Photosynthetic eukaryotes unite: endosymbiosis connects the dots

Debashish Bhattacharya,* Hwan Su Yoon, and Jeremiah D. Hackett

Summary

The photosynthetic organelle of algae and plants (the plastid) traces its origin to a primary endosymbiotic event in which a previously non-photosynthetic protist engulfed and enslaved a cyanobacterium. This eukaryote then gave rise to the red, green and glaucophyte algae. However, many algal lineages, such as the chlorophyll *c*-containing chromists, have a more complicated evolutionary history involving a secondary endosymbiotic event, in which a protist engulfed an existing eukaryotic alga (in this case, a red alga). Chromists such as diatoms and kelps then rose to great importance in aquatic habitats. Another algal group, the dinoflagellates, has undergone tertiary (engulfment of a secondary plastid) and even quaternary endosymbioses. In this review, we examine algal diversity and show endosymbiosis to be a major force in algal evolution. This area of research has advanced rapidly and long-standing issues such as the chromalveolate hypothesis and the extent of endosymbiotic gene transfer have recently been clarified.

***BioEssays* 26:50–60, 2004.**

Introduction

Accounting for the remarkable diversity of photosynthetic protists (algae) remains an important challenge in evolutionary biology. The photosynthetic organelle, termed the plastid, is found in forms as diverse as microscopic diatoms, giant kelps, toxic dinoflagellates, edible red algae, and land plants (see Table 1). A critical question is whether all photosynthetic forms are united on a single branch of the tree of life or whether plastids have been spread throughout the tree through multiple independent lateral transfers. If the latter, then how many lateral transfers are needed to explain extant algal lineages? Answers to these questions have been rapidly accumulating recently and they provide surprising insights into the course of eukaryotic evolution.

It has now been clearly documented that plastids originate through endosymbiosis,[1–6] whereby a single-celled protist engulfs and retains a foreign photosynthetic cell (see Fig. 1). Over time, the foreign cell is reduced to a plastid and is vertically transmitted to subsequent generations. Endosymbiosis comes in three major types, primary, secondary and tertiary endosymbiosis. The first results from the engulfment of a photosynthetic prokaryote (cyanobacterium, Fig. 1A) and gives rise to a plastid bound by two membranes (the inner, 1st, and outer, 2nd, membranes of the cyanobacterium, see Ref. 7). The outer, phagosomal membrane of this primary "host" cell is lost. Primary endosymbiosis is believed to have occurred once in evolution, giving rise to the proto-alga that is the ultimate root of all plastids.[4] The plastid became fully established in the proto-algal population and many genes of both photosynthetic and non-photosynthetic function were transferred to the nucleus of the primary host. The remarkable achievement of these cells was to engineer a way of reintroducing all the plastid proteins, whose genes were transferred to the nucleus and translated in the cytoplasm, back into the plastid compartment to express their function. The establishment of this protein import system after primary endosymbiosis required the evolution of a "transit" sequence (about 24–100 amino acids in length)[8] at their N termini. This extra sequence probably originated through mutations that extended the original open reading frame at the 5′ terminus, and the encoded proteins were selected for because of their targeting capacity. Once the sequences encoding these transit peptides were established in a few genes, then they may have spread through exon shuffling into other photosynthetic genes.[9,10] It appears that, once these unlikely events had occurred and the first algal populations were established, the path was paved for the rise of photosynthetic eukaryotes.

Department of Biological Sciences and Center for Comparative Genomics, University of Iowa, Iowa City

Funding agency: This work was supported by grants by the National Science Foundation. (NSF) awarded to D. B. (DEB 01-07754, MCB 01-10252, MCB 02-36631). J. D. H. was supported by an Institutional NRSA (T 32 GM98629) from the National Institutes of Health and H. S. Y. was partially supported by the Post Doctoral Fellows Program of the Korean Science and Engineering Foundation.

*Correspondence to: Debashish Bhattacharya, Department of Biological Sciences and Center for Comparative Genomics, University of Iowa, 210 Biology Building, Iowa City, Iowa 52242-1324.
E-mail: dbhattac@blue.weeg.uiowa.edu

DOI 10.1002/bies.10376
Published online in Wiley InterScience (www.interscience.wiley.com).

Abbreviations: COXII, mitochondrial cytochrome c oxidase subunit II; *gnd*, 6-phosphogluconate dehydrogenase gene; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; Ma, millions of years ago; rRNA, ribosomal RNA.

Table 1. A list of the different types of algae and related protists, their classification, and key characteristics with an emphasis on plastid features*

Algal group	# Species	Plastid		Well - Known Representatives	Distinctive Features	Tree of Life Classification
		Origin	Mem.			
Red algae (Rhodophyta)	5000–6000	Primary - CB	2	*Porphyra*, *Gracilaria*, nori	Lack of flagellate stages, widely cultivated, used to make sushi, and source of agar and carrageenan. (http://www.ucmp.berkeley.edu/protista/rhodophyta.html)	Plantae
Green algae/land plants (Viridiplantae)	>500,000	Primary - CB	2	*Chlamydomonas*, flowering plants	The green algae are common in intertidal regions and are the sister to land plants that are dominant in the terrestrial ecosystem. (http://www.ucmp.berkeley.edu/greenalgae/greenalgae.html)	Plantae
Glaucophytes (Glaucocystophyceae)	13	Primary - CB	2	*Cyanophora*	A small group of algae best known for their primary plastid (cyanelle) that retains the ancestral peptidiglycan wall of cyanobacterial origin between its two plastid membranes. (http://protist.i.hosei.ac.jp/taxonomy/Others/Glaucophyta.html)	Plantae
Cryptophytes (Cryptophyta)	200	Secondary - RA	4	*Guillardia*	Cold water-loving algae that retain the nucleomorph of the red algal secondary endosymbiont. (http://www.uni-koeln.de/math-nat-fak/botanik/bot1/Phylogenie-Webseite/cryptophyta.html)	**Chromista**
Haptophytes (Haptophyceae)	500	Secondary - RA	4	*Emiliania*, coccolithophore	Bloom-forming algae that can form coccolith scales that are a global carbon sink in oceans. (http://www.soes.soton.ac.uk/staff/tt/)	**Chromista**
Stramenopiles	>10,000	Secondary - RA	4	Kelps, diatoms, water molds, downy mildews, late blight of potato	A diverse group of both unicellular and macroscopic protists that includes plastid-containing (often called chromophyte algae) and plastid-less forms (oomycetes) that have presumably lost the organelle. (http://www.ucmp.berkeley.edu/chromista/chromista.html)	**Chromista**
Dinoflagellates (Dinophyceae)	>4,000	Secondary - RA Tertiary - HA*	3	*Alexandrium*, *Karenia*, *Pfiesteria*, red tide	Predominantly unicellular algae that are often mixotrophic and contain a diversity of different plastids. They cause harmful algal blooms such as toxic red tides and are a serious health threat (e.g., paralytic and diarrhetic shellfish poisoning) in coastal waters. (http://www.ucmp.berkeley.edu/protista/dinoflagellata.html)	**Alveolata**
Apicomplexans (Apicomplexa)	>2,400	Secondary - RA*	4	*Plasmodium*, *Toxoplasma*, *Cryptosporidium*	Non-photosynthetic, obligate intracellular parasites that are causes of serious human diseases such as malaria and toxoplasmosis. Their remnant plastid (apicoplast) contains a reduced (35 kb) genome. (http://www.ucmp.berkeley.edu/protista/apicomplexa.html)	**Alveolata**
Ciliates (Ciliophora)	>7,500	Loss	—	*Paramecium*, *Tetrahymena*	Plastid-less protists whose cell bodies are covered with cilia, often fused together in rows or tufts called cirri. The ciliate nucleus is differentiated into macro- and micronuclei. (http://www.ucmp.berkeley.edu/protista/ciliata.html)	**Alveolata**
Chlorarachniophytes (Chlorarachniophyceae)	6	Secondary - GA	4	*Chlorarachnion*	Unicellular marine amoebae containing a green plastid. The plastid retains the remnant nucleus (nucleomorph) of the green algal secondary endosymbiont. (http://tolweb.org/tree?group=Chlorarachniophytes&contgroup=The_Other_Protists)	Cercozoa

