What Was the Real Contribution of Endosymbionts to the Eukaryotic Nucleus? Insights from Photosynthetic Eukaryotes

David Moreira and Philippe Deschamps

Unité d'Ecologie, Systématique et Evolution, UMR CNRS 8079, Université Paris-Sud, 91405 Orsay Cedex, France

Correspondence: david.moreira@u-psud.fr

Cold Spring Harbor Perspectives in Biology
www.cshperspectives.org

Eukaryotic genomes are composed of genes of different evolutionary origins. This is especially true in the case of photosynthetic eukaryotes, which, in addition to typical eukaryotic genes and genes of mitochondrial origin, also contain genes coming from the primary plastids and, in the case of secondary photosynthetic eukaryotes, many genes provided by the nuclei of red or green algal endosymbionts. Phylogenomic analyses have been applied to detect those genes and, in some cases, have led to proposing the existence of cryptic, no longer visible endosymbionts. However, detecting them is a very difficult task because, most often, those genes were acquired a long time ago and their phylogenetic signal has been heavily erased. We revisit here two examples, the putative cryptic endosymbiosis of green algae in diatoms and chromerids and of Chlamydiae in the first photosynthetic eukaryotes. We show that the evidence sustaining them has been largely overestimated, and we insist on the necessity of careful, accurate phylogenetic analyses to obtain reliable results.

Today it is widely accepted that photosynthesis originated in eukaryotes by the endosymbiosis of a cyanobacterium within a heterotrophic eukaryotic host. This occurred in a lineage that subsequently diversified to give rise to the three contemporary groups of primary photosynthetic eukaryotes: Viridiplantae (including green algae and land plants), Rhodophyta and Glaucophyta, grouped collectively within a unique eukaryotic superphylum called Archaeplastida (Adl et al. 2005) or Plantae (Cavalier-Smith 1982). Recently, a second case of primary endosymbioses has been unveiled thanks to the characterization of *Paulinella chromatophora*, a filose amoeba that hosts a cyanobacterium with a reduced genome that has been described as "a plastid in the making" (Marin et al. 2005; Keeling and Archibald 2008; Nowack et al. 2008). Primary endosymbioses resulted in the establishment of plastids with two membranes. However, a vast variety of eukaryotes possess plastids with three or more membranes. They derive from the endosymbioses of primary photosynthetic eukaryotes within other eukaryotic cells (Delwiche 1999; Keeling 2013). Such secondary endosymbioses have spread photosynthesis across the eukaryotic tree, either by the endosymbiosis of red or of green algae. Whereas

Editors: Patrick J. Keeling and Eugene V. Koonin
Additional Perspectives on The Origin and Evolution of Eukaryotes available at www.cshperspectives.org.

 doi: 10.1101/cshperspect.a016014
Cite this article as *Cold Spring Harb Perspect Biol* 2014;6:a016014

it is almost certain that secondary endosymbioses of green algae occurred twice (in euglenids and chlorarachniophytes), secondary red algal plastids are found in a variety of alveolates, stramenopiles, cryptophytes, and haptophytes, and the number of red algal endosymbioses at the origin of these groups has been matter of intense debate (Baurain et al. 2010; Keeling 2010, 2013; Burki et al. 2012b). Moreover, the existence of tertiary endosymbioses (namely, the symbiosis of a secondary photosynthetic eukaryote within another eukaryotic cell) and of plastid replacements makes the picture of plastid evolution in eukaryotes even more complex. Dinoflagellates, some of which have replaced their ancestral red algal plastids by green algae, diatoms, haptophytes, or cryptophytes, are paradigmatic examples of such complex situations (Keeling 2013).

The evolution of plastids has been studied using genes from the plastid genome as well as typical eukaryotic nuclear genes, which allow inferring the phylogenies of both the plastids and their hosts. The use of those markers has led to interesting discoveries, such as the monophyly of the Archaeplastida (Moreira et al. 2000; Rodríguez-Ezpeleta et al. 2005) or the difficulties in reconciling the plastid and host histories in eukaryotes with red algal plastids (Baurain et al. 2010; Burki et al. 2012b). However, a third class of genes can also provide useful complementary information: the genes of plastid origin retrieved within the nuclear genome of the host. In fact, contemporary plastids have small genomes, which is due to the fact that most of the original cyanobacterial symbiont genes were lost or transferred to the host nucleus (by a process called endosymbiotic gene transfer, EGT) during the evolution of plastids (Weeden 1981; Martin et al. 1998). These transfer events are not restricted to plastid endosymbioses—the same phenomenon occurred during the endosymbiosis that gave rise to the mitochondria (Gray et al. 1999; Burger et al. 2003).

