Opinion TRENDS in Plant Science Vol.12 No.5

Shopping for plastids

Anthony W.D. Larkum[1], Peter J. Lockhart[2] and Christopher J. Howe[3]

[1] School of Biological Sciences, University of Sydney, Sydney, NSW 2006, Australia
[2] Allan Wilson Centre, Institute of Molecular Biosciences, Massey University, Palmerston North, New Zealand
[3] Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UK

Recent suggestions that endosymbionts in a diatom and an amoeba represent independent origins of plastids from those in plants and algae raise again the question of how many times plastids have evolved. In this Opinion article, we review the evidence for a single origin or multiple origins of primary plastids. Although the data are widely taken as supporting a single origin, we stress the assumptions underlying that view, and argue for a more cautious interpretation. We also suggest that the implicit view of plastids being acquired from single ancestors at a single point (or points) in time is an over-simplification.

History – the concept of chloroplasts as endosymbionts

It is more than 100 years since Andreas Schimper boldly suggested that chloroplasts are the result of a symbiosis between a photosynthetic organism and a non-photosynthetic host [1], an idea developed by Constantin Mereschkowsky in the early 20th century [2,3]. The demonstration in 1962 by Hans Ris and Walter Plaut [4] that chloroplasts contain DNA provided further support, and by the late 1960s, the idea that chloroplasts derived from cyanobacteria or something closely related to them was generally accepted [5–7]. Further confirmation came with chloroplast nucleotide sequence data from the late 1970s onwards [8].

The variation in plastid structure and light-harvesting pigments across plants and algae emphasized the question whether multiple primary endosymbioses should be invoked (a polyphyletic origin [9]), or whether a single one was sufficient (a monophyletic origin). The most widely held view today is that plastids have a monophyletic origin [10]. However, other organisms have recently been suggested as containing primary endosymbionts on their way to becoming permanent organelles, and which are not closely related to known plastids. These endosymbionts are the intracellular spheroid bodies of the diatom *Rhopalodia* [11] and photosynthetic inclusions within the filose amoeba *Paulinella* [12] (Figure 1). In this Opinion article, we review the evidence for the different models of plastid origin, and consider the position of *Rhopalodia* and *Paulinella*.

Generally accepted facts

The features of plastids in plants and algae are summarized in Table 1. At the algal level, there are three occurrences of plastids with only two envelope membranes – in Chlorophyta, Rhodophyta and Glaucophyta. The plastids of plants (Plantae) were inherited when they evolved from a chlorophyte ancestor. All these are the primary plastids and are the ones to which the monophyly/polyphyly debate refers. The secondary plastids all have three or four envelope membranes. They result from secondary endosymbiotic acquisition of a photosynthetic eukaryote. There is little doubt that this has happened at least twice, generating the chlorophyll *a*,*c*-containing lineages and the chlorophyll *a*,*b*-containing chlorarachniophytes, although whether it happened more than twice is controversial [13,14]. We will not discuss secondary endosymbioses in detail. In some dinoflagellate lineages, the secondary plastid has been lost and replaced, generating tertiary plastids [15,16].

Accepting monophyly or polyphyly

Before looking at the evidence for the two models, we describe what would be needed to accept one or other, and the associated pitfalls. To establish a monophyletic origin using sequence data, we need to show that sequences from plastid-containing organisms all form a single group in rooted phylogenetic trees. This needs to be true both with sequences derived from the endosymbiont and with sequences derived from the host. For example, it is possible that one particular group of closely related organisms was particularly good at acquiring endosymbionts and did so on multiple occasions. Trees based solely on host sequences would therefore incorrectly indicate monophyly [17]. What if we base our analysis of plastid origins on cell biological characters rather than sequences? Then we need to show that all plastids share derived characters unique to the plastid lineage.

Although this might seem straightforward, there are difficulties. The first is the problem of incomplete sampling. Suppose phylogenetic trees based on endosymbiont sequences all form a single group nested within the cyanobacteria. To interpret this as irrefutable evidence for monophyly requires an assumption that no free-living cyanobacterium exists (or has ever existed) that would split up that single group in a tree. Little information is available to evaluate this assumption. Second, several problems have been identified with the construction of phylogenetic trees where taxa diverged anciently. These are discussed in more detail in Box 1. Third, we know little about the features of cyanobacteria over a billion years ago, when primary plastids first appeared. The cyanobacteria we see today have almost certainly been through several population bottlenecks, associated with Global Snowballs (Box 2). Can we assume that present-day cyanobacteria

Corresponding author: Larkum, A.W.D. (alark@mail.usyd.edu.au).
Available online 9 April 2007.

