Sizing up the genomic footprint of endosymbiosis

Marek Elias,[1] *and John M. Archibald*[2]*

[1]Faculty of Science, Department of Botany, Charles University in Prague, Benatska 2, Prague, Czech Republic
[2]Integrated Microbial Biodiversity Program, Department of Biochemistry and Molecular Biology, The Canadian Institute for Advanced Research, Dalhousie University, Halifax, Nova Scotia, Canada

A flurry of recent publications have challenged consensus views on the tempo and mode of plastid (chloroplast) evolution in eukaryotes and, more generally, the impact of endosymbiosis in the evolution of the nuclear genome. Endosymbiont-to-nucleus gene transfer is an essential component of the transition from endosymbiont to organelle, but the sheer diversity of algal-derived genes in photosynthetic organisms such as diatoms, as well as the existence of genes of putative plastid ancestry in the nuclear genomes of plastid-lacking eukaryotes such as ciliates and choanoflagellates, defy simple explanation. Collectively, these papers underscore the power of comparative genomics and, at the same time, reveal how little we know with certainty about the earliest stages of the evolution of photosynthetic eukaryotes.

Keywords: endosymbiosis; endosymbiotic gene transfer; evolution; genomics; organelles; plastids

Introduction

With its roots dating back to the dawn of the last century,[1] endosymbiotic theory has been part of the lexicon of mainstream biology for more than 30 years.[2,3] Mitochondria and plastids clearly evolved from bacterial endosymbionts (alpha-proteobacteria and cyanobacteria, respectively) and tremendous progress has been made towards a comprehensive understanding of the evolution of these textbook eukaryotic organelles.[4,5] Much remains to be understood but one thing we know for certain is that the genetic integration of endosymbiont and host involves the massive flux of genes from the former to the latter, a process known as endosymbiotic gene transfer (EGT) (Fig. 1).[6,7] Simply put, EGT is the reason why contemporary mitochondrial and plastid genomes encode at most a hundred or so of the thousand plus proteins they need for proper function. The bulk of their proteomes are in fact nucleus encoded and targeted to the organelles post-translationally *via* complex protein import machineries.[5,8]

Straightforward as it may be, a simple gene transfer protein re-import model is insufficient to fully account for the complexity of eukaryotes and their endosymbiotically derived organelles, for two main reasons. First, by no means all of the nucleus-encoded proteins that function in mitochondria and plastids are of endosymbiotic origin: numerous plastid-related processes have now been shown to involve proteins of diverse evolutionary origins (*e.g.* see Refs[9–15]). Second, and conversely, genomic analyses show that the cyanobacterial contribution to the proto-algal nuclear genome appears to have been much greater than simply donating genes whose products faithfully home to their compartment of origin. Some estimates, such as a 2002 analysis of the nuclear genome of the flowering plant *Arabidopsis*,[16] peg the cyanobacterial 'footprint' in the nuclear genome at almost 20% of all genes. Remarkably, fewer than half of the cyanobacterial-derived genes in *Arabidopsis* were predicted to encode plastid-targeted proteins, with the majority instead predicted to function in everything from transcription to cell division. The precise numbers vary depending on organism and analytical method (*e.g.* see Refs[17–20]) but, overall, the picture emerging is one of extensive genome and proteome mosaicism, whereby the proteins functioning in a given eukaryotic cellular compartment are of mixed ancestry.

Here we discuss recently published articles that push the envelope of nuclear genome mosaicism even further and complicate our interpretation of the evolution of plastids in a large fraction of algal diversity. We also discuss several recent papers purporting to detect a genomic footprint of past endosymbioses in the genomes of plastid-lacking eukaryotes, and the challenges associated with such interpretations.

Red and green algal footprints in diatom genomes

The first set of papers relate to the evolution of plastids in diatoms, ubiquitous marine microbes of profound environmental significance.[21] The ancestor's of diatoms, like many other algae, acquired photosynthesis secondarily by engulfing

*Correspondence to: J. M. Archibald, The Canadian Institute for Advanced Research, Integrated Microbial Biodiversity Program, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia, Canada B3H 1X5.
E-mail: john.archibald@dal.ca
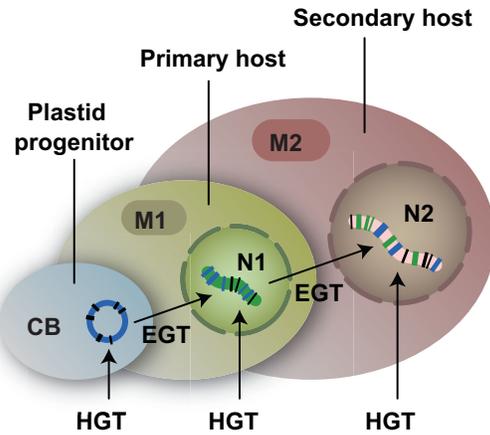
Figure 1. Endosymbiosis and gene flow in photosynthetic eukaryotes. Diagram depicts movement of genes in the context of primary and secondary endosymbiosis, beginning with the cyanobacterial endosymbiont (CB) that gave rise to modern-day plastids. Acquisition of genes by horizontal (or lateral) gene transfer is always a possibility, and such genes can be difficult to distinguish from those acquired by endosymbiotic gene transfer (EGT). CB, cyanobacterium; HGT, horizontal gene transfer; MT, mitochondrion.