EGT genes may serve to study the evolutionary history of plastids and, in particular, the presence of cryptic endosymbioses. In fact, species that had a plastid in the past but lost photosynthesis may have conserved genes of plastid origin in their nuclear genomes. This has been shown for a variety of nonphotosynthetic eukaryotes, such as, for example, apicomplexan parasites (Fast et al. 2001; Roos et al. 2002; Williams and Keeling 2003; Huang et al. 2004), perkinsids (Stelter et al. 2007; Matsuzaki et al. 2008; Fernández Robledo et al. 2011) or nonphotosynthetic dinoflagellates (Sanchez-Puerta et al. 2007; Slamovits and Keeling 2008), and green algae (de Koning and Keeling 2004). Although much more controversial, potential EGTs have also been used to propose a photosynthetic ancestry for ciliates (Reyes-Prieto et al. 2008) or that algae with secondary plastids of red algal origin, such as diatoms and chromerids, may have contained green algal endosymbionts in their past (Moustafa et al. 2009; Woehle et al. 2011). Likewise, several dozens of potential EGTs have been detected in algae and plants that appear to have been acquired from Chlamydiae, a group of parasitic bacteria (Huang and Gogarten 2007; Becker et al. 2008; Moustafa et al. 2008), which led to proposing that cryptic chlamydial endosymbionts may have helped to establish the first plastids, in particular, by providing essential functions for plastid activity (Greub and Raoult 2003; Ball et al. 2013; Baum 2013).

We revise here some of these cases of cryptic endosymbiosis, with special attention on the difficulties in accurately detecting EGT and the importance of proper phylogenetic analysis and of an adequate taxonomic sampling to achieve that task.

CRYPTIC GREEN ALGAL ENDOSYMBIOSES IN DIATOMS AND CHROMERIDS

Diatoms are a speciose group of unicellular algae belonging to the phylum Stramenopila or Heterokonta, whereas chromerids are a recently discovered group of algae closely related to the Apicomplexa (Moore et al. 2008). Thanks to the availability of complete genome sequences or of transcriptome data, recent studies have tried to identify EGTs in these organisms. In the case of diatoms, a phylogenomic survey performed by Moustafa et al. (2009) detected 4956 putative EGTs in the genomes of the species *Thalassiosira*

pseudonana (2533 cases) and *Phaeodactylum tricornutum* (2423 cases). Such a large number of EGTs was not necessarily surprising because, for example, thousands of genes of presumed cyanobacterial origin have been seen in *Arabidopsis thaliana* (Martin et al. 2002). However, the origin of most of those EGTs in diatoms was completely unexpected. More than 70% of them appeared to be more closely related to green algal and plant homologs than to red algal ones. This was astonishing because diatom plastids are widely accepted to be derived from a red algal endosymbiont, and, thus, EGTs should be related to red algal homologs. To explain their results, Moustafa et al. proposed that diatoms, and perhaps other related phyla, originally acquired a plastid by the endosymbiosis of a green alga, which was secondarily replaced by the red algal plastid found today.

In the case of chromerids, Woehle et al. (2011) analyzed expressed sequence tags (ESTs) from the species *Chromera velia*. From 3151 ESTs, 513 appeared to be of EGT origin and had homologs in red and green algae. Two hundred sixty-three of them were more similar to red algal sequences, as expected, because chromerids, as the closely related Apicomplexa, are supposed to have secondary plastids of red algal origin (Moore et al. 2008). However, a similar number of EGTs, 250, appeared to have an unexpected green algal ancestry. Although inferior to the 70% found in diatoms, this represented ~50% of putative "green" EGT genes in chromerids. However, Woehle et al. interpreted their results in a very different way from Moustafa and coworkers. In fact, they considered the green signal found in *Chromera* not as evidence of a cryptic green algal endosymbiosis (which, in their own words, "leads to worryingly complicated evolutionary scenarios") but most likely as phylogenetic error, probably caused by poor taxonomic sampling for red algae. In fact, when Woehle et al. repeated their analysis eliminating all of the red algal sequences with the exception of only one species (the highly derived thermophile *Cyanidioschyzon merolae*, incidentally the single one used by Moustafa and coworkers), the number of potential green EGTs in *Chromera* slightly increased. This supported the idea that the green signal did not reveal true green algal ancestry but just the lack of a sampling of red algal genes as rich as the one available for green algae and plants. Woehle et al. (2011) thus concluded that improving the taxonomic sampling of red algae should continue to erase the green signal observed in *Chromera* and, likely, also in diatoms.