 1360-1385/$ – see front matter doi:10.1016/j.tplants.2007.03.011

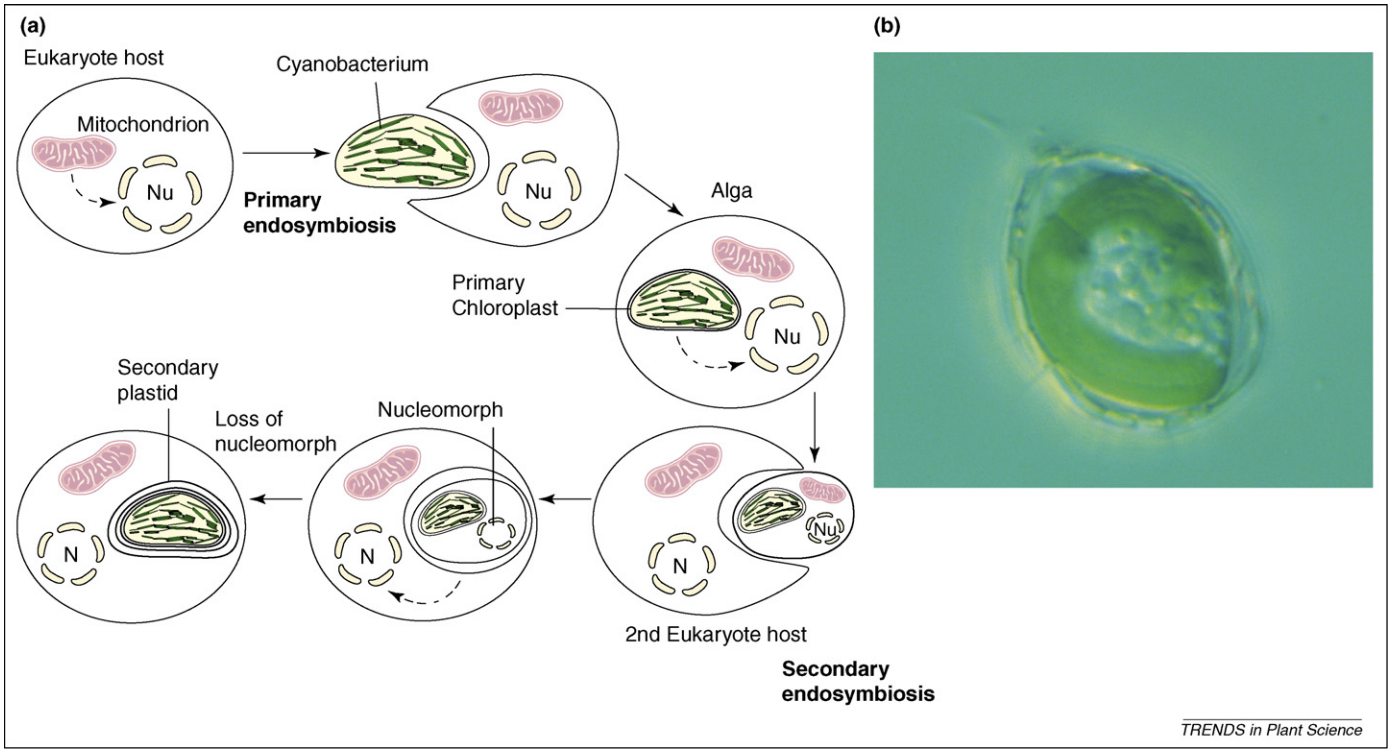

Figure 1. Origin of plastids by primary and secondary endosymbiosis. **(a)** Acquisition of a cyanobacterium by primary endosymbiosis, and subsequent secondary endosymbiotic acquisition of the resulting eukaryotic alga. The intermediate algal nucleus (Nu) forms the nucleomorph, which is subsequently reduced. N indicates the nucleus of the second eukaryote host. Broken arrows indicate gene transfer. **(b)** Photomicroscope image of a cell of *Paulinella*. Cell length is ~25 μm. The photograph shows the scales of the theca, a filopodium, and a large, dividing photosynthetic body or 'chromatophore'. Photograph kindly supplied by Birger Marin.

exhibit ancestral features of early oxygenic photosynthetic taxa? We have little information on which to evaluate this assumption either.

Establishing polyphyly is equally difficult. With phylogenetic trees, we need to show that endosymbiont (or host) genes from different lineages are located in different places in phylogenetic trees. But this is subject to the same technical difficulties with tree-building as already outlined in Box 1. Alternatively, we could show that individual plastid lineages share derived characters

Table 1. Features of different plastid types[a]

