DIVERSITY AND EVOLUTIONARY HISTORY OF PLASTIDS
AND THEIR HOSTS

PATRICK J. KEELING

By synthesizing data from individual gene phylogenies, large concatenated gene trees, and other kinds of molecular, morphological, and biochemical markers, we begin to see the broad outlines of a global phylogenetic tree of eukaryotes. This tree is apparently composed of five large assemblages, or “supergroups.” Plants and algae, or more generally eukaryotes with plastids (the photosynthetic organelle of plants and algae and their nonphotosynthetic derivatives) are scattered among four of the five supergroups. This is because plastids have had a complex evolutionary history involving several endosymbiotic events that have led to their transmission from one group to another. Here, the history of the plastid and of its various hosts is reviewed with particular attention to the number and nature of the endosymbiotic events that led to the current distribution of plastids. There is accumulating evidence to support a single primary origin of plastids from a cyanobacterium (with one intriguing possible exception in the little-studied amoeba Paulinella), followed by the diversification of glaucophytes, red and green algae, with plants evolving from green algae. Following this, some of these algae gave rise to the chromalveolates, a diverse group including cryptomonads, haptophytes, heterokonts, and alveolates. Dinoflagellates (alveolates) have since taken up other algae in serial secondary and tertiary endosymbioses, raising a number of controversies over the origin of their plastids, and by extension, the recently discovered cryptic plastid of the closely related apicomplexan parasites.

Key words: algae; endosymbiosis; phylogeny; plastid; tree of eukaryotes.

THE TREE OF EUKARYOTES

To understand an evolutionary process fully, it is important to understand the phylogenetic relationships among the organisms in which that process is manifested. Accordingly, our interpretation of the history of photosynthesis has been strongly influenced by changes in our understanding of the tree of life and, in particular, the tree of eukaryotes.

Early attempts to reconstruct an evolutionary tree of eukaryotes were based on morphological and biochemical similarities. These were highly successful at identifying many major groups of related organisms, but not as successful at identifying the relationships between these large groups (Brugerolle and Taylor, 1977). This is due in part to the fact that the majority of eukaryotic diversity lies in the microbial world (protists and algae), in which morphological characters uniting a particular group are often obvious, but characters for between-group comparisons are less easy to come by. Certain highly conserved characters, such as features of mitochondrial ultrastructure (Taylor, 1978), were useful in extending our understanding of large-scale relationships, but for some time the overall tree of eukaryotes resembled more a bush than a hand of broken branches than it did a tree.

The introduction of molecular data to phylogenetic analyses held great promise to solve this problem by providing a seemingly limitless supply of unambiguously comparable amino acid and nucleotide characters. The first large-scale molecular analyses based on small subunit ribosomal RNA (SSU rRNA) yielded a fully resolved tree, integrating microbial diversity and placing the plants, animals, and fungi in the context of their microbial cousins (Sogin, 1991). While some relationships on the rRNA tree were arguable (Hasegawa et al., 1993), overall it was the accepted model of eukaryotic evolution, and much of the interpretation of the macroevolutionary history of eukaryotes was based on this model. In time, however, a broader taxonomic sampling of a number of other gene sequences showed that an accurate tree of eukaryotes was not to be had quite so simply. A variety of analyses based on protein-coding genes contradicted various aspects of the rRNA tree, and it soon became clear that each molecular analysis (both rRNA and proteins) included problematic taxa and artificial relationships, which were sometimes well supported (Embley and Hirt, 1998; Philippe et al., 2000). Some of these contradictions were dramatic; in particular the erstwhile “deep-branching” eukaryotes microsporidia were shown to be fungi and slime molds were shown to be relatives of animals and fungi (Keeling and Doolittle, 1996; Baldauf, 1999). These cases have attracted considerable attention, so in fairness, I point out that a great number of the previously unproven relationships revealed by the rRNA tree continue to be supported by subsequent analyses of protein-coding genes (for instance, the monophyly of the alveolates and the close relationship between animals and fungi; Baldauf, 1999; Fast et al., 2002).

This period of deconstruction has been followed by a strategy of synthesizing information from a variety of sources to try to build a more robust view of eukaryotic phylogeny. A tree is now emerging, based on analysis of concatenated genes, synthesis of many individual trees, incorporation of discrete characters such as insertions, deletions, and gene fusion events, and consideration of morphology and biochemistry. A tree representing these diverse types of evidence currently consists of five eukaryotic “supergroups” (Fig. 1, top), all of which contain microbial members. Some of these five are better supported than others, and aspects of this scheme are bound to change with new data. The bikont group (informal names are used here because some of these supergroups do not have formal names that are universally accepted) includes animals and fungi and is well supported by numerous phylogenies.
Fig. 1. Tree of eukaryotes and diversity of plastid-bearing eukaryotes. Top: an unrooted hypothetical phylogeny of eukaryotes based on a synthesis of many gene trees, protein insertions and deletions, and cellular and biochemical characters. In this tree, eukaryotes are divided into five large groups, or “supergroups,” within which representatives of the major lineages are shown with their interrelationships as we know them. Dotted lines are plausible but more weakly supported parts of the tree. All groups in which plastids are known from at least a large number of species are indicated by white text on black. Bottom: a small taste of the diversity of plastid-bearing eukaryotes can be seen from one representative of each of the major “algal” lineages. Outside photographs, clockwise from
Plastids and Algae

