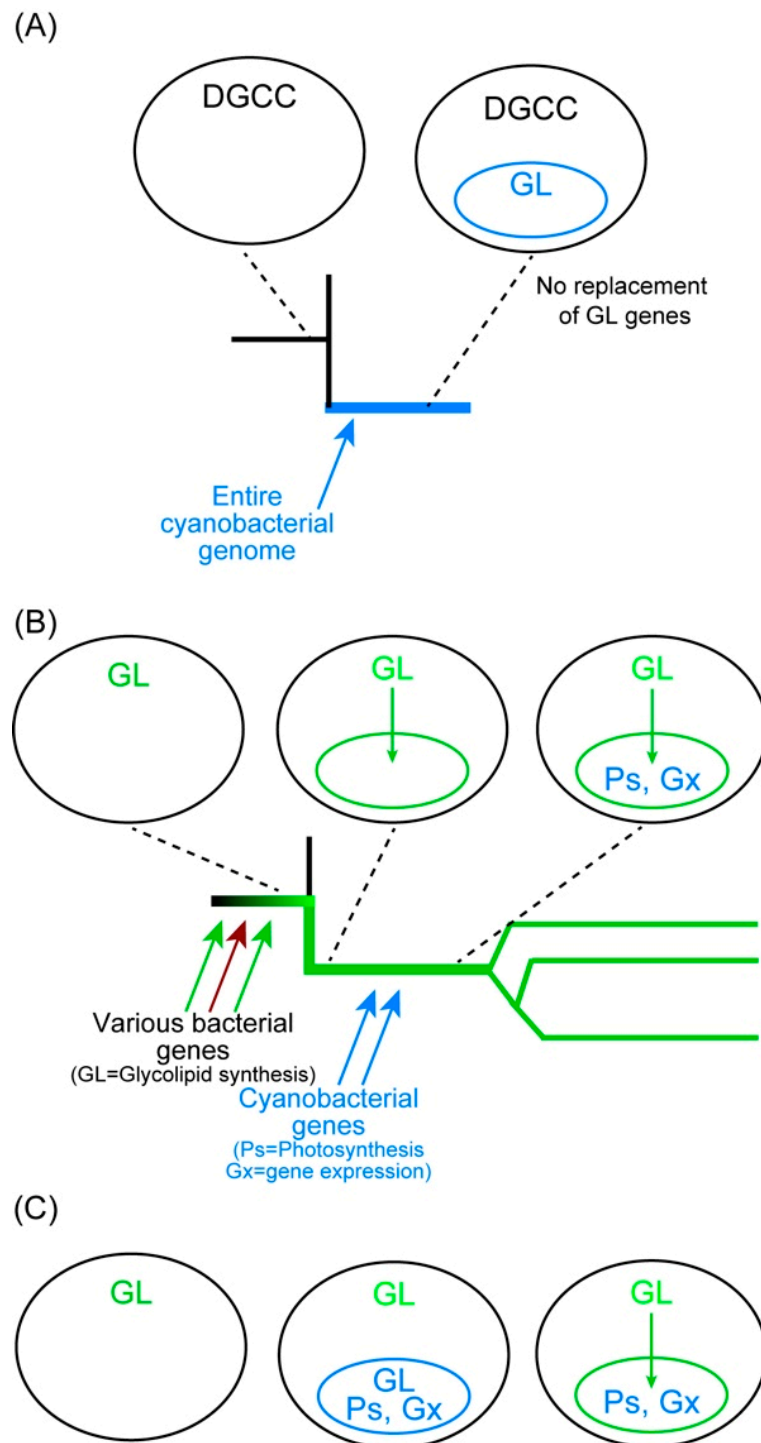
In microorganisms that do not synthesize glycolipids, phosphate limitation results in the accumulation of betaine lipids [134,135]. Betaine lipids are supposed to substitute zwitterionic phospholipids such as phosphatidylcholine and are widely distributed in algae, mosses, and pteridophytes [136]. In this respect, it is interesting to note that *P. micropora* contains a high level of betaine lipid called diacylglyceryl 3-*O*-carboxyhydroxymethylcholine (DGCC) [74]. The accidental choice of glycolipids or betaine lipids in an ancestral eukaryotic cell under phosphate starvation probably determined the fate of photosynthetic organelles in their offsprings.

#### 6.4. Unified Model of Chromatophore and Chloroplast Formation

Imagine a proto-algal cell living in seawater, in which the concentration of inorganic phosphate is very low. There are two alternative lipids to overcome phosphate limitation: glycolipids or betaine lipids. If the proto-alga acquired the biosynthetic ability of glycolipids, then, glycolipids must have been beneficial to life before the beginning of photosynthesis. In this “pre-adaptation” scenario (Figure 4D), glycolipids were used to form various cellular membranes as known in extant plants. The site of glycolipid synthesis might be a small vesicle, which then, became the site of photosynthesis by acquiring various genes from cyanobacteria. We do not have to suppose acquisition of a cyanobacterial cell as a whole, but the acquisition of cyanobacterial genes is sufficient. A cyanobacterial cell could be incorporated, but the membranes were replaced by the host-derived membranes consisting of glycolipids. This is an alternative scenario for the formation of the chloroplast.

The case of *Paulinella* was different because betaine lipid was adopted to overcome phosphate limitation. Currently, the mechanism of DGCC synthesis is unknown, and no genes involved in its synthesis are found. Lipid analysis of non-photosynthetic *Paulinella* has not been performed. Nevertheless, we can assume that non-photosynthetic *Paulinella* has DGCC because this is a eukaryotic component of the cell that is synthesized in the ER [74]. If glycolipids but not betaine lipids are required for photosynthesis, then, *Paulinella* had to keep the complete machinery of glycolipid synthesis after the acquisition of a cyanobacterial endosymbiont which became chromatophores (Figure 5A).

In contrast, the situation of chloroplast may be viewed as Figure 5B or Figure 5C. The scenario in panel B is briefly described above. The host cell acquired glycolipid synthesis systems as an adaptation to phosphate limitation and formed a small compartment surrounded by a membrane which is the site of glycolipid synthesis. Then, various genes for photosynthesis and gene expression systems were acquired and converted the membrane compartment into a chloroplast. In a slightly different scenario in Figure 5C, a cyanobacterial cell was incorporated within the glycolipid-synthesizing compartment. At the initial stage, the whole cyanobacterial genome is kept in the compartment. The two membranes surrounding the cyanobacterial cell were soon lost, and the glycolipids necessary for the endosymbiont were provided by the host-derived membrane. Accordingly, the genes for the glycolipid synthesis enzymes were lost from the endosymbiont genome. An explanation of two membranes in the chloroplast envelope could be invagination of the original glycolipid-forming membrane at the moment of incorporation of cyanobacterial materials. This is a point to clarify in future studies.