Phylum	Plastid type	Pigments	Light-harvesting complex	Thylakoid characters	Envelope membranes and cell placement
Plantae	Primary	Chl *a*, *b* Zeaxanthin	CAB	Grana	2, in cytosol
Chlorophyta	Primary	Chl *a*, *b* Zeaxanthin	CAB	Thylakoids can be appressed; grana present only in streptophytes	2, in cytosol
Rhodophyta	Primary	Chl *a* Phycobiliproteins Lutein	PBS, ψLHC	Thylakoids non-appressed	2, in cytosol
Glaucophyta	Primary	Chl *a* Phycobiliproteins Lutein	PBS, ψLHC	Thylakoids non-appressed	2, in cytosol
Cryptophyta	Secondary	Chl *a*, c_2 Phycobiliproteins Alloxanthin	CAC PB	Thylakoids in groups of two	4, RER lumen
Ochrophyta and heterokont algae	Secondary	Chl *a*, c_1, c_2 Fucoxanthin	CAC	Thylakoids in groups of three	4, RER lumen
Haptophyta	Secondary	Chl *a*, c_1, c_2 Fucoxanthin	CAC	Thylakoids in groups of three	4, RER lumen
Pyrrophyta (Dinoflagellates)	Secondary	Chl *a*, c_2 Peridinin	CAC PCP	Thylakoids in groups of three	3, in cytosol
Apicomplexa	Remnant	None	None	No thylakoids	4, in cytosol
Euglenophyta	Secondary	Chl *a*, *b* Diadinoxanthin	CAB	Thylakoids in groups of three	3, in cytosol
Chlorarachniophyta	Secondary	Chl *a*, *b* Violaxanthin	CAB	Thylakoids many, appressed	4, in cytosol

Abbreviations: CAB, chlorophyll *a*,*b*; CAC, chlorophyll *a*,*c*; PB, phycobiliprotein; PBS, phycobilisome; PCP, peridinin and chlorophyll; RER, rough endoplasmic reticulum; ψLHC, putative chlorophyll-binding protein homologous to CAB [47].
[a]Data taken from Ref. [72].

Box 1. The troubles with trees

In the past few years, technical concerns have emerged over the extent to which gene trees record phylogenetic history for anciently diverged taxa. A significant component of what is measured as phylogenetic signal for anciently diverged orthologues is often the result of lineage-specific patterns of substitution and evolving structural and functional constraints [34,61–65]. Many of the molecules used to construct phylogeny interact with other proteins, and co-evolution of these interactions occurs over time [62,66]. The nature of the interactions will constrain which sites in a target sequence are free to vary, and the differences between bacterial and organelle interactions needs to be better evaluated in terms of this component of phylogenetic signal.

Understanding the spatial pattern of substitutions in different lineages is also important in understanding and correcting for a second problem that confounds phylogenetic inference. This is the problem of compositional heterogeneity between evolutionary lineages [63]. It has been stressed, but often under-appreciated, that the impact of this evolutionary property of data needs to be understood within the context of the sites free to vary in molecules [34,67]. The interaction of effects due to lineage-specific substitution patterns and compositional heterogeneity is still poorly understood, but it is important for understanding the extent and limit of genetic divergence between sequences, and the reliability of tree building. Currently, lineage-specific evolutionary properties of data are not adequately modelled in phylogenetic analyses. Many authors were initially optimistic that lineage-specific patterns generally help to reinforce the true underlying phylogeny; however, more recently there has been increasing recognition that this might not be so [64,68].

Box 2. Global snowballs and the cyanobacterial radiation

A crucial part of the debate on plastid origins is whether extant cyanobacteria are representative of their ancestors and, therefore, what were the organisms available to form chloroplasts. Today we are so used to thinking of cyanobacteria as an ancient line that stretches back into deep time – at least 2.3 billion years ago and perhaps more [69] – that it is difficult to separate in our minds the character states of extant cyanobacteria from those of their ancient ancestors. For this reason we, and others, have advocated the name chloroxybacteria for the ancient group of photosynthetic organisms, able to split water and liberate oxygen in photosynthesis based on a pigment system dominated by chlorophylls, preceded by pro-chloroxybacteria, without water-splitting [67]. The chloroxybacteria might have been significantly different from extant cyanobacteria, particularly given that they are separated by successive periods of climatic devastation caused by global ice cover ('global snowballs'), which probably caused major population bottlenecks.

The Earth has experienced global snowballs at several times in its history. The most recent of these episodes occurred ~635 million and 710 million years ago, and others occurred earlier (2.3 billion years ago) [69]. They were probably precipitated by the arrangement of landmasses increasing the Earth's albedo, reducing the energy absorbed from the sun. Cooling of the oceans led to increased carbon dioxide solubility, reducing the amount of carbon dioxide in the atmosphere and reducing the Earth's temperature further. As the ice sheets spread towards the equator, the Earth's albedo would have increased further, compounding the cooling. The effect was probably reversed by the loss of carbon dioxide sinks, including photosynthesis, leading to increased atmospheric carbon dioxide levels as more was pumped out by volcanoes. The rapid reversal of the snowball led to extraordinarily violent wind and wave activity, documented in the geological record. The large-scale icing over of the land and oceans almost certainly resulted in large-scale extinctions of photosynthetic organisms. So present-day cyanobacterial species might well have diverged as recently as 635 million years ago, following the last global snowball. Given that photosynthetic eukaryotes were probably in existence by 1.5 billion years ago [70,71], features common to present-day cyanobacteria might have evolved in the intervening 900 million years, becoming fixed in a global snowball, and not being representative of the organisms present when endosymbiosis first happened.

with different cyanobacterial lineages, but this will require more detailed taxon sampling than we have at present.

The evidence

We will now look at various lines of evidence that have been used to try to decide in favour of one or other hypothesis.