Plastids are the organelles of plants and algae that harbor photosynthetic and other biochemical pathways for compounds such as aromatic amino acids, heme, isopenoids, and fatty acids. Among eukaryotes (cyanobacteria will not be covered in this review), "algae" might loosely be defined as cells that undertake photosynthesis and/or possess plastids because a number of nonphotosynthetic plastids are known from plants and virtually all algal groups (Williams and Keeling, 2003). However, algae are not a group of organisms in the way that plants or animals are—instead "algae" is a term of convenience that refers to a collection of unrelated organisms that possess this organelle. In the hypothetical phylogeny of eukaryotes shown in Fig. 1, the phylogenetic diversity of plastid-bearing groups is clear: they are scattered across four of the five major groups of eukaryotes.

Historically, this has not always been obvious: because plastids are complex structures that share much in common, it was not unreasonable to assume that the cells that possessed them were closely related (Christen, 1962; Leedale, 1967). Early ultrastructural investigations complicated this view by showing that similarities between the plastids of two algal groups often conflicted with similarities between cytoplasmic features shared by algae and nonalgal eukaryotes. This was perhaps most obvious in the euglenids, whose plastids were quickly recognized to share a number of characteristics with green algae, whereas their flagellar apparatus was more akin to that found in the parasitic and nonphotosynthetic kinetoplastid protozoa (Leedale, 1967; Kivic and Walne, 1984). Molecular systematics further reinforced these contradictions by demonstrating that euglenids and trypanosomes were closely related at the nuclear level to the exclusion of green algae (Sogin et al., 1986).

We now know that this complexity is the result of endosymbiosis. The recognition that plastids and mitochondria were derived from endosymbiotic bacteria was a major turning point in our understanding of the history of the eukaryotic cell (Gray and Doolittle, 1982), but the role of endosymbiosis in plastid evolution did not end with their origin. Instead, endosymbiosis has played an extensive and ongoing role during the elaborate evolutionary history of plastids (Archibald and Keeling, 2002), which has made determining this history especially difficult and intriguing. In the context of the tree of eukaryotes, endosymbiosis implies that several phylogenetic trees are superimposed over one another: there is not only the phylogeny of the organisms themselves, but also a somewhat different phylogeny of their plastids and perhaps even different phylogenies of individual plastid genes.

Primary Plastids

Primary endosymbiosis refers to the original uptake and retention of a cyanobacterium by a eukaryote (Fig. 2A, B). These plastids are bound by two membranes, which are derived from the inner and outer membranes of the Gram-negative cyanobacterium (Jarvis and Soll, 2001). The once free-living prokaryote was reduced and transformed into the organelles we see today, partly by the loss of much of its genome and the transfer of most of the remaining genes to the nucleus of its host. The protein products of most of these nuclear genes continue to function in the plastid, and they are post-translationally targeted there by means of an N-terminal leader called a transit peptide. Transit peptides are recognized by protein complexes in the inner and outer membranes of the plastid (called TICs and TOCs) that direct the translocation of proteins across the membranes (McFadden, 1999). Primary plastids are found in three major lineages: glaucophytes, red algae, and green algae (including plants).

Glaucophytes (glaucocystophytes) are a small group of microscopic algae found in freshwater environments. There are only about 13 species of glaucophytes, and although not particularly common in nature they are important because they occupy a pivotal position in the evolution of photosynthesis in eukaryotes (see later). They also represent an intermediate in the transition from endosymbiotic plastid to plastid that they are unique among plastids in retaining the prokaryotic peptidoglycan layer between their two membranes. Glaucophyte plastids contain photosynthetic pigments chlorophyll a, phycobilins, and phycobilisomes, small particles organized on the outer face of thylakoid membranes that are also found in cyanobac-
Fig. 2. Primary and secondary endosymbiosis. A–B. Primary endosymbiosis. A heterotrophic eukaryote eats a Gram-negative cyanobacterium (A), which is retained rather than being digested (B). The cyanobacterial endosymbiont is substantially reduced, and a large number of genes are transferred to the nuclear genome of the host. The protein products of these genes are targeted to the plastid by way of a transit peptide. The primary plastid is bounded by two membranes derived from the inner and outer membranes of the cyanobacterium. The presumed phagosomal membrane is lost, as is the peptidoglycan wall (except in glaucophyte algae). B–C. Secondary endosymbiosis. A primary alga (either a red or green alga) is eaten but not digested by a second eukaryote (C). This eukaryotic endosymbiont degenerates and genes encoding plastid-targeted proteins are moved from its nucleus to the secondary host nuclear genome. Some genes may also move from the plastid genome to the secondary host nucleus. These plastids would originally be bounded by four membranes derived as indicated. In euglenids and dinoflagellates, the plastid is bounded by three membranes, and the primary algal cytoplasmic membrane (second from outside) is presumed to have been lost. In cryptomonads and chlorarachniophytes, the primary algal nucleus is retained in a highly divergent form, called a nucleomorph, between the second and third membrane (in the space corresponding to the primary algal cytoplasm). Plastid-targeted proteins encoded in the secondary host nucleus make their way to the plastid using a bipartite leader consisting of a signal peptide followed by a transit peptide (in dinoflagellates and euglenids a third region is also encoded; see text).