A PROBLEM OF TAXONOMIC SAMPLING . . . AND METHODS

The work by Woehle et al. showed the importance of taxonomic sampling for the accurate characterization of EGT events. This was a clear problem in the work of Moustafa and colleagues on diatoms. They compared all proteins encoded by two diatom genomes (*T. pseudonana* and *P. tricornutum*) against a local sequence database that, concerning the Archaeplastida, contained 11,356 protein sequences of red algae and 193,394 from green algae and plants, namely, an overwhelming 1:17 red:green ratio. This disproportion could explain many of the cases in which diatom sequences were close to green ones just because of the absence of red algal homologs, an absence that could reflect incomplete sampling (i.e., the red algal genomes containing the corresponding genes have simply not been sequenced yet) rather than a true general absence in all red algal species. Moreover, balanced taxonomic sampling is essential for accurate phylogenetic reconstruction (Lecointre et al. 1993), and many trees where diatoms branched with green sequences even in the presence of red algal homologs should be tested with a richer sampling of red algal species. This is what we tried in a study in which we reanalyzed all of the putative EGTs in diatoms using a database with as many red algal sequences as possible (Deschamps and Moreira 2012). We managed to double the number of red algal sequences, and, even if the red:green ratio remained extremely biased, this had a great impact on the inference of EGT events (see below).

In addition to the problem of unbalanced taxonomic sampling, the diatom and *Chromera* studies had a second source of potential bias: the use of automatic tools to filter phylogenetic

trees. In both cases, trees were screened to look for the nearest neighbor of diatoms and *Chromera* using PhyloSort (Moustafa and Bhattacharya 2008) and the Newick Utilities package (Junier and Zdobnov 2010), respectively. This approach was very limited because it only considers a pair of branches (the target—diatoms or *Chromera*—and its closest relative) without taking into account the general topology of trees and their global support or the taxonomic representation of the different proteins across the eukaryotic diversity. In fact, that simplistic approach ignores very common problems such as unresolved phylogenies (typical for single-marker phylogenetic analyses), hidden paralogies, or very unbalanced taxonomic samplings. The incidence of these problems on the inference of EGTs in diatoms and *Chromera* was well illustrated by the reanalyses conducted using a manual inspection of phylogenetic trees instead of the automatic filtering. In the case of *Chromera*, Burki et al. (2012a) showed that, from the 513 genes reported to have originated from red and green algae in an ~1:1 ratio, only 51 appeared to actually have an EGT origin, and, among them, 23 were more closely related to red algae, only nine supported a green algal origin, and 19 had an ambiguous signal. Using a similar manual approach to the study of diatom EGTs, we observed that only 286 cases (<10% of the EGTs originally proposed by Moustafa and coworkers) actually supported an EGT origin (Deschamps and Moreira 2012). Moreover, only 30 of these 286 cases appeared to have a well-supported green algal origin, with more cases supporting a red algal origin or just an ambiguous attribution. It is remarkable that the two manual reanalyses (Burki et al. 2012a; Deschamps and Moreira 2012) converged to a similar degree of validation of only ~10% of the EGTs detected by automatic filtering. For diatoms, among the >90% of cases rejected, half of them corresponded to genes with a very poor representation of eukaryotic species, and the other half to genes that produced phylogenetic trees compatible with a vertical inheritance from an ancient eukaryotic ancestor rather than a more recent EGT from a green or red algal donor.

These studies weakened the support for the hypothesis of cryptic endosymbioses of green algae in diatoms and chromerids. They also, as Burki et al. (2012a) clearly stated, "emphasize the lack of congruence and the subjectivity resulting from independent phylogenomic screens for EGT, which appear to call for extreme caution when drawing conclusions for major evolutionary events."

DID CRYPTIC CHLAMYDIAE PLAY A ROLE AT THE ORIGIN OF PHOTOSYNTHETIC EUKARYOTES?

Plastids import ATP from the cell cytosol thanks to their ATP/ADP translocases, which are enzymes that catalyze the transport of ATP across a membrane in exchange with ADP. Among prokaryotes, these enzymes are very rare and have only been identified in a few obligate intracellular bacteria belonging to the Chlamydiales and the Rickettsiales. Both are parasitic bacteria that steal ATP from their infected hosts. The first phylogenetic analysis of ATP/ADP translocases that included a good representation of these bacteria suggested that these translocases originated from a chlamydial ancestor and were transferred horizontally to rickettsiae and plants (Schmitz-Esser et al. 2004). This was in agreement with an earlier study showing a high proportion of chlamydial proteins similar to plant proteins, which was interpreted as the reflection of an unappreciated evolutionary relationship between the Chlamydiae and the Cyanobacteria-chloroplast lineage (Brinkman et al. 2002). Some years subsequently, a massive phylogenomic analysis of the red alga *Cyanidioschyzon merolae* identified at least 21 additional genes that appeared to have been transferred to the primary photosynthetic eukaryotes from Chlamydiae similar to the genus *Protochlamydia* (Huang and Gogarten 2007). A similar analysis including more genomes (*C. merolae* as well as the green algae *Chlamydomonas reinhardtii*, *Ostreococcus lucimarinus*, and *Ostreococcus tauri*, and the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*) retrieved 39 proteins of putative chlamydial origin, not only in primary photosynthetic eukary-

