

# Endosymbiotic and horizontal gene transfer in microbial eukaryotes

## Impacts on cell evolution and the tree of life

Cheong Xin Chan,<sup>1</sup> Debashish Bhattacharya<sup>2,\*</sup> and Adrian Reyes-Prieto<sup>3</sup>

<sup>1</sup>The University of Queensland; Institute for Molecular Bioscience and ARC Centre of Excellence in Bioinformatics; Brisbane, Queensland Australia; <sup>2</sup>Department of Ecology and Evolution and Institute for Marine and Coastal Sciences; Rutgers University; New Brunswick, NJ USA; <sup>3</sup>Canadian Institute for Advanced Research and Department of Biology; University of New Brunswick; Fredericton, New Brunswick Canada

**T**he evolution of microbial eukaryotes, in particular of photosynthetic lineages, is complicated by multiple instances of endosymbiotic and horizontal gene transfer (E/HGT) resulting from plastid origin(s). Our recent analysis of diatom membrane transporters provides evidence of red and/or green algal origins of 172 of the genes encoding these proteins (ca. 25% of the examined phylogenies), with the majority putatively derived from green algae. These data suggest that E/HGT has been an important driver of evolutionary innovation among diatoms (and likely other stramenopiles), and lend further support to the hypothesis of an ancient, cryptic green algal endosymbiosis in “chromalveolate” lineages. Here, we discuss the implications of our findings on the understanding of eukaryote evolution and inference of the tree of life.

Single-celled photosynthetic eukaryotes are significant contributors to total biomass production in oceanic and freshwater environments.<sup>1,2</sup> The origin of photoautotrophic metabolism in eukaryotes is explained by a primary endosymbiotic event with captured cyanobacteria that evolved into the plastids in the common ancestor of the Plantae (i.e., red algae, green algae and plants, and glaucophytes<sup>3–5</sup>). The plastids have subsequently spread into a multitude of other eukaryote lineages such as euglenids, chlorarachniophytes, and the diverse chlorophyll *c*-containing algae (often referred to as “chromalveolates”) through secondary and tertiary eukaryote-eukaryote endosymbiosis.<sup>6,7</sup> The complex evolutionary history

of algae with secondary and tertiary plastids also includes enigmatic non-photosynthetic related groups that presumably have lost photosynthetic capacity several times independently and, in other cases, replaced their original plastid via a subsequent endosymbiosis.<sup>8</sup> Analysis of the recently sequenced genome of the glaucophyte *Cyanophora paradoxa* substantiates Plantae monophyly and the single origin of primary plastids.<sup>4</sup> Importantly, the evidence for this long-sought after result comes not from multigene trees but instead from the more convincing analysis of groups of genes that are involved in complex processes such as fermentation, plastid solute transport, and plastid protein translocation.<sup>4</sup> In contrast, even with the availability of genome data from diverse “chromalveolates,” the phylogenetic relationships between these groups (e.g., the ubiquitous diatoms, haptophytes, and cryptophytes) remain murky and multigene trees again do not provide unambiguous evidence for their union as a supergroup. These types of analyses are either too difficult to interpret or suggest more complex evolutionary scenarios to explain “chromalveolate” interrelationships.<sup>9–11</sup> The persistent problem of resolving the phylogenetic relationships of “chromalveolates” is likely not due to the failure of phylogenetic approaches in general to resolve evolutionary history (although confounding issues such as long-branch attraction abound). Rather, these taxa share a feature that may grossly mislead phylogenetic inference: extensive horizontal gene transfer (HGT). Early genome-wide studies of “chromalveolates”<sup>12</sup> and chlorarachniophytes<sup>13</sup> revealed a substantial contribution

**Keywords:** horizontal gene transfer, endosymbiotic gene transfer, diatoms, membrane transporters, eukaryote evolution, tree of life

Submitted: 02/09/12

Accepted: 03/22/12

<http://dx.doi.org/10.4161/mge.20110>

\*Correspondence to: Debashish Bhattacharya;  
Email: [bhattacharya@aesop.rutgers.edu](mailto:bhattacharya@aesop.rutgers.edu)