a distantly related photosynthetic eukaryote whose plastid evolved directly from the cyanobacterial plastid progenitor. Inferring how many times the 'primary' plastids of red algae, green algae (and plants) and glaucophyte algae evolved into 'secondary' plastids is an area of active investigation and debate.[22–25] No secondary plastids derived from glaucophytes are known, but both green and red algae have, each at the very least on one occasion, been captured and converted into a secondary plastid (Fig. 2). This process involves a second round of EGT, this time from the endosymbiont nucleus to that of the secondary host (Fig. 1), as well as the evolution of another protein import pathway built on top of that used by primary plastids.[5,26] For these reasons, successful integration of a secondary endosymbiont is thought by many to be difficult to achieve, and secondary endosymbiosis is thus usually invoked only sparingly.

Secondary plastids of green algal origin occur in euglenophytes and chlorarachniophytes, whereas most plastids in so-called 'chromalveolates' are derived from red algae. Chromalveolates include cryptophytes, haptophytes, dinoflagellates, apicomplexans, the newly discovered coral alga *Chromera velia* and stramenopiles (or heterokonts), the latter being the group to which diatoms belong (Fig. 2). The chromalveolate hypothesis was put forth as an attempt to
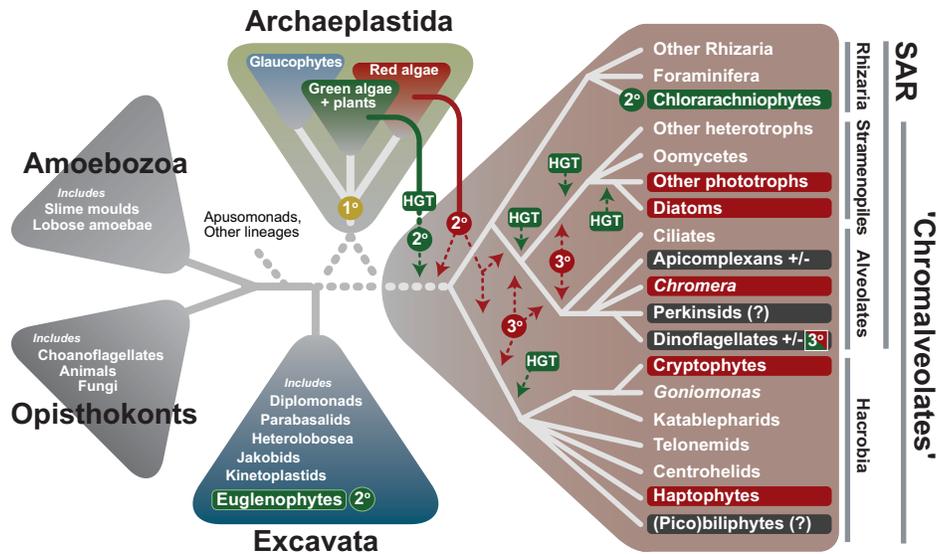
Figure 2. Origin and spread of photosynthesis across the eukaryotic tree of life. Diagram shows hypothesised 'supergroups' with emphasis on those containing photosynthetic lineages. Some (but not all) lineages within the different supergroups are provided for context. The tree topology shown within the 'chromalveolate' + SAR (Stramenopiles + Alveolates + Rhizaria) clade represents a synthesis of phylogenetic and phylogenomic data published as of 31 August 2009. Branch lengths do not correspond to evolutionary distance. Dashed lines indicate uncertainties with respect to the timing and/or directionality of secondary (2°) or tertiary (3°) endosymbiotic events, question marks (?) indicate uncertainty as to the presence of a plastid and '+/−' indicates that both plastid-bearing (+) and plastid-lacking (−) dinoflagellates and apicomplexans exist. Examples of green and red algal-derived tertiary plastids in dinoflagellates are known (see text for discussion). HGT, horizontal (or lateral) gene transfer.

explain, in the most economical way, the origin and distribution of the chlorophyll *c*-containing plastids found in many of the above-mentioned lineages: chromalveolate taxa are hypothesised to stem from an ancestor that emerged as a result of an ancient union between an as-yet-obscure non-photosynthetic eukaryote and a red algal endosymbiont.[23,27]