Red algae are a very large and diverse group of microscopic algae and macroalgae that are present in freshwater and common in marine environments. Red algal plastids contain chlorophyll a, phycobilins, and phycobilisomes. For a review of red algae, see Saunders and Hommersand (2004) in this issue.

Green algae are another large and diverse group of predominantly freshwater algae whose plastids harbor chlorophylls a and b. Green algae are roughly divided into chlorophytes and charophytes. Charophytes are the ancestors of land plants, which share a great number of similarities to charophytes and green algae as a whole. For a review of green algae, see Lewis and McCourt (2004) in this issue.

ORIGIN OF PRIMARY PLASTIDS

The single vs. multiple origins of primary plastids and the relationships between the three major lineages of primary algae are issues that have both been extensively debated, but the bulk of the evidence has now converged on the simple conclusion that all primary plastids trace back to a single endosymbiotic event.

Molecular phylogenetic data from plastid-encoded genes have generally supported a single origin by showing a monophyletic plastid clade associated with the cyanobacteria (Bhattacharya and Melkonian, 1995; Delwiche et al., 1995; Turner et al., 1999; Archibald et al., 2003). This is also supported by several characteristics of plastid genome structure (McFadden and Waller, 1997) and by the presence of common light harvesting complex proteins in the green and red algae (Durnford et al., 1999). Unfortunately, the relationship of plastids to the various cyanobacterial groups has proved more difficult to resolve, so it is impossible to say exactly what kind of cyanobacterium might have given rise to plastids; plastid gene analyses are still open to the (unlikely) possibility that plastids are derived from distinct but closely related cyanobacteria (Turner, 1997; Turner et al., 1999).

In contrast to the plastid genes, early analyses based on nucleus-encoded cytoplasmic proteins from primary algae cast some doubts on the single origin of primary plastids. In general, nuclear gene trees failed to resolve a monophyletic clade including red, green, and glaucophyte algae (Bhattacharya et al., 1995; Bhattacharya and Weber, 1997; Van de Peer and De Wachter, 1997; Keeling et al., 1998, 1999). The lack of primary algal monophyly was sometimes interpreted as evidence for multiple independent plastid origins, or more complex events. Most of these trees are very poorly supported, however, and a failure to show a relationship between primary algal lineages is distinctly different from actually demonstrating that they are not related to one another. One gene that originally appeared to show strong support for separating red and green algae was the largest subunit of RNA polymerase II (Stiller and Hall, 1997). Increased sampling eroded support.
for their separation (Dacks et al., 2002), and in some recent trees, the two groups are united with weak support (Longet et al., 2003). Moreover, mitochondrial genes from red and green algae have now demonstrated support for their monophyly (Burger et al., 1999), and some nuclear gene phylogenies have begun to show a relationship between the host components of primary algae (Moreira et al., 2000; Keeling and Palmer, 2001). These data are generally restricted to red and green algae, but in concatenated gene trees, glaucophytes also fall with the other primary algal lineages (Moreira et al., 2000).

Therefore, there is evidence from all three genomes for a monophyletic origin of red, green, and glaucophyte algae and no strong evidence against such a relationship, supporting the single origin of primary plastids, which is reflected in the "plant" supergroup (Fig. 1).

The single origin of primary plastids and a common ancestor for primary algae raises the question, "Which lineage came first?" Molecular analyses have, in their turn, supported each of the three lineages as the first primary algal group to have diverged (e.g., Valentin and Zetsche, 1990; Delwiche et al., 1995; Helmen et al., 1995; Kowallik, 1997; Martin et al., 1998; Turner et al., 1998). In general, however, the best supported molecular trees tend to favor the glaucophytes branching prior to the divergence of red and green algae from one another, which is specifically supported by two recent analyses of concatenated plastid- and nuclear-encoded genes (Martin et al., 1998; Moreira et al., 2000). In addition, the red and green algal nucleus-encoded, plastid-targeted fructose bisphosphate aldolase (FBA) gene is the product of an endosymbiotic gene replacement event in which the nuclear gene for the cytosolic, glycolytic enzyme was duplicated and replaced the gene for the plastid-targeted Calvin cycle enzyme (Gross et al., 1999). This is significant because this is not true of the glaucophyte FBA (Nickol et al., 2000), and the tree of the red and green algal enzymes indicates that this replacement event took place once in their common ancestor, providing further support for the ancient divergence of the glaucophytes (Rogers and Keeling, 2003).

