The endosymbiotic origin, diversification and fate of plastids

Patrick J. Keeling*

Botany Department, Canadian Institute for Advanced Research, University of British Columbia, 3529-6270 University Boulevard, Vancouver, BC, Canada V6T 1Z4

Plastids and mitochondria each arose from a single endosymbiotic event and share many similarities in how they were reduced and integrated with their host. However, the subsequent evolution of the two organelles could hardly be more different: mitochondria are a stable fixture of eukaryotic cells that are neither lost nor shuffled between lineages, whereas plastid evolution has been a complex mix of movement, loss and replacement. Molecular data from the past decade have substantially untangled this complex history, and we now know that plastids are derived from a single endosymbiotic event in the ancestor of glaucophytes, red algae and green algae (including plants). The plastids of both red algae and green algae were subsequently transferred to other lineages by secondary endosymbioses. Green algal plastids were taken up by euglenids and chlorarachniophytes, as well as one small group of dinoflagellates. Red algae appear to have been taken up only once, giving rise to a diverse group called chromalveolates. Additional layers of complexity come from plastid loss, which has happened at least once and probably many times, and replacement. Plastid loss is difficult to prove, and cryptic, non-photosynthetic plastids are being found in many non-photosynthetic lineages. In other cases, photosynthetic lineages are now understood to have evolved from ancestors with a plastid of different origin, so an ancestral plastid has been replaced with a new one. Such replacement has taken place in several dinoflagellates (by tertiary endosymbiosis with other chromalveolates or serial secondary endosymbiosis with a green alga), and apparently also in two rhizarian lineages: chlorarachniophytes and Paulinella (which appear to have evolved from chromalveolate ancestors). The many twists and turns of plastid evolution each represent major evolutionary transitions, and each offers a glimpse into how genomes evolve and how cells integrate through gene transfers and protein trafficking.

Keywords: plastids; endosymbiosis; evolution; algae; protist; phylogeny

1. THE ORIGIN OF PLASTIDS: A SINGLE EVENT OF GLOBAL SIGNIFICANCE

Endosymbiosis has played many roles in the evolution of life, but the two most profound effects of this process were undoubtedly the origins of mitochondria and plastids in eukaryotic cells. There are many parallels in how these organelles originated and how their subsequent evolution played out, for example, the reduction of the bacterial symbiont genome and the development of a protein-targeting system (Whatley et al. 1979; Douglas 1998; Gray 1999; Gray et al. 1999; Gould et al. 2008). There are also many differences, however, and one of the more striking is the ultimate fate of the organelle once established: where mitochondria were integrated into the host and the two were seemingly never again separated (Williams & Keeling 2003; van der Giezen et al. 2005), plastid evolution has seen many more twists, turns and dead ends (for other reviews that cover various aspects of this history, see Delwiche 1999; McFadden 1999, 2001; Archibald & Keeling 2002; Stoebe & Maier 2002; Palmer 2003; Williams & Keeling 2003; Keeling 2004; Archibald 2005; Keeling in press).

Ultimately, mitochondria and plastids (with the small but interesting exception detailed in §2) have each been well established to have evolved from a single endosymbiotic event involving an alphaproteobacterium and cyanobacterium, respectively (Gray 1999). This conclusion has not come without considerable debate, which has stemmed from several sources. First, the ancient nature of the event makes reconstructing its history difficult because a great deal of change has taken place since the origin of this system, and all of this change masks ancient history. In addition, however, the complexity of subsequent plastid evolution has made for special and less expected problems. Specifically, plastids were originally established in a subset of eukaryotes by a so-called 'primary' endosymbiosis with an ancient cyanobacterial lineage. Once established, primary plastids then spread from that lineage to other eukaryotes by additional rounds of endosymbiosis between two eukaryotes (secondary and tertiary endosymbioses, which are each discussed in detail within their own section below). This led to a very confusing
picture of plastid diversity and distribution, as the plastids and their hosts can have different evolutionary histories. Until this was realized, the tendency was, reasonably enough, to treat all photosynthetic lineages as close relatives, which created a long list of paradoxical observations where the plastid of some lineage seemed similar to that of one type of eukaryote, but the host component appeared more similar to another (Dougherty & Allen 1960; Leedale 1967; Brugerolle & Taylor 1977). For example, euglenids had plastids like those of green algae but their cytosolic features were like trypanosomes, which in fact proved to be exactly the case.

Once these added layers of complexity were resolved, or at least understood to exist (Gibbs 1978, 1981; Ludwig & Gibbs 1989), the multiple versus single origin of plastids boiled down to resolving the origin of primary plastids. Primary plastids are surrounded by two bounding membranes and are only found in three lineages. Glaucophytes are a small group of microbial algae with plastids that contain chlorophyll $a$ and are distinguished by the presence of a relic of the peptidoglycan wall that would have been between the two membranes of the cyanobacterial symbiont (Bhattacharya & Schmidt 1997; Steiner & Loffelhardt 2002). Red algae are considerably more diverse and conspicuous than glaucophytes, with 5000–6000 species ranging from tiny, non-flagellated coccoid cells in extreme environments to marine macroalgae that are known to anyone that has walked in the rocky intertidal zone (figure 1). Their plastids also contain chlorophyll $a$ and phycobilisomes and are distinguished by the presences of phycerothrin (Graham & Wilcox 2000). The green algae are also a diverse group and are abundant in both marine and freshwater environments (figure 1), and on dry land where a few green algal lineages have proliferated (e.g. Trentopahales and many Trebouxio-phytes that interact with fungi in lichens), as well as giving rise to the land plants, a major lineage with global terrestrial impacts. Their plastids contain chlorophyll $a$ and $b$ and are also distinguished by the storage of starch within the plastid (Lewis & McCourt 2004).

The plastids in all three of these lineages share a good deal in common with one another, and genome sequences from all three lineages also share several features that are not known in any cyanobacterial lineage, for example, the presence of inverted repeats surrounding the rRNA operon (Palmer 1985; McFadden & Waller 1997). Molecular phylogenies of plastid-encoded genes have consistently shown them to be monophyletic to the exclusion of all currently known cyanobacterial lineages (Delwiche et al. 1995; Helmchen et al. 1995; Turner et al. 1999; Yoon et al. 2002; Hagopian et al. 2004; Khan et al. 2007). Further analysis of plastid-targeted proteins also showed that plastids contain a unique light harvesting complex protein that is not found in any known cyanobacteria (Wolfe et al. 1995; Durnford et al. 1999). All these observations are most consistent with the monophyly of plastids, but debate persisted, primarily because of analyses of nuclear sequences. In early molecular phylogenies, the three primary plastid lineages were not monophyletic (Bhattacharya et al. 1993; Bhattacharya & Weber 1997), and in some cases gene trees seemed to show well-supported evidence against this monophyly (Stiller et al. 2001, 2003; Kim & Graham 2008). The early analyses have not held up, however, and more compellingly analysed large datasets of concatenated genes have most consistently demonstrated the monophyly of the nuclear lineages, and in those analyses with the most data this is recovered with strong support (Moreira et al. 2000; Rodriguez-Espeleta et al. 2005; Hackett et al. 2007; Burki et al. 2009).