Genome organization and content

Early on it was recognized that almost all plastid genomes included inverted repeated regions containing ribosomal rRNA genes and separating single copy regions. Although this might appear to support monophyly [18], with the advent of genome sequences it is now clear that not only do some cyanobacteria (such as *Synechocystis* sp. PCC6803) have rRNA genes in an inverted repeat configuration, other bacteria (such as *Chlorobium tepidum*) do too (see CyanoBase, http://www.kazusa.or.jp/cyanobase/). So the existence of inverted repeated rRNA genes is not phylogenetically informative, and does not favour either monophyly or polyphyly.

The majority of the original genome of the endosymbiont has been lost, with typically only 100–200 genes left in the plastid. Some of the original genes have been lost altogether, but the majority have been relocated to the nucleus [19]. Some of these encode proteins that still function in the plastid and are translocated there after synthesis. Remarkably, several thousand encode proteins that are now used elsewhere in the cell [19]. Transfer to the nucleus has been shown to be surprisingly frequent in tobacco, and probably involves plastid lysis [20–22]. A core of genes for proteins involved in photosynthesis and protein synthesis has been retained in the plastid in almost all photosynthetic organisms [19,23]. Although the retention of a similar gene set across plastids could be argued as supporting monophyly, John Stiller has argued that the pattern is not statistically significant and the similarities reflect similar (i.e. convergent) patterns of gene transfer from chloroplast to nucleus in different lineages [24]. There are powerful arguments as to why certain genes should be retained in plastids, indicating that convergent evolution of gene location is to be expected [25,26]. So the location of individual genes is not a phylogenetically reliable character, and does not allow us to distinguish rigorously between monophyly and polyphyly.

The occurrence of gene clusters throughout plastids that are absent from cyanobacteria has also been taken as evidence for monophyly. Thus, the cluster containing the *rps2* gene and the *atpA* operon has not been reported from cyanobacteria but is present in multiple plastid lineages [27]. This observation therefore supports monophyly, but rests on the key assumptions that present-day cyanobacteria are representative of the ancestral state, that other uncharacterized cyanobacteria do not exist with the same gene organization as plastids, and that changes in genome reorganizations are not convergent (see above, Box 2 and [24]). The existence of a cluster containing *psbB*, *psbN* and *psbH* provides a cautionary tale. Its existence in many plastid lineages but absence from cyanobacteria was taken

to support monophyly [28]. However, when the complete genome sequence of *Gloeobacter violaceus* (a cyanobacterium that lacks thylakoids and is often regarded as diverging early from other cyanobacterial lineages) became available, the *psbB*, *psbN* and *psbH* genes were found clustered in this organism, as in chloroplasts [29].

Sequence-based trees

Countless phylogenetic trees have been published, and some have attempted to deal with the technical problems indicated in Box 1 [30]. The results range from placing plastids as a sister group to cyanobacteria, to indications of a monophyletic origin within the cyanobacteria or indication of a polyphyletic origin (e.g. Refs [30–37]). The conclusions depend on which organisms are included in the datasets (and whether the unusual cyanobacterium *Gloeobacter* genuinely diverged early from the others or has been erroneously placed in phylogenetic reconstructions) [30], the categories of genes [38], and the methods and models used. Although our understanding of tree-building artefacts has improved in recent years (Box 1), sequence-based trees should be treated with great caution until we have rigorous methods of dealing with these artefacts.

Import machinery

There are several similarities among the protein import machinery of different plastid types. For example, precursors of proteins from the cyanelles (plastids) of *Cyanophora* can be imported into chloroplasts, and chloroplast protein precursors can be imported into cyanelles after slight modification of the transit sequence [39]. This suggests a common origin of the import pathways of *Cyanophora* plastids and chloroplasts. Although the result is striking, import experiments are not always accurate, and chloroplast protein precursors can be imported into mitochondria *in vitro* and *in vivo* [40,41], even though these organelles do not have a common origin. Heterologous import experiments need to be interpreted cautiously.

Genome surveys and immunochemical analyses have indicated that several polypeptides of the protein import machinery are found across a wide range of plastids [39,42]. Some of these are also present in cyanobacteria, consistent with the likely development of the import machinery from a pre-existing function in cyanobacteria [43]. Other proteins are specific to plastids. The Tic110 protein is represented in all lineages containing primary plastids, although it is absent from the apicoplast of *Plasmodium* [39,42] and parts of it show some similarity to other bacterial proteins (C.J. Howe, unpublished). The Toc34 protein is present in chlorophyte and rhodophyte lineages, but is apparently absent from secondary plastids, and no data are available for glaucophytes [42]. The evidence of similar import proteins in red and green plastids is more consistent with monophyly than with polyphyly, but even under a monophyly hypothesis independent losses need to be invoked in different plastid lineages (as well as the assumptions made about the ancestral cyanobacterium outlined in Box 2).