**Figure 5.** Unified schemes explaining the formation of membrane systems of (A) chromatophore and (B,C) chloroplast. The key components are phosphorus-free lipids. DGCC is a betaine lipid, and GL represents glycolipid trio (MGDG, DGDG, and SQDG). In scenario (B), glycolipids that were introduced as an adaptation to phosphate limitation form a membrane system, in which cyanobacterial photosynthesis and gene expression systems were incorporated. In scenario (C), the host cell acquired glycolipids before the cyanobacterial endosymbiosis, but then, the cyanobacterial glycolipid synthesis system was lost. Colors are used to indicate different lineages, such as cyan for cyanobacteria and green for chloroplast.



## 6.5. Additional Notes on the Models

### 6.5.1. DGDG Synthesis in Red Algae

The models in Figures 4 and 5 are drawn with a focus on membrane lipid biosynthesis. The origins of most enzymes for the synthesis of chloroplast lipids are not cyanobacteria, but red algal DgdA should be commented on. The galactosyltransferase that produces DGDG is DgdA in cyanobacteria, which is not homologous to DGD1 in plants and algae. The chloroplast genome of Cyanidiales (or Cyanidiophytina) red algae encodes the *dgdA* gene (*ycf82*). These algae do not have the *DGD1* gene in the nuclear genome. In contrast, other red algae (Rhodophytina) have the nuclear *DGD1* gene but not the chloroplast *dgdA* gene. However, all red algae are monophyletic [137], and green algae and glaucophytes have the *DGD1* gene. As the three lineages are monophyletic, we have to assume that the ancestral red algae must have had both *DGD1* and *dgdA* genes [95]. The *dgdA* gene must have been introduced into the ancestral red algae (already possessing *DGD1*) from cyanobacteria with various other genes involved in photosynthesis and gene expression. The co-existence of *dgdA* and *DGD1* was resolved with diversification into the two lineages of red algae: The *dgdA* gene was lost in Rhodophytina, whereas the *DGD1* gene was lost in Cyanidiophytina.

### 6.5.2. Multiple Gene Transfers from Cyanobacteria and Other Bacteria

For simplicity, the models in Figures 4 and 5 are drawn with a minimal number of arrows that indicate gene transfers to the chloroplast. As explained in Section 3, gene transfers must have occurred many times, either from cyanobacteria or from other bacteria. Even the transfers from cyanobacteria occurred at least three times, from the base of the clade C1 (+C2), the base of the clades A-B1-B2, and the common base of cyanobacteria (Figure 3A). Type 2 phylogeny could be explained by gene transfer from an ancestor of cyanobacteria or a parallel, extinct lineage of cyanobacteria. Alternatively, gene transfer could occur from various bacteria to cyanobacteria (or other prokaryotes) and held within the pangenomes before being introduced into chloroplasts [138]. This could be an explanation of the diverse origins of peptidoglycan synthesis enzymes, which is otherwise difficult to explain.

### 6.5.3. Eukaryotic Nature of Chloroplast Envelope

Microscopic observation of a plant or algal cell gives us an impression that chloroplast is a single entity, namely, a package of thylakoid membranes wrapped with envelope membranes. A chloroplast seems to correspond to a microorganism living within a eukaryotic cell. However, another view may be possible. Recent morphological studies revealed that chloroplasts are not always spherical. Many holes or invaginations were found in *Chlamydomonas* chloroplast [139,140]. Holes or pockets were also found in plant chloroplasts [141,142]. The chloroplast surface looks more like the dynamically undulating ER membranes rather than a smooth spherical plasma membrane of cyanobacteria. I propose to consider the envelope membranes as part of the eukaryotic compartment, or the interface between the eukaryotic cytosol and chloroplast stroma. In this respect, the envelope membranes could be considered as a specialized eukaryotic membrane system, in which the glycolipids are synthesized by nuclear-encoded, non-cyanobacterial enzymes (Figure 5B,C).

## 7. Conclusions: Flexible Views on the Origin

The present article focuses on the phylogenetic origin of the similarity of cyanobacteria and chloroplasts. As is well known, essentially all components of primary reactions of photosynthesis (photosystems, electron transfer, and ATP synthase) and gene expression systems (transcription and translation) of chloroplast undoubtedly originated from cyanobacteria. This led to the prevailing view of the endosymbiotic origin of chloroplasts from cyanobacteria. However, many of the common traits in cyanobacteria and chloroplasts are not attributed to common origins (Table 2). We found different biosynthetic

pathways or phylogenetically unrelated enzymes in the synthesis of glycolipids and peptidoglycan. Non-cyanobacterial enzymes are also known to function in various metabolisms of chloroplasts. The currently accepted view of chloroplast origin assumes a single event of cyanobacterial endosymbiosis, probably a cyanobacterium related to *Gloeomargarita* [54], which is assumed as the common ancestor of the chloroplasts of Archaeplastida (green algae/plants, red algae, and glaucophytes). However, the single cyanobacterial endosymbiosis and the monophyly of the three lineages are different problems. Monophyly of the three lineages just supposes the common event of chloroplast formation, which could be more complicated than a single event of cyanobacterial endosymbiosis. We found at least three gene flows from cyanobacteria to the common ancestor of chloroplasts. This plurality of gene flow, as well as non-cyanobacterial enzymes involved in the traits shared by cyanobacteria and chloroplasts, suggests complexity in chloroplast formation. Biosynthesis of glycolipids could have preceded chloroplast formation as an adaptation to phosphate limitation, and it could have prepared the formation of membranes made of glycolipids, which then facilitated the incorporation of photosynthetic machinery from cyanobacteria (pre-adaptation). I call this “host-directed chloroplast formation” hypothesis [9]. This may be qualified as a “quasi-autogenous” model of chloroplast origin, but I prefer to characterize it in terms of “pre-adaptation”, which can explain the formation of both chloroplasts and chromatophores. For many researchers working on chloroplasts and cyanobacteria, the origin of membrane lipids might not be a central question. But I believe that lipid origin is crucial in all illustrations of chloroplast origin.