otes but also spreading through secondary plastid endosymbioses to other photosynthetic eukaryotes (Becker et al. 2008). The number of genes of potential chlamydial origin continued to increase when a similar analysis with a larger taxonomic sampling of photosynthetic eukaryotes led to the detection of 55 candidate genes in algae and plants. Moreover, 37 were putatively plastid-targeted, reinforcing the potential role of these genes in plastid functioning (Moustafa et al. 2008). Those studies agreed to hypothesize that a chlamydial endosymbiont or parasite may have facilitated the subsequent establishment of the cyanobacterial endosymbiont that give rise to the plastids.

More recently, it has been proposed that chlamydial cells infecting the host of that primary cyanobacterial endosymbiont secreted into the host cytosol effector proteins that allowed the host to use carbohydrates exported from the incipient plastid (Ball et al. 2013). Thus, Chlamydiae would have been essential not only to stabilize the plastid endosymbiosis through the long term by providing the ATP/ADP translocases, but also to allow the very first installation of the cyanobacterial endosymbiont thanks to those effector proteins. Such necessity of the presence of chlamydial cells would also explain why primary plastids have evolved so rarely (Ball et al. 2013; Baum 2013). Although this is a very interesting hypothesis, one potential caveat is that Chlamydiae have been detected as parasites of a variety of animal phyla and different amoebae, but never of plants or algae (Horn 2008).

We have addressed this question by reanalyzing the largest set of Chlamydiae-like proteins, 55 sequences, using Bayesian inference phylogenetic reconstruction, which is less sensitive than other methods to several tree reconstruction artifacts. As in the case of the study of the putative diatom green EGTs, we retrieved a variety of situations (P Deschamps and D Moreira, unpubl.). Whereas seven proteins did not retrieve any close relationship between Chlamydiae and primary photosynthetic eukaryotes, 17 others appeared to support a Chlamydiae–Archaeplastida relationship (two of them being found only in these two lineages). Two of those 17 trees did not support the monophyly of the Archaeplastida, with green algae and plants branching close to the Chlamydiae but red algae branching with Cyanobacteria. Ten trees showed a surprising pattern because Archaeplastida emerged in a group containing the Chlamydiae and the Cyanobacteria and, in some cases, a few other bacteria (e.g., Fig. 1A). A similar case, but including Archaeplastida, Chlamydiae, and Alphaproteobacteria, was supported by two proteins (Fig. 1B). Nevertheless, the most common class of trees was the one in which Chlamydiae branched close to a variety of eukaryotes, not only including Archaeplastida but also nonphotosynthetic eukaryotes that are supposed to never have hosted plastids. In five cases, this concerned just a single eukaryotic lineage (such as the Diplomonadida—*Giardia* and *Hexamita* [Fig. 1C] or the Mycetozoa—*Dictyostelium*) and in 14 other cases a larger eukaryotic diversity, with animals, fungi, and other lineages (e.g., Fig. 1D). Thus, in >50% of the putative chlamydial genes in Archaeplastida, we either did not observe a real phylogenetic relationship with this group or it concerned also nonphotosynthetic eukaryotes.

Among the proteins that appeared to support a putative chlamydial origin of the archaeplastid sequences, the 12 phylogenies showing a relationship also with cyanobacterial or alphaproteobacterial sequences were difficult to interpret. This was notably the case for the chlamydial-like ADP-Glc–specific starch synthases postulated to have been crucial to the installation of the first plastids in eukaryotes (see Fig. 2 in Ball et al. 2013). Phylogenetic analyses have suggested that Chlamydiae are closely related to Planctomycetes, Verrucomicrobia, and some other less-known lineages, forming the "PVC superphylum" (see Wagner and Horn 2006), which has no particular affinity with the Cyanobacteria or the Proteobacteria. Thus, the close relationship of the Archaeplastida, Chlamydiae, and Cyanobacteria and/or Alphaproteobacteria observed in 12 of our phylogenetic trees did not reflect the expected bacterial phylogeny (i.e., the PVC relationship). Two alternative possibilities may explain those 12 phylogenies: either that Chlamydiae are not the donors