Light harvesting machinery

The two major light harvesting systems in primary plastids are (i) the membrane-intrinsic light-harvesting proteins (lhc), and (ii) phycobiliproteins assembled into phycobilisomes that are extrinsic to the thylakoid membrane (Table 1). The lhc family is not present in any extant cyanobacteria (although possible evolutionary precursors have been identified [44]), and cyanobacteria with Chl *b* (or Chl *d*) have a different chlorophyll-binding protein [45]. Red algae and glaucophytes have polypeptides with some similarities (and differences) to the lhc family [46,47]. If the red algal and glaucophyte lhc proteins are indeed related to the green plant lhc to the exclusion of any cyanobacterial homologues, this would add support to a monophyletic hypothesis. However, this highlights the need for better sampling among cyanobacteria, as well as the difficulties highlighted in Box 2.

Metabolic pathways

The existence in rhodophyte and chlorophyte plastids (although not *Cyanophora* plastids) of a Class I fructose *bis*phosphate aldolase (FBA) apparently more closely related to those in the cytosol of eukaryotes than to those in cyanobacteria (Class II) would seem to indicate monophyly, at least of rhodophytes and chlorophytes, because it requires a single retargeting of a cytosolic enzyme to the plastid, rather than multiple retargeting required under polyphyly [48,49]. However, the grouping of red algal plastid FBA with green chloroplast enzymes is not robust in phylogenetic trees [50]. Furthermore, multiple replacements of FBA have occurred elsewhere in evolution, because independent acquisition of FBA has taken place in the secondary plastids of chromalveolates [49]. Some bacteria also possess Class I FBA enzymes [50], so there is also the possibility (Box 2) that ancient cyanobacteria might have had this form of the enzyme. The observed grouping of rhodophyte plastid FBPase with the chlorophyte plastid enzymes should also be treated with caution, given poor bootstrap support [50].

Miroslav Obornik and Beverley Green analysed the origin of haem biosynthesis enzymes in chlorophytes, the rhodophyte *Cyanidioschyzon merolae* and the diatom *Thalassiosira pseudonana* using genomic data [51]. They concluded that although most of the enzymes had a cyanobacterial origin, the porphobilinogen deaminase was of mitochondrial origin, and the glutamyl-tRNA synthetase was derived from the host nucleus. The similar origins of these genes in different lineages support monophyly. However, it is difficult to assess the strength of the support because (i) the number of sources of genes available is limited, making chance convergence more likely (and the apparent origin of ferrochelatase was not consistent across lineages), and (ii) there might be biochemical advantages for using a particular form of the gene (e.g. the host isoform) in any plastid, whatever its origin, which would also lead to convergence.

Evaluating the evidence

So several lines of evidence give some support to monophyly. However, we need to be clear about the assumptions underlying this interpretation. They include assumptions that (i) no cyanobacteria remain to be sampled that will disprove monophyly, (ii) at least some of the phylogenetic trees indicating monophyly deal adequately with

well-documented phylogenetic artefacts, and (iii) extant cyanobacteria adequately represent the plastid ancestor. Questioning the support for monophyly does not argue that the evidence supports polyphyly instead; the analyses might be unable to distinguish between these evolutionary models. Furthermore, we should look for more than one line of evidence in picking a model. For example, the diatom *Thalassiosira oceanica* shares the use of plastocyanin as a redox carrier with chlorophytes [52], but this is not (yet) taken as indicating that diatoms are derived from the green chloroplast lineage.

So what of the new plastid contenders *Rhopalodia* and *Paulinella*? Phylogenetic analysis indicates that their 'plastids' have separate origins from those of the plants and algae we have considered. But what makes an endosymbiont a plastid? Endosymbionts are common in eukaryotic cells, and have been catalogued for decades [53]. The features that distinguish a genuine plastid from an endosymbiont intuitively include indefinite stable maintenance (albeit not necessarily as a photosynthetic organelle [26]) and the transfer of DNA from the organelle to the nucleus with the concomitant development of a protein import machinery [54]. Although *Rhopalodia* and *Paulinella* retain their 'plastids' stably, there is no evidence yet on the extent of gene loss to the nucleus for *Rhopalodia* [11], and evidence that it has not occurred for *Paulinella* [55]. If we relax the definition of a plastid, and do not require gene loss and an import system (as others have argued [56]), these organisms could be regarded as having a plastid with origins independent of the other plastid lineage(s).

Is life that simple? The shopping bag model

So far we have talked of endosymbiosis (whether happening once or many times) as involving in each case a single host engulfing a single target organism. But is this accurate? We know DNA can be transferred from plastid to nucleus surprisingly easily by plastid lysis. We can readily envisage a situation where a host forms a transient relationship in which the endosymbiont persists for a while, and is then lysed. Some of the DNA might enter the host nucleus and be retained. The host might form a series of transient symbioses before a stable relationship is finally achieved, in which the resulting organelle is a chimaera of products from different predecessors. Although the physical compartment that marks the organelle might be ascribable to a single endosymbiont, its contents are not so clear-cut. This is the basis of the 'promiscuous hypothesis' of plastid evolution, or 'polysymbiosis' [57]. It could also be described as the 'shopping bag model'. Your shopping might all be in a bag that came from an identifiable store, and some of the contents of the bag might have come from the same place. But some came from elsewhere, and you cannot ascribe a single origin to all your shopping.