As would be expected from their plastid genomes, which are demonstrably red algal in nature,[28–32] the complete nuclear genomes of the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana* revealed the presence of at least 170 genes presumably introduced by EGT from the red algal endosymbiont.[33] Where investigated, genes of red algal ancestry are also readily apparent in the nuclear genomes of a wide variety of other photosynthetic chromalveolates (*e.g*. see Refs[34–37]). What has come as a surprise, however, is the presence of genes showing green algal ancestry in chromalveolate nuclear genomes. For example, the enzyme phosphoribulokinase (a critical component of the plastid-localised Calvin cycle) is in chromalveolates encoded by a nuclear gene more closely related to green algal and plant genes than to red algal genes: a horizontal (or lateral) gene transfer (HGT) from a green alga into a chromalveolate ancestor was suggested as the explanation.[38] More recently, the list of 'green' genes in chromalveolates has been greatly expanded. First, Frommolt *et al*.[39] analysed the phylogeny of carotenoid biosynthetic enzymes and found that in chromalveolates five of them appear to have been recruited from green algae. Second, Becker *et al*[40] investigated the evolutionary history of non-cyanobacterial genes present in the ancestor of all photosynthetic eukaryotes, specifically those coming from Chlamydiae. They concluded that ten such genes were passed into diatoms from red algae by secondary endosymbiosis, whereas another ten chlamydial genes in diatoms were donated by green algae. These findings led both teams to consider the possibility that the green genes speak to the existence of a chromalveolate ancestor that harboured a green algal endosymbiont.

Most recently, a comprehensive phylogenetic analysis of genes in the nuclear genome of the diatoms *T. pseudonana* and *P. tricornutum* revealed more than 1,700 genes believed to be derived from a green alga.[41] Amazingly, 'green' genes in diatoms outnumber 'red' genes by more than three-to-one. In addition, the authors found that a considerable fraction of the diatom 'green' genes are shared with other chromalveolate lineages, notably including more than 400 genes in common with haptophytes. What do these results mean? At face value, the data provide compelling evidence for the presence of a cryptic secondary endosymbiosis involving a green alga early in the evolution of chromalveolates, prior to red algal plastid acquisition (Fig. 2). The study is, however, not without potentially serious technical limitations (see Ref.[42] for discussion) and the results are open to interpretation. Below we discuss some additional pitfalls of this study that

complicate the evidence for the hypothesised green algal endosymbiosis early in the evolution of chromalveolates.

Endosymbiosis and the evolving tree of eukaryotes

The chromalveolate hypothesis has long been controversial, in part because of the failure of traditional phylogenetic markers to recover the predicted chromalveolate clade (see Refs[43–45] for recent analyses). However, the recent explosion in partial and complete genomic data available from diverse eukaryotes has made it possible to perform phylogenomic analyses based on 100 or more genes, resulting in unprecedented resolution of the eukaryotic tree.[46–53] Such analyses have revealed unsuspected relationships between specific subsets of 'traditional' chromalveolate lineages and non-chromalveolate taxa (Fig. 2). For example, the 'SAR' clade[47] is comprised of stramenopiles and alveolates (both chromalveolates) plus Rhizaria, the latter originally conceived as an independent eukaryotic 'supergroup'.[54,55] Hacrobia[56] comprise haptophytes and cryptomonads (the remaining 'traditional' chromalveolates), plus a growing list of previously orphan lineages, including katablepharids,[45,56,57] telonemids,[53,58] and centrohelids.[53] Although the position of SAR and Hacrobia on the eukaryotic tree is still equivocal, at least some phylogenomic analyses show them as mutual sister branches,[47,49,53] suggesting that chromalveolates may indeed form a coherent, although now greatly expanded, supergroup.

At the same time, however, the latest phylogenomic analyses indicate that chromalveolates are potentially more closely related to Archaeplastida (=Plantae) – the lineage from which they acquired their red (and green?) secondary plastid – than they are to any other eukaryotes. This chromalveolate-Archaeplastida 'megagroup'[50] is consistently recovered, often with considerable statistical support.[47,50,52,53] In addition, the monophyly of Archaeplastida is still contentious[59–63] and it is possible that some chromalveolate lineages,[52] or chromalveolates as a whole,[60,64] actually nest within Archaeplastida, specifically as sister to green algae and plants (Chloroplastida). Moustafa *et al*.[41] dismiss this possibility, but if true, genes corresponding to the chromalveolate secondary host lineage would in fact be expected *a priori* to branch close to homologues from Archaeplastida, a pattern that has hitherto generally been interpreted as evidence for EGT/HGT from red or green algae into chromalveolates.[41,65] The best candidates for chromalveolate genes obtained by EGT/HGT are those nested with reasonable support within red or green algae. Unfortunately, it is not clear from the report of Moustafa *et al*. how many of the >1,700 presumably 'green' genes in diatoms show this pattern.