**Table 2.** The similarity of chloroplasts and cyanobacteria turned out to be explained by different mechanisms.

Apparent Similarity	Revealed Difference	Cyanobacteria	Chloroplasts	Reference
Chlorophylls and tetrapyrroles	Partially different biosynthesis pathway (PPO) <sup>1</sup>	HemJ (or HemY)	HemY	[143]
Galactolipids (MGDG and DGDG) <sup>2</sup>	Different biosynthesis pathway	MgdA- > MgdE- > DgdA	MGD1- > DGD1	[50]
Peptidoglycan <sup>3</sup>	Different origins of biosynthesis enzymes	Cyanobacterial enzymes	Most enzymes originated from bacteria other than cyanobacteria	[77,78]
DNA replication	Different enzyme	PolIII	POP <sup>4</sup>	[108]
Transcription	Additional enzyme in chloroplast	Prokaryotic RNAP	Prokaryotic RNAP + NEP <sup>5</sup>	[110]
Division by binary fission	Additional components of division machinery in chloroplast	MinC, MinD, MinE, FtsZ, etc.	MinD, MinE, FtsZ, dynamin, etc (no MinC)	[111–113]
Oxygen producing photosynthesis	Partially different components of oxygen-evolving complex	PsbO, PsbU, PsbV	PsbO, PsbP, PsbQ (Red algae retain cyanobacteria-like system)	[144]

<sup>1</sup> PPO, protoporphyrinogen IX oxidase. <sup>2</sup> MGDG, monogalactosyl diacylglycerol; DGDG, digalactosyl diacylglycerol. <sup>3</sup> Peptidoglycan synthesis system is found in glaucophytes, some green algae, bryophytes, pteridophytes, and some gymnosperms. <sup>4</sup> POP, plant/protist organellar DNA polymerase. <sup>5</sup> RNAP, RNA polymerase; NEP, nuclear-encoded RNA polymerase.

**Supplementary Materials:** Supplementary tables and figures are available online at <https://www.mdpi.com/article/10.3390/genes12060823/s1>, Table S1: Summary of phylogenetic analysis of PsaA/PsaB, Table S2: List of organisms used for the phylogenetic analysis of PsaA and PsaB, Figure S1: Phylogenetic trees of RbcL, Figure S2: Phylogenetic trees of RpoA, Figure S3: Phylogenetic trees of PsaA, Figure S4: Phylogenetic trees of PsaB, Figure S5: Phylogenetic trees of Tic20.

**Funding:** This research was funded by Grant-in-Aid for Scientific Research (17H03715) from the Japan Society for the Promotion of Science.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are available from related publications and supplements.

**Acknowledgments:** The author thanks Koichi Kobayashi, Osaka Prefectural University, for fruitful discussions. The super-computing resource was provided by Human Genome Center, The Institute of Medical Science, The University of Tokyo.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Mereschkowsky, C. Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol. Centralblatt* **1905**, *25*, 593–604.
2. Schimper, A.F.W. Über die Entwicklung der Chlorophyllkörper und Farbkörper. *Bot. Z.* **1883**, *41*, 105–162.
3. Martin, W.; Kowallik, K. Annotated English translation of Mereschkowsky's 1905 paper 'Über Natur und Ursprung der Chromatophoren im Pflanzenreiche'. *Eur. J. Phycol.* **1999**, *34*, 287–295. [[CrossRef](#)]
4. McFadden, G.I. Origin and evolution of plastids and photosynthesis in eukaryotes. *Cold Spring Harb. Perspect. Biol.* **2014**, *6*, a016105. [[CrossRef](#)] [[PubMed](#)]
5. Archibald, J. *One Plus One Equals One. Symbiosis and the Evolution of Complex Life*; Oxford University Press: Oxford, UK, 2014.
6. Archibald, J.M. Endosymbiosis and eukaryotic cell evolution. *Curr. Biol.* **2015**, *25*, R911–R921. [[CrossRef](#)] [[PubMed](#)]
7. Martin, W.F. Physiology, anaerobes, and the origin of mitosing cells 50 years on. *J. Theor. Biol.* **2017**, *434*, 2–10. [[CrossRef](#)] [[PubMed](#)]
8. Sato, N. Revisiting the theoretical basis of the endosymbiotic origin of plastids in the original context of Lynn Margulis on the origin of mitosing, eukaryotic cells. *J. Theor. Biol.* **2017**, *434*, 104–113. [[CrossRef](#)] [[PubMed](#)]
9. Sato, N. Endosymbiotic Theories of Organelles Revisited. In *Retrospects and Prospects*; Springer: Singapore, 2020.