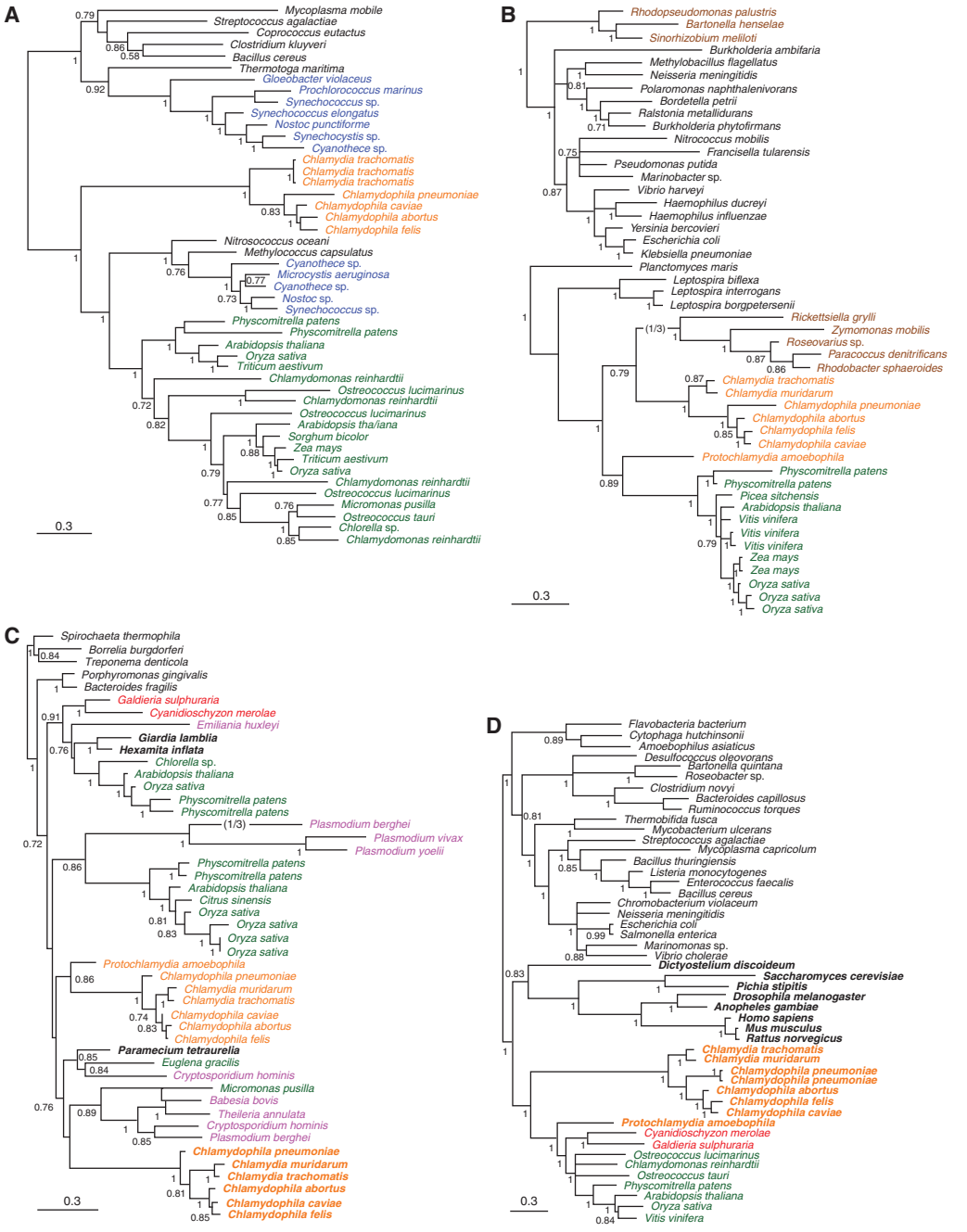

Figure 1. Bayesian phylogenetic trees of proteins suggested to have been acquired by photosynthetic eukaryotes from Chlamydiae. (*A*) Starch synthase/glycogen synthase. (*B*) Glycosyltransferase. (*C*) Pyrophosphate-dependent phosphofructokinase. (*D*) Tyrosyl-tRNA synthase. (Bold orange) Species names correspond to chlamydiae; (blue) cyanobacteria; (brown) alphaproteobacteria; (green) green algae and plants and eukaryotes with secondary green algal plastids; (red) red algae; (pink) eukaryotes with secondary red algal plastids. (Bold black) Species are nonphotosynthetic eukaryotes. Some branches shortened to one-third of their actual length are indicated by (1/3). Numbers at branches are posterior probabilities. The scale bar indicates the number of substitutions per position.