The current consensus is that there is a single lineage, called Plantae or Archaeplastida (Adl et al. 2005), in some probably biflagellated heterotrophic ancestor of which the primarily endosymbiotic uptake of a cyanobacterium took place. The cyanobacterium was reduced by loss of genes and their corresponding functions, and also genetically integrated with its host. A complex mechanism for targeting nucleus-encoded proteins to the endosymbiont was progressively established, resulting in the outer and inner membrane complexes today known as translocon outer (TOC) and inner (ITC) chloroplast membranes (McFadden 1999; van Dooren et al. 2001; Wickner & Scheckman 2005; Hormann et al. 2007; Gould et al. 2008). The targeted proteins mostly acquired amino terminal leaders called transit peptides, which are recognized by the TOC and used to drag the protein across the membranes, and which are subsequently cleaved in the plastid stroma by a specific peptidase (McFadden 1999; Wickner & Scheckman 2005; Hormann et al. 2007; Patron & Waller 2007; Gould et al. 2008), a system remarkably similar to the protein-targeting system used by mitochondria (Lithgow 2000). This system probably coevolved with the transfer of a few genes to the nucleus, and once the system was established its presence would make it relatively easy for many genes to move to the host genome. Today, plastid genomes are a small fraction of the size of cyanobacterial genomes (Douglas 1998; Douglas & Raven 2003), and there is relatively little diversity in the size and overall structure of the genome, at least in comparison with mitochondrial genomes, and gene content is relatively stable, although many genes are known to have moved to the nucleus independently in multiple lineages (Martin et al. 1998; Gould et al. 2008). There has been some debate over whether the endosymbiont also donated a substantial number of genes whose products are not now targeted to the plastid (Martin et al. 1998, 2002; Martin 1999; Reyes-Prieto et al. 2006), but these will not be discussed here.

Once the endosymbiont was established and integrated with its host, the three major lineages of Archaeplastida diverged (figure 2). Again, there is considerable controversy about the branching order of these three groups, but current data lean towards the glaucophytes branching first. Glaucophyte plastids are unique in retaining the peptidoglycan wall between the two membranes, which would seem to support this (Bhattacharya & Schmidt 1997; Steiner & Loffelhardt 2002). However, all three groups have been proposed to be most basal at one time or another (Cavalier-Smith 1982; Kowalik 1997; Martin et al. 1998).

Molecular phylogenies have not provided a particularly strong answer to this question, but many recent analyses based on large datasets of both nuclear
and plastid genes often show glaucophytes branching first, although some also show red algae branching earlier (Moreira et al. 2000; Rodriguez-Ezpeleta et al. 2005; Hackett et al. 2007). It has also been shown that glaucophytes alone retain the ancestral cyanobacterial fructose bisphosphate aldolase, and that in both green and red algae the original gene has been replaced by a duplicate of the non-homologous (analogous) nuclear-encoded cytosolic enzyme (Gross et al. 1999; Nickol et al. 2000). Phylogenetic analysis of the red and green algal genes has shown that the plastid and cytosolic paralogues each form a (weak) clade including both red and green algal genes (Rogers & Keeling 2003), which suggests that the gene replacement must have been ancestral to both red and green algae, and that the glaucophytes therefore diverged prior to this event, making them the first-diverging lineage of the Archaeplastida (figure 2).

2. PAULINELLA AND THE POSSIBILITY OF A SECOND ORIGIN OF PLASTIDS
All the data outlined above relate to the ultimate origin of plastids in nearly all major lineages of eukaryotes.
There is, however, a large number of endosymbiotic relationships seemingly based on photosynthesis that are less well understood and vary across the entire spectrum of integration, from passing associations to long term and seemingly well-developed partnerships (e.g. Rumpho et al. 2008). Indeed, the line between what is an organelle and what is an endosymbiont is an arbitrary one. There are a few different, specific criteria that have been argued to distinguish the two, the most common being the genetic integration of the two partners, and the establishment of a protein-targeting system. Most photosynthetic endosymbionts probably
are not integrated with their host at this level, but one case has attracted considerable recent attention for possibly having done so, and this is *P. chromatophora*.

Paulinella is a genus of euglyphid amoeba, which are members of the Cercozoa (Bhattacharya et al. 1995). These are testate amoebae with shells built from silicous scales of great intricacy. Most euglyphids, including most members of the genus *Paulinella*, are non-photosynthetic heterotrophs that feed using granular filopods that emerge from an opening in their test (Johnson et al. 1988). However, one species, *P. chromatophora*, has lost its feeding apparatus and instead acquired a cyanobacterial endosymbiont that allows it to live without feeding: each *P. chromatophora* cell contains two kidney-bean-shaped cyanobacterial symbionts, called chromatophores (figure 2), and their division is synchronized with that of the host cell cycle so that each daughter amoeba retains the symbiont (Kies 1974; Kies & Kremer 1979). Early observations led to the suggestion that chromatophores were cyanobacteria related to *Synechococcus*, and molecular phylogenies later confirmed that they are related to the *Synechococcus*/Prochlorococcus lineage (Marin et al. 2005; Yoon et al. 2006). Early work also demonstrated that the symbiont transferred photosynthetic genes to the amoeba host (Kies & Kremer 1979), altogether suggesting that the symbiont, called a chromatophore, was at least functionally equivalent to a plastid.

There has been a lot of debate over whether the chromatophore should be called an ‘organelle’ or a ‘plastid’, or an ‘endosymbiont’ (Theissen & Martin 2006; Yoon et al. 2006; Bhattacharya et al. 2007; Budylov et al. 2007). To some extent, this is semantic, but in other ways the distinction in what we call it is important because it does affect the way we think about organelles. If we define ‘organelles’ in a narrow way, for instance restricted to cases in which a protein-targeting system has evolved, then we will inevitably come to the conclusion that ‘organelles’ can only originate or integrate in certain ways. All other cellular bodies will be given a different name and will have less impact on our thinking on organellogenesis. It could be beneficial to leave the definition a little more open, as two cells can become highly integrated in more ways than protein targeting. If, for example, an endosymbiont becomes dependent on its host for control over the division and segregation of the endosymbiont, then one might consider the endosymbiont to be an organelle, which allows us to see the potential variation in the way that organellogenesis can take place, and two cells become integrated.

Returning to the chromatophore, the debate over whether or not it should be called an organelle has been stoked by the complete sequence of its genome (Nowack et al. 2008). As previously indicated by phylogenetic analysis of single genes, the genome clearly shows the chromatophore to be a member of the *Synechococcus*/Prochlorococcus lineage (Marin et al. 2005; Yoon et al. 2006), but rather than the roughly 3 Mbp and 3300 genes common to members of the genus *Synechococcus*, the chromatophore genome is a mere 1 Mbp in size and encodes only 867 genes (Nowack et al. 2008). Reduction has mostly taken place by the elimination of whole pathways and functional classes of genes. This reduction is different from that seen in plastids, because not only is it less severe (the chromatophore has more than four times the number of genes encoded by even the largest known plastid), but also in the tendency for whole pathways to be lost or kept: in plastids, most pathways that are retained are incomplete because many or most genes have moved to the nucleus (Nowack et al. 2008). This, at face value, suggested that there was no protein targeting or gene transfer (Keeling & Archibald 2008). However, analysis of expressed sequence tags from the *Paulinella* nuclear genome found a copy of *psaE* (Nakayama & Ishida 2009). The gene is phylogenetically related to the *Synechococcus*/Prochlorococcus lineage, and is missing from the chromatophore genome, so is most likely a case of transfer from the symbiont to the host (Nakayama & Ishida 2009). Whether the *PsaE* protein is targeted back to the chromatophore is not known, and intriguingly there is no evidence for an amino-terminal protein-targeting extension. It is possible that the *psaE* gene is non-functional, functions in the host (which seems unlikely given the gene), or that targeting has evolved by a completely new mechanism, which would not be surprising given it is an independent endosymbiosis. *Paulinella* is a fascinating system that will doubtless receive much more attention, and if it does prove to be a fully integrated ‘organelle’, determining how its evolution has paralleled that of canonical plastids and how it has differed will both provide valuable comparisons.
3. SECONDARY ENDOSYMBIOSIS AND THE RISE OF PLASTID DIVERSITY