Commentary to: Chan CX, Reyes-Prieto A, Bhattacharya D. Red and green algal origin of diatom membrane transporters: insights into environmental adaptation and cell evolution. PLoS One 2011; 6:e29138; PMID:22195008; <http://dx.doi.org/10.1371/journal.pone.0029138>

of HGT to their nuclear genomes and, importantly, hundreds of transferred genes that were recruited via endosymbiotic gene transfer (EGT). In “chromalveolates,” EGT is a specific case of HGT resulting from secondary and tertiary algal endosymbiosis. The potential phylogenetic confusion spawned by E/HGT is magnified with each round of secondary endosymbiosis that has occurred, with each event contributing novel sequences to the host genome that add to or replace existing genes.<sup>14,15</sup>

The extent of EGT can be relatively easily discerned in the case of primary endosymbiosis (i.e., Plantae lineages), whereby the transferred genes are of prokaryotic provenance (i.e., cyanobacterial and  $\alpha$ -proteobacterial origin, respectively, for the plastid and mitochondrion) and differ greatly in sequence from the resident eukaryotic nuclear genes. Therefore, the finding of hundreds of nuclear genes in algae and other eukaryotes with a prokaryotic phylogenetic signature is considered unambiguous evidence for EGT associated with the primary endosymbiotic origin of organelles.<sup>16,17</sup> This level of resolution may not occur, however, in the case of serial, eukaryote-eukaryote endosymbioses. In these instances, the closer the phylogenetic affiliation between the donor (the endosymbiont) and recipient (the host) lineages the less likely that standard phylogenetic methods can distinguish bona fide EGT events from diverged host genes. This expectation of highly reticulate relationships driven by serial eukaryotic endosymbioses involving closely related lineages has been used to explain the great difficulties in establishing a robust “chromalveolate” tree of life using multigene phylogenetic inference.<sup>9,11</sup>

Demonstrable evidence for the existence of serial eukaryotic endosymbiosis was recently provided with the analysis of complete genome data from diatoms. Moustafa et al.<sup>18</sup> showed the presence of hundreds of genes of green algal origin in *Thalassiosira pseudonana* and *Phaeodactylum tricoratum*, two species that possess a plastid of unambiguous red algal origin. The “green genes” reside alongside genes of red algal origin in these diatom nuclear genomes and are postulated to derive from an ancient, cryptic green algal secondary

endosymbiosis (with associated EGT) that predated the capture of the red algal plastid. This initial “controversial” finding has now been supported with analysis of complete genome data from other stramenopiles (e.g., *Ectocarpus siliculosus*<sup>19</sup>). Given the growing evidence suggesting multiple algal endosymbiosis in some lineages, the key questions become: what specific functions do these green genes have and with a putative residence time of hundreds of millions of years, how have they contributed to the large-scale evolution of the host lineage?

### Evolution of the Diatom Permeome and the “Green-Algal” Footprints

We recently completed a phylogenomic analysis of genes encoding membrane transporters in the nuclear genomes of diatoms *T. pseudonana* and *P. tricoratum*.<sup>20</sup> In addition to examples of long-term vertical inheritance of membrane transporters in these and other eukaryotes (Fig. 1A), we also found putative red and green algal origins (Fig. 1B) of many of these genes. Specifically, 172 (ca. 25%) of the 697 encoded membrane transporters were found to be of Plantae origin. Although 50 of these show an unresolved affiliation with red and/or green algae, interestingly, the majority (103) are of putative green algal provenance compared with 19 of putative red algal origin. To explain their long-term retention, we hypothesized that these genes, presumably derived from different algal secondary endosymbionts, contributed important membrane transport functions to diatoms.<sup>20</sup>