of those genes but the recipients in hypothetical Archaeplastida-to-Chlamydiae horizontal gene transfer (HGT) events, or that the ancestral cyanobacterium at the origin of the first plastid had acquired those genes from chlamydial donors and then transferred them to the eukaryotic host by EGT (Wolf et al. 1999). The fact that, as explained before, no chlamydial species is known to infect plants or algae could be seen as evidence against the first hypothesis. In contrast, the second hypothesis could also explain the observed Archaeplastida–Chlamydiae–Alphaproteobacteria relationships by Chlamydiae-to-Alphaproteobacteria HGTs followed by EGTs to eukaryotes through the mitochondrial endosymbiosis. This same idea has been proposed in more general grounds to explain the many non-alphaproteobacterial genes of bacterial origin found in a phylogenomic study of the yeast nuclear genome (Esser et al. 2004). Likewise, the cases in which Viridiplantae branched close to Chlamydiae but not the other archaeplastid lineages (which branched, as expected, close to Cyanobacteria) could also argue against the idea of a very early implication of Chlamydiae in the origin of Archaeplastida. This phylogenetic pattern has already been observed for the 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase, involved in carotenoid biosynthesis (Frommolt et al. 2008). Thus, the real contribution of Chlamydiae to the acquisition of the first plastids has to be seen as a not yet close, difficult evolutionary question.

CONCLUDING REMARKS

Comparative genomic and phylogenomic analyses have shown that eukaryotic genomes are intricate patchworks of genes of different evolutionary origins. This is a general situation because all known eukaryotes have, at least, integrated genes from the mitochondrial endosymbiont into their nuclear genomes. The case of photosynthetic eukaryotes is certainly the most complex one, because their nuclear genomes have been bombarded by bacterial genes coming from the primary plastids but also by eukaryotic genes provided by the nuclei of red or green algae in the case of secondary photosynthetic eukaryotes. The extent of the contribution of these different endosymbionts to the host genome remains uncertain, but some researchers consider that it may have been massive with, for example, ~75% of *Saccharomyces cerevisiae* genes of putative bacterial origin (Esser et al. 2004) and ~18% of *A. thaliana* genes of potential cyanobacterial origin (Martin et al. 2002). However, most of these analyses were performed several years ago with the limited taxonomic sampling available at that time, which may have induced an overestimation of the number of EGT events.

In addition, these are very ancient events that took place many millions of years ago, so that the phylogenetic signal conserved in individual genes has been considerably erased. In addition, the genes acquired by EGT had to adapt to a new genomic environment and then, often, accelerated their evolutionary rates (Baurain et al. 2010). The combination of these two factors makes the identification of those EGT genes very difficult. This is especially the case for secondary endosymbiotic events involving red and green algae. These two groups of algae are separated by a single node and a very short evolutionary distance in most phylogenetic trees, so even small biases in tree reconstruction can be enough to shift a sequence from one group to the other. Moreover, the distance between the Archaeplastida and other eukaryotic groups seems also to be very short, so that tree reconstruction artifacts can easily misplace artificially a sequence close to or far from them, not to speak about the numerous gene, or even complete genome, duplications that most eukaryotes have experienced during their evolution and that very frequently lead to hidden paralogy problems. Because of all these problems, it is naïve to believe that rigid automatic procedures can accurately identify the complete set of EGT genes in a given genome. On the contrary, these methods have been shown to greatly overestimate EGT to levels approaching not less than ~90% of false positives (Burki et al. 2012a; Deschamps and Moreira 2012). Not only careful analyses trying to control for as many of those biases as possible are unavoidable to address these questions, but, repeating

Burki et al.'s words, it is also important to avoid as much as possible the subjectivity in those analyses and remember to keep "extreme caution when drawing conclusions for major evolutionary events."

ACKNOWLEDGMENTS

We thank the French Centre National de la Recherche Scientifique (CNRS) for support.

REFERENCES

Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, et al. 2005. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J Eukaryot Microbiol* **52:** 399–451.

Ball SG, Subtil A, Bhattacharya D, Moustafa A, Weber AP, Gehre L, Colleoni C, Arias MC, Cenci U, Dauvillée D. 2013. Metabolic effectors secreted by bacterial pathogens: Essential facilitators of plastid endosymbiosis? *Plant Cell* **25:** 7–21.

Baum D. 2013. The origin of primary plastids: A pas de deux or a ménage à trois? *Plant Cell* **25:** 4–6.