As mentioned previously, primary plastids are found in glaucophytes, red algae and green algae (from which plants are derived). These groups represent a great deal of diversity and are collectively significant ecologically, but they only represent a fraction of eukaryotic phototrophs. Most algal lineages acquired their plastids through secondary endosymbiosis, which is the uptake and retention of a primary algal cell by another eukaryotic cell (Delwiche 1999; McFadden 2001; Archibald & Keeling 2002; Stoebel & Maier 2002; Palmer 2003; Keeling 2004, in press; Archibald 2005; Gould et al. 2008). The plastids in most algal lineages can be attributed to this process, namely those of chlorarachniophytes and euglenids, which acquired plastids from green algae, and haptophytes, cryptomonads, heterokonts, dinoflagellates and apicomplexans, which acquired a plastid from a red alga (figure 2). Overall, this secondary spread of plastids had a major impact on eukaryotic diversity, evolution and global ecology (Falkowski et al. 2004). Many of the lineages with secondary plastids have grown to dominate primarily production in their environment, and collectively they represent a significant fraction of known eukaryotic diversity: it has been estimated that the lineage encompassing all the red algal secondary plastids alone represents over 50 per cent of the presently described protist species (Cavalier-Smith 2004).

Our understanding of how such a process might have taken place is actually somewhat more clear than our understanding of how primary endosymbiosis might have played out, in part because the events were more recent, but more importantly because it happened more than once so parallels and differences can be compared. As was the case for the cyanobacterium in primary endosymbiosis, the secondary endosymbiotic algae progressively degenerated until all that remained in most cases was the plastid, and one or two additional membranes around it. Essentially, no trace of mitochondria, flagella, Golgi, endoplasmic reticulum (ER) or many other cellular features remain in these endosymbionts. In most lineages (see §4 for exceptions), the algal nucleus is completely absent as well, and any genes for plastid proteins that it once encoded have moved once again to the nuclear genome of the secondary host (it seems likely that many other genes also moved to that genome, but this has not been investigated very thoroughly). In general, it is thought that secondary endosymbiosis takes place through endocytosis, so the eukaryotic algae is taken up into a vacuole derived from the host endomembrane system (Delwiche 1999; McFadden 2001; Archibald & Keeling 2002; Keeling 2004; Archibald 2005; Gould et al. 2008). After reduction, such a plastid would be predicted to be surrounded by four membranes, corresponding to (from the outside inward) the host endomembrane, the plasma membrane of the engulfed alga and the two membranes of the primary plastid (McFadden 2001; Archibald & Keeling 2002; Keeling 2004; Gould et al. 2008). Most secondary plastids are indeed surrounded by four membranes, and there is abundant evidence that the outermost membrane is derived from the host endomembrane system. Indeed, in cryptomonads, haptophytes and heterokonts, the outermost membrane is demonstrably contiguous with the host ER and nuclear envelope (Gibbs 1981). In other cases, the same situation is inferred from the way proteins are targeted (Foth et al. 2003; Gould et al. 2008; Bolte et al. 2009). Many plastid proteins are encoded in the new host nucleus, and these are targeted to the plastid post-translationally, but the process requires an additional step compared with the analogous process in primary plastids. Recall that targeting in primary plastids typically relies on the recognition of an N-terminal transit peptide by the TOC and TIC complexes on the outer and inner plastid membranes (McFadden 1999; Wickner & Schekman 2005; Hormann et al. 2007; Patron & Waller 2007; Gould et al. 2008). Secondary plastids are surrounded by additional membranes and are in effect situated within the endomembrane system of the secondary host (as opposed to the primary plastid, which is situated in the cytosol of its host), so any protein expressed in the cytosol is not directly exposed to the plastid outer membrane and could not be imported by the TIC–TOC system alone (McFadden 1999; Gould et al. 2008; Bolte et al. 2009; Kalanon et al. 2009). This has led to an additional leg in the journey from the cytosol to the plastid, and because these plastids are located in the endomembrane of their host, travel to the plastid is initially through the secretory system. As a result, proteins targeted to secondary plastids have a more complex, bipartite leader, which consists of a signal peptide followed by a transit peptide (Patron & Waller 2007). The signal peptide is recognized by the signal recognition particle, which stops translation and directs the protein to the rough ER (RER), where translation resumes and the nascent protein crosses the membrane as it is elongated. How the protein crosses the next membrane is only just emerging in some groups (Hempel et al. 2007; Sommer et al. 2007; Bolte et al. 2009), and it is not yet certain if the same mechanism is used by all secondary plastids, but once exposed to the inner pair of membranes, the transit peptide may be finally recognized by the TOC and TIC complexes and import completed.

The origin of membranes in four-membrane secondary plastids presents few mysteries and corresponds well with what is known about protein trafficking across those membranes. However, in euglenids and dinoflagellates, the secondary plastid is bounded by only three membranes (figure 2). This raised questions as to whether one membrane had been lost, and if so which one, or whether these plastids were derived by a different process. Specifically, it has been proposed that these plastids were acquired by myzocytosis, a feeding strategy whereby a predator attaches to a food cell and sucks its contents into a feeding vacuole, rather than engulfing the whole cell (Schnepf & Deichgraber 1984). Myzocytosis is known in dinoflagellates and euglenids, and such a process could indeed explain the membrane topology, but it remains unclear how an alga taken up by...
myzocyto
tic and division with divide its plasma membrane. Moreover, it is now abundantly clear that the three-membrane dinoflagellate plastid is orthologous to the four-membrane plastid in api
complexes (Moore et al. 2008; Keeling in press), so the three-membrane topology of dinoflagellate and eu
genid plastids must be the result of some other con
tergent evolutionary pathway (Lukes et al. 2009). If the process of plastid origins is the same for three-
and four-membrane secondary plastids (as must be the case in dinoflagellates), then which membrane was lost and why? There is no certain answer to this, but by a process of reduction it must have been the second membrane, corresponding to the plasma mem
brane of the engulfed alga. The rationale for this conclusion is that all other membranes are crossed by a known mechanism in protein targeting, and removing any one of them would have predictable and disastrous effects on trafficking. For example, as the outermost membrane is part of the host endomembrane system and protein trafficking uses the first steps in the secretion pathway to target plastid proteins to this compartment, the loss of this membrane would mean plastid-targeted proteins would be diverted to the secretory pathway. Similarly, the two inner membranes are involved in plastid function, and necessary to transit peptide recognition, so probably cannot be lost. In contrast, no satisfying explanation for why the second membrane is retained has been proposed (and we can therefore more readily imagine losing it), and the mechanism thought to be used to cross it in some groups is not clearly incompatible with loss (Bolte et al. 2009). This is not to say that the loss has not impacted targeting, because it has. Intrigu
ingly, the plastid-targeted proteins of both engul
dinoflagellates have targeting peptides that are different in some respects from those of other second
ary plastids (Sulli et al. 1999; Nassoury et al. 2003), and they share similar differences (Patron et al. 2005; Durnford & Gray 2006), despite having acquired their plastids independently from a green and red alga, respectively.