How do we test the shopping bag model? It predicts (i) that the nuclear genes for different plastid proteins should map to different places in a phylogenetic tree and (ii) that organisms that have not yet established a stable symbiosis can nevertheless have acquired symbiont genes. For point (i) with organisms bearing primary plastids, endosymbiosis probably happened so long ago that the detailed phylogenetic resolution we would have to extract from the sequences has been obliterated by other factors such as shifts in nucleotide composition, heterotachy (the situation where, for a given site, there are lineage-specific differences in the rate of evolution), and so on – particularly if transient endosymbionts were closely related. Where we might have a chance of detecting the process is in organisms that formed new endosymbioses more recently. Thus, as predicted by the shopping bag model, a large fraction of the chloroplast-targeted proteins from the chlorarachniophyte *Bigelowiella natans* appear to have been derived from a range of other organisms, including red algae and bacteria [58]. Likewise the tertiary plastid of the dinoflagellate *Karlodinium* uses genes from two different endosymbionts [59]. What about point (ii), the presence of genes in organisms that do not yet have symbionts? Few data are yet available, but the report that the genome of the sea slug *Elysia crispata* contains gene(s) for chloroplast proteins derived from the algae that it eats is consistent with this proposal [60].

Conclusions and further work

What can we conclude? Although there are lines of evidence in favour of monophyly (and some in favour of polyphyly) it is crucial to recognize the assumptions underlying this evidence. In the future, we need a better evaluation of these assumptions – how far we can trust phylogenetic trees for anciently diverged sequences, and how realistic it is to take present-day cyanobacteria as representative of the plastid ancestor(s). More information on the extent of gene loss to the nucleus from endosymbionts that might be on their way to becoming organelles should help us assess how far they have gone down that track, and data on host genes might provide traces of endosymbionts that tried to tread the same path but failed.

Acknowledgements

We thank the Australian Research Council (A.W.D.L.), the Marsden Fund of New Zealand (P.J.L.), Leverhulme Trust and the BBSRC (C.J.H.) for financial support. We thank anonymous referees and our collaborators for helpful comments on the manuscript.

References

1 Schimper, A.F.W. (1883) Ueber die entwickelung der chlorophyllkoerner und farbkoerper. *Botanische Zeitung* 41, 105–113
2 Mereschowsky, C. (1905) Ueber Natur und Ursprung der Chromatophoren im Pflanzereiche. *Biol. Centralblatt* 25, 593–604
3 Martin, W. and Kowalik, K.V. (1999) Annotated English translation of Mereschkowsky's 1905 paper 'Ueber Natur und Ursprung der Chromatophoren im Pflanzenreiche'. *Eur. J. Phycol.* 34, 287–295
4 Ris, H. and Plaut, W. (1962) Ultrastructure of DNA-containing areas in the chloroplast of *Chlamydomonas*. *J. Cell Biol.* 13, 383–391
5 Echlin, P. (1966) The cyanophytic origin of higher plant chloroplasts. *Brit. Phycol. Bull.* 3, 150–151
6 Goksoyr, J. (1967) Evolution of eucaryotic cells. *Nature* 214, 1161
7 Margulis, L. (1970) *Origin of Eukaryotic Cells.* Yale University Press
8 Schwarz, Z. and Koessel, H. (1979) Sequencing of the 3′-terminal region of a 16S rRNA gene from *Zea mays* chloroplast reveals homology with *E. coli* 16S rRNA. *Nature* 279, 520–522
9 Raven, P.H. (1970) A multiple origin for plastids and mitochondria. *Science* 169, 641–646
10 Palmer, J.D. (2003) The symbiotic birth and spread of plastids: how many times and whodunit? *J. Phycol.* 39, 4–11
11 Prechtl, J. *et al.* (2004) Intracellular spheroid bodies of *Rhopalodia gibba* have nitrogen-fixing apparatus of cyanobacterial origin. *Mol. Biol. Evol.* 21, 1477–1481