Considering all the data, it seems likely that the glaucophytes diverged before red and green algae diverged from one another, but it is worth noting that the concatenated analyses rely on limited taxon sampling, and the evolution of FBA is marked by complex events that are difficult to interpret, so the debate is not a dead one. Nevertheless, these results should raise new awareness of this intriguing but poorly understood group of algae.

Having concluded that primary plastids evolved only once, there is an exceptional case that warrants further attention. The ( cercozoan) euglyphid amoeba *Paulinella chromatophora* is a marine testate amoeba that contains two kidney-shaped cyanobacterial endosymbionts, which are similar to *Synechococcus* and could be classified as independently acquired plastids. *Paulinella chromatophora* is closely related to the heterotrophic *P. ovalis*, which lacks endosymbionts but preferentially feeds on *Synechococcus* using the flose pseudopods characteristic of this group of amoebae (Johnson et al., 1988). *Paulinella chromatophora*, on the other hand, does not feed, and its endosymbionts are known to actively photosynthesize and transfer photosynthetic products to the host. Moreover, the endosymbionts cannot be cultivated independently and were shown to divide synchronously with the host (Kies, 1974; Kies and Kremmer, 1979). The exact criteria that distinguishes endosymbiont from organelle are largely a matter of opinion, but the relationship between *P. chromatophora* and its endosymbionts is strongly indicative that the two are sufficiently integrated for the endosymbionts to be comfortably considered organelles and, by extension, independently acquired plastids.

**SECONDARY PLASTIDS**

Although red, green, and glaucophyte algae (and plants) represent tremendous diversity, they still only account for a fraction of the diversity of photosynthetic eukaryotes. The remainder of plastid-bearing eukaryotes acquired their plastids by secondary endosymbiosis, which is the uptake and retention of a primary plastid-containing alga by a second eukaryote (Fig. 2C, D). Secondary plastids are characterized by additional membranes surrounding the plastid; rather than the two membranes bounding a primary plastid, most secondary plastids are surrounded by four or sometimes three membranes. The extra membranes are a consequence of the phagocytosis of the primary alga: the inner two membranes correspond to the two membranes of the primary plastid (and by extension, the two membranes of the Gram-negative cyanobacterium), the third membrane from the inside in plastid with four membranes corresponds to the cytoplasmic envelope of the primary algal endosymbiont, whereas the outermost membrane is part of the secondary host endomembrane system, derived from the phagocytic vacuole in which the algal prey was taken up (Archibald and Keeling, 2002). Both red and green algae have been taken up in secondary endosymbiotic events, but a secondary plastid derived from a glaucophyte has never been observed, probably because they are comparatively rare in nature and therefore less likely to be involved in such an event. In secondary plastids of both red and green origin, the primary algal nucleus is highly reduced or more commonly lost altogether, so the majority of the plastid-targeted proteins that were encoded in the primary algal nucleus have been transferred to the secondary host nucleus. Targeting these proteins to the secondary plastid involves an extra step because the secondary plastid is not physically located in the cytoplasm, as are the primary plastids, but rather is in the lumen of the endomembrane system. Accordingly, plastid-targeted proteins in algae with secondary plastids encode bipartite leaders with a signal peptide to direct them to the endomembrane system, followed by a transit peptide as is found in primary algae (McFadden, 1999). The exact way in which the proteins travel though the endomembrane system to the plastid varies from group to group, in particular with plastids bounded by three membranes (Sulli et al., 1999; Nassoury et al., 2003), but in general all secondary algae use this basic strategy to target proteins. This represents an interesting case of convergence, in which several groups have independently co-opted the same existing cellular machine (the secretion pathway) to get around the tricky protein-trafficking problem.

There are presently seven major lineages recognized to possess secondary plastids. Some distinguishing features of each are reviewed very briefly here.

**Euglenids** are a diverse group of common marine and freshwater flagellates, about half of which contain a green algal plastid (*chlorophyll a* and *b*) bounded by three membranes. The remainder of the group are osmotrophs or heterotrophs that feed on bacteria or other eukaryotes. Euglenids are closely related to the parasitic *trypanosomes*, together with diplomonds, making up the Euglenozoa. Euglenozoa, in turn are related to heterolobosean amoeboflagellates, which are generally con-
Chlorarachniophytes are a relatively rare group of marine amoeboflagellates and flagellates that contain a green algal plastid (chlorophyll $a$ and $b$) bounded by four membranes. Chlorarachniophyte plastids have received some attention because they are one of only two groups in which the primary algal nucleus has not been completely lost; they retain a small relic nucleus called a nucleomorph. Chlorarachniophyte nucleomorphs contain highly reduced genomes on the order of 380–710 kbp, but are still highly reduced and divided onto three small chromosomes. These genomes are tightly compacted with highly divergent sequences, but many signatures of their green algal origin remain. The host component of chlorarachniophytes has been an enigma until recently, but is now known to be part of the growing "rhizaria" supergroup (yellow in Fig. 1). For a review of chlorarachniophytes, see McFadden et al. (1997).