Table 1. (*Continued*)

Algal group	# Species	Plastid: Origin	Plastid: Mem.	Well - Known Representatives	Distinctive Features	Tree of Life Classification
Euglenids (Euglenida)	>800	Secondary - GA	3	*Euglena*, *Astasia*	Algae that are common in organic-rich freshwater habitats. They contain a green plastid of secondary endosymbiotic origin and are known for their distinctive flowing movement (metaboly) and cell surface (pellicle). Euglenids are closely related to the parasitic trypanosomes. (http://bio.rutgers.edu/euglena/)	Euglenozoa

*Mem. is for the number of membranes surrounding plastids, CB is cyanobacterium, RA is red alga, GA is green alga, and HA is haptophyte alga. The taxonomic classification of the algal groups follows NCBI (http://www.ncbi.nlm.nih.gov/), whereas the Tree of Life classifications indicates their position in the eukaryotic tree and follows different authors (see text for details) - the groups shown in bold are the chromalveolates and share a single red algal secondary endosymbiont, whereas the underlined groups may share a single green algal endosymbiont and are termed cabozoans. The asterisks indicate cases of possible plastid replacements. The web sites are good sources for additional information and/or images of the different algae.

Recent phylogenetic analyses using nuclear and mitochondrial loci[11–13] suggest that the proto-alga split into two lineages. The first contains the glaucophyte algae, which failed to rise to any great taxonomic importance,[14,15] whereas the second gave rise to the highly successful red algae[16] and their sister group the green algae[17] plus land plants.[18] These three lineages are classified as the Plantae (see Table 1). Endosymbiosis left a sizeable mark on the Plantae that goes well beyond the lateral transfer of photosynthetic capacity. A recent analysis[19] of the complete *Arabidopsis* nuclear genome suggests that up to 18% of this plant's genes, many of non-photosynthetic function (e.g., disease resistance, intracellular protein routing), originated from the cyanobacterium through endosymbiotic gene transfer. Endosymbiosis had, therefore, a considerable influence in the early evolution of algae by significantly enriching their nuclear genomes with cyanobacterial, often duplicated genes. Selection could act on these divergent sequences to explore new functions[13,19] or to replace existing host genes with those from the endosymbiont.[20,21]

Secondary endosymbiosis and the rise of algae

Once the Plantae had been established (Fig. 1A), the stage was set for secondary endosymbiosis, whereby a protist engulfed an existing alga (Fig. 1B).[22,23] This type of eukaryote–eukaryote endosymbiosis explains the vast majority of algal diversity. Secondary plastids are found in algae containing chlorophyll *c* (chromophytic algae)[24] and non-green algal/plant forms containing chlorophyll *b* (euglenids and chlorarachniophytes), and are unambiguously identified by the presence of three or four bounding membranes (Table 1). The extra membranes, from inside out, are thought to be the plasma membrane of the engulfed alga (3rd) and the phagosomal membrane of the host cell of the secondary plastid (4th).[25] The 4th membrane has presumably been lost in some groups (e.g., euglenids and most dinoflagellates, see below) and two retain a smoking gun of eukaryotic enslavement, the remnant nucleus of the algal symbiont. This reduced genome, the nucleomorph, is found between the 2nd and 3rd plastid membranes in the periplastid space (the former cytoplasm of the algal endosymbiont) and has been sequenced in a cryptophyte and a chlorarachniophyte (see Fig. 1B). Phylogenies using nucleomorph genes reveal the identity of the algal symbiont (red in cryptophytes and green in chlorarachniophytes).[26–28] Interestingly, both reduced genomes have converged on a similar size (551 kb for cryptophytes, 380 kb for chlorarachniophytes), set of retained genes, and a compact organization.[29,30]

A final important marker of secondary endosymbiosis is the presence of an extended N terminus in nuclear-encoded plastid proteins that encodes a bipartite targeting sequence. The first extension is a signal sequence that leads the plastid-bound proteins to the secretory pathway via the endoplasmic reticulum-derived outer membrane of these plastids, whereas the second sequence is the typical transit peptide used to cross the inner two plastid membranes.[8,25] Therefore, like primary endosymbiosis, secondary endosymbiosis also required at least one significant evolutionary innovation to make the process work: the independent origin of a signal sequence that is N-terminal of the existing transit peptide for each nuclear-encoded plastid protein to ensure correct passage across the three to four membranes.

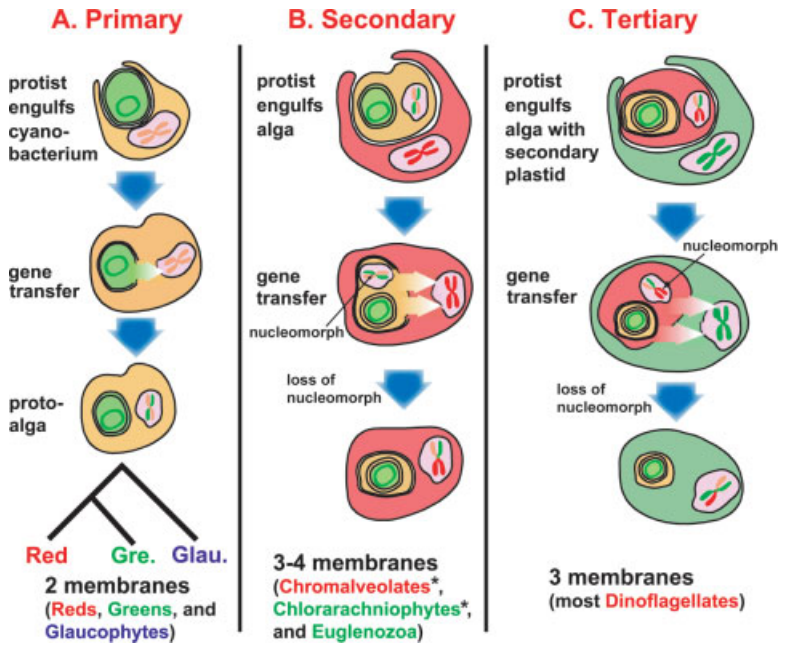

Figure 1. Origins of plastids via **A:** primary, **B:** secondary, and **C:** tertiary endosymbiosis. Gene transfer from the endosymbiont to the host nucleus is shown with the arrows and the colored chromosomes. The mitochondrion has been omitted from these figures. The remnant algal nucleus (nucleomorph) in secondary and tertiary endosymbioses is shown. This genome has been lost in all algae but the cryptophytes (member of the chromalveolates) and chlorarachniophytes (marked with asterisks).