An additional uncertainty involves an at-present inability to effectively discriminate between alternative scenarios for how the foreign 'green' genes would have come to reside in the genomes of diatoms and other chromalveolates. The problem is the very 'patchy' sampling of genomes currently available from the Chloroplastida lineage, which is restricted to three major groups: land plants, volvocalean chlorophytes (namely *Chlamydomonas* and *Volvox*) and the order Mamiellales of the paraphyletic class 'Prasinophyceae' (represented by *Ostreococcus* spp. and *Micromonas* spp.). It is therefore possible that some of the 'green' genes that appear to be shared by distant chromalveolate lineages (*e.g.* the >400 genes shared by diatoms and haptophytes) in fact exist as clusters of exclusively chromalveolate genes only because intervening green algal lineages have not yet been sampled. It is likely that when additional green algal, and in particular, prasinophyte, genomes are sequenced, some genes of this category will prove to have been independently introduced, presumably *via* HGT, into different chromalveolate lineages from different sources, rather than as the result of a unique acquisition by an early chromalveolate (Fig. 2).

Algal genes in plastid-lacking eukaryotes: How many and how did they get there?

Even if the notion of a cryptic green algal endosymbiosis in chromalveolates proves true, it need not necessarily mean that the endosymbiosis occurred before extant chromalveolates radiated. The counterargument is the same as that often levelled against the traditional chromalveolate hypothesis, *i.e.* the notion of a single red algal endosymbiosis in the common ancestor of all chromalveolate taxa. Although there is now a considerable body of evidence supporting the common ancestry of chromalveolate plastids from red algae,[23,26,28,30–32,66] a significant number of taxa known to be related to individual plastid-bearing chromalveolate lineages appear to – or are known to – lack this plastid (Fig. 2). Multiple secondary losses of the red algal-derived secondary plastid must therefore be invoked if a plastid was truly an ancestral feature of chromalveolates. Some researchers see this requirement as a serious complication and suggest alternative explanations, including engulfment of a red algal endosymbiont by a specific chromalveolate lineage (cryptophytes or a common ancestor of cryptophytes and haptophytes) followed by spread of the secondary plastid (and nuclear genes essential for its function) into additional chromalveolate lineages *via* tertiary or even quaternary endosymbioses.[25,67] If such endosymbioses indeed took place, they could in principle also serve as routes for the horizontal movement of 'green' genes from one chromalveolate lineage to another.

A potentially powerful way to discriminate between these competing explanations is to look for the genomic footprints of past endosymbioses in plastid-lacking chromalveolates in the same way that Moustafa *et al.* searched for green genes in diatoms. An early example was an analysis by Huang *et al.*[68] of the genome of *Cryptosporidium*, an apicomplexan parasite related to the *Plasmodium* and *Toxoplasma*, both of which possess remnant non-photosynthetic plastids.[5] *Cryptosporidium* does not possess a plastid but its genome was found to possess 7 genes identified as being of 'plastid/endosymbiont' origin, as well as 24 prokaryotic genes believed to be the product of HGT.[68] More recently, algal or 'plastid-like' genes have been found in the genome of the plastid-lacking oomycete stramenopile *Phytophthora*,[69] as well as in the ciliate genome.[65] [Interestingly, only 11 red algal-like genes in *Phytophthora* appear to be shared with diatoms,[33] to which they are specifically related (Fig. 2), and how many 'green' genes in diatoms are shared with *Phytophthora* is unclear from the results of Moustafa *et al.*[41]] Such results are often cited as evidence for secondary plastid loss in these lineages (*e.g.* see Refs[23,70,71]), but how confident should we be?

In the case of ciliates, it is important to consider whether the 16 algal/plastid genes identified by Reyes-Prieto *et al.*[65] rise above the background of what would be expected as a result of HGT, something which would not be surprising for a group of organisms known to harbour symbiotic algae (*e.g.* see Ref.[72]) – and thus presumably exposed to DNA from lysed algal prey on a regular basis – and for which examples of HGT are already known.[73] The same general methodological concerns raised above in the case of the diatom green genes (*e.g.* patchy genomic sampling) apply to the ciliate analyses, as well as to the discovery of a handful of cyanobacteria-related genes in some plastid-lacking non-chromalveolate eukaryotes, including the heterolobosean *Naegleria* and the choanoflagellate *Monosiga*.[74] Whether the genomic footprints described in the above-mentioned organisms truly speak to the past existence of plastids, or are the result of algal/cyanobacterial genes coming into the genomes by HGT, is, in our view, an open question. Stiller *et al.*[75] have recently argued for rigorous testing of *a priori* hypotheses in such cases, and have performed a comprehensive re-assessment of cyanobacterial/algal genes in the nuclear genomes of two *Phytophthora* species. Their results suggest that, despite previous suggestions,[69] the number of algal-like genes in these genomes is not significant and does not rise above background, *i.e.* when compared to 'algal' genes found in the reference genome of the plastid-lacking amoebozoans *Dictyostelium* and *Entamoeba*, and when diatom genomes are used as a 'positive control'.[75]