Cryptomonads are an abundant group of marine and freshwater flagellates that contain a red algal plastid with chlorophyll $a$ and $c$ surrounded by four membranes. Cryptomonads are the other group with a nucleomorph and share many superficial similarities with that of chlorarachniophytes. Cryptomonad nucleomorph genomes are slightly larger, between 450–710 kbp, but are still highly reduced and divided onto three small chromosomes. The presence of nucleomorphs derived from both green and red algal nuclei gives us a remarkable opportunity to study parallel events of genome reduction. One cryptomonad nucleomorph genome is now complete, and a chlorarachniophyte is underway (McFadden et al., 1997; Douglas et al., 2001). The host component of cryptomonads has also been a matter of debate, but evidence is accumulating that it and the remaining groups are all members of a single supergroup, the "chromalveolates" (light blue in Fig. 1). For a review of cryptomonads, see Fraunholz et al. (1997).

Haptophytes (prymnesiophytes or coccolithophorids) are an abundant group of marine flagellates with four-membrane plastids (chlorophyll $a$ and $c$) derived from a red alga. The outermost membrane is continuous with the endoplasmic reticulum and nuclear envelope. Haptophytes are ecologically important as significant primary producers—their blooms can be large enough to be distinctly visible from space. For a review of haptophytes, see Andersen (2004) in this issue.

Heterokonts (stramenopiles) are an extremely diverse group of photosynthetic and nonphotosynthetic groups that were once classified separately as protozoa, algae, and fungi. Heterokont plastids, where they occur, are structurally similar to those of haptophytes and also contain chlorophyll $a$ and $c$. Heterokont algae include microscopic forms of great ecological significance (e.g., diatoms) as well as macroalgae (e.g., kelps). For a review of heterokont algae, see Andersen (2004) in this issue.

Dinoflagellates are another large and diverse group, about half of which possess recognized plastids that are derived from red algae and that contain chlorophyll $a$ and $c$. Like euglenids, the major plastid type of dinoflagellates (distinguished by the carotenoid pigment peridinin) is bounded by three membranes. Peridinin-containing plastids have unusual genomes consisting of only a handful of genes, each encoded on a single-gene minicircle. Dinoflagellates are also unusual in that they have lost and replaced their plastid on several occasions (see later). For a review of dinoflagellates, see Bhattacharya (2004) in this issue.

Apicomplexa are a large group composed entirely of obligate intracellular parasites, including several that cause significant diseases such as malaria. As intracellular parasites, the discovery of a relic plastid (or apicoplast) in apicomplexa bounded by four membranes has drawn a great deal of attention as an evolutionary novelty and possible drug target. The origin of this plastid has also been the focus of considerable debate (discussed later), but most evidence now indicates it is derived from a red alga and is probably related to that of their sister group, the dinoflagellates. For a review of apicomplexa relating to their plastid, see Foth and McFadden (2003).

A HISTORY OF SECONDARY ENDOSYMBIOSIS

Because secondary plastids are known to be derived from both green and red algae, it is clear that this process has taken place more than once, but exactly how many secondary endosymbiotic events explain the diversity of plastids has been debated extensively. A current hypothesis of the history of plastids in eukaryotes is shown in Fig. 3. On the green side, it has been proposed that euglenid and chlorarachniophyte plastids are derived from a common endosymbiosis (Cavalier-Smith, 1999, 2000), but the majority of evidence indicates otherwise. There is no evidence that these two groups share a recent common ancestry because chlorarachniophytes are certainly members of the cercozoa (Bhattacharya et al., 1995; Cavalier-Smith and Chao, 1997; Keeling et al., 1998; Archibald et al., 2002), whereas euglenids are clearly members of the discicristates related to trypanosomes (Sogin et al., 1986; Baldauf et al., 2000; Simpson, 2003) (see Fig. 1). Similarly, the few genes from the plastid genome that have been analyzed do not show a common ancestry between these two groups (e.g., McFadden et al., 1995), and an analysis of euglenid phylogeny, cytoskeleton, and feeding strategies indicates that their plastids arose after the euglenids diverged from other euglenozoa (Leander et al., 2001). It has recently been suggested that trypanosomes contain a number of plastid-derived genes (Hannaert et al., 2003), which could support a more ancient origin of euglenid plastids; however, the strongest examples of "plant-like" genes in Trypanosoma are both artifacts arising from lack of sampling in the analyses (Rogers and Keeling, 2003), so there is no real evidence for a plastid in the history of trypanosomes. While it seems unlikely that they will prove to be closely related, the true test will be a comprehensive phylogeny of green plastid genes, including representation from various green algal lineages as well as euglenids and chlorarachniophytes. If their plastids are shown to be related to the exclusion of all other green plastids, then it is possible they do derive from a common endosymbiosis. If they are not closely related, then they must have been acquired independently (Fig. 3).