The secondary plastids in brief

The chromalveolate hypothesis

The chromalveolates group was postulated primarily on the basis of molecular phylogenetic analyses that unite particular members of these morphologically disparate lineages,[32] and the hypothesis that all taxa containing chlorophyll *c* (i.e., a chromophytic plastid)[33,34] share a common origin. The alveolates share two characters that unambiguously unite them as a lineage, namely tubular mitochondrial cristae and sacs or alveoli under the plasma membrane. The ecologically and economically important chromalveolates (see Table 1)[35] defines a broadly diverse group that includes the Chromista, comprising the cryptophyte,[36] haptophyte,[37] and stramenopile[16] algae, and the Alveolata, comprising the parasitic apicomplexans,[38] the plastid-less ciliates,[39] and the dinoflagellate algae.[40] The chromalveolate plastid is believed to have originated from a red algal secondary endosymbiosis with the ensuing evolution of chlorophyll c_2. The plastid was putatively lost in ciliates and parasitic/saprobic stramenopiles like oomycetes (e.g., the water mold *Achlya*) and its genome reduced to a 35 kb DNA circle in the apicomplexan plastid (the apicoplast).[41,42] Evidence in support of these ideas has been slow in coming but recent data now appear to have established the chromalveolates as a monophyletic entity.[42,43]

Phylogenies of nuclear genes have until now, however, only provided marginal support for the chromalveolate hypothesis. Analysis of small subunit rRNA[44] and a combined data set of EF-1α, actin, α-tubulin, and β-tubulin amino acid sequences (that did not include haptophytes)[45] are consistent with a sister group relationship between two chromalveolate groups, the stramenopiles and alveolates. The limited taxon sampling and phylogenetic power of these data sets are a weakness that needs to be addressed in future studies. In contrast, strong support for chromalveolate monophyly has come from analyses of plastid genes[43] and from the finding of a unique gene duplication shared by members of this assemblage.[42]

We have recently analyzed a concatenated data set of five plastid-encoded genes (5827 nt in total) from 36 taxonomically diverse members of the red and chromist algae. In this study, we addressed one cornerstone of the chromalveolate hypothesis that had not yet been adequately tested using multiple nuclear loci, the monophyly of the Chromista (see Table 1).[33] This group previously had little justification on morphological[24] or phylogenetic grounds (nuclear and mitochondrial loci do not resolve their positions)[4,11,13] for being united in a single lineage. They primarily share a four-membrane-bound chlorophyll *c*-containing plastid that is located within the lumen of the rough endoplasmic reticulum. The haptophytes and stramenopiles, however, share characters such as tubular mitochondrial cristae, similar storage products, and fucoxanthin that suggest a specific relationship between these taxa.[24] The combined plastid gene tree (Fig. 2) appears to have settled the issue of chromist monophyly by providing

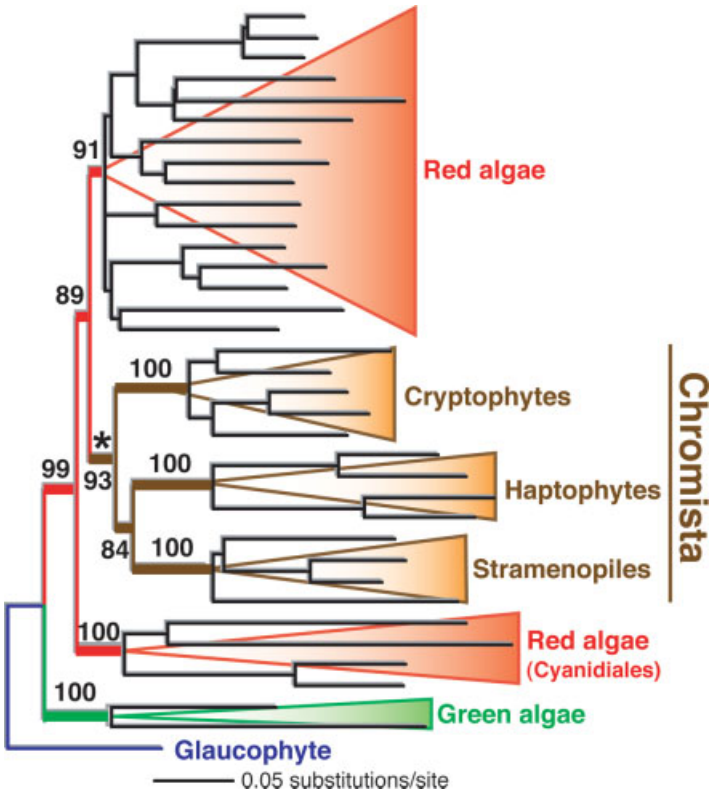

Figure 2. Phylogenetic relationships of red algal and chromist plastids inferred from a minimum evolution (ME) analysis of the combined DNA sequences of 16S rRNA, *psa*A, *psb*A, *rbc*L, and *tuf*A (5827 nt). The general time reversible model[79] with estimations of nucleotide frequencies, the proportion of invariant sites that are unable to accept substitutions, and the shape parameter of the gamma distribution to accommodate rate variations across sites (GTR + I + Γ [gtr] model) was used in this analysis. Results of a ME-gtr bootstrap analysis (2000 replicates) are shown above the branches. The thick branches are of significant evolutionary importance and all have >95% posterior probability in a Bayesian analysis of the DNA data (for details, see Yoon et al.).[43] This tree is rooted on the branch leading to the glaucophyte sequence. The branch lengths are proportional to the number of substitutions per site (see scale in figure). The asterisk marks the branch where the chromist plastid originated through a red algal secondary endosymbiosis (ca. 1260 Ma).[43]

strong support for three important ideas: (1) as previously suggested,[46] the chromist plastids are all of red algal origin, (2) chromist plastids share a monophyletic origin and, by extension, so do the host cells containing these plastids, and (3) the basal position of the cryptophytes in the Chromista suggests that retention of the red algal nucleomorph in the periplastid space, the presence of phycobilin pigments, and the storage of photosynthates as starch (all are absent in haptophytes and stramenopiles) are ancestral endosymbiont characters that were likely lost after the divergence of the cryptophytes (see character evolution model in Fig. 3).