Future directions

The notion that two consecutive endosymbioses gave rise to a major and highly successful eukaryotic assemblage is a provocative working hypothesis that will undoubtedly inspire further intensive investigations. Whether it is ultimately corroborated or refuted depends on a number of critical tests yet to be performed. Most importantly, a far better sampling of genomes of both the plastid donor lineages (*i.e.* green and, in particular, red algae) and chromalveolate taxa is needed to produce more reliable gene trees and to more confidently interpret gene presence/absence data. It will be important to search for green and/or red algal genes in the genomes of poorly studied organisms such as *Telonema*, centrohelids, katablepharids, cryptophytes and Rhizaria (Fig. 2) using rigorous methodologies, and to re-assess previous analyses in light of this new information. Fortunately, the first cryptophyte and rhizarian genomes will soon be available (*Guillardia theta* and *Bigelowiella natans*, respectively), thanks to the Joint Genome Institute's Community Sequencing Programme. In the case of the chlorarachniophyte rhizarian *B. natans*, inferring the origin of its genes may prove especially challenging: under the model proposed by Moustafa *et al.*,[41] the chlorarachniophyte lineage would have hosted no fewer than three different secondary endosymbionts, *i.e.* the hypothetical prasinophyte-related green alga, the 'classical' red algal endosymbiont purported to have been present in the ancestor of all chromalveolates and finally the green algal ancestor that gave rise to the secondary plastid that chlorarachniophytes currently possess.

We also need a more detailed picture of the actual functions of the genes acquired in the context of the putative ancient green algal endosymbiosis. For example, what proportion of the 'green' genes in chromalveolates has contributed innovations to the ancestral host genome and what proportion represents replacements of original host genes by foreign genes with the same function? Of the >1,700 'green' genes in diatoms, only ∼14% were predicted to encode plastid-targeted proteins, yet diatoms appear to share many more 'green' genes with their distant photosynthetic relatives the haptophytes than with the more closely related, aplastidic and/or non-photosynthetic organisms such as apicomplexans and ciliates.[41] Does this mean that, although the majority of proteins encoded by the 'green' genes function outside the plastid, they still serve processes related to photosynthesis, for example cytosolic or mitochondrial metabolic pathways directly connected to pathways taking place within the plastid?

Interestingly, a potential unrecognised connection between chromalveolates and prasinophyte green algae may actually have been known for a long time. A feature that differentiates the chromalveolate red algal-derived plastids from those of free-living red algae is the presence of several forms of chlorophyll *c*.[76,77] While little is currently known about chlorophyll *c* biosynthesis, it is potentially significant that chlorophyll *c*-like pigments have been reported from some prasinophyte green algae.[76,78] An evolutionary connection between these pigments and chlorophyll *c* in chromalveolates was previously considered implausible,[78] but in light of the significant prasinophyte genomic footprint in diatoms and other chromalveolates discussed here, it is possible that at least some enzymes of chlorophyll *c* biosynthesis in chromalveolates were obtained from the elusive prasinophyte endosymbiont. This is just one of many questions that deserves an answer as the genomes continue to pour in and the biology of the organisms in which they reside is better understood. Despite all the above-mentioned uncertainties, few would argue against the notion that the degree of mosaicism in the genomes of modern-day eukaryotes vastly exceeds that which could have been predicted just a few years ago, and the picture is likely to grow even more complex in the years to come.

References

1. **Mereschkowsky C.** 1905. Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol Centralbl* **25**: 593–604.
2. **Gray MW.** 1992. The endosymbiont hypothesis revisited. *Int Rev Cytol* **141**: 233–357.
3. **Margulis L.** 1970. Origin of Eukaryotic Cells. New Haven: Yale University Press.
4. **Embley TM, Martin W.** 2006. Eukaryotic evolution, changes and challenges. *Nature* **440**: 623–30.
5. **Gould SB, Waller RF, McFadden GI.** 2008. Plastid evolution. *Annu Rev Plant Biol* **59**: 491–517.
6. **Martin W, Brinkmann H, Savonna C,** *et al.* 1993. Evidence for a chimeric nature of nuclear genomes: eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. *Proc Natl Acad Sci USA* **90**: 8692–6.
7. **Timmis JN, Ayliffe MA, Huang CY,** *et al.* 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet* **5**: 123–35.
8. **Dolezal P, Likic V, Tachezy J,** *et al.* 2006. Evolution of the molecular machines for protein import into mitochondria. *Science* **313**: 314–18.
9. **Reumann S, Inoue K, Keegstra K.** 2005. Evolution of the general protein import pathway of plastids. *Mol Membr Biol* **22**: 73–86.
10. **Richards TA, Dacks JB, Campbell SA,** *et al.* 2006. Evolutionary origins of the eukaryotic shikimate pathway: gene fusions, horizontal gene transfer, and endosymbiotic replacements. *Eukaryot Cell* **5**: 1517–31.
11. **Weber AP, Linka M, Bhattacharya D.** 2006. Single, ancient origin of a plastid metabolite translocator family in Plantae from an endomembrane-derived ancestor. *Eukaryot Cell* **5**: 609–12.