The history of red algal secondary endosymbiosis is more complicated, largely due to the number of groups involved. Indeed, the diversity of organisms with red algal secondary plastids is great, and many different endosymbiotic events have been postulated to explain this diversity. It has also been suggested that the diversity of red algal secondary plastids can be traced to the relatively gene-rich nature of red algal plastid genomes: because red algal plastid genomes encode many more genes than those of green algae, they are hypothesized to be more adept at integrating into a new host environment (Grzybek et al., 2003). Despite the great diversity of red algal secondary plastids, evidence is mounting that they can all be traced to a single endosymbiosis of a red alga in the ancestor

considered part of the supergroup “excavates” (pink in Fig. 1). For a review of euglenids, see Leedale and Vickerman (2000).
of all chromalveolates (Cavalier-Smith, 1999; Fig. 3), which would mean that red algae have been involved in secondary endosymbiotic events less often than green algae.

This conclusion is supported by a variety of observations. There are a few physical and chemical characteristics shared by some or all of these plastids, including the unique presence of chlorophyll $c$ in all but the nonphotosynthetic plastids of apicomplexa and the similarities in membrane organization in cryptomonad, heterokont, and haptophytes plastids (Cavalier-Smith, 1981, 1982; Andersen, 1991). The most compelling evidence, however, comes from molecular sequence analysis. At first, molecular trees failed to show a close relationship between either nuclear or plastid genes from these organisms (Bhattacharya et al., 1995; Medlin et al., 1995; Daugbjerg and Andersen, 1997; Keeling et al., 1999; Oliveira and Bhattacharya, 2000; Muller et al., 2001), with the exception of the alveolate apicomplexa and dinoflagellates, which are clearly closely related (Van de Peer et al., 1996; Bhattacharya et al., 2000; Fast et al., 2002). More recently, however, analyses of both plastid and nuclear genomes have begun to support a relationship between members of the chromalveolates. Nucleus-encoded genes for cytosolic proteins and RNAs have never supported the chromalveolates as a whole, but SSU rRNA has consistently demonstrated a relationship between alveolates and heterokonts (Van de Peer et al., 1996; Van de Peer and De Wachter, 1997), as have several protein-coding genes individually and concatenated (Baldauf et al., 2000; Saldarriaga et al., 2003). These analyses do not support a specific position for haptophytes or cryptomonads (for which there are few data), although an analysis of six concatenated genes supports the alveolates and heterokonts strongly and the haptophytes and cryptomonads weakly (Harper and Keeling, 2004). Plastid-encoded genes have also begun to indicate such a relationship: a recent analysis of five concatenated genes showed strong support for the union of cryptomonads, haptophytes, and heterokonts (Yoon et al., 2002). While none of these analyses show all chromalveolate groups together (either due to unavailability of data or lack of resolution), taken together they do support the chromalveolates as a whole. The only data that address the entire group at once are nucleus-encoded, plastid-targeted genes, most conspicuously glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Like many plastid metabolic genes, GAPDH is present in two copies in plant and algal nuclear genomes, one expressed in the cytosol and one targeted to the plastid. Normally, the plastid-targeted copy would be cyanobacterial in origin, but the plastid GAPDH genes of cryptomonads, haptophytes, heterokonts, dinoflagellates, and apicomplexa are all derived from a duplication of the cytosolic gene (Fagan et al., 1998; Liaud et al., 2000; Fast et al., 2001; Harper and Keeling, 2003). Such an event, called endosymbiotic gene replacement, has been seen in other genes, but is quite rare, and the common origin of all chromalveolates plastid GAPDH genes is compelling evidence that their plastids are derived from a single common endosymbiotic event (Fast et al., 2001; Harper and Keeling, 2003).

Interestingly, the same result has now been observed for another plastid enzyme, fructose bisphosphate aldolase (FBA). The heterokont plastid-targeted FBA has been shown to be an entirely different class of enzyme than that found in red algal plastids (Rogers and Keeling, 2003), and the plastid FBA genes of haptophytes, cryptomonads, and dinoflagellates have now been characterized and are related to the heterokont genes (Patron et al., 2004). GAPDH and FBA bolster support for the common ancestry of chromalveolates and the single origin of their plastids. However, more compelling data are needed, especially to show a relationship between the host lineages using concatenated nuclear gene analyses. Moreover, the apicomplexan and dinoflagellate plastids continue to spark controversy, as described later.

**TERTIARY AND SERIAL SECONDARY ENDOSYMBIOSIS**

The dinoflagellates tend to do things a little differently, so it is perhaps not surprising that they also have an unusual diversity of plastids types, now known to be the result of their ability to lose, replace, or gain new plastids. As mentioned earlier, most dinoflagellates with plastids have a three-membrane, peridinin-containing plastid, but a few groups of dinoflagellates have other types of plastids (Saldarriaga et al., 2001), now recognized to be derived from serial secondary endosymbiosis (the uptake of a new primary plastid-containing endosymbiont) or tertiary endosymbiosis (the uptake of a secondary plastid-containing endosymbiont). Tertiary plastids are found in the toxic Dinophysis, some of which have two-membrane, cryptomonad-derived plastids (Hallegaard and Lucas, 1988; Schnepf and Elbrächter, 1988); the toxic Karenia, which has a haptophyte-derived plastid surrounded by two to four membranes (Steidinger et al., 1978; Kite and Dodge, 1988; Tengs et al., 2000); and Kryptoperidinium and its close relatives, which have a five-membrane, diatom-derived plastid that includes a diatom nucleus of unknown complexity (Dodge, 1971; Tomas et al., 1973; Inagaki et al., 2000). Kryptoperidinium has apparently not lost its original three-membrane peridinin plastid, but rather converted it into an eyespot (Dodge, 1969). Alternatively, a serial secondary plastid is found in Lepidodinium and its close relatives, in which the peridinin plastid has apparently been replaced by a secondary plastid derived from a green alga, which is now surrounded by two membranes (Watanabe et al., 1987, 1990). Each of these cases, should they prove to be fully integrated organelles, presents a fascinating problem for plastid targeting and could represent several novel solutions to protein trafficking.