Furthermore, in contrast to the haptophytes and stramenopiles that have tubular mitochondrial cristae, the cryptophytes have flattened cristae, suggesting that this was the ancestral condition in the Chromista. The alveolates, however, contain tubular cristae and are presumably sister to the chromists (Fig. 3). If flattened cristae was ancestral in the chromalveolates, then this character was lost twice, once in the common ancestor of alveolates and once after the divergence of the cryptophytes in the Chromista (see Fig. 3). A more parsimonious explanation for the cristae data would emerge if the cryptophytes were to branch at the base of the chromalveolates and the chromists would become paraphyletic. Analysis of an endosymbiotic replacement involving the enolase gene in cryptophytes is consistent with this idea.[20] Verification of the early divergence of the cryptophytes will have to wait, however, until alveolate plastid sequences are included in the analysis shown in Fig. 2. This analysis could have two likely outcomes: (1) alveolate photosynthetic genes of red algal origin diverge from the branch uniting all chromists (as expected = 2 losses of flattened cristae, or (2) they diverge on the branch uniting the haptophytes and stramenopiles (= 1 loss). Such target genes are most likely to be found in the nucleus of dinoflagellates because the non-photosynthetic apicomplexans and ciliates may not have maintained these coding regions.

Another important result from the Yoon et al.[43] study was the dating of the chromist secondary endosymbiotic event. Usage of two red algal fossil constraint dates and a molecular clock method that does not assume a uniform mutation rate in different lineages suggests that the earliest date (asterisk in Fig. 2) for the origin of the chromist plastid is 1260 ± 30 million years ago (Ma).[43] The endosymbiosis marks, therefore, the birth of the chromist algae. This date is substantially earlier than some estimates (e.g., 850 Ma)[31] but agrees well with the fossil record, which shows the appearance of a diversity of algae and protists near the Mesoproterozoic/Neoproterozoic boundary about 1000 Ma.[47,48] A recent analysis of the

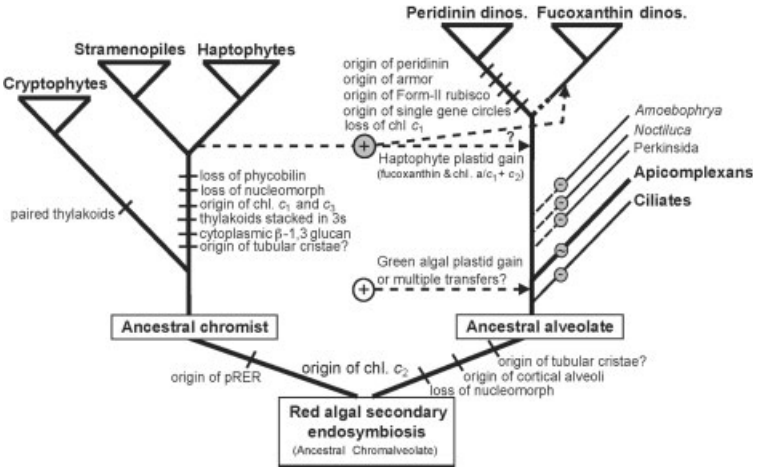

Figure 3. Proposed evolutionary tree of the chromalveolates showing the origin and loss of important (mostly plastid) characters. The evidence for chromist plastid monophyly is described in Yoon et al.[43] The timing of the tertiary endosymbiosis that gave rise to the haptophyte-type plastid in either the pre-dinoflagellate[66,80] or only in fucoxanthin-containing taxa[69] remains an open question (large filled circle with a cross). The possible green algal endosymbiosis in the common ancestor of apicomplexans and dinoflagellates (large open circle with a cross) is based on Funes et al.,[21,76] Kohler et al.,[75] and our preliminary results (see text for details). Gain or loss of character states is shown with the slashes across the branches. The branches with broken lines at the origin represent uncertainty about the position of these taxa in the host tree, whereas the small filled circles with dashes or the wavy line inside represent plastid loss and plastid degeneration, respectively. pRER is the plastid rough endoplasmic reticulum.

Earth's geochemical record, in particular with respect to the levels of the oxygen and nitrogen that are critical to algal growth, is consistent with the known eukaryotic fossil record[49] and our molecular dating results. Furthermore, a subsequent analysis (H.S.Y., J.D.H., and D.B. unpublished data) using a six plastid gene data set that includes green algae and land plants to incorporate additional robust fossil constraints in the gene phylogeny (e.g., origin of land plants and the monocot–dicot split) substantiates the findings of Yoon et al.[43]

Now that Chromista monophyly appears to be established, what evidence is there to include them in the chromalveolates? The strongest evidence comes from phylogenetic analyses that show a specific phylogenetic relationship between the nuclear-encoded plastid-targeted glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene in the cryptophytes and stramenopiles (chromists) on one hand and the apicomplexans and dinoflagellates (alveolates) on the other.[42] Importantly, this gene appears to have originated through a unique duplication of the existing cytosolic gene in these taxa, and the subsequent replacement of the original plastid-targeted sequence (of cyanobacterial origin) by the gene duplicate. This unusual shared derived character unites the studied chromalveolates and, as corroboration, the phylogenies show the plastid-targeted genes to be phylogenetically distinct from the homolog of cyanobacterial origin in non-chromalveolates such as green algae/land plants and red algae. These latter taxa have not undergone the gene duplication–replacement event.[42] These findings all point to a critical event in evolution, a single red algal secondary endosymbiosis that gave birth to common ancestor of a super-assemblage, the chromalveolates. The endosymbiotic event is ancient (ca. 1200 Ma) and the host cell of the secondary plastid has diverged extensively during this long period of evolution. It is a testimony to the power of modern molecular evolutionary methods and the slow evolutionary rate of many plastid genes that we are able to look back so far in evolutionary time and still detect the signal of shared ancestry (see Fig. 2).

The major insights into eukaryotic evolution that come from the finding of chromalveolate monophyly are:

- secondary plastid loss is common (e.g., in plastid-less stramenopiles such as oomycetes, in ciliates, and in some dinoflagellates).[42,50] In support of this idea, Andersson and Roger[51] have recently found a 6-phosphogluconate dehydrogenase (*gnd*) gene of cyanobacterial (i.e., plastid) origin in the parasitic stramenopile, *Phytophthora infestans*. This suggests that *Phytophthora* was likely once photosynthetic because its *gnd* gene is closely related to the homologue in photosynthetic members of this lineage.
- Chromalveolates share a homologous plastid protein import system that evolved once in their common ancestor.