12. **Reyes-Prieto A, Bhattacharya D.** 2007. Phylogeny of Calvin cycle enzymes supports Plantae monophyly. *Mol Phylogenet Evol* **45**: 384–91.
13. **Huang J, Gogarten JP.** 2007. Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? *Genome Biol* **8**: R99.
14. **Moustafa A, Reyes-Prieto A, Bhattacharya D.** 2008. Chlamydiae has contributed at least 55 genes to Plantae with predominantly plastid functions. *PLoS One* **3**: e2205.
15. **Park AK, Kim H, Jin HJ.** 2009. Comprehensive phylogenetic analysis of evolutionarily conserved rRNA adenine dimethyltransferase suggests diverse bacterial contributions to the nucleus-encoded plastid proteome. *Mol Phylogenet Evol* **50**: 282–9.
16. **Martin W, Rujan T, Richly E,** *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 USA* **99**: 12246–51.
17. **Archibald JM.** 2006. Algal genomics: examining the imprint of endosymbiosis. *Curr Biol* **16**: R1033–5.
18. **Bhattacharya D, Archibald JM, Weber AP,** *et al.* 2007. How do endosymbionts become organelles? Understanding early events in plastid evolution. *BioEssays* **29**: 1239–46.
19. **Deusch O, Landan G, Roettger M,** *et al.* 2008. Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. *Mol Biol Evol* **25**: 748–61.
20. **Reyes-Prieto A, Hackett JD, Soares MB,** *et al.* 2006. Cyanobacterial contribution to algal nuclear genomes is primarily limited to plastid functions. *Curr Biol* **16**: 2320–5.
21. **Armbrust EV.** 2009. The life of diatoms in the world's oceans. *Nature* **459**: 185–92.
22. **Archibald JM.** 2009. The puzzle of plastid evolution. *Curr Biol* **19**: R81–8.
23. **Keeling PJ.** 2009. Chromalveolates and the evolution of plastids by secondary endosymbiosis. *J Eukaryot Microbiol* **56**: 1–8.
24. **Reyes-Prieto A, Weber AP, Bhattacharya D.** 2007. The origin and establishment of the plastid in algae and plants. *Annu Rev Genet* **41**: 147–68.
25. **Sanchez-Puerta MV, Delwiche CF.** 2008. A hypothesis for plastid evolution in chromalveolates. *J Phycol* **44**: 1097–107.
26. **Hempel F, Bullmann L, Lau J,** *et al.* 2009. ERAD-derived preprotein transport across the second outermost plastid membrane of diatoms. *Mol Biol Evol* **26**: 1781–90.
27. **Cavalier-Smith T.** 1999. Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. *J Eukaryot Microbiol* **46**: 347–66.
28. **Khan H, Parks N, Kozera C,** *et al.* 2007. Plastid genome sequence of the cryptophyte alga *Rhodomonas salina* CCMP1319: lateral transfer of putative DNA replication machinery and a test of chromist plastid phylogeny. *Mol Biol Evol* **24**: 1832–42.
29. **Oudot-Le Secq MP, Grimwood J, Shapiro H,** *et al.* 2007. Chloroplast genomes of the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*: comparison with other plastid genomes of the red lineage. *Mol Genet Genomics* **277**: 427–39.
30. **Sanchez-Puerta MV, Bachvaroff TR, Delwiche CF.** 2007. Sorting wheat from chaff in multi-gene analyses of chlorophyll c-containing plastids. *Mol Phylogenet Evol* **44**: 885–97.
31. **Yoon HS, Hackett JD, Van Dolah FM,** *et al.* 2005. Tertiary endosymbiosis driven genome evolution in dinoflagellate algae. *Mol Biol Evol* **22**: 1299–308.
32. **Iida K, Takishita K, Ohshima K,** *et al.* 2007. Assessing the monophyly of chlorophyll-c containing plastids by multi-gene phylogenies under the unlinked model conditions. *Mol Phylogenet Evol* **45**: 227–38.
33. **Bowler C, Allen AE, Badger JH,** *et al.* 2008. The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* **456**: 239–44.
34. **Li S, Nosenko T, Hackett JD,** *et al.* 2006. Phylogenomic analysis identifies red algal genes of endosymbiotic origin in the chromalveolates. *Mol Biol Evol* **23**: 663–74.
35. **Waller RF, Patron NJ, Keeling PJ.** 2006. Phylogenetic history of plastid-targeted proteins in the peridinin-containing dinoflagellate *Heterocapsa triquetra*. *Int J Syst Evol Microbiol* **56**: 1439–47.