Baurain D, Brinkmann H, Petersen J, Rodríguez-Ezpeleta N, Stechmann A, Demoulin V, Roger AJ, Burger G, Lang BF, Philippe H. 2010. Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles. *Mol Biol Evol* **27:** 1698–1709.

Becker B, Hoef-Emden K, Melkonian M. 2008. Chlamydial genes shed light on the evolution of photoautotrophic eukaryotes. *BMC Evol Biol* **8:** 203.

Brinkman FS, Blanchard JL, Cherkasov A, Av-Gay Y, Brunham RC, Fernandez RC, Finlay BB, Otto SP, Ouellette BF, Keeling PJ, et al. 2002. Evidence that plant-like genes in *Chlamydia* species reflect an ancestral relationship between Chlamydiaceae, cyanobacteria, and the chloroplast. *Genome Res* **12:** 1159–1167.

Burger G, Gray MW, Lang BF. 2003. Mitochondrial genomes: Anything goes. *Trends Genet* **19:** 709–716.

Burki F, Flegontov P, Oborník M, Cihlár J, Pain A, Lukes J, Keeling PJ. 2012a. Re-evaluating the green versus red signal in eukaryotes with secondary plastid of red algal origin. *Genome Biol Evol* **4:** 626–635.

Burki F, Okamoto N, Pombert JF, Keeling PJ. 2012b. The evolutionary history of haptophytes and cryptophytes: Phylogenomic evidence for separate origins. *Proc Biol Sci* **279:** 2246–2254.

Cavalier-Smith T. 1981. Eukaryote kingdoms: Seven or nine? *Biosystems* **14:** 461–481.

de Koning AP, Keeling PJ. 2004. Nucleus-encoded genes for plastid-targeted proteins in *Helicosporidium*: Functional diversity of a cryptic plastid in a parasitic alga. *Eukaryot Cell* **3:** 1198–1205.

Delwiche CF. 1999. Tracing the thread of plastid diversity through the tapestry of life. *Am Nat* **154:** S164–S177.

Deschamps P, Moreira D. 2012. Reevaluating the green contribution to diatom genomes. *Genome Biol Evol* **4:** 683–688.

Esser C, Ahmadinejad N, Wiegand C, Rotte C, Sebastiani F, Gelius-Dietrich G, Henze K, Kretschmann E, Richly E, Leister D, et al. 2004. A genome phylogeny for mitochondria among α-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. *Mol Biol Evol* **21:** 1643–1660.

Fast NM, Kissinger JC, Roos DS, Keeling PJ. 2001. Nuclear-encoded, plastid-targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids. *Mol Biol Evol* **18:** 418–426.

Fernández Robledo JA, Caler E, Matsuzaki M, Keeling PJ, Shanmugam D, Roos DS, Vasta GR. 2011. The search for the missing link: A relic plastid in *Perkinsus*? *Int J Parasitol* **41:** 1217–1229.

Frommolt R, Werner S, Paulsen H, Goss R, Wilhelm C, Zauner S, Maier UG, Grossman AR, Bhattacharya D, Lohr M. 2008. Ancient recruitment by chromists of green algal genes encoding enzymes for carotenoid biosynthesis. *Mol Biol Evol* **25:** 2653–2667.

Gray MW, Burger G, Lang BF. 1999. Mitochondrial evolution. *Science* **283:** 1476–1481.

Greub G, Raoult D. 2003. History of the ADP/ATP-translocase-encoding gene, a parasitism gene transferred from a Chlamydiales ancestor to plants 1 billion years ago. *Appl Environ Microbiol* **69:** 5530–5535.

Horn M. 2008. Chlamydiae as symbionts in eukaryotes. *Annu Rev Microbiol* **62:** 113–131.

Huang J, Gogarten JP. 2007. Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? *Genome Biol* **8:** R99.

Huang J, Mullapudi N, Lancto CA, Scott M, Abrahamsen MS, Kissinger JC. 2004. Phylogenomic evidence supports past endosymbiosis, intracellular and horizontal gene transfer in *Cryptosporidium parvum*. *Genome Biol* **5:** R88.

Junier T, Zdobnov EM. 2010. The Newick utilities: High-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* **26:** 1669–1670.

Keeling PJ. 2010. The endosymbiotic origin, diversification and fate of plastids. *Philos Trans R Soc Lond B Biol Sci* **365:** 729–748.

Keeling PJ. 2013. The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. *Annu Rev Plant Biol* **64:** 583–607.

Keeling PJ, Archibald JM. 2008. Organelle evolution: What's in a name? *Curr Biol* **18:** R345–R347.