- Chlorophyll c_2 appears to have evolved only once in the common ancestor of the chromalveolates. Phylogenetic analysis of chlorophyll *a/b* and chlorophyll *a/c* light harvesting complex proteins supports this idea showing that the chlorophyll *a/c*-binding proteins in chromists (and dinoflagellates) form a monophyletic group that traces its origin to a red algal-like ancestor.[46]

Given these impressive data, one may be led to believe that the story of chromalveolate evolution has been convincingly solved. This is not, however, entirely the case as the section on dinoflagellates below will explain. But first, we will review current knowledge about the secondary plastids of green algal origin.

The green secondary plastids

Secondary endosymbiosis also gave rise to the chlorophyll *b*-containing plant-like plastid in distinctly non-plant-like taxa, the euglenids[52] and the chlorarachniophytes. The Euglenida are sister to the Kinetoplastida (which includes the parasitic trypanosomatids), and together form the Euglenozoa. The presence of a three-membrane green plastid in photosynthetic euglenids and the absence of a plastid in the trypanosomes had previously suggested that the algal secondary endosymbiosis occurred at the base of the euglenid lineage and that trypanosomatids ancestrally lacked a plastid. The remarkable finding of plant-like genes in trypanosomatids now suggest, however, that all Euglenozoa may have at one time been photosynthetic and that kinetoplastids lost their photosynthetic organelle secondarily.[53] The footprint of the secondary endosymbiosis in the trypanosomatids is metabolic enzymes localized in specialized peroxisomes (glycosomes [not derived from plastids]) that trace their origin to the green alga through endosymbiotic gene transfer. This story underlines the surprising ancestral distribution of photosynthesis in eukaryotes and suggests that plastids may have been more widespread than we imagine. The distribution of these organelles has been decimated by secondary losses due often to the evolution of a parasitic or saprobic life-style (i.e., trypanosomatids were presumably once free-living algae),[54] a theme that was repeated in the chromalveolates.

The second group of non-green algal protists that contain chlorophyll *b* is the chlorarachniophytes. These are amoeboflagellate members of the protist assemblage Cercozoa[55] that includes euglyphids, formanifera, and plasmodiophorid plant pathogens.[56–58] This broad array of protists contains two photosynthetic groups, the chlorarachniophytes with their chlorophyll *b*-containing secondary plastid and the filose amoeba, *Paulinella chromatophora*, with its cyanelle that superficially resembles the plastids of glaucophytes.[59,60] Whereas the origin of the *Paulinella* cyanelle still needs to be resolved, the phylogenetic evidence based on plastid and nucleomorph sequences has established the green algal origin of the secondary plastid in chlorarachniophytes.[27,28] A recent analysis of cDNAs from the chloroarachniophyte *Bigelowiella natans* indicates that a large number of plastid genes in this species were transferred from the green algal endosymbiont to the host nucleus. More remarkably, about 21% of the photosynthetic genes were derived from lateral transfers involving non-green sources such as red or stramenopile algae.[61] These results underline the importance of endosymbiosis to facilitating intergenome gene transfer and suggest that mixotrophic species such as *Bigelowiella* (and dinoflagellates, see below) may be particularly adept at scavenging genes from different prey. In contrast, taxa such as the green alga *Chlamydomonas reinhardtii*, which is autotrophic and for which the complete nuclear genome is known, do not show evidence of lateral transfer of photosynthetic genes from different sources (i.e., all are derived from the original cyanobacterial primary endosymbiont).[61]

What still remains to be determined with regard to green algal secondary plastids is whether chlorarachniophytes and euglenids share the same green algal endosymbiont in a common ancestor. The available nuclear and plastid gene data do not support this hypothesis,[62–64] although the high divergence rate and base composition bias of chlorarachniophyte and euglenid sequences leaves open the possibility of Cercozoa–Euglenozoa monophyly (together the Cabozoa [Table 1]).[35] If true, this latter scenario would reduce the number of secondary endosymbioses involving a chlorophyll *b*-containing endosymbiont to one, render secondary plastid loss a widespread force in eukaryotic evolution, and place algae near the root of much of the tree of life.

The dinoflagellates and their tertiary endosymbioses

As if the genomic gymnastics described above were not enough, nature has devised yet another way of distributing plastids among eukaryotes, tertiary endosymbiosis. This process entails the engulfment of an alga with a secondary plastid (Fig. 1C). Tertiary endosymbiosis has, until now, been limited to the dinoflagellates and, in this group, involves the replacement of the existing red algal secondary plastid (shared with the chromists)[50,64–66] with another of secondary origin (so-called tertiary plastid replacement). The putative ancestral (and most common) plastid in this group is bound by three membranes, contains chlorophyll c_2 and the unique accessory pigment peridinin as the main carotenoid. The peridinin plastid does not, however, contain a typical genome because its genes have been reduced to single- or two-gene minicircles.[65,67,68] Until now, only 15 plastid protein genes have been found in peridinin dinoflagellates, leaving in question the location of the sequences that encode the remaining components of the photosynthetic apparatus. Although it has been presumed that the minicircle genes trace their origin to

the red algal secondary endosymbiosis in the chromalveolate ancestor[65], it is also possible that these sequences may have multiple origins from different endosymbiotic events (see below).

Given the ancestral origin of the red algal secondary plastid in dinoflagellates, tertiary replacement has been used to explain the plastid in taxa such as *Karenia* spp. and *Karlodinium micrum* that contain chlorophylls $c_1 + c_2$ and 19′-hexanoyloxy-fucoxanthin and/or 19′-butanoyloxy-fucoxanthin but lack peridinin, similar to the haptophyte algae.[66,69] Tertiary endosymbiosis also would account for plastid origin in other dinoflagellates such as *Dinophysis* spp. (cryptophyte origin),[70] *Peridinium foliaceum* (stramenopile origin),[71] and *Lepidodinium viride* (green algal origin).[72] Give these complex series of events, understanding the impact of endosymbiosis on dinoflagellate evolution will require a genomics approach. In this regard, we are currently generating a data set of up to 10,000 unique 3′ reads from normalized and subtracted cDNA libraries of the toxic dinoflagellate *Alexandrium tamarense*. These data will help us quantify the genomic contribution of multiple endosymbioses to dinoflagellate evolution and, as we and others[73] predict, likely identify the missing dinoflagellate plastid genes as components of the *Alexandrium* nuclear genome. Preliminary data from the *Alexandrium* cDNA data set have in fact confirmed the presence of many plastid genes in the nucleus of this dinoflagellate (J.D.H., H.S.Y., and D.B. unpublished results).