4. NUCLEOMORPHS

Another exception that has emerged more than once is the retention of a relict nucleus of the secondary endo
symbiotic alga, structures called nucleomorphs (Gilson & McFadden 2002; Archibald 2007). In most cases, the secondary endosymbiont nucleus is completely lost, presumably owing to the movement of all genes necessary for the upkeep of the plastid to the nucleus of the new secondary host. In cryptomo
nads and chlorarachniophytes, however, this algal nucleus has persisted (figures 1 and 2) and has been the focus of much attention. When these tiny structures were first described using transmission electron microscopy (TEM) (Greenwood et al. 1977; Hibberd & Norris 1984), the process of secondary endosymbiosis was not appreciated and the distri
bution of plastids correspondingly difficult to understand. The demonstration that cryptomonad and chlorarachniophyte plastid compartments were associated with a eukaryotic nucleus and residual cytoplasm with 80S ribosomes (Ludwig & Gibbs 1989; Douglas et al. 1991; McFadden et al. 1994) was a galvanizing discovery that ushered in widespread acceptance that eukaryote-eukaryote symbiosis was an important part of plastid evolution.

Not surprisingly, attention soon turned to the gen
omes retained in nucleomorphs. As the cryptomonad plastid is clearly derived from a red alga and the chlor
arachniophyte plastid from a green alga, their nucleomorphs must have evolved independently from fully-fledged algal nuclei, but they were soon found to share a number of superficial similarities. In all spe
cies examined to date, the nucleomorph genome is composed of three small chromosomes, for a total genome size from as little as 373 kbp to over 650 kbp (Rensing et al. 1994; McFadden et al. 1997a; Gilson & McFadden 1999, 2002; Gilson et al. 2006; Archibald 2007; Silver et al. 2007). In nearly all cases, the rRNA operons are found as subtelomeric repeats on all six chromosome ends, although they face in opposite directions in some species (Gilson & McFadden 1996, 2002; Archibald 2007). Between these repeats, the chromosomes are ‘jam-packed’ with genes, the sequenced genomes have gene densities of about one gene per kilobase, the highest known density for a nuclear genome. This tight organization has apparently affected gene expression in nucleomorphs in both cryptomonads and chlorarachniophytes, so that there is now a high frequency of overlapping tran
scription: in Guillardia theta, nearly 100 per cent of characterized transcripts either begin in an upstream gene, terminate within or beyond a downstream gene, or both (Williams et al. 2005).

While these similarities certainly suggest that gen
omes in both lineages have been under similar pressures and constraints or been affected by similar modes of evolution, there are more differences the deeper one digs. Most importantly, there is no real pattern to the actual genes retained in the nucleomorphs. When nucleomorphs were first discovered, it was thought that they might harbour an extensive collection of genes for plastid-targeted proteins, but in the complete genomes sequenced to date, plastid-targeted protein genes are relatively scarce: only 17 in Bigelowiella natans (Gilson et al. 2006) and 30 in the cryptomonads (Douglas et al. 2001; Lane et al. 2007) have been identified. Moreover, there is no significant overlap in the identity of these genes, rather they seem to be two random subsets of possible plastid proteins (Gilson et al. 2006). Other interesting differences have been found in how these genomes have reacted to whatever process led to their severe reduction and compaction. For example, introns in cryptomonad nucleomorphs are not unusual in size or sequence, but they are extremely rare in number: the G. theta genome has only 18 introns (Douglas et al. 2001; Williams et al. 2005), and the nucleomorph of Hemiselmis anderseni has lost them altogether (Lane et al. 2007). In contrast, chlorarachniophyte nucleo
morph genes are riddled with introns: the B. natans genome retains over 800 identified introns, and most seem to be ancient introns conserved with green algae and other chlorarachniophytes (Gilson et al. 2006). These introns, however, have dramatically
reduced in size so that all known cases in *B. natans* are between 18 and 21 bp, and the majority are 19 bp, a situation more or less conserved in other chlorarachniophytes (Slamovits & Keeling 2009).

The majority of genes in nucleomorph genomes are housekeeping genes responsible for the maintenance and expression of the genome itself, although in no case is the complement of genes in a nucleomorph genome sufficient for all necessary functions (Douglas et al. 2001; Gilson et al. 2006; Lane et al. 2007). Many genes are inferred to have moved to the host nucleus, and are presumably targeted back to the cytoplasm surrounding the nucleomorph (the periplastid compartment or PPC), or the nucleomorph itself. PPC-targeted proteins are interesting from a protein trafficking perspective, because in most secondary plastids there is little activity between the two pairs of membranes, and probably few genes targeted to this compartment. In cells with a nucleomorph, on the other hand, many genes appear to be targeted to this compartment and, as this is half way to the plastid, how they are targeted is an interesting question that could help resolve how proteins cross the second membrane (see above). The first such genes to be identified were in the cryptomonad, *G. theta*, and the leaders were shown to include a signal peptide and a transit peptide (Gould et al. 2006a,b). Intriguingly, the transit peptides share a single common feature, which was the lack of a phenylalanine residue immediately downstream of the signal peptide. This residue is part of a motif common to transit peptides of glaucophytes, red algae and all red-algal-derived plastids (Patron & Waller 2007). Adding a phenylalanine to this position of the PPC-targeted proteins led to their re-targeting to the plastid, suggesting that this position plays an important role in distinguishing PPC proteins from plastid proteins (Gould et al. 2006a,b), a function that seems to be directed by a derivative of the ERAD complex (Sommer et al. 2007; Bolte et al. 2009).

Interestingly, targeting to secondary green plastids does not rely on the F-residue (Patron et al. 2005; Patron & Waller 2007), so this information cannot participate even partially in the distinction between PPC- and plastid-targeting. Only two putative PPC-targeted proteins have been identified in chlorarachniophytes, for the translation factors EFL and eIF1 (Gile & Keeling 2008). Once again, the targeting information on these proteins appears to consist of a signal peptide followed by a sequence with all the characteristics of a transit peptide. Interestingly, the only difference between PPC-targeted and plastid-targeted protein leaders is the presence of an acid-rich domain at the C-terminus of the transit peptide (Gile & Keeling 2008). The *B. natans* genome is presently being sequenced, and will presumably yield a long list of PPC-targeted proteins whose targeting information, together with the ability to transform *Lotharella amoeboidformis* (Hirakawa et al. 2008), will shed additional light on this problem.

The nucleomorphs of cryptomonads and chlorarachniophytes are relatively well-studied components of a eukaryotic endosymbiont, and they have evolved along remarkably similar lines. But this does not mean such a path is inevitable. Indeed, in the tertiary endosymbiotic partnership between a diatom and dinoflagellates such as *Kryptoperidinium* and *Durinschia* (see below), the endosymbiont nucleus and its genome have not reduced at all, but grossly expanded (Kite et al. 1988). These endosymbionts are fully integrated in the host cell cycle, but are poorly studied at the molecular level and it is not known if any genetic exchange has taken place. Because of their enormous size, the nuclei are not typically referred to as nucleomorphs, but in some ways the situation is similar to the arguments about the *Paulinella* chromatophore and whether it should be called a plastid: the interesting point is not so much the name, but that the organelle has followed a different evolutionary path than have those in cryptomonad and chlorarachniophyte endosymbionts.

5. HOW MANY TIMES HAVE PLASTIDS MOVED BETWEEN EUKARYOTES?

We know secondary endosymbiosis has happened on multiple occasions, because both green and red algal endosymbionts are known, but exactly how many times secondary endosymbiosis has taken place has been a subject of ongoing debate.