Lecointre G, Philippe H, Vân Lê HL, Le Guyader H. 1993. Species sampling has a major impact on phylogenetic inference. *Mol Phylogenet Evol* **2:** 205–224.

Marin B, Nowack EC, Melkonian M. 2005. A plastid in the making: Evidence for a second primary endosymbiosis. *Protist* **156:** 425–432.

Martin W, Stoebe B, Goremykin V, Hapsmann S, Hasegawa M, Kowallik KV. 1998. Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* **393:** 162–165.

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D. 2002. Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands

of cyanobacterial genes in the nucleus. *Proc Natl Acad Sci* **99:** 12246–12251.

Matsuzaki M, Kuroiwa H, Kuroiwa T, Kita K, Nozaki H. 2008. A cryptic algal group unveiled: A plastid biosynthesis pathway in the oyster parasite *Perkinsus marinus*. *Mol Biol Evol* **25:** 1167–1179.

Moore RB, Oborník M, Janouskovec J, Chrudimský T, Vancová M, Green DH, Wright SW, Davies NW, Bolch CJ, Heimann K, et al. 2008. A photosynthetic alveolate closely related to apicomplexan parasites. *Nature* **451:** 959–963.

Moreira D, Le Guyader H, Philippe H. 2000. The origin of red algae and the evolution of chloroplasts. *Nature* **405:** 69–72.

Moustafa A, Bhattacharya D. 2008. PhyloSort: A user-friendly phylogenetic sorting tool and its application to estimating the cyanobacterial contribution to the nuclear genome of *Chlamydomonas*. *BMC Evol Biol* **8:** 6.

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.

Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D. 2009. Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science* **324:** 1724–1726.

Nowack EC, Melkonian M, Glöckner G. 2008. Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. *Curr Biol* **18:** 410–418.

Reyes-Prieto A, Moustafa A, Bhattacharya D. 2008. Multiple genes of apparent algal origin suggest ciliates may once have been photosynthetic. *Curr Biol* **18:** 956–962.

Rodríguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Löffelhardt W, Bohnert HJ, Philippe H, Lang BF. 2005. Monophyly of primary photosynthetic eukaryotes: Green plants, red algae, and glaucophytes. *Curr Biol* **15:** 1325–1330.

Roos DS, Crawford MJ, Donald RG, Fraunholz M, Harb OS, He CY, Kissinger JC, Shaw MK, Striepen B. 2002. Mining the *Plasmodium* genome database to define organellar function: What does the apicoplast do? *Philos Trans R Soc Lond B Biol Sci* **357:** 35–46.

Sanchez-Puerta MV, Lippmeier JC, Apt KE, Delwiche CF. 2007. Plastid genes in a non-photosynthetic dinoflagellate. *Protist* **158:** 105–117.

Schmitz-Esser S, Linka N, Collingro A, Beier CL, Neuhaus HE, Wagner M, Horn M. 2004. ATP/ADP translocases: A common feature of obligate intracellular amoebal symbionts related to Chlamydiae and Rickettsiae. *J Bacteriol* **186:** 683–691.

Slamovits CH, Keeling PJ. 2008. Plastid-derived genes in the nonphotosynthetic alveolate *Oxyrrhis marina*. *Mol Biol Evol* **25:** 1297–1306.

Stelter K, El-Sayed NM, Seeber F. 2007. The expression of a plant-type ferredoxin redox system provides molecular evidence for a plastid in the early dinoflagellate *Perkinsus marinus*. *Protist* **158:** 119–130.

Wagner M, Horn M. 2006. The Planctomycetes, Verrucomicrobia, Chlamydiae and sister phyla comprise a superphylum with biotechnological and medical relevance. *Curr Opin Biotechnol* **17:** 241–249.

Weeden NF. 1981. Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. *J Mol Evol* **17:** 133–139.

Williams BA, Keeling PJ. 2003. Cryptic organelles in parasitic protists and fungi. *Adv Parasitol* **54:** 9–68.

Woehle C, Dagan T, Martin WF, Gould SB. 2011. Red and problematic green phylogenetic signals among thousands of nuclear genes from the photosynthetic and apicomplexa-related *Chromera velia*. *Genome Biol Evol* **3:** 1220–1230.

Wolf YI, Aravind L, Koonin EV. 1999. Rickettsiae and Chlamydiae: Evidence of horizontal gene transfer and gene exchange. *Trends Genet* **15:** 173–175.

Cold Spring Harbor Perspectives in Biology

What Was the Real Contribution of Endosymbionts to the Eukaryotic Nucleus? Insights from Photosynthetic Eukaryotes

David Moreira and Philippe Deschamps

Cold Spring Harb Perspect Biol 2014; doi: 10.1101/cshperspect.a016014