10. Kowallik, K.V.; Martin, W.F. The origin of symbiogenesis: An annotated English translation of Mereschkowsky's 1910 paper on the theory of two plasma lineages. *Biosystems* **2021**, *199*, 104281. [[CrossRef](#)]
11. Stiller, J.W. Toward an empirical framework for interpreting plastid evolution. *J. Phycol.* **2014**, *50*, 462–471. [[CrossRef](#)]
12. Sagan, L. On the origin of mitosing cells. *J. Theor. Biol.* **1967**, *14*, 225–274. [[CrossRef](#)]
13. Margulis, L. *Origin of Eukaryotic Cells. Evidence and Research Implications for a Theory of the Origin and Evolution of Microbial, Plant, and Animal Cells on the Precambrian Earth*; Yale University Press: New Haven, CT, USA, 1970.
14. Gray, M.W. The pre-endosymbiont hypothesis: A new perspective on the origin and evolution of mitochondria. *Cold Spring Harb. Perspect. Biol.* **2014**, *6*, a016097. [[CrossRef](#)]
15. Gray, M.W. Mosaic nature of the mitochondrial proteome: Implications for the origin and evolution of mitochondria. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 10133–10138. [[CrossRef](#)]
16. Keeling, P.J. The impact of history on our perception of evolutionary events: Endosymbiosis and the origin of eukaryotic complexity. *Cold Spring Harb. Perspect. Biol.* **2014**, *6*, a016196. [[CrossRef](#)]
17. Sapp, J. *Evolution by Association*; Oxford University Press: Oxford, UK, 1994.
18. Pascher, A. Studien über Symbiosen. Über einige Endosymbiosen von Blaualgen in Einzellern. *Jahrb. Wiss. Bot.* **1929**, *71*, 386–462.
19. Buchner, P. *Endosymbiose der Tiere mit Pflanzlichen Mikroorganismen*; Birkhäuser: Basel, Switzerland, 1953.
20. Lederberg, J. Cell genetics and hereditary symbiosis. *Physiol. Rev.* **1952**, *32*, 403–430. [[CrossRef](#)] [[PubMed](#)]
21. Kozo-Polyanski, B.M. *Symbiogenesis: A New Principle of Evolution*; Translated from Russian to English by Fet, V.; Harvard University Press: Cambridge, MA, USA, 2010.
22. Myer-Abich, A. Beiträge zur Theorie der Evolution der Organismen: II. Typensynthese durch Holobiose. *Bibl. Biotheor.* **1950**, *5*, 1–206.
23. Echlin, P. The cyanophytic origin of higher plant chloroplasts. *Br. Phycol. Bull.* **1966**, *3*, 150–151.
24. Goksøyr, J. Evolution of Eucaryotic cells. *Nature* **1967**, *214*, 1161. [[CrossRef](#)]
25. Taylor, F.J.R. Implications and extensions of the serial endosymbiosis theory of the origin of eukaryotes. *Taxon* **1974**, *23*, 229–258. [[CrossRef](#)]
26. Uzzell, T.; Spolsky, C. Mitochondria and plastids as endosymbionts: A revival of special creation? *Am. Sci.* **1974**, *62*, 334–343.
27. Cavalier-Smith, T. The origin of nuclei and eukaryotic cells. *Nature* **1975**, *256*, 463–468. [[CrossRef](#)] [[PubMed](#)]
28. Bogorad, L. Evolution of organelles and eukaryotic genomes. *Science* **1975**, *188*, 891–898. [[CrossRef](#)]
29. Stanier, R.Y. Some aspects of the biology of cells and their possible evolutionary significance. *Symp. Soc. Gen. Microbiol.* **1970**, *20*, 1–38.
30. Cohen, S.S. Are/Were mitochondria and chloroplasts microorganisms? *Am. Sci.* **1970**, *58*, 281–289. [[PubMed](#)]
31. Cohen, S.S. Mitochondria and chloroplasts revisited. *Am. Sci.* **1973**, *61*, 437–445. [[PubMed](#)]
32. Stanier, R.Y. The origins of photosynthesis in eukaryotes. *Symp. Soc. Gen. Microbiol.* **1974**, *24*, 219–242.
33. Schwartz, R.M.; Dayhoff, M.O. Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. A perspective is derived from protein and nucleic acid sequence data. *Science* **1978**, *199*, 395–403. [[CrossRef](#)]
34. Gray, M.W.; Doolittle, W.F. Has the endosymbiont hypothesis been proven? *Microbiol. Rev.* **1982**, *46*, 1–42. [[CrossRef](#)] [[PubMed](#)]
35. Wallace, D.C. Structure and evolution of organelle genomes. *Microbiol. Rev.* **1982**, *46*, 208–240. [[CrossRef](#)]