36. **Nosenko T, Lidie KL, Van Dolah FM,** *et al.* 2006. Chimeric plastid proteome in the Florida "red tide" dinoflagellate *Karenia brevis*. *Mol Biol Evol* **23**: 2026–38.
37. **Bhattacharya D, Nosenko T.** 2008. Endosymbiotic and horizontal gene transfer in chromalveolates. *J Phycol* **44**: 7–10.
38. **Petersen J, Teich R, Brinkmann H,** *et al.* 2006. A "green" phosphoribulokinase in complex algae with red plastids: evidence for a single secondary endosymbiosis leading to haptophytes, cryptophytes, heterokonts, and dinoflagellates. *J Mol Evol* **62**: 143–57.
39. **Frommolt R, Werner S, Paulsen H,** *et al.* 2008. Ancient recruitment by chromists of green algal genes encoding enzymes for carotenoid biosynthesis. *Mol Biol Evol* **25**: 2653–67.
40. **Becker B, Hoef-Emden K, Melkonian M.** 2008. Chlamydial genes shed light on the evolution of photoautotrophic eukaryotes. *BMC Evol Biol* **8**: 203.
41. **Moustafa A, Beszteri B, Maier UG,** *et al.* 2009. Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science* **324**: 1724–6.
42. **Dagan T, Martin W.** 2009. Microbiology. Seeing green and red in diatom genomes. *Science* **324**: 1651–2.
43. **Moreira D, von der Heyden S, Bass D,** *et al.* 2007. Global eukaryote phylogeny: combined small- and large-subunit ribosomal DNA trees support monophyly of Rhizaria, Retaria and Excavata. *Mol Phylogenet Evol* **44**: 255–66.
44. **Yoon HS, Grant J, Tekle YI,** *et al.* 2008. Broadly sampled multigene trees of eukaryotes. *BMC Evol Biol* **8**: 14.
45. **Reeb VC, Peglar MT, Yoon HS,** *et al.* 2009. Interrelationships of chromalveolates within a broadly sampled tree of photosynthetic protists. *Mol Phylogenet Evol* **53**: 202–11.
46. **Patron NJ, Inagaki Y, Keeling PJ.** 2007. Multiple gene phylogenies support the monophyly of cryptomonad and haptophyte host lineages. *Curr Biol* **17**: 887–91.
47. **Burki F, Shalchian-Tabrizi K, Minge M,** *et al.* 2007. Phylogenomics reshuffles the eukaryotic supergroups. *PLoS One* **8**: e790.
48. **Rodriguez-Ezpeleta N, Brinkmann H, Burger G,** *et al.* 2007. Toward resolving the eukaryotic tree: the phylogenetic positions of jakobids and cercozoans. *Curr Biol* **17**: 1420–5.
49. **Hackett JD, Yoon HS, Li S,** *et al.* 2007. Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of Rhizaria with chromalveolates. *Mol Biol Evol* **24**: 1702–13.
50. **Burki F, Shalchian-Tabrizi K, Pawlowski J.** 2008. Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes. *Biol Lett* **4**: 366–9.
51. **Minge MA, Silberman JD, Orr RJ,** *et al.* 2009. Evolutionary position of breviate amoebae and the primary eukaryote divergence. *Proc Biol Sci* **276**: 597–604.
52. **Hampl V, Hug L, Leigh JW,** *et al.* 2009. Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups". *Proc Natl Acad Sci USA* **106**: 3859–64.
53. **Burki F, Inagaki Y, Brate J,** *et al.* 2009. Large-scale phylogenomic analyses reveal that two enigmatic protist lineages, Telonemia and Centroheliozoa, are related to photosynthetic chromalveolates. *Genome Biol Evol* **1**: 231–8.
54. **Keeling PJ, Burger G, Durnford DG,** *et al.* 2005. The tree of eukaryotes. *Trends Ecol Evol* **20**: 670–766.
55. **Simpson AGB, Roger AJ.** 2004. The real 'kingdoms' of eukaryotes. *Curr Biol* **14**: R693–6.
56. **Okamoto N, Chantangsi C, Horak A,** *et al.* 2009. Molecular phylogeny and description of the novel katablepharid *Roombia truncata* gen. et sp. nov., and establishment of the Hacrobia taxon nov. *PLoS One* **4**: e7080.
57. **Okamoto N, Inouye I.** 2005. The katablepharids are a distant sister group of the Cryptophyta: a proposal for Katablepharidophyta Divisio Nova/Katablepharida Phylum Novum based on SSU rDNA and beta-tubulin phylogeny. *Protist* **156**: 163–79.
58. **Shalchian-Tabrizi K, Eikrem W, Klaveness D,** *et al.* 2006. Telonemia, a new protist phylum with affinity to chromist lineages. *Proc Biol Sci* **273**: 1833–42.
59. **Stiller JW.** 2007. Plastid endosymbiosis, genome evolution and the origin of green plants. *Trends Plant Sci* **12**: 391–6.