On the green side, the question is more or less settled. Secondary green algal plastids are known in euglenids and chlorarachniophytes, and there is no strong similarity between the two. Euglenids have three-membrane plastids and store paramylon in the cytosol, whereas chlorarachniophytes have four-membrane plastids with a nucleomorph and store beta-1–3-glucans in the cytosol (McFadden et al. 1997b; Leedale & Vickerman 2000; Ishida et al. 2007). The hosts are also different, and although the chlorarachniophytes have only recently found a home in the tree of eukaryotes with Cercozoa (Bhattacharya et al. 1995; Cavalier-Smith & Chao 1997; Keeling 2001; Nikolaev et al. 2004; Burki et al. 2009), it was clear early on that euglenids were similar to trypanosomes (Leedale & Vickerman 2000), and not chlorarachniophytes. In keeping with this, early molecular phylogenies did not support a close relationship between the two groups. Nevertheless, it was proposed that their plastids shared a common ancestor, in the so-called Caboza hypothesis (Cavalier-Smith 1999). The rationale behind this hypothesis was that secondary endosymbiosis and the evolution of a protein-targeting system in particular is a complex process, and hypotheses for plastid evolution should severely limit the number of times this would have taken place. However, analysis of complete plastid genomes has refuted this proposal, because chlorarachniophytes and euglenid plastid genomes have now been shown to be specifically related to different green algal lineages (Rogers et al. 2007a; Turmel et al. 2009), and so could not have been derived from a single endosymbiosis (figure 2).

The evolution of secondary plastids from red algae is far more complex and a much greater number of lineages are involved. Secondary red algal plastids are found in cryptomonads, haptophytes, heterokonts, dinoflagellates and apicomplexans. Although these
organisms represent a great deal of plastid diversity (some have nucleomorphs, some are non-photosynthetic, some have three membranes, etc.), and although molecular phylogenies originally grouped neither nuclear nor plastid lineages together, again it was proposed that their plastids arose from a single common endosymbiosis, an idea known as the chromalveolate hypothesis (Cavalier-Smith 1999). In contrast to the Cabozoa hypothesis, data eventually emerged to support the chromalveolate hypothesis and, although still contentious, a variation on this hypothesis is gaining some general acceptance (figure 2). The evidence for this relationship and controversies surrounding it was recently reviewed elsewhere (Keeling in press), but as data are emerging quickly and because it does represent a major fraction of plastid diversity, the evidence will be summarized briefly here.

Plastid gene phylogenies have supported the relationship of heterokonts, haptophytes and cryptomonads to the exclusion of what few red algal lineages have been sampled (Yoon et al. 2002; Hagopian et al. 2004; Khan et al. 2007; Rogers et al. 2007a), but the strange plastid genomes of apicomplexans and dinoflagellates have essentially excluded them from such analyses. In addition, two nucleus-encoded genes for plastid-targeted proteins, glyceraldehyde-3-phosphate dehydrogenase and fructose bisphosphate aldolase, have also supported the common origin of chromalveolate plastids because the chromalveolate genes have a unique evolutionary history that differs from homologues in other plastids (Fast et al. 2001; Harper & Keeling 2003; Patron et al. 2004). Nuclear gene phylogenies have typically not supported a monophyletic chromalveolate (e.g. Kim & Graham 2008), including analyses of some large multi-gene datasets (Patron et al. 2004; Hampl et al. 2007, 2008). However, other large multi-gene analyses (e.g. Hackett et al. 2007; Hampl et al. 2009; Burki et al. 2009) have recovered a monophyletic chromalveolate, with one important provision: that it also includes Rhizaria. In the relatively short time since the monophyly of Rhizaria was discovered (Cavalier-Smith & Chao 1997; Keeling 2001; Archibald et al. 2002; Nikolaev et al. 2004), they have been considered a supergroup in their own right, but a string of large-scale analyses have consistently shown that they are closely related to alveolates and heterokonts (Burki et al. 2007, 2008, 2009; Hackett et al. 2007; Hampl et al. 2009), and were recently shown to share a novel class of Rab GT-Pase (Elias et al. 2009). If all of these observations are correct, it also means the ancestor of Rhizaria possessed the red algal plastid that is still present in many chromalveolates. This is interesting, not least because two rhizarian lineages are today photosynthetic: chlorarachniophytes and Paulinella, but they have acquired their plastids more recently and from different sources (see above). If Rhizaria are derived from an ancestrally photosynthetic chromalveolate, then these two lineages have reverted to phototrophy by entering into new symbioses (figure 2).

Although the single origin of all chromalveolate plastids remains in question, the ancestral state of two large and diverse subgroups has become much more clear in recent years. First, many large multi-gene phylogenetic analyses, as well as the shared presence of a rare horizontal gene transfer in the plastid genome, have shown that cryptomonads and haptophytes are sisters (Rice & Palmer 2006; Burki et al. 2007, 2008, 2009; Hackett et al. 2007; Hampl et al. 2009; Okamoto et al. 2009). At the same time, they have been shown to be related to a number of more poorly studied lineages, including several that are non-photosynthetic (Okamoto & Inouye 2005; Not et al. 2007; Cavalier et al. 2008; Burki et al. 2009; Okamoto et al. 2009). This fast-growing group of increasing importance and diversity has recently been called the Hacrobia (Okamoto et al. 2009). The other group where the ancestral state is now well established is the apicomplexans and dinoflagellates. Since the discovery of the apicomplexan plastid, there has been a long-running debate over the ancestry of plastids in these two lineages, and indeed whether the apicomplexan plastid is derived from a red or green alga (Williamson et al. 1994; Köhler et al. 1997; Roberts et al. 1998; Keeling et al. 1999; Fast et al. 2001; Funes et al. 2002; Waller et al. 2003). This debate is a difficult problem, because the plastid genomes of these two lineages are both highly derived and share almost no genes in common, so they are nearly impossible to compare directly (Keeling 2008). However, the discovery of Chromera velia, the photosynthetic sister to Apicomplexa (figure 1), changes this completely (Moore et al. 2008). Chromera forms a link between these two unusual lineages so that the initial characterization of molecular data (Moore et al. 2008) and now the complete sequence of its plastid genome (J. Janousková, A. Horák, M. Obernink, J. Luks & P. J. Keeling 2009, unpublished data) completely eliminate any basis for the green origin of the apicomplexan plastid, and firmly supports the presence of a plastid in the common ancestor of apicomplexans and dinoflagellates.

6. PLASTID LOSS AND CRYPTIC PLASTIDS

One of the outcomes of the chromalveolate hypothesis has been an increased interest in the process of plastid loss and the prevalence of cryptic, non-photosynthetic plastids in heterotrophic lineages. This is because the early origin of the secondary red algal plastid central to the chromalveolate hypothesis requires that a number of currently non-photosynthetic lineages must have had photosynthetic ancestors. To understand the implications of this aspect of the hypothesis, it is necessary to consider several things carefully, in particular the difference between plastid loss and the loss of photosynthesis, and also how well we actually understand the distribution of cryptic plastids.