36. Delwiche, C.F.; Kuhsel, M.; Palmer, J.D. Phylogenetic analysis of *tufA* sequences indicates a cyanobacterial origin of all plastids. *Mol. Phylogenet. Evol.* **1995**, *4*, 110–128. [[CrossRef](#)]
37. Moreira, D.; Philippe, H. Sure facts and open questions about the origin and evolution of photosynthetic plastids. *Res. Microbiol.* **2001**, *152*, 771–780. [[CrossRef](#)]
38. Delwiche, C.F. Tracing the thread of plastid diversity through the tapestry of life. *Amer. Nat.* **1999**, *154*, S164–S177. [[CrossRef](#)]
39. Palmer, J.D. The symbiotic birth and spread of plastids: How many times and whodunit? *J. Phycol.* **2003**, *39*, 4–11. [[CrossRef](#)]
40. Stiller, J.W.; Reel, D.C.; Johnson, J.C. A single origin of plastids revisited: Convergent evolution in organellar genome content. *J. Phycol.* **2003**, *39*, 95–105. [[CrossRef](#)]
41. Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; Hirosawa, M.; Sugiura, M.; Sasamoto, S.; et al. 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.* **1996**, *3*, 109–136. [[CrossRef](#)]
42. McFadden, G.J. Primary and secondary endosymbiosis and the origin of plastids. *J. Phycol.* **2001**, *37*, 951–959. [[CrossRef](#)]
43. Cavalier-Smith, T. The origin of plastids. *Biol. J. Linn. Soc.* **1982**, *17*, 289–306. [[CrossRef](#)]
44. Cavalier-Smith, T. Membrane heredity and early chloroplast evolution. *Trends Plant Sci.* **2000**, *5*, 174–182. [[CrossRef](#)]
45. Martin, W.F.; Garg, S.; Zimorski, V. Endosymbiotic theories for eukaryote origin. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2015**, *370*, 20140330. [[CrossRef](#)] [[PubMed](#)]
46. Campbell, N.A.; Reece, J.B.; Urry, L.A.; Cain, M.L.; Wasserman, S.A.; Minorsky, P.V.; Jackson, R.B. *Biology*, 10th ed.; Pearson Education: San Francisco, CA, USA, 2016.
47. Martin, W.; Stoebe, B.; Goremykin, V.; Hansmann, S.; Hasegawa, M.; Kowallik, K.V. Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* **1998**, *393*, 162–165. [[CrossRef](#)] [[PubMed](#)]
48. Huang, C.Y.; Ayliffe, M.A.; Timmis, J.N. Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* **2003**, *422*, 72–76. [[CrossRef](#)]
49. Timmis, J.N.; Ayliffe, M.A.; Huang, C.Y.; Martin, W. Endosymbiotic gene transfer: Organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* **2004**, *5*, 123–135. [[CrossRef](#)] [[PubMed](#)]
50. Sato, N. Complex origins of chloroplast membranes with photosynthetic machineries: Multiple transfers of genes from divergent organisms at different times or a single endosymbiotic event? *J. Plant Res.* **2020**, *133*, 15–33. [[CrossRef](#)]
51. Ohta, N.; Sato, N.; Nozaki, H.; Kuroiwa, T. 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. *J. Mol. Evol.* **1997**, *45*, 688–695. [[CrossRef](#)] [[PubMed](#)]
52. Stoebe, B.; Kowallik, K.V. Gene-cluster analysis in chloroplast genomics. *Trends Genet.* **1999**, *15*, 344–347. [[CrossRef](#)]
53. Shih, P.M.; Wu, D.; Latifi, A.; Axen, S.; Fewer, D.P.; Talla, E.; Calteau, A.; Cai, F.; de Marsac, N.T.; Rippka, R.; et al. Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1053–1058. [[CrossRef](#)] [[PubMed](#)]
54. Ponce-Toledo, R.I.; Deschamps, P.; López-García, P.; Zivanovic, Y.; Benzerara, K.; Moreira, D. An early-branching freshwater cyanobacterium at the origin of plastids. *Curr. Biol.* **2017**, *27*, 386–391. [[CrossRef](#)]
55. Rodríguez-Ezpeleta, N.; Brinkmann, H.; Burey, S.C.; Roure, B.; Burger, G.; Löffelhardt, W.; Bohnert, H.J.; Philippe, H.; Lang, B.F. Monophyly of primary photosynthetic eukaryotes: Green plants, red algae, and glaucophytes. *Curr. Biol.* **2005**, *15*, 1325–1330. [[CrossRef](#)] [[PubMed](#)]
56. Matsuzaki, M.; Misumi, O.; Shin-I, T.; Maruyama, S.; Takahara, M.; Miyagishima, S.-Y.; Mori, T.; Nishida, K.; Yagisawa, F.; Nishida, K.; et al. Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* **2004**, *428*, 653–657. [[CrossRef](#)] [[PubMed](#)]
57. Reyes-Prieto, A.; Bhattacharya, D. Phylogeny of Calvin cycle enzymes supports Plantae monophyly. *Mol. Phyl. Evol.* **2007**, *45*, 384–391. [[CrossRef](#)] [[PubMed](#)]
58. Stiller, J.W. Plastid endosymbiosis, genome evolution and the origin of green plants. *Trends Plant Sci.* **2007**, *12*, 391–396. [[CrossRef](#)] [[PubMed](#)]
59. Nozaki, H.; Iseki, M.; Hasegawa, M.; Misawa, K.; Nakada, T.; Sasaki, N.; Watanabe, M. Phylogeny of primary photosynthetic eukaryotes as deduced from slowly evolving nuclear genes. *Mol. Biol. Evol.* **2007**, *24*, 1592–1595. [[CrossRef](#)] [[PubMed](#)]
60. Gawryluk, R.M.R.; Tikhonenkov, D.V.; Hehenberger, E.; Husnik, F.; Mylnikov, A.P.; Keeling, P.J. Non-photosynthetic predators are sister to red algae. *Nature* **2019**, *572*, 240–243. [[CrossRef](#)]
61. Adl, S.M.; Simpson, A.G.B.; Farmer, M.A.; Andersen, R.A.; Anderson, O.R.; Barta, J.; Bowser, S.S.; Brugerolle, G.; Fensome, R.A.; Fredericq, S.; et al. The new higher level classification of Eukaryotes with emphasis on the taxonomy of protists. *J. Eukaryot. Microbiol.* **2005**, *52*, 399–451. [[CrossRef](#)] [[PubMed](#)]
62. Adl, S.M.; Bass, D.; Lane, C.E.; Lukeš, J.; Schoch, C.L.; Smirnov, A.; Agatha, S.; Berney, C.; Brown, M.W.; Burki, F.; et al. Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes. *J. Eukaryot. Microbiol.* **2019**, *66*, 4–119. [[CrossRef](#)] [[PubMed](#)]
63. Pittis, A.A.; Gabaldón, T. Late acquisition of mitochondria by a host with chimaeric prokaryotic ancestry. *Nature* **2016**, *531*, 101–104. [[CrossRef](#)]
64. Gabaldón, T. Relative timing of mitochondrial endosymbiosis and the “pre-mitochondrial symbioses” hypothesis. *IUBMB Life* **2018**, *70*, 1189–1196. [[CrossRef](#)]
65. Harish, A.; Kurland, C.G. Mitochondria are not captive bacteria. *J. Theor. Biol.* **2017**, *434*, 88–98. [[CrossRef](#)]