Most of these organisms are only poorly studied at the molecular level, so the impacts of these events on the host and plastid genomes are only starting to emerge. The first plastid-targeted gene from one such organism has recently been described for Karenia (Ishida and Green, 2002). Not surprisingly, it is haptophyte in nature, suggesting that plastid replacement probably includes replacing many or most of the nucleus-encoded genes for plastid-targeted proteins. Moreover, a number of dinoflagellates contain short-term plastids stolen from their food source (kleptochloroplasts), and it is sometimes difficult to distinguish these and permanently replaced plastids. It seems clear that the Kryptoperidinium and Karenia plastids are permanent, but there is debate over whether the Dinophysis plastid is truly integrated or is a kleptochloroplast (Takishita et al., 2002; Hackett et al., 2003). The Lepidodinium plastid is virtually uncharacterized. Altogether the dinoflagellates stand out as possessing an unusual ability to take up new plastids, although even in dinoflagellates these events are still quite rare.

**APICOMPLEXAN AND DINOFLAGELLATE PLASTIDS—MAGNETS FOR CONTROVERSY**

The apicomplexa and dinoflagellates are closely related sister groups within the tree of eukaryotes, as members of the
Fig. 3. Endosymbiosis in the history of plastid evolution. All primary (top), secondary (middle), serial secondary and tertiary (bottom) endosymbiotic associations mentioned in the text are represented here. Cells are color-coded so that the cytoplasmic color matches the color of the supergroup in Fig. 1 to which that eukaryote belongs (Cercozoa are yellow, plants are green, excavates are purple, and chromalveolates are blue). Plastids are color-coded to distinguish the three primary plastid lineages (cyanobacteria and glaucophyte plastids are both blue-green, red algal plastids are red, and green algal plastids are dark green). Primary endosymbiosis: At the top left, the cercozoan euglyphid amoeba *Paulinella* takes up a *Synechococcus*-like cyanobacterium and retains two apparently permanent endosymbionts, losing its feeding pseudopods. This may represent an independent primary endosymbiosis. At the top center, a cyano-
alveolates (Fig. 1). Accordingly, when the plastid was discovered in apicomplexa, it immediately raised the question of whether they acquired their plastids from the same endosymbiosis as dinoflagellates. The single origin of all chromalveolate plastids discussed indicates that they most likely have a common origin; however, some data have been interpreted as showing that the story is more complex. To a large extent, the debate stems from the fact that apicomplexan and dinoflagellate plastids are both very odd and, more importantly, are odd in different ways that make them difficult to compare. 

The four-membrane, apicomplexan plastid is the most recently discovered plastid (McFadden et al., 1996; Wilson et al., 1996) and was something of a surprise, given the obligate intracellular nature of these parasites (Wilson, 2002). The characterization of function and trafficking systems in this plastid have rapidly outstripped our understanding of many other secondary plastids, driven by the novelty of the organelle and the potential importance of the plastid in treating parasites such as Plasmodium (Foth and McFadden, 2003; Foth et al., 2003; Ralph et al., 2004). The medical and commercial importance of these parasites also means that many tools exist to study the apicomplexan plastid that do not exist for other groups. Ironically, the malaria parasite Plasmodium was the first completely sequenced nuclear genome of any “alga” (Gardner et al., 2002). Several apicomplexan plastid genomes have also been fully sequenced (Wilson et al., 1996; Köhler et al., 1997; Cai et al., 2003). They have lost all genes related to photosynthesis, but otherwise resemble reduced red algal plastid genomes in structure and content (McFadden and Walker, 1997). Phylogenetic analyses based on several plastid-targeted genes have also weakly supported a red algal origin for apicomplexan plastids (Gardner et al., 1994; Blanchard and Hicks, 1999), but analyses of three of these genes have been used to argue for a green algal origin. Plastid tufA and RNA polymerase gene trees show a weak affinity between the apicomplexa and green alga (Köhler et al., 1997; Cai et al., 2003), but in both cases the apicomplexan homologue is highly divergent and branches with other divergent homologues. The mitochondrial cox2 has also been suggested to indirectly show a green algal origin of the apicomplexan plastid because the apicomplexan and chlorophyte green algae both have a nuclear-encoded cox2 that is split into two subunits, whereas other eukaryotes have a single mitochondrion-encoded gene (Funes et al., 2002). The conclusion that the apicomplexan plastid is derived from a green alga is therefore based on two assumptions: that cox2 split and moved to the nuclear genome once and that the donor was the same alga that gave rise to the plastid (Palmer, 2003). However, other analyses indicate that the split and nuclear location may not be independent events and that they may have taken place in parallel in chlorophytes and apicomplexa (Waller et al., 2003). Moreover, cox2 II is not a plastid protein, so even if there were unambiguous evidence for a green algal origin of the apicomplexan cox2, it would not necessarily mean the plastid was also green algal. 