The difference between losing photosynthesis and losing a plastid may seem straightforward, but the distinction is surprisingly often ignored. This is problematic, because the available evidence suggests that the likelihoods of the two processes are quite different. Photosynthesis has been lost many times, and there is at least one case of this in nearly all photosynthetic groups (William & Hirt 2004; Krause 2008). Land plants have lost photosynthesis at least a dozen times (Nickrent et al. 1998), and dinoflagellates
are equally likely to dispense with the process (Saldarriaga et al. 2001; Hackett et al. 2004a). In many of these lineages, the plastid is readily detectable (e.g. Sevpens-vol 1973; Siu et al. 1976; Nickrent et al. 1998; Williams & Keeling 2003; Krause 2008), whereas in others it has been harder to detect (see below). Plastid loss, on the other hand, could be viewed as an extreme subset of cases where photosynthesis has been lost, and it is an apparently rare subset that is very difficult to demonstrate: non-photosynthetic plastids can be challenging to detect, and showing they are absent is substantially harder again (Williams & Hirt 2004; Krause 2008). No plastid has unambiguously been demonstrated to exist without a genome (but see Nickrent et al. 1997), but it remains a possibility (mitochondria have lost their genome many times; Williams & Keeling 2003; van der Giezen et al. 2005; Hjort et al. 2010); so one could argue that a complete nuclear genome sequence is required to confidently conclude that the cell lacks a cryptic plastid, and all associated plastid-targeted proteins. Moreover, for plastid loss to have occurred, the ancestor of a lineage must have once had a plastid. While this may seem trivial, it is an important aspect of the argument in many cases. For example, within the chromalveolates, we can concretely state that there is no plastid in those ciliates and oomycetes with complete genomes (Aury et al. 2006; Eisen et al. 2006; Tyler et al. 2006), but whether their ancestors had a plastid is still open to debate, and we cannot be certain of plastid loss until this is demonstrated clearly (although a case has been made for both groups that relict, endosymbiont-derived genes have been retained; Tyler et al. 2006; Reyes-Prieto et al. 2008). Conversely, there is strong evidence that the ancestors of katablepharids contained a plastid (Patron et al. 2007), but we lack sufficient data from these non-photosynthetic heterotrophs to confidently conclude that a cryptic plastid does not still exist in this lineage. Indeed, Cryptosporidium is arguably the only lineage of eukaryotes that can be concluded to have lost its plastid outright. It is presently unique among all eukaryotes in that the structural and genomic evidence is available to conclude that this organism does not contain a plastid (Abrahamsen et al. 2004; Xu et al. 2004), and at the same time the evolutionary evidence is also available to state that its ancestors did contain a plastid (Keeling 2008; Moore et al. 2008). This is not to say that this is the only lineage in which plastids have been lost: it is likely that many non-photosynthetic lineages have completely lost plastids, as well as many dinoflagellates with tertiary plastids (see below).

In all remaining non-photosynthetic lineages where some evolutionary argument for a plastid-bearing ancestry has been made, cryptic plastids have either now been found, or the presence or absence of a plastid is simply unknown. Indeed, evidence for cryptic plastids is now emerging in several lineages where phylogenetic relationships had suggested a possible plastid ancestry. The best characterized example is the so-called apicoplast of apicomplexan parasites, where the plastid may have been discovered relatively recently (McPadden et al. 1996; Wilson et al. 1996; Köhler et al. 1997), but is already among the best studied plastids (for reviews, see Wilson 2002; Ralph et al. 2004; Lim & McFadden 2010). In another previously enigmatic lineage of parasites, the Helicosporidia, a similar story has unfolded: while these were previously believed to be related to parasites such as Apicomplexa, molecular phylogenetic analysis showed they are in fact green algae (Tartar et al. 2002). This led to the suggestion that they contain a cryptic plastid, the presence of which was confirmed by identifying the plastid genome (Tartar et al. 2003; de Koning & Keeling 2006), and several nuclear genes for plastid-targeted proteins (de Koning & Keeling 2004). Despite these advances, the organelle itself has yet to be identified.

A less complete picture is also forming for two non-photosynthetic sister lineages to dinoflagellates. In Perkinsus, genes for several plastid-derived proteins have been characterized and there is also evidence for the presence of an actual organelle (Grauvogel et al. 2007; Stelter et al. 2007; Teles-Grilo et al. 2007; Matsuzaki et al. 2008). In Oxyrrhis marina, another non-photosynthetic sister to dinoflagellates, several plastid-derived genes have also been found, and some of these have leaders which suggest that the proteins are targeted to an as-yet unidentified organelle, although interestingly other plastid-derived proteins no longer appear to be targeted (Slamovits & Keeling 2008).

The functions of these organelles are also known or partially known in many cases, and there are many parallels. Where the genome of a cryptic plastid is known (i.e. those of apicomplexans and helicosporidia), they give few clues as to the function of the organelle, and most of the functionally significant genes are nucleus-encoded genes whose products are targeted to the plastid. The cryptic plastid that is best characterized functionally is the apicoplast, which is involved in synthesis of fatty acids, isoprenoids and haem, although there is some variability between different species (Wilson 2002; Foth & McFadden 2003; Ralph et al. 2004; Sato et al. 2004; Goodman & McFadden 2008; Gould et al. 2008). "In Helicosporidium, representative genes involved in all these pathways have also been found, as well as genes in several amino acid biosynthetic pathways and genes involved in controlling redox potential (de Koning & Keeling 2004), and a similar complement of putative plastid-targeted proteins has also been found in Prototheca (Borza et al. 2005), a non-photosynthetic green alga that is closely related to Helicosporidium. Smaller numbers of genes are known from Oxyrrhis and Perkinsus, but they too represent subsets of these pathways (Grauvogel et al. 2007; Stelter et al. 2007; Teles-Grilo et al. 2007; Matsuzaki et al. 2008; Slamovits & Keeling 2008), suggesting that these represent the core functions of diverse cryptic plastids.

The genomes of non-photosynthetic plastids are now known from a variety of lineages, and in many cases can be compared with close relatives that are photosynthetic. The genomes of the non-photosynthetic plastids are predictably reduced in size (so far, they are consistently smaller than those of their closest photosynthetic relatives: figures 2 and 3), but they generally retain some or all of the common features

Phil. Trans. R. Soc. B (2010)
of a plastid genome, such as the inverted repeat or the ribosomal protein operons. As they have lost photosynthesis, one would expect all the genes related to this process to be gone as well, and in several cases they are \( \text{Wilson et al. 1996; de Koning & Keeling 2006.} \]

In other cases, however, some of the genes that have been retained are of interest given the absence of photosynthesis, such as ATP synthase genes in \( \text{Prototheca} \) or rubisco subunits in non-photosynthetic heterokonts and plants \( \text{(Wolfe \\& dePamphilis 1997; Knauf \\& Hachtel 2002; Sekiguchi et al. 2002; McNeal et al. 2007; Barrett \\& Freudenstein 2008; Krause 2008).} \]

**7. TERTIARY ENDOSYMBIOSIS: A DINOFLAGELLATE ODDITY**

In dinoflagellates, another layer of endosymbiotic complexity has been added to the evolutionary history of plastids: tertiary endosymbiosis. The ancestor of dinoflagellates already had a secondary endosymbiotic plastid of red algal origin; however, in a few dinoflagellate lineages, this plastid is either gone or at least no longer photosynthetically active, and primary production is carried out by a new plastid that has been derived from another lineage with a secondary plastid of red algal origin. So far, dinoflagellates are known to have acquired such tertiary plastids from cryptomonads \( \text{(Schnepf \\& Elbracht 1988), haptophytes (Tengs et al. 2000) and diatoms (Dodge 1969; Chesnick et al. 1997), and they have also acquired a new secondary endosymbiont from the green algal lineage in one case called 'serial secondary endosymbiosis' (Watanabe et al. 1990).} \]

Relatively little is known about the process of tertiary endosymbiosis, and the level of genetic integration appears to be variable. For example, the haptophyte endosymbiont of \( \text{Karolodinium and Karenia} \)
has been reduced to a similar extreme as most secondary plastids, so that all that remains is the plastid itself, and perhaps some additional membranes (Tengs et al. 2000). There is no evidence for the retention of the original dinoflagellate plastid, and the new plastid is serviced by many nucleus-encoded genes and a plastid-targeting system (Nosenko et al. 2006; Patron et al. 2006); so in this case the process appears to have completely substituted one fully integrated plastid for another. In contrast, the diatom endosymbiont found in another lineage of dinoflagellates (including the genera *Kryptoperidinium*, *Durinskia* and others) is far less reduced (figure 1): this endosymbiont retains a large nucleus with protein coding genes, a substantial amount of cytosol and even its original mitochondria—a condition unique among eukaryotes (Dodge 1969; Chesnick et al. 1997; McEwan & Keeling 2004; Imanian & Keeling 2007).