66. Pfanzagl, B.; Zenker, A.; Pittenauer, E.; Allmaier, G.; Martinez-Torrecuadrada, J.; Schmid, E.R.; De Pedro, M.A.; Löffelhardt, W. Primary structure of cyanelle peptidoglycan of *Cyanophora paradoxa*: A prokaryotic cell wall as part of an organelle envelope. *J. Bacteriol.* **1996**, *178*, 332–339. [[CrossRef](#)] [[PubMed](#)]
67. Stanier, R.Y.; Adelberg, E.A.; Ingraham, J.L. *The Microbial World*, 4th ed.; Prentice-Hall: Englewood Cliffs, NJ, USA, 1976.
68. Takano, H.; Takechi, K. Plastid peptidoglycan. *Biochim. Biophys. Acta* **2010**, *1800*, 144–151. [[CrossRef](#)]
69. Lin, X.; Li, N.; Kudo, H.; Zhang, Z.; Li, J.; Wang, L.; Zhang, W.; Takechi, K.; Takano, H. Genes sufficient for synthesizing peptidoglycan are retained in gymnosperm genomes, and *MurE* from *Larix gmelinii* can rescue the albino phenotype of Arabidopsis *MurE* mutation. *Plant Cell Physiol.* **2017**, *58*, 587–597. [[CrossRef](#)] [[PubMed](#)]
70. Garcia, M.; Myouga, F.; Takechi, K.; Sato, H.; Nabeshima, K.; Nagata, N.; Takio, S.; Shinozaki, K.; Takano, H. An Arabidopsis homolog of the bacterial peptidoglycan synthesis enzyme MurE has an essential role in chloroplast development. *Plant J.* **2008**, *53*, 924–934. [[CrossRef](#)] [[PubMed](#)]
71. Nowack, E.C.; Melkonian, M.; Glöckner, G. Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. *Curr. Biol.* **2008**, *18*, 410–418. [[CrossRef](#)]
72. Nowack, E.C.; Price, D.C.; Bhattacharya, D.; Singer, A.; Melkonian, M.; Grossman, A.R. Gene transfers from diverse bacteria compensate for reductive genome evolution in the chromatophore of *Paulinella chromatophora*. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 12214–12219. [[CrossRef](#)] [[PubMed](#)]
73. Lhee, D.; Yang, E.C.; Kim, J.I.; Nakayama, T.; Zuccarello, G.; Andersen, R.A.; Yoon, H.S. Diversity of the photosynthetic *Paulinella* species, with the description of *Paulinella micropora* sp. nov. and the chromatophore genome sequence for strain KR01. *Protist* **2017**, *168*, 155–170. [[CrossRef](#)] [[PubMed](#)]
74. Sato, N.; Yoshitomi, T.; Mori-Moriyama, N. Characterization and biosynthesis of lipids in *Paulinella micropora* MYN1. Evidence for efficient integration of chromatophores into cellular lipid metabolism. *Plant Cell Physiol.* **2020**, *61*, 869–881. [[CrossRef](#)] [[PubMed](#)]
75. Hirano, T.; Tanidokoro, K.; Shimizu, Y.; Kawarabayasi, Y.; Ohshima, T.; Sato, M.; Tadano, S.; Ishikawa, H.; Takio, S.; Takechi, K.; et al. Moss chloroplasts are surrounded by a peptidoglycan wall containing D-amino acids. *Plant Cell* **2016**, *28*, 1521–1532. [[CrossRef](#)]
76. Sato, N.; Toyoshima, M.; Tajima, N.; Takechi, K.; Takano, H. Single-pixel densitometry revealed the presence of peptidoglycan in the intermembrane space of moss chloroplast envelope in conventional electron micrographs. *Plant Cell Physiol.* **2017**, *58*, 1743–1751. [[CrossRef](#)]
77. Sato, N.; Takano, H. Diverse origins of enzymes involved in the biosynthesis of chloroplast peptidoglycan. *J. Plant Res.* **2017**, *130*, 635–645. [[CrossRef](#)]
78. Price, D.C.; Goodenough, U.W.; Roth, R.; Lee, J.-H.; Kariyawasam, T.; Mutwil, M.; Ferrari, C.; Facchinelli, F.; Ball, S.G.; Cenci, U.; et al. Analysis of an improved *Cyanophora paradoxa* genome assembly. *DNA Res.* **2019**, *26*, 287–299. [[CrossRef](#)]
79. Zepke, H.D.; Heinz, E.; Radunz, A.; Linscheid, M.; Pesch, R. Combination and positional distribution of fatty acids in lipid from blue-green algae. *Arch. Microbiol.* **1978**, *119*, 157–162. [[CrossRef](#)]
80. Sato, N.; Murata, N. Lipid biosynthesis in the blue-green alga, *Anabaena variabilis* I. Lipid classes. *Biochim. Biophys. Acta* **1982**, *710*, 271–278. [[CrossRef](#)]
81. Sato, N.; Murata, N. Lipid biosynthesis in the blue-green alga (cyanobacterium), *Anabaena variabilis* III. UDP-glucose:diacylglycerol glucosyltransferase activity in vitro. *Plant Cell Physiol.* **1982**, *23*, 1115–1120. [[CrossRef](#)]
82. Shimojima, M.; Ohta, H.; Iwamatsu, A.; Masuda, T.; Shioi, Y.; Takamiya, K. Cloning of the gene for monogalactosyldiacylglycerol synthase and its evolutionary origin. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 333–337. [[CrossRef](#)] [[PubMed](#)]
83. Dörmann, P.; Balbo, I.; Benning, C. Arabidopsis galactolipid biosynthesis and lipid trafficking mediated by DGD1. *Science* **1999**, *284*, 2181–2184. [[CrossRef](#)]
84. Awai, K.; Kakimoto, T.; Awai, C.; Kaneko, T.; Nakamura, Y.; Takamiya, K.-I.; Wada, H.; Ohta, H. Comparative genomic analysis revealed a gene for monoglucosyldiacylglycerol synthase, an enzyme for photosynthetic membrane lipid synthesis. *Plant Physiol.* **2006**, *141*, 1120–1127. [[CrossRef](#)]
85. Awai, K.; Watanabe, H.; Benning, C.; Nishida, I. Digalactosyldiacylglycerol is required for better photosynthetic growth of *Synechocystis* sp. PCC6803 under phosphate limitation. *Plant Cell Physiol.* **2007**, *48*, 1517–1523. [[CrossRef](#)]
86. Sakurai, I.; Mizusawa, N.; Wada, H.; Sato, N. Digalactosyldiacylglycerol is required for stabilization of the oxygen-evolving complex in photosystem II. *Plant Physiol.* **2007**, *145*, 1361–1370. [[CrossRef](#)]
87. Awai, K.; Ohta, H.; Sato, N. Oxygenic photosynthesis without galactolipids. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 13571–13575. [[CrossRef](#)]
88. Sato, N. Is monoglucosyl diacylglycerol a precursor to monogalactosyl diacylglycerol in all cyanobacteria? *Plant Cell Physiol.* **2015**, *56*, 1890–1899. [[CrossRef](#)]
89. Sato, N.; Okazaki, Y.; Saito, K. Isotopic combinator analysis provides in vivo evidence of the direct epimerization of monoglucosyl diacylglycerol in cyanobacteria. *Biochemistry* **2016**, *55*, 5689–5701. [[CrossRef](#)]
90. Maeda, K.; Narikawa, R.; Ikeuchi, M. CugP is a novel ubiquitous non-GalU-type bacterial UDP-glucose pyrophosphorylase found in cyanobacteria. *J. Bacteriol.* **2014**, *196*, 2348–2354. [[CrossRef](#)]
91. Mori, N.; Moriyama, T.; Sato, N. Uncommon properties of lipid biosynthesis of isolated plastids in the unicellular red alga *Cyanidioschyzon merolae*. *FEBS Open Bio* **2019**, *9*, 114–128. [[CrossRef](#)]