60. **Nozaki H, Iseki M, Hasegawa M,** *et al.* 2007. Phylogeny of primary photosynthetic eukaryotes as deduced from slowly evolving nuclear genes. *Mol Biol Evol* **24**: 1592–5.
61. **Elias M.** 2008. The guanine nucleotide exchange factors Sec2 and PRONE: candidate synapomorphies for the Opisthokonta and the Archae-plastida. *Mol Biol Evol* **25**: 1526–9.
62. **Kim E, Graham LE.** 2008. EEF2 analysis challenges the monophyly of Archaeplastida and Chromalveolata. *PLoS One* **3**: e2621.
63. **Inagaki Y, Nakajima Y, Sato M,** *et al.* 2009. Gene sampling can bias multi-gene phylogenetic inferences: the relationship between red algae and green plants as a case study. *Mol Biol Evol* **26**: 1171–8.
64. **Nozaki H, Maruyama S, Matsuzaki M,** *et al.* 2009. Phylogenetic positions of Glaucophyta, green plants (Archaeplastida) and Haptophyta (Chromalveolata) as deduced from slowly evolving nuclear genes. *Mol Phylogenet Evol* (in press).
65. **Reyes-Prieto A, Moustafa A, Bhattacharya D.** 2008. Multiple genes of apparent algal origin suggest ciliates may once have been photosynthetic. *Curr Biol* **18**: 956–62.
66. **Sommer MS, Gould SB, Lehmann P,** *et al.* 2007. Der1-mediated pre-protein import into the periplastid compartment of chromalveolates? *Mol Biol Evol* **24**: 918–28.
67. **Bodyl A, Stiller JW, Mackiewicz P.** 2009. Chromalveolate plastids: Direct descent or multiple endosymbioses? *Trends Ecol Evol* **24**: 119–21.
68. **Huang J, Mullapudi N, Lancto CA,** *et al.* 2004. Phylogenomic evidence supports past endosymbiosis, intracellular and horizontal gene transfer in *Cryptosporidium parvum*. *Genome Biol* **5**: R88.
69. **Tyler BM, Tripathy S, Zhang X,** *et al.* 2006. *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogen-esis. *Science* **313**: 1261–6.
70. **Lane CE, Archibald JM.** 2008. The eukaryotic tree of life: endosymbiosis takes its TOL. *Trends Ecol Evol* **23**: 268–75.
71. **Slamovits CH, Keeling PJ.** 2008. Plastid-derived genes in the non-photosynthetic alveolate *Oxyrrhis marina*. *Mol Biol Evol* **25**: 1297–306.
72. **Summerer M, Sonntag B, Sommaruga R.** 2007. An experimental test of the symbiosis specificity between the ciliate *Paramecium bursaria* and strains of the unicellular green alga *Chlorella*. *Environ Microbiol* **9**: 2117–22.
73. **Ricard G, McEwan NR, Dutilh BE,** *et al.* 2006. Horizontal gene transfer from bacteria to rumen ciliates indicates adaptation to their anaerobic, carbohydrates-rich environment. *BMC Genomics* **7**: 22.
74. **Maruyama S, Matsuzaki M, Misawa K,** *et al.* 2009. Cyanobacterial contribution to the genomes of the plastid-lacking protists. *BMC Evol Biol* **9**: 197.
75. **Stiller JW**, **Huang J**, **Ding Q**, *et al.* 2009. Are algal genes in nonphoto-synthetic protists evidence of historical plastid endosymbioses? *BMC Genomics* **10**: 484.
76. **Zapata M**, **Garrido JL**, **Jeffrey SW.** 2006. Chlorophyll c pigments: current status. In Grimm B, Porra RJ, Rüdiger W, Scheer H, ed; Chlorophylls and Bacteriochlorophylls: Biochemistry, Biophysics, Functions and Applica-tions. Berlin: Springer.
77. **Kim E**, **Archibald JM.** 2009. Diversity and evolution of plastids and their genomes. In Aronsson H, Sandelius AS, ed; The Chloroplast—Interactions with the Environment. Berlin: Springer-Verlag.
78. **Six C**, **Worden AZ**, **Rodriguez F**, *et al.* 2005. New insights into the nature and phylogeny of prasinophyte antenna proteins: *Ostreococcus tauri*, a case study. *Mol Biol Evol* **22**: 2217–30.