Conclusions

Endosymbiosis has created a plethora of photosynthetic eukaryotes that have significantly shaped the Earth's history. The known fossil record and our molecular dating analysis suggest that photosynthetic eukaryotes have been around for well over a billion years. The super-assemblage, chromalveolates, appears to be a monophyletic group and the ancient red algal secondary endosymbiosis that we have recently described[43] and the GAPDH duplication found by Fast et al.[42] unite them as a cohesive lineage. The analysis of multi-gene nuclear and mitochondrial data sets that accommodate a broad taxonomic sampling should verify this hypothesis in the coming years.[74] The present state of knowledge regarding plastid and host relationships among algae is summarized in Fig. 4. The plastid sequence data have significantly added to our understanding of the algal tree of life both within lineages

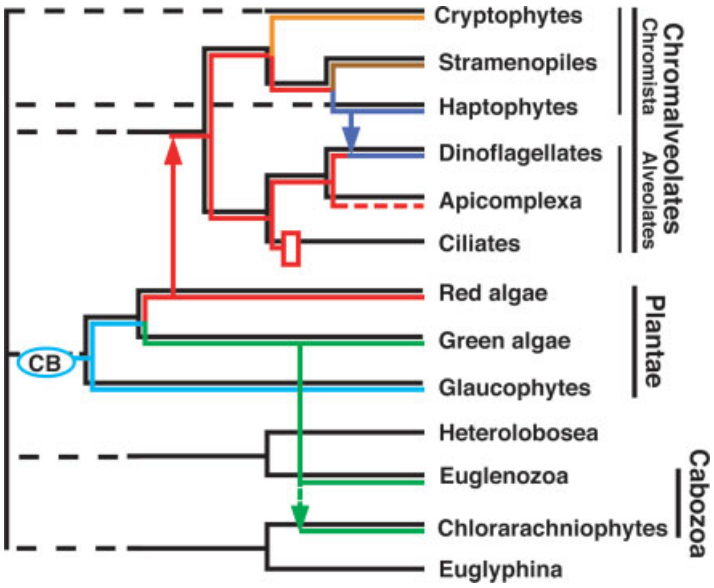

Figure 4. Present understanding of the phylogenetic relationships of algae. The host relationships that have resulted from analyses of nuclear or mitochondrial loci are shown with the black line. The branches with broken lines represent uncertainty about the position of these taxa in the host tree. Phylogenetic relationships inferred from plastid gene comparisons are shown with the colored lines: red for the red algal secondary endosymbiosis in chromalveolates, green for the putative green algal secondary endosymbiosis in cabozoans, and blue for the haptophyte tertiary replacement either in the dinoflagellate common ancestor or only in fucoxanthin-containing taxa. Cabozoan monophyly has not been substantiated with molecular phylogenetic data therefore the line leading from the Euglenozoa to the chlorarachniophytes is dashed. Cases of potential gene transfer from green algae in the common ancestor of apicomplexans and dinoflagellates are not shown in this figure. Plastid loss in the ciliates is shown with the open red box and plastid degeneration in the apicomplexans is shown with the dashed red line. CB denotes the primary endosymbiosis that gave rise to the proto-algal ancestor of the red, green, and glaucophyte algae.

and through the union of groups such as the Chromista that had yet to be proven using other loci.

Perhaps most exciting is the mounting evidence[21,75] for a green algal influence in the alveolates. Kohler et al.[75] originally suggested a green algal origin of the apicoplast based on phylogenetic analyses of the highly divergent *tuf*A gene. This relatively weakly supported result was recently buttressed by the findings of Funes et al.[21] who identified a unique nuclear-encoded *cox2* (mitochondrial cytochrome c oxidase subunit II gene) in apicomplexans and in some green algae that is split into two coding regions (*cox*2a, cox2b). Phylogenetic analysis supports the monophyly of the COXIIA + COXIIB sequence in greens and apicomplexans, implying a single origin of these split genes in a green alga, and its lateral transfer into the apicomplexans presumably after endosymbiosis.[21,64] Funes et al.[76] and Waller et al.[77] have recently provided differing interpretations of the COXII data and its support for a green algal endosymbiosis in the apicomplexans. Interestingly, analysis of our initial *Alexandrium* cDNA set shows the presence of multiple green algal genes in this taxon that is sister to apicomplexans (J.D.H., H.S.Y., and D.B. unpublished results), supporting the findings of Funes et al.[21,76] In our estimation, these exciting results are either due to multiple gene transfers from different green algae or a "hidden" green algal endosymbiosis that occurred at least in the common ancestor of apicomplexans and dinoflagellates. Addition of other completed genome sequences (e.g., *Plasmodium falciparum*)[78] to the analyses should provide the data to test this idea. We also predict that genomics projects will, in the near future, turn up other examples of gene transfer in chromalveolates and in other protists that have had a photosynthetic ancestry. Groups that contain, or at one time contained, a secondary or tertiary plastid will likely encode nuclear genes that have arisen through lateral gene transfer from the eukaryotic endosymbiont(s), much like the case found for the genes of cyanobacterial origin in *Arabidopsis*.(19)

These events may best be interpreted under the "Russian Doll Paradigm" in which each round of endosymbiosis leads to large-scale gene transfer that significantly reshapes the nuclear genome of the host. This process, schematically portrayed in the nuclear chromosomes shown in Fig. 1, is akin to nested Russian dolls and is predicted to be a driving force in generating eukaryotic biodiversity by providing a major influx of novel genes into protistan genomes.

Acknowledgments

Figure 1 is based on an image created by NSF graphic artist Kirk D. Woellert.