92. Lu, Y.J.; Zhang, Y.M.; Grimes, K.D.; Qi, J.; Lee, R.E.; Rock, C.O. Acyl-phosphates initiate membrane phospholipid synthesis in Gram-positive pathogens. *Mol. Cell.* **2006**, *23*, 765–772. [[CrossRef](#)]
93. Sato, N.; Awai, K. “Prokaryotic Pathway” is not prokaryotic: Noncyanobacterial origin of the chloroplast lipid biosynthetic pathway revealed by comprehensive phylogenomic analysis. *Genome Biol. Evol.* **2017**, *9*, 3162–3178. [[CrossRef](#)]
94. Matsumoto, T.; Awai, K. Adaptations in chloroplast membrane lipid synthesis from synthesis in ancestral cyanobacterial endosymbionts. *Biochem. Biophys. Res. Commun.* **2020**, *528*, 473–477. [[CrossRef](#)]
95. Sato, N.; Awai, K. Diversity in biosynthetic pathways of galactolipids in the light of endosymbiotic origin of chloroplasts. *Front. Plant Sci.* **2016**, *7*, 117. [[CrossRef](#)] [[PubMed](#)]
96. Sato, N.; Moriyama, T. Genomic and biochemical analysis of lipid biosynthesis in the unicellular rhodophyte *Cyanidioschyzon merolae*: Lack of plastidic desaturation pathway results in mixed pathway of galactolipid synthesis. *Eukaryot. Cell* **2007**, *6*, 1006–1017. [[CrossRef](#)] [[PubMed](#)]
97. Essigmann, B.; Güler, S.; Narang, R.A.; Linke, D.; Benning, C. Phosphate availability affects the thylakoid lipid composition and the expression of *SQD1*, a gene required for sulfolipid biosynthesis in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 1950–1955. [[CrossRef](#)] [[PubMed](#)]
98. Yu, B.; Xu, C.; Benning, C. Arabidopsis disrupted in *SQD2* encoding sulfolipid synthase is impaired in phosphate-limited growth. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5732–5737. [[CrossRef](#)] [[PubMed](#)]
99. Tabita, F.R. Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: A different perspective. *Photosynth. Res.* **1999**, *60*, 1–28. [[CrossRef](#)]
100. Delwiche, C.F.; Palmer, J.D. Rampant horizontal transfer and duplication of rubisco genes in eubacteria and plastids. *Mol. Biol. Evol.* **1996**, *13*, 873–882. [[CrossRef](#)]
101. Gross, J.; Meurer, J.; Bhattacharya, D. Evidence of a chimeric genome in the cyanobacterial ancestor of plastids. *BMC Evol. Biol.* **2008**, *8*, 117. [[CrossRef](#)]
102. Badger, M.R.; Bek, E.J. Multiple Rubisco forms in proteobacteria: Their functional significance in relation to CO<sub>2</sub> acquisition by the CBB cycle. *J. Exp. Bot.* **2008**, *59*, 1525–1541. [[CrossRef](#)] [[PubMed](#)]
103. Cannon, G.C.; Heinhorst, S.; Kerfeld, C.A. Carboxysomal carbonic anhydrase: Structure and role in microbial CO<sub>2</sub> fixation. *Biochim. Biophys. Acta* **2010**, *1804*, 382–392. [[CrossRef](#)] [[PubMed](#)]
104. Fong, A.; Archibald, J.M. Evolutionary dynamics of light-independent protochlorophyllide oxidoreductase genes in the secondary plastids of cryptophyte algae. *Eukaryot. Cell* **2008**, *7*, 550–553. [[CrossRef](#)] [[PubMed](#)]
105. van der Staay, G.W.M.; Moon-van der Staay, S.Y.; Garczarek, L.; Partensky, F. Rapid evolutionary divergence of Photosystem I core subunits PsaA and PsaB in the marine prokaryote *Prochlorococcus*. *Photosynth. Res.* **2000**, *65*, 131–139. [[CrossRef](#)] [[PubMed](#)]
106. Qiu, H.; Yang, E.C.; Bhattacharya, D.; Yoon, H.S. Ancient gene paralogy may mislead inference of plastid phylogeny. *Mol. Biol. Evol.* **2012**, *29*, 3333–3343. [[CrossRef](#)]
107. Moriyama, T.; Terasawa, K.; Sato, N. Conservation of POPs, the plant organellar DNA polymerases, in eukaryotes. *Protist* **2011**, *162*, 177–187. [[CrossRef](#)] [[PubMed](#)]
108. Moriyama, T.; Sato, N. Enzymes involved in organellar DNA replication in photosynthetic eukaryotes. *Front. Plant Sci.* **2014**, *5*, 480. [[CrossRef](#)]
109. Moriyama, T.; Tajima, N.; Sekine, K.; Sato, N. Localization and phylogenetic analysis of enzymes related to organellar genome replication in the unicellular rhodophyte *Cyanidioschyzon merolae*. *Genome Biol. Evol.* **2014**, *6*, 228–237. [[CrossRef](#)]
110. Sato, N. Was the evolution of plastid genetic machinery discontinuous? *Trends Plant Sci.* **2001**, *6*, 151–156. [[CrossRef](#)]
111. Miyagishima, S.Y.; Nakanishi, H.; Kabeya, Y. Structure, regulation, and evolution of the plastid division machinery. *Int. Rev. Cell Mol. Biol.* **2011**, *291*, 115–153. [[CrossRef](#)]
112. Yoshida, Y. Insights into the Mechanisms of Chloroplast Division. *Int. J. Mol. Sci.* **2018**, *19*, 733. [[CrossRef](#)] [[PubMed](#)]
113. Fujiwara, M.T.; Sato, N. Plastid replication in Arabidopsis: Complexity of the molecular components for the control of division. *Recent Res. Devel. Plant Sci.* **2004**, *2*, 219–248.
114. Martin, W.; Schnarrenberger, C. The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: A case study of functional redundancy in ancient pathways through endosymbiosis. *Curr. Genet.* **1997**, *32*, 1–18. [[CrossRef](#)] [[PubMed](#)]
115. Reyes-Prieto, A.; Moustafa, A. Plastid-localized amino acid biosynthetic pathways of Plantae are predominantly composed of non-cyanobacterial enzymes. *Sci. Rep.* **2012**, *2*, 955. [[CrossRef](#)]
116. Richardson, L.G.L.; Schnell, D.J. Origins, function, and regulation of the TOC–TIC general protein import machinery of plastids. *J. Exp. Bot.* **2020**, *71*, 1226–1238. [[CrossRef](#)]
117. Gagat, P.; Bodył, A.; Mackiewicz, P. How protein targeting to primary plastids via the endomembrane system could have evolved? A new hypothesis based on phylogenetic studies. *Biol. Direct.* **2013**, *8*, 18. [[CrossRef](#)]
118. Gross, J.; Bhattacharya, D. Reevaluating the evolution of the Toc and Tic protein translocons. *Trends Plant Sci.* **2009**, *14*, 13–20. [[CrossRef](#)]
119. Chotewutmontri, P.; Holbrook, K.; Bruce, B.D. Plastid protein targeting: Preprotein recognition and translocation. *Int. Rev. Cell Mol. Biol.* **2017**, *330*, 227–294. [[CrossRef](#)] [[PubMed](#)]
120. Lauterborn, R. Protozoenstudien. II. Paulinella chromatophora nov. gen. nov. spec., ein beschalter Rhizopode des Sü.wassers mit blaugrünen chromatophorenartigen Einschlüssen. *Z. Wiss. Zool.* **1895**, *59*, 537–544.