12 Marin, B. *et al.* (2005) A plastid in the making: evidence for a second primary endosymbiosis. *Protist* 156, 425–432
13 Bodyl, A. (2005) Do plastid-related characters support the chromalveolate hypothesis? *J. Phycol.* 41, 712–719
14 Not, F. *et al.* (2007) Picobiliphytes: a marine picoplanktonic algal group with unknown affinities to other eukaryotes. *Science* 315, 253–255
15 Ishida, K. and Green, B.R. (2002) Second- and third-hand chloroplasts in dinoflagellates: phylogeny of oxygen-evolving enhancer 1 (PsbO) protein reveals replacement of a nuclear-encoded plastid gene by that of a haptophyte tertiary endosymbiont. *Proc. Natl. Acad. Sci. U. S. A.* 99, 9294–9299
16 Yoon, H.S. *et al.* (2005) Tertiary endosymbiosis driven genome evolution in dinoflagellate algae. *Mol. Biol. Evol.* 22, 1299–1308
17 Howe, C.J. *et al.* (2003) Evolution of the chloroplast genome. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 99–106
18 Turmel, M. *et al.* (1999) The compete chloroplast DNA sequence of the green alga *Nephroselmis olivacea*: insights into the architecture of ancestral chloroplast genomes. *Proc. Natl. Acad. Sci. U. S. A.* 96, 10248–10253
19 Martin, W. *et al.* (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci. U. S. A.* 99, 12246–12251
20 Lister, D.L. *et al.* (2003) DNA transfer from chloroplast to nucleus is much rarer in *Chlamydomonas* than in tobacco. *Gene* 316, 33–38
21 Huang, C.Y. *et al.* (2003) Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* 422, 72–76
22 Stegemann, S. *et al.* (2003) High-frequency gene transfer from the chloroplast genome to the nucleus. *Proc. Natl. Acad. Sci. U. S. A.* 100, 8828–8833
23 Martin, W. *et al.* (1998) Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* 393, 162–165
24 Stiller, J.W. *et al.* (2003) A single origin of plastids revisited: convergent evolution in organellar genome content. *J. Phycol.* 39, 95–105
25 Allen, J.F. (2003) The function of genomes in bioenergetic organelles. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 19–38
26 Barbrook, A.C. *et al.* (2006) Why are plastid genomes retained in non-photosynthetic organisms? *Trends Plant Sci.* 11, 101–108
27 Stoebe, B. and Kowallik, K.V. (1999) Gene-cluster analysis in chloroplast genomics. *Trends Genet.* 15, 344–347
28 Reith, M. and Munholland, J. (1993) A high-resolution gene map of the chloroplast genome of the red alga *Porphyra purpurea*. *Plant Cell* 5, 465–475
29 Nakamura, Y. *et al.* (2003) Complete genome structure of Gloeobacter violaceus PCC7421, a cyanobacterium that lacks thylakoids. *DNA Res.* (Suppl.) 10, 181–201
30 Rodriguez-Ezpelata, N. *et al.* (2005) Monophyly of primary photosynthetic eukaryotes: green plants, red algae and glaucophytes. *Curr. Biol.* 15, 1325–1330
31 Lockhart, P.J. *et al.* (1992) Substitutional bias confounds inference of cyanelle origins from sequence data. *J. Mol. Evol.* 34, 153–162
32 Morden, C.W. *et al.* (1992) Gene phylogenies and the endosymbiotic origin of plastids. *Biosystems* 28, 75–90
33 Delwiche, C.F. *et al.* (1995) Phylogenetic analysis of *tufA* sequences indicates a cyanobacterial origin of all plastids. *Mol. Phylogenet. Evol.* 4, 110–128
34 Lockhart, P. *et al.* (1998) A covariotide model explains apparent phylogenetic structure of oxygenic photosynthetic lineages. *Mol. Biol. Evol.* 15, 1183–1188
35 Barbrook, A.C. *et al.* (1998) Phylogenetic analysis of plastid origins based on secA sequences. *Curr. Genet.* 34, 336–341
36 Valentin, K. and Zetsche, K. (1990) Structure of the Rubisco operon from the unicellular red alga *Cyanidium caldarium* – evidence for a polyphyletic origin of the plastids. *Mol. Gen. Genet.* 222, 425–430
37 Stiller, J.W. and Hall, B.D. (1997) The origin of red algae: implications for plastid evolution. *Proc. Natl. Acad. Sci. U. S. A.* 94, 4520–4525
38 Rodriguez-Ezpelata, N. *et al.* (2007) Phylogenetic analyses of nuclear, mitochondrial and plastid multi-gene datasets support the placement of *Mesostigma* in the Streptophyta. *Mol. Biol. Evol.* 24, 723–731
39 Steiner, J.M. *et al.* (2005) Homologous protein import machineries in chloroplasts and cyanelles. *Plant J.* 44, 646–652
40 Cleary, S.P. *et al.* (2002) Isolated plant mitochondria import chloroplast precursor proteins *in vitro* with the same efficiency as chloroplasts. *J. Biol. Chem.* 277, 5562–5569
41 Hurt, E.C. *et al.* (1986) The cleavable pre-sequence of an imported chloroplast protein directs attached polypeptides into yeast mitochondria. *EMBO J.* 5, 1343–1350
42 McFadden, G.I. and van Dooren, G. (2004) Evolution: red algal genome affirms a common origin of all plastids. *Curr. Biol.* 14, R514–R516