It is unknown whether the dinoflagellate nucleus encodes genes for proteins targeted to the diatom plastid, but it may not be necessary given that the diatom nucleus is present and shows no signs of reduction. Moreover, in at least some of these genera, the original plastid is also thought to have been retained in the form of a three-membrane-bounded eyespot (Dodge 1969). Unlike *Karenia* and *Karenia*, where tertiary endosymbiosis led to an outright substitution of organelles, in this case the process has resulted in a considerable degree of redundancy and sub-functionalization (Imanian & Keeling 2007).

Tertiary endosymbiosis also presents another wrinkle in the evolution of plastid proteins because, in these events, a new plastid is introduced into a lineage that is already photosynthetic, or at least had photosynthetic ancestors and might have retained a cryptic plastid. During the integration of the new plastid, if and when a protein-targeting system was established and genes moved from the tertiary endosymbiont nucleus to the host nucleus, there might already be a number of plastid-derived genes in the host nucleus. What happens to the old genes? Are they ‘overwritten’ by the incoming genes that are better suited to the compartment where they have always functioned, or do some of the dinoflagellate plastid genes survive by re-compartmentalizing to the new plastid? Expressed sequence tag surveys of both *Karlodinium* and *Karenia* show that both kinds of events take place: most plastid-targeted proteins are derived from the haptophyte plastid lineage, and therefore probably came from the tertiary endosymbiont (although see below), but a significant fraction are derived from the dinoflagellate plastid lineage, and therefore have been re-directed to the new plastid (Nosenko et al. 2006; Patron et al. 2006). This also suggests that the old plastid coexisted with the new one, or plastid-targeted genes might have been lost before they could be targeted to the new plastid. Interestingly, the targeting peptides in these lineages bear little similarity to those of either dinoflagellates or haptophytes, so the integration of the tertiary plastid must have led to upheaval in the protein trafficking system (Patron et al. 2006). It is possible that this upheaval allowed dinoflagellate proteins to supplant the incoming haptophyte proteins that were presumably better adapted to the organelle because both classes of proteins were equally likely to coadapt with the changing trafficking system.

Tertiary endosymbiosis is also important because the events are much less ancient than secondary and primary endosymbiotic plastid origins, and can offer a window into the process that shows not only diverse outcomes, but also perhaps evidence of transient states that have vanished in more ancient events. One of the most interesting of these relates to the shift between a transient host–symbiont association and the fixation of a permanent association. Although it is often characterized as a sudden event, endosymbiotic associations are more probably integrated gradually over long periods of time. It is likely that a host becomes adapted to associations with particular kinds of symbionts, and after a long period of transient associations, which may even include gene transfer and protein targeting, some event took place which made the association permanent (perhaps a shift in the control over cell division of the endosymbiont). The exact order of events was quite probably different in the fixation of different endosymbiotic organelles, but most cases probably involved an intermediate stage based on long-term but non-permanent associations. Although this seems likely, there is no direct evidence for this state in the highly integrated primary and secondary endosymbiotic organelles that remain today. In the case of tertiary endosymbionts, however, important evidence of this critical period does remain in at least two cases. The dinoflagellate hosts with haptophyte plastids are all closely related, as are the dinoflagellate hosts that took up diatoms. However, in both cases, the endosymbionts are derived from different lineages of haptophyte and diatom, respectively. At least two different haptophyte endosymbionts have been fixed, one in *Karlodinium* and a different one in *Karenia* (Gast et al. 2007). Similarly, at least three different diatom endosymbionts have been fixed, one centric and two distantly related pinnate diatoms (Horiguchi & Takano 2006; Takano et al. 2008). This implies that there was a period where these dinoflagellate lineages were forming transient relationships with particular kinds of algae, and that in multiple subgroups these associations were fixed with different symbionts. From these lineages, we ought to be able to detect events relating to integration that come from the illusive transient period (centric diatom-derived plastid-targeted genes in a lineage with a pennate diatom, for example), but currently there are few data of a comparative nature from these groups.

**8. HORIZONTAL TRANSFER OF PLASTID GENES**

Endosymbiosis and organelle evolution obviously involve a great deal of movement of genetic information, but the movement of genes between a host and symbiont is only one special variety of horizontal gene transfer. More generally, transfers can also take place between more transiently associated cells, food, vector transmission or perhaps just DNA from the environment (Gogarten et al. 2009). Whatever the mode of transfer, it is growing clear that eukaryotes can and have acquired genes from a variety of sources beyond their organelles (Keeling & Palmer 2008),

*Phil. Trans. R. Soc. B* (2010)
which raises an interesting question: have the organelles themselves been affected by horizontal gene transfer? In the case of the mitochondrion of plants, a great deal of transfer has been described, making it perhaps the most promiscuous class of cellular genomes known (Berghorsson et al. 2003, 2004). This means that, for plants at least, the mechanism exists to transfer genes between organelles from distinct species, so what about the plastid? Curiously, analysis of plant plastids not only failed to show the same high level of transfer as found in mitochondria, but revealed no evidence of transfer whatsoever (Rice & Palmer 2006). (One genus of parasitic plants has been found to encode a fragment of the plastid genome from another plant genus, but it is not known to reside in the plastid itself; Park et al. 2007.) This is thought to reflect the fact that plant mitochondria fuse when cells fuse, but plastids do not, a reasonable limitation of transfer in plants (Berghorsson et al. 2003, 2004; Rice & Palmer 2006; Richardson & Palmer 2007). However, extending the search to all known plastids revealed a remarkable lack of evidence for transfer in general (Rice & Palmer 2006). Of all known plastid genes and genomes, a good case for horizontal acquisition of genes can only be made for a few genes and several introns. Intron transfers have been found to affect many plastid lineages, and involve several types of introns, including group I introns (Cho et al. 1998; Besendahl et al. 2000), group II introns (Sheveleva & Hallick 2004) and a subclass of group II introns called group III introns, in this case nested within a group II intron in a situation called ‘twintrons’. Twintrons and group III introns were originally thought to be unique to euglenid plastids (Hallick et al. 1993), but have been found to have transferred to the cryptomonad *Rhodomonas salina* (Maier et al. 1995).