In contrast, there has never been any question that dinoflagellates contain a plastid because about half are photosynthetic and are well studied as primary producers, photosynthetic endosymbionts, and producers of toxic blooms (Taylor, 1987). Nevertheless, the three-membrane, peridinin-containing plastid found in most dinoflagellates has remained mysterious because of the lack of molecular data from the organelle. Recently, the first genes from the dinoflagellate plastid have been characterized, and these were found to reside on unusual single-gene minicircles rather than the typical plastid chromosome (Zhang et al., 1999). Phylogenetic analyses of several of these genes (Zhang et al., 2000) confirmed the already widely believed red algal origin of the peridinin-containing plastids, based on the presence of chlorophyll c. However, analyses of photosystem genes led to the hypothesis that the peridinin-containing plastid was not ancestral to dinoflagellates, but was instead acquired by tertiary endosymbiosis with a haptophyte (Yoon et al., 2002). The precedent for tertiary endosymbiosis in dinoflagellates (see earlier) certainly lends credibility to this possibility, but a recent reanalysis of the data shows that the tree is the result of a base-composition bias rather than tertiary endosymbiosis (Inagaki et al., 2004). 

So, on one hand, we have the chromalveolate hypothesis, which states that all red algal secondary plastids are derived from a common endosymbiosis, while on the other hand, we have the suggestions that the apicomplexan plastid is derived from a green alga and that the peridinin-containing plastid is derived from a haptophyte. These alternative suggestions are not incompatible with the chromalveolate hypothesis. Indeed, all have been woven together by the suggestion that the apicomplexan and dinoflagellates ancestrally contained chromalveolate plastids, but that the plastids of one or both have been replaced (e.g., see Yoon et al., 2002; Palmer, 2003). The most direct test of these alternatives would be to analyze their plastid genomes to see if they are closely related. Unfortunately, the plastid genomes of these two groups are virtually impossible to compare. Only 16 genes have been characterized in the plastid genome of any peridinin-containing dinoflagellate (for a review, see Green, 2004); the remainder appears to have been transferred to the nuclear genome (Bachvaroff et al., 2004; Hackett et al., 2004; N. Patron and P. J. Keeling, unpublished data). This transfer is significant because nearly all
of the few genes that do remain encode proteins related to photosynthesis, which the apicomplexan plastid has lost. As a result, only two genes, SSU and LSU rRNA, have been analyzed from both apicomplexa and dinoflagellates (Zhang et al., 2000). Both the rRNAs are highly divergent in both groups and not easily compared, although it should be noted that they do weakly support a common origin (Zhang et al., 2000).

At present, there are no definitive answers to the questions that surround the origin of apicomplexan and dinoflagellate plastids, but the simplest explanation is that the chromalveolates and their plastids share a common ancestor. A corollary of this would be that the nucleus-encoded genes (for both cytosolic and plastid-targeted proteins) in these groups should tend to be closely related (within the limits of phylogenetic reconstruction). We know little about the fate of nucleus-encoded genes for plastid-targeted proteins when a plastid is replaced, but in the one instance in which it has been studied the protein comes from the new plastid (Ishida and Green, 2002). Therefore, if either dinoflagellate or apicomplexan plastids have evolved through replacement, then a substantial fraction of their plastid-targeted proteins should differ from the aforementioned prediction: they should be distantly related phylogenetically. Moreover, these genes should all come from a single other source because it has recently been shown that genes for plastid-targeted proteins in a chlorarachniophyte can come from a variety of sources, and there is no indication that this has anything to do with plastid replacement (Archibald et al., 2003). Although few genes have been analyzed with this question in mind, there is currently no evidence for such a trend in data from either apicomplexa or dinoflagellates.

CONCLUDING REMARKS

The evolutionary history of plastids and their hosts has not been a straightforward story of one tree, but instead has been an intertwining of many overlapping trees brought about by many rounds of endosymbiosis, lateral gene transfers, and gene replacements, all leading to the spread of genes and organelles across a substantial proportion of eukaryotic diversity. Despite this complexity, the history and the processes that underlie plastid evolution are remarkably well understood, thanks to a large body of work on plastid ultrastructure and biochemistry synthesized with a rapidly expanding body of molecular data from plants and algae. The general position of all algal groups in the tree of eukaryotes is now coming into focus, and of the five supergroups identified in Fig. 1, plants are the most developed. The ancient green plastid lineage. Journal of Eukaryotic Microbiology 46: 367–375.


Daugbjerg, N., and R. A. Andersen. 1997. Phylogenetic analyses of the rbcL sequences from haplophytes and heterokont algae suggest their...