References

1. Mereschkowsky C. Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. Biol Zentralbl 1905;25:593–604.
2. Margulis L. Origin of Eucaryotic Cells. New Haven: Yale University Press; 1970. 349 pp.
3. Gray MW. The endosymbiont hypothesis revisited. Int Rev Cytol 1992; 141:233–357.
4. Bhattacharya D, Medlin L. The phylogeny of plastids: A review based on comparisons of small subunit ribosomal RNA coding regions. J Phycol 1995;31:489–498.
5. Delwiche CF. Tracing the thread of plastid diversity through the tapestry of life. Am Nat 1999;154(S4):S164–S177.
6. McFadden GI. Primary and secondary endosymbiosis and the origin of plastids. J Phycol 2001;37:951–959.
7. Reumann S, Davila-Aponte J, Keegstra K. The evolutionary origin of the protein-translocating channel of chloroplastic envelope membranes: identification of a cyanobacterial homolog. Proc Natl Acad Sci USA 1999; 96:784–789.
8. Foth BJ, Ralph SA, Tonkin CJ, Struck NS, Fraunholz M, Roos DS, Cowman AF, McFadden GI. Dissecting apicoplast targeting in the malaria parasite *Plasmodium falciparum*. Science 2003;299:705–708.
9. Martin W, Herrmann RG. Gene transfer from organelles to the nucleus: how much, what happens, and why? Plant Physiol 1998;118:9–17.
10. Blanchard JL, Lynch M. Organellar genes: why do they end up in the nucleus? Trends Genet 2000;16:315–320.
11. Gray MW, et al. Genome structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res 1998;26:865–878.
12. Moreira D, Le Guyader H, Philippe H. The origin of red algae and the evolution of chloroplasts. Nature 2000;405:69–72.
13. Stibitz T, Keeling P, Bhattacharya D. Symbiotic origin of a novel actin gene in the cryptophyte *Pyrenomonas helgolandii*. Mol Biol Evol 2000; 17:1731–1738.
14. Kies L. Zur systematischen Einordnung von *Cyanophora paradoxa*, *Gloeochaete wittrockiana* und *Glaucocystis nostochinearum*. Ber Deutsch Bot Ges Bd 1979;92:445–454.
15. Kies L, Kremer BP. Typification of the Glaucocystophyta. Taxon 1986; 35:128–133.
16. Graham LD, Wilcox LW. Algae. New Jersey: Prentice-Hall Incorporated; 2000. 700 pp.
17. Melkonian M. Phylum Chlorophyta: introduction to the Chlorophyta. In: Margulis L, Corliss JO, Melkonian M, Chapman DJ, editors. Handbook of Protoctista. Boston: Jones and Bartlett Publishers Incorporated; 1990. p 597–599.
18. Raven P, Evert RF, Eichhorn SE. Biology of Plants. New York: Bedford, Freeman, and Worth Publishing Group; 1998. 875 pp.
19. Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D. Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci USA 2002;99:12246–12251.
20. Keeling PJ, Palmer JD. Lateral transfer at the gene and subgenic levels in the evolution of eukaryotic enolase. Proc Natl Acad Sci USA 2001; 98:10745–10750.
21. Funes S, Davidson E, Reyes-Prieto A, Magallón S, Herion P, King MP. González-Halphen D. A green algal apicoplast ancestor. Science 2002; 298:2155.
22. Ludwig M, Gibbs SP. Are the nucleomorphs of cryptomonads and *Chlorarachnion* the vestigial nuclei of eukaryotic endosymbionts? Ann Acad Sci NY 1987;503:198–211.
23. Gibbs S. The evolution of algal chloroplasts. In: Lewin RA, editor. Origins of Plastids. New York: Chapman and Hall Company; 1993. p 107–121.
24. Bhattacharya D, Medlin L, Wainright PO, Ariztia EV, Bibeau C, Stickel SK, Sogin ML. Algae containing chlorophylls *a*+*c* are polyphyletic: Molecular evolutionary analysis of the Chromophyta. Evolution 1992;46:1801–1817.
25. McFadden GI. Plastids and protein targeting. J Euk Microbiol 1999; 46:339–346.
26. Douglas SE, Murphy CA, Spencer DF, Gray MW. Molecular evidence that cryptomonad algae are chimaeras of two phylogenetically distinct unicellular eukaryotes. Nature 1991;350:148–151.
27. McFadden GI, Gilson PR, Waller RF. Molecular phylogeny of chlorarachniophytes based on plastid rRNA and *rbcL* sequences. Arch Protistenkd 1995;145:231–239.