The first transfer of plastid gene sequences to be described was the large and small subunits of rubisco (rbcS and rbcL). Plastids in green algae, plants and glaucophytes use a cyanobacterium-derived type I rubisco that seems to be native to the plastid. In green plastids, only rbcL is encoded in the plastid, and rbcS is nucleus-encoded. In contrast, red algal plastid genomes encode both subunits, but they appear to be derived from proteobacteria by horizontal gene transfer (Valentin & Zetsche 1989; Delwiche & Palmer 1996). An early analysis of rubisco sequences revealed this transfer, and further analysis based on hundreds of organelle and bacterial genomes has lent additional support to the original interpretation (Rice & Palmer 2006). The second gene for which strong evidence for horizontal gene transfer exists is rpl36, where the copy found in both haptophyte and cryptomonad plastid genomes is clearly a parologue of the copy found in all other plastids, and seems to be derived from an undefined bacterial lineage (Rice & Palmer 2006). Similarly, a bacterial *dhaX* gene was found in the plastid genome of *R. salina*, but not in any other plastid genome, including that of its close relative *G. theta* (Khan et al. 2007).

Plastid genomes may not be likely to acquire genes by horizontal transfer, but this does not mean the plastid proteome is not affected by the process, because most plastid proteins are encoded in the nucleus. Indeed, one of the first and most famous cases of lateral gene transfer to a eukaryote was the plastid-targeted rubisco in dinoflagellates. Dinoflagellates would have originally used the proteobacterial rubisco acquired by horizontal transfer by the ancestor of red algae, but in dinoflagellates the proteobacterial gene has itself been replaced by another horizontal transfer of a single-subunit type II rubisco from another proteobacterium (Morse et al. 1995; Whitney et al. 1995; Palmer 1996). The possible impact of horizontal transfer on the plastid proteome has now been explored in large-scale analyses in a few lineages. Analysis of expressed sequence tags from the chlorarachniophyte *B. natans* showed that about 20 per cent of genes for which the phylogeny was resolvable were not evidently derived from a chlorophyte green alga, the source of the secondary plastid in this species (Archibald et al. 2003). Instead, several genes were related to red algae or red algal secondary plastids, streptophyte green algae (once again, the rubisco small subunit (SSU) was found to have been transferred), or non-cyanobacterial lineages of bacteria, although the history of transfer of some of these genes is more complex than originally envisioned (Rogers et al. 2007b). Similarly, genome-wide analyses of genes for plastid-targeted proteins in several dinoflagellates have found several cases of such transfer in this lineage (Bachvaroff et al. 2004; Hackett et al. 2004b; Waller et al. 2006a,b; Nosenko & Bhattacharya 2007), and a variety of individual gene transfer events have also been recorded in haptophytes and diatoms (Armbrust et al. 2004; Obornik & Green 2005; Patron et al. 2006). Overall, it appears that some lineages are prone to replacing genes for plastid-targeted proteins, and others are not, perhaps a reflection of mixotrophic or purely autotrophic ancestry, as eating other algae would provide an obvious ongoing source for potential new genes.

9. CONCLUDING REMARKS—THE CONTRASTING HISTORIES OF TWO ORGANELLES

Returning to the contrasts between plastid and mitochondrial evolution raised in the Introduction, the different paths travelled by these organelles through time and eukaryotic biodiversity are now becoming clearer than ever. Mitochondrial evolution has undeniably led to a rich diversity of highly derived organelles, but the evolution of this organelle appears to have been entirely vertical and without dead-ends. In contrast, plastid evolution has resulted in an equally rich diversity of organelles, but has added layers of complexity because plastids have moved between host lineages on a number of occasions—a conservative estimate based on current data would be seven times: two green secondary endosymbioses, one red secondary endosymbiosis, one green serial secondary endosymbiosis and three tertiary endosymbioses. In addition, there is mounting evidence of at least one other independent origin of a ‘plastid’ in *Paulinella*, and clear evidence for complete and outright plastid loss in at least one (*Cryptosporidium*) and probably many other
lineages. These deviations and dead-ends have no analogue in mitochondrial evolution, where secondary transfer and outright loss have never been observed. This is probably due to a variety of contributing factors: while it is certainly likely that the degree of metabolic integration of the two organelles differs and more highly restricts the evolution of mitochondria, it also seems likely that the original distribution of mitochondria and plastids sets the two organelles up for differing histories. Plastids originated within one of the eukaryotic supergroups, so most eukaryotes would not ancestrally have possessed a plastid. This would, at least intuitively, seem to support the likelihood of secondary endosymbiosis as it suggests that there is some point to the process (i.e. the spread of a useful organelle to new lineages). Mitochondria, in contrast, originated in the common ancestor of all eukaryotes; all eukaryotes ancestrally possessed a mitochondrion. This questions the point of and, by extension, the likelihood of secondary transfer because the only possible recipient cells would already have a homologous organelle. Such a transfer would either replace one organelle with another very similar organelle, or would require that the host lost or substantially reduced its mitochondrion, and then ‘recovered’ it through secondary endosymbiosis. Even if a mechanism for such an event were known, it vastly reduces the possible range of recipient lineages as reduced mitochondria are a relative rarity overall and tend to be restricted to habitats in which aerobic mitochondria are equally rare, and mitochondrial loss is completely unknown.

I wish to thank Kevin Carpenter and Ken Ishida for providing micrographs, Ken Ishida for sharing unpublished data and Geoff McDadden and Brian Leander for providing valuable comments on figure 2. Work in the Keeling laboratory on plastid evolution is supported by grants from the Natural Sciences and Engineering Research Council of Canada, the Canadian Institutes for Health Research and a grant to the Center for Microbial Diversity at the Tula Foundation. P.J.K. is a fellow of the Canadian Institute for Advanced Research and a Senior Scholar of the Michael Smith Foundation for Health Research.

REFERENCES


rhizopod, Paulinella ovalis n. comb. (Filosea: Euglyphina).
Characterization of two putative protein translocation components in the apicoplast of Plasmodium falciparum. 
Keeling, P. J. 2001 Bonminieria and Cercopita are related in actin phylogeny: two orphans find a home?
Molecular Biology and Evolution 18, 1551–1557.
Keeling, P. J. 2004 The diversity and evolutionary history of plastids and their hosts. 
Nature Reviews Genet. 9, 605–618. (doi:10.1038/nrg2386)
Kim, E. & Graham, L. E. 2008 EEP2 analysis challenges the monophyly of Archaeplastida and Chromalveolata. 
Kite, G. C., Rothschild, L. J. & Dodge, J. D. 1988 Nuclear components in the apicoplast of 
Knauf, U. & Hachtel, W. 2002 The genes encoding subunits of ATP synthase are conserved in the reduced plastid genome of the heterotrophic alga Chlorella variabilis. 
Martin, W. 1999 A briefly argued case that mitochondria and plastids are descendants of endosymbionts, but that the nuclear compartment is not. Proceedings of the Royal Society of London B 266, 1387–1395. (doi:10.1098/rspb.1999.0792)


Sato, S., Clough, B., Coates, L. & Wilson, R. J. 2004

Saldarriaga, J. F., Taylor, F. J. R., Keeling, P. J. & Cavalier-Schnepf, E. & Deichgraber, G. 1984 ‘Myzocytosis,’ a kind of


Sato, S., Clough, B., Coates, L. & Wilson, R. J. 2004 Enzymes for heme biosynthesis are found in both the mitochondrial and plastid of the malaria parasite Plasmodium falciparum. Protist 155, 117–125. (doi:10.1078/1434461000169)


Phil. Trans. R. Soc. B (2010)
Phylogenetic analyses indicate that the 19th hexanoyloxy-


Wilson, R. J. 2002 Progress with parasite plastids. J. Mol. Biol. 319, 257--274. (doi:10.1016/S0022-2836(02)00303-0)