121. Pascher, A. Über die Natur der blaugrünen Chromatophoren des Rhizopoden *Paulinella chromatophora*. *Zool. Anz.* **1929**, *81*, 189–194.
122. Marin, B.; Nowack, E.C.; Melkonian, M. A plastid in the making: Evidence for a second primary endosymbiosis. *Protist* **2005**, *156*, 425–432. [[CrossRef](#)]
123. Maréchal, E. Primary Endosymbiosis: Emergence of the Primary Chloroplast and the Chromatophore, Two Independent Events. Chapter 1. In *Plastids: Methods and Protocols*; Methods in Molecular Biology; Maréchal, E., Ed.; Humana Press: New York, NY, USA, 2018; Volume 1829, pp. 3–16.
124. Delaye, L.; Valadez-Cano, C.; Pérez-Zamorano, B. How really ancient is *Paulinella chromatophora*? *PLoS Curr.* **2016**, *8*. [[CrossRef](#)] [[PubMed](#)]
125. Parfrey, L.W.; Lahr, D.J.; Knoll, A.H.; Katz, L.A. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 13624–13629. [[CrossRef](#)]
126. Eme, L.; Sharpe, S.C.; Brown, M.W.; Roger, A.J. On the age of eukaryotes: Evaluating evidence from fossils and molecular clocks. *Cold Spring Harb. Perspect. Biol.* **2014**, *6*, 8. [[CrossRef](#)]
127. Härtel, H.; Benning, C. Can digalactosyldiacylglycerol substitute for phosphatidylcholine upon phosphate deprivation in leaves and roots of Arabidopsis? *Biochem. Soc. Trans.* **2000**, *28*, 729–732. [[CrossRef](#)]
128. Andersson, M.X.; Stridh, M.H.; Larsson, K.E.; Liljenberg, C.; Sandelius, A.S. Phosphate-deficient oat replaces a major portion of the plasma membrane phospholipids with the galactolipid digalactosyldiacylglycerol. *FEBS Lett.* **2003**, *537*, 128–132. [[CrossRef](#)]
129. Jouhet, J.; Maréchal, E.; Baldan, B.; Bligny, R.; Joyard, J.; Block, M.A. Phosphate deprivation induces transfer of DGDG galactolipid from chloroplast to mitochondria. *J. Cell Biol.* **2004**, *167*, 863–874. [[CrossRef](#)]
130. Russo, M.A.; Quartacci, M.F.; Izzo, R.; Belligno, A.; Navari-Izzo, F. Long- and short-term phosphate deprivation in bean roots: Plasma membrane lipid alterations and transient stimulation of phospholipases. *Phytochemistry* **2007**, *68*, 1564–1571. [[CrossRef](#)] [[PubMed](#)]
131. Botté, C.Y.; Deligny, M.; Rocchia, A.; Bonneau, A.-L.; Saïdani, N.; Hardré, H.; Aci, S.; Yamaryo-Botté, Y.; Jouhet, J.; Dubots, E.; et al. Chemical inhibitors of monogalactosyldiacylglycerol synthases in Arabidopsis thaliana. *Nat. Chem. Biol.* **2011**, *7*, 834–842. [[CrossRef](#)]
132. Kobayashi, K.; Fujii, S.; Sato, M.; Toyooka, K.; Wada, H. Specific role of phosphatidylglycerol and functional overlaps with other thylakoid lipids in Arabidopsis chloroplast biogenesis. *Plant Cell Rep.* **2015**, *34*, 631–642. [[CrossRef](#)]
133. Semeniuk, A.; Sohlenkamp, C.; Duda, K.; Hölzl, G.A. bifunctional glycosyltransferase from *Agrobacterium tumefaciens* synthesizes monoglucosyl and glucuronosyl diacylglycerol under phosphate deprivation. *J. Biol. Chem.* **2014**, *289*, 10104–10114. [[CrossRef](#)]
134. Geiger, O.; Röhrs, V.; Weissenmayer, B.; Finan, T.M.; Thomas-Oates, J.E. The regulator gene *phoB* mediates phosphate stress-controlled synthesis of the membrane lipid diacylglycerol-*N,N,N*-trimethylhomoserine in *Rhizobium (Sinorhizobium) meliloti*. *Mol. Microbiol.* **1999**, *32*, 63–73. [[CrossRef](#)]
135. Riekhof, W.R.; Naik, S.; Bertrand, H.; Benning, C.; Voelker, D.R. Phosphate starvation in fungi induces the replacement of phosphatidylcholine with the phosphorus-free betaine lipid diacylglycerol-*N,N,N*-trimethylhomoserine. *Eukaryot. Cell* **2014**, *13*, 749–757. [[CrossRef](#)]
136. Sato, N. Betaine lipids. *Bot. Mag. Tokyo* **1992**, *105*, 185–197. [[CrossRef](#)]
137. Tajima, N.; Sato, S.; Maruyama, F.; Kurokawa, K.; Ohta, H.; Tabata, S.; Sekine, K.; Moriyama, T.; Sato, N. Analysis of the complete plastid genome of the unicellular red alga *Porphyridium purpureum*. *J. Plant Res.* **2014**, *127*, 389–397. [[CrossRef](#)]
138. Ku, C.; Nelson-Sathi, S.; Roettger, M.; Garg, S.; Hazkani-Covo, E.; Martin, W.F. Endosymbiotic gene transfer from prokaryotic pangenomes: Inherited chimerism in eukaryotes. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 10139–10146. [[CrossRef](#)] [[PubMed](#)]
139. Schötz, F. Dreidimensionale, maßstabgetreue Rekonstruktion einer grünen Flagellatenzelle nach Elektronenmikroskopie von Serienschritten. *Planta* **1972**, *102*, 152–159. [[CrossRef](#)] [[PubMed](#)]
140. Moriyama, T.; Toyoshima, M.; Saito, M.; Wada, H.; Sato, N. Revisiting the algal “chloroplast lipid droplet”: The absence of an entity that is unlikely to exist. *Plant Physiol.* **2018**, *176*, 1519–1530. [[CrossRef](#)] [[PubMed](#)]
141. Oi, T.; Enomoto, S.; Nakao, T.; Arai, S.; Yamane, K.; Taniguchi, M. Three-dimensional intracellular structure of a whole rice mesophyll cell observed with FIB-SEM. *Ann. Bot.* **2017**, *120*, 21–28. [[CrossRef](#)] [[PubMed](#)]
142. Yamane, K.; Oi, T.; Enomoto, S.; Nakao, T.; Arai, S.; Miyake, H.; Taniguchi, M. Three-dimensional ultrastructure of chloroplast pockets formed under salinity stress. *Plant Cell Environ.* **2018**, *41*, 563–575. [[CrossRef](#)]
143. Kobayashi, K.; Masuda, T.; Tajima, N.; Wada, H.; Sato, N. Molecular phylogeny and intricate evolutionary history of the three isofunctional enzymes involved in the oxidation of Protoporphyrinogen IX. *Genome. Biol. Evol.* **2014**, *6*, 2141–2155. [[CrossRef](#)] [[PubMed](#)]
144. Sato, N.; Moriyama, T. *Photosynthesis*. Chapter 17. In *Cyanidioschyzon merolae. A New Model Eukaryote for Cell and Organelle Biology*; Kuroiwa, T., Miyagishima, S., Matsunaga, S., Sato, N., Nozaki, H., Tanaka, K., Misumi, O., Eds.; Springer: Singapore, 2018; pp. 263–281.