28. Van de Peer Y, Rensing SA, Maier U-G, De Wachter R. Substitution rate calibration of small subunit rRNA identifies chlorarachniophyte endosymbionts as remnants of green algae. Proc Natl Acad Sci USA 1996; 93:7732–7736.
29. Gilson PR, McFadden GI. Good things in small packages: the tiny genomes of chlorarachniophyte endosymbionts. BioEssays 1997;19: 167–173.
30. Douglas S, Zauner S, Fraunholz M, Beaton M, Penny S, Deng LT, Wu X, Reith M, Cavalier-Smith T, Maier UG. The highly reduced genome of an enslaved algal nucleus. Nature 2001;410:1091–1096.
31. Cavalier-Smith T. Chloroplast evolution: secondary symbiogenesis and multiple losses. Curr Biol 2002;12:R62–R64.
32. Gajadhar AA, Marquardt WC, Hall R, Gunderson J, Ariztia-Carmona EV, Sogin ML. Ribosomal RNA sequences of *Sarcocystis muris*, *Theileria annulata* and *Crypthecodinium cohnii* reveal evolutionary relationships among apicomplexans, dinoflagellates, and ciliates. Mol Biochem Parasitol 1991;45:147–154.
33. Cavalier-Smith T. The kingdom Chromista: origin and systematics. In: Round FE, Chapman DJ, editors. Progress in Phycological Research (Vol. 4). Bristol: Biopress Limited; 1986. p 309–347.
34. Cavalier-Smith T. Membrane heredity and early chloroplast evolution. Trends Plant Sci 2000;5:174–182.
35. Cavalier-Smith T. Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. J Eukaryot Microbiol 1999;46:347–366.
36. Gillott M. Phylum Cryptophyta (Cryptomonads). In: Margulis L, Corliss JO, Melkonian M, Chapman DJ, editors. Handbook of Protoctista. Boston: Jones and Bartlett Publishers Incorporated; 1990. p 139–151.
37. Green JC, Perch-Nielsen K, Westbroek P. Phylum Prymnesiophyta. In: Margulis L, Corliss JO, Melkonian M, Chapman DJ, editors. Handbook of Protoctista. Boston: Jones and Bartlett Publishers Incorporated; 1990. p 293–317.
38. Vivier E, Desportes I. Phylum Apicomplexa. In: Margulis L, Corliss JO, Melkonian M, Chapman DJ, editors. Handbook of Protoctista. Boston: Jones and Bartlett Publishers Incorporated; 1990. p 549–573.
39. Lynn DH, Small EB. Phylum Ciliophora. In: Margulis L, Corliss JO, Melkonian M, Chapman DJ, editors. Handbook of Protoctista. Boston: Jones and Bartlett Publishers Incorporated; 1990. p 498–523.
40. Taylor FJR. Phylum Dinoflagellata. In: Margulis L, Corliss JO, Melkonian M, Chapman DJ, editors. Handbook of Protoctista. Boston: Jones and Bartlett Publishers Incorporated; 1990. p 419–437.
41. Wilson RJ, et al. Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. J Mol Biol 1996;261:155–172.
42. Fast NM, Kissinger JC, Roos DS, Keeling PJ. Nuclear-encoded, plastid-targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids. Mol Biol Evol 2001;18:418–426.
43. Yoon HS, Hackett J, Pinto G, Bhattacharya D. The single, ancient origin of chromist plastids. Proc Natl Acad Sci USA 2002;99:15507–15512.
44. Van de Peer Y, De Wachter R. Evolutionary relationships among the eukaryotic crown taxa taking into account site-to-site rate variation in 18S rRNA. J Mol Evol 1997;45:619–630.
45. Baldauf SL, Roger AJ, Wenk-Siefert I, Doolittle WF. A kingdom-level phylogeny of eukaryotes based on combined protein data. Science 2000;290:972–977.
46. Durnford DG, Deane JA, Tan S, McFadden GI, Gantt E, Green BR. A phylogenetic assessment of the eukaryotic light-harvesting antenna proteins, with implications for plastid evolution. J Mol Evol 1999;48:59–68.
47. Knoll AH. The early evolution of eukaryotes: a geological perspective. Science 1992;256:622–627.
48. Xiao S, Zhang Y, Knoll AH. Three-dimensional preservation of algae and animal embryos in a Neoproterozoic phosphorite. Nature 1998;391:553–558.
49. Anbar AD, Knoll AH. Proterozoic ocean chemistry and evolution: a bioinorganic bridge. Science 2002;297:1137–1142.
50. Saldarriaga JF, Taylor FJ, Keeling PJ, Cavalier-Smith T. Dinoflagellate nuclear SSU rRNA phylogeny suggests multiple plastid losses and replacements. J Mol Evol 2001;53:204–213.
51. Andersson JO, Roger AJ. A cyanobacterial gene in nonphotosynthetic protists-an early chloroplast acquisition in eukaryotes? Curr Biol 2002; 12:115–119.
52. Walne PL, Kivic PA. Phylum Euglenida. In: Margulis L, Corliss JO, Melkonian M, Chapman DJ, editors. Handbook of Protoctista. Boston: Jones and Bartlett Publishers Incorporated; 1990. p 270–287.
53. Hannaert V, Saavedra E, Duffieux F, Szikora J-P, Rigden DJ, Michels PAM, Opperdoes FR. Plant-like traits associated with metabolism of *Trypanosoma* parasites. Proc Natl Acad Sci USA 2003;100:1067–1071.
54. Martin W, Borst P. Secondary loss of chloroplasts in trypanosomes. Proc Natl Acad Sci USA 2003;100:765–767.
55. Cavalier-Smith T. A revised six-kingdom system of life. Biol Rev Camb Philos Soc 1998;73:203–266.
56. Bhattacharya D, Helmchen T, Melkonian M. Molecular evolutionary analyses of nuclear-encoded small subunit ribosomal RNA identify an independent rhizopod lineage containing the Euglyphidae and the Chlorarachniophyta. J Euk Microbiol 1995;42:65–69.
57. Cavalier-Smith T, Chao EE. Molecular phylogeny of the free-living archezoan *Trepomonas agilis* and the nature of the first eukaryote. J Mol Evol 1996;43:551–562.
58. Keeling PJ. Foraminifera and Cercozoa are related in actin phylogeny: two orphans find a home? Mol Biol Evol 2001;18:1551–1557.
59. Bhattacharya D, Helmchen T, Bibeau C, Melkonian M. Comparisons of small subunit ribosomal RNAs reveal the evolutionary position of the Glaucocystophyta. Mol Biol Evol 1995;12:415–420.
60. Helmchen T, Bhattacharya D, Melkonian M. Analyses of ribosomal RNA sequences from glaucocystophyte cyanelles provide new insights into the evolutionary relationships of plastids. J Mol Evol 1995;41:203–210.
61. Archibald JM, Rogers MB, Toop M, Ishida K-I, Keeling PJ. Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga *Bigelowiella natans*. Proc Natl Acad Sci USA 2003;100:7678–7683.
62. Oliveira M, Bhattacharya D. Phylogeny of the Bangiophycidae (Rhodophyta) and the secondary endosymbiotic origin of algal plastids. Am J Bot 2000;87:482–492.
63. Archibald JM, Keeling PJ. Recycled plastids: a 'green movement' in eukaryotic evolution. Trends Genet 2002;18:577–584.
64. Palmer JD. The symbiotic birth and spread of plastids: how many times and whodunit? J Phycol 2003;39:4–12.
65. Zhang Z, Green BR, Cavalier-Smith T. Single gene circles in dinoflagellate chloroplast genomes. Nature 1999;400:155–159.
66. Yoon HS, Hackett J, Bhattacharya D. A single origin of the peridinin-, and fucoxanthin-containing plastids in dinoflagellates through tertiary endosymbiosis. Proc Natl Acad Sci USA 2002b;99:11724–11729.
67. Barbrook AC, Howe CJ. Minicircular plastid DNA in the dinoflagellate *Amphidinium operculatum*. Mol Gen Genet 2000;263:152–158.
68. Barbrook AC, Symington H, Nisbet RE, Larkum A, Howe CJ. Organization and expression of the plastid genome of the dinoflagellate *Amphidinium operculatum*. Mol Genet Genomics 2001;266:632–638.
69. Hansen G, Daugbjerg N, Henriksen P. Comparative stidy of *Gymnodinium mikimotoi* and *Gymnodinium aureolum*, comb. nov. (=*Gyrodinium aureolum*) based on morphology, pigment composition, and molecular data. J Phycol 2000;36:394–410.
70. Hackett JD, Maranda L, Yoon HS, Bhattacharya D. Phylogenetic evidence for the cryptophyte origin of the plastid of *Dinophysis* (Dinophysiales, Dinophyceae). J Phycol 2003;39:440–448.
71. Chesnick JM, Morden CW, Schmieg AM. Identity of the endosymbiont of *Peridinium foliaceum* (Pyrrophyta): analysis of the *rbc*LS operon. J Phycol 1996;32:850–857.
72. Watanabe MM, Takeda Y, Sasa T, Inouye I, Suda S, Sawaguchi T, Chihara M. A green dinoflagellate with chlorophylls *a* and *b* morphology fine structure of the chloroplast and chlorophyll composition. J Phycol 1987;23:382–389.
73. Zhang Z, Cavalier-Smith T, Green BR. Evolution of dinoflagellate unigenic minicircles and the partially concerted divergence of their putative replicon origins. Mol Biol Evol 2002;19:489–500.
74. Nozaki H, Matsuzaki M, Takahara M, Misumi O, Kuroiwa H, Hasegawa M, Shin-I T, Kohara Y, Ogasawara N, Kuroiwa T. The phylogenetic position of red algae revealed by multiple nuclear genes from mitochondria-containing eukaryotes and an alternative hypothesis on the origin of plastids. J Mol Evol 2003;56:485–497.

75. Kohler S, Delwiche CF, Denny PW, Tilney LG, Webster P, Wilson RJ, Palmer JD, Roos DS. A plastid of probable green algal origin in Apicomplexan parasites. Science 1997;275:1485–1489.
76. Funes S, Davidson E, Reyes-Prieto A, Magallon S, Herion P, King MP, Gonzalez-Halphen D. Response to Comment on ''A Green Algal Apicoplast Ancestor''. Science 2003;301:49.
77. Waller RF, Keeling PJ, Van Dooren GG, McFadden GI. Comment on ''A Green Algal Apicoplast Ancestor''. Science 2003;301:49.
78. Gardner MJ, et al. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 2002;419:498–511.
79. Rodríguez F, Oliver JL, Marin A, Medina JR. The general stochastic model of nucleotide substitution. J Theor Biol 1990;142: 485–501.
80. Morden CW, Sherwood AR. Continued evolutionary surprises among dinoflagellates. Proc Natl Acad Sci USA 2002;99:11558–11560.