43 Bolter, B. *et al.* (1998) Origin of a chloroplast protein importer. *Proc. Natl. Acad. Sci. U. S. A.* 95, 15831–15836
44 Dolganov, N.A.M. *et al.* (1995) Cyanobacterial proteins with similarity to the chlorophyll *a*/*b* binding proteins of higher plants –evolution and regulation. *Proc. Natl. Acad. Sci. U. S. A.* 92, 636–640
45 Chen, M.W. *et al.* (2005) Unique origin and lateral transfer of prokaryotic chlorophyll-*b* and chlorophyll *d*-light-harvesting systems. *Mol. Biol. Evol.* 22, 21–28
46 Wolfe, G.R. *et al.* (1994) Evidence for a common origin of chloroplasts with light-harvesting complexes of different pigmentation. *Nature* 367, 566–568
47 Rissler, H.M. and Durnford, D.G. (2005) Isolation of a novel carotenoid-rich protein in *Cyanophora paradoxa* that is immunologically related to the light-harvesting complexes of photosynthetic eukaryotes. *Plant Cell Physiol.* 46, 416–424
48 Gross, W. *et al.* (1999) Characterization, cloning, and evolutionary history of the chloroplast and cytosolic class I aldolases of the red alga *Galdieria sulphuraria*. *Gene* 230, 7–14
49 Patron, N.J. *et al.* (2004) Gene replacement of fructose-1,5-bisphosphate aldolase supports the hypothesis of a single photosynthetic ancestor of chromalveolates. *Eukaryot. Cell* 3, 1169–1175
50 Rogers, M. and Keeling, P.J. (2004) Lateral transfer and recompartmentalization of Calvin cycle enzymes of plants and algae. *J. Mol. Evol.* 58, 367–375
51 Obornik, M. and Green, B.R. (2005) Mosaic origin of the heme biosynthesis pathway in photosynthetic eukaryotes. *Mol. Biol. Evol.* 22, 2343–2353
52 Peers, G. and Price, N.M. (2006) Copper-containing plastocyanin used for electron transport by an oceanic diatom. *Nature* 441, 341–344
53 Buchner, P. (1965) *Endosymbiosis of Animals with Plant Microorganisms.* John Wiley
54 Theissen, U. and Martin, W. (2006) The difference between organelles and endosymbionts. *Curr. Biol.* 16, R1016–R1017
55 Yoon, H.S. *et al.* (2006) Minimal plastid genome evolution in the *Paulinella* endosymbiont. *Curr. Biol.* 16, R670–R672
56 Bhattacharya, D. and Archibald, J.M. (2006) The difference between organelles and endosymbionts. *Curr. Biol.* 16, R1017–R1018
57 Larkum, A.W.D. Evolution of the reaction centers and photosystems. In *Primary Processes of Photosynthesis: Principles and Apparatus* (Vol. 2) (Renger, G., ed.), Royal Society of Chemistry, Cambridge (in press)
58 Archibald, J.M. *et al.* (2003) Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga *Bigelowiella natans*. *Proc. Natl. Acad. Sci. U. S. A.* 100, 7678–7683
59 Patron, N.J. *et al.* (2006) A tertiary plastid uses genes from two endosymbionts. *J. Mol. Biol.* 357, 1371–1382
60 Pierce, S.K. *et al.* (2003) Horizontal transfer of functional nuclear genes between multicellular organisms. *Biol. Bull.* 204, 237–240
61 Lockhart, P.J. *et al.* (2000) How molecules evolve in eubacteria. *Mol. Biol. Evol.* 17, 835–838
62 Lopez, P. *et al.* (2002) Heterotachy, an important process of protein evolution. *Mol. Biol. Evol.* 18, 1–7
63 Jermiin, L.S. *et al.* (2004) The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. *Syst. Biol.* 53, 639–643
64 Lockhart, P. *et al.* (2006) Heterotachy and tree building: a case study with plastids and eubacteria. *Mol. Biol. Evol.* 23, 40–45
65 Buck, M.J. and Atchley, W.R. (2005) Networks of coevolving sites in structural and functional domains of serpin proteins. *Mol. Biol. Evol.* 22, 1627–1634
66 Guo, Z. and Stiller, J.W. (2005) Comparative genomics and evolution of proteins associated with RNA polymerase II C-terminal domain. *Mol. Biol. Evol.* 22, 2166–2178

67 Lockhart, P.J. *et al.* (1996) Evolution of chlorophyll and bacteriochlorophyll: the problem of invariant sites in sequence analysis. *Proc. Natl. Acad. Sci. U. S. A.* 93, 1930–1934
68 Shalchian-Tabrizi, K. *et al.* (2006) Heterotachy processes in rhodophyte-derived secondhand plastid genes: implications for addressing the origin and evolution of dinoflagellate plastids. *Mol. Biol. Evol.* 23, 1504–1515
69 Kopp, R.E. *et al.* (2005) The Paleoproterozoic snowball Earth: a climate disaster triggered by the evolution of oxygenic photosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 102, 11131–11136
70 Butterfield, N.J. (2000) *Bangiomorpha pubescens* n.gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiology* 26, 386–404
71 Falkowski, P.G. *et al.* (2004) The evolution of modern eukaryotic phytoplankton. *Science* 305, 354–360
72 Larkum, A.W.D. *et al.* (2003) *Photosynthesis in Algae* (Vol. 14) (*Advances in Photosynthesis and Respiration*), (Govindjee, Series ed.), Kluwer