The fluctuations of redox-sensitive transition metals over evolutionary time-scales driven most notably by the rise of oxygen ca. 2.5 BYA, would have introduced immense selective pressure for cell survival in aquatic environments. Such environmental pressure is particularly relevant to the trafficking of molecules across cellular membrane, whether to acquire useful molecular compounds or to expel molecules that cause harm to the cell (i.e., cell detoxification). Free-living unicellular algae are known to possess efficient transport systems, particularly of metal cations such as sodium and potassium, to maintain cell homeostasis with

respect to the inter- and extra-cellular environments.<sup>21</sup> Therefore, the acquisition of red and/or green algal genes via E/HGT could have been a crucial selective force for the survival of ancient, and hence extant diatom lineages. Our observations are also in general agreement with the hypothesis of a cryptic green algal endosymbiosis in diatoms,<sup>18</sup> and likely more generally in “chromalveolates.”<sup>22,23</sup> And finally, these data<sup>20</sup> suggest that the cryptic endosymbiont contributes gene functions that extend beyond photosynthetic capacity to roles related to cell adaptation to the external environment.

### Implications for Inferring the Tree of Life

A well-sampled phylogenetic tree of eukaryotes (i.e., the eukaryote tree of life) is essential to understand the tempo and mode of endosymbiotic events that have occurred during algal evolution, and to test different hypotheses about the relationships between the major algal groups. The recent findings of extensive EGT from putative serial endosymbioses,<sup>9,18,20,23</sup> in combination with HGT events from diverse sources,<sup>24</sup> stresses the significant impact that gene transfer has had on algal and eukaryote evolution in general. Gene transfer events are not only of significance to cell evolution, but may also mislead inference of the tree of life. Some of our previous studies<sup>3,25</sup> were aimed at identifying genes that do not “follow” the presumed vertical (e.g., host lineage) history of algal groups and their non-photosynthetic relatives (e.g., Plantae, Alveolata and stramenopiles) to study the endosymbiotic origin of plastids, and, importantly, the impact of E/HGT on algal genome evolution. It has now become clear that the identification of bona fide cases of EGT or HGT is often saddled by well-recognized stochastic errors (e.g., insufficient data) and systematic biases (e.g., base composition, high mutation rates) inherent to molecular phylogenetics that can lead to equivocal support for phylogenetic hypotheses.<sup>26,27</sup>

In spite of these difficulties, independent analyses using multiple single-locus phylogenetic trees provide support for the existence of EGT during the evolution of



In recent years, our understanding of algal evolution and diversification has benefited greatly from the availability of genome data from non-photosynthetic groups such as picobiliphytes<sup>39</sup> and katablepharids.<sup>10</sup> The inclusion of these data in phylogenetic analyses has necessitated the invocation of complex evolutionary scenarios to explain relationships between the major algal and protist lineages. The original idea of a united “chromalveolate”

lineage is no longer tenable after the identification of the SAR group<sup>40</sup> and the finding of the intermingled positions of picobiliphytes,<sup>39</sup> telonemids, and katablepharids with photosynthetic “chromalveolates.”<sup>9,10</sup> Genome-scale data from many of these algal and protist groups is still lacking. Major sequencing efforts of key algal groups (e.g., red algae, glaucophytes, prasinophytes) and non-photosynthetic taxa (e.g., picobiliphytes, telonemids,

katablepharids) are therefore required both for exploring the relationships between these lineages and for elucidating the number and tempo of endosymbiotic events that have generated the astonishing diversity of microbial eukaryotes.

#### Acknowledgments

We acknowledge generous support from Rutgers University and valuable discussions with colleagues in the Bhattacharya lab.

#### References

- Nelson DM, Tréguer P, Brzezinski MA, Leynaert A, Quéguiner B. Production and dissolution of biogenic silica in the ocean: revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biogeochem Cycles* 1995; 9:359-72; <http://dx.doi.org/10.1029/95GB01070>
- Cuvelier ML, Allen AE, Monier A, McCrow JP, Messié M, Tringe SG, et al. Targeted metagenomics and ecology of globally important uncultured eukaryotic phytoplankton. *Proc Natl Acad Sci U S A* 2010; 107:14679-84; PMID:20668244; <http://dx.doi.org/10.1073/pnas.1001665107>
- Chan CX, Yang EC, Banerjee T, Yoon HS, Martone PT, Estevez JM, et al. Red and green algal monophyly and extensive gene sharing found in a rich repertoire of red algal genes. *Curr Biol* 2011; 21:328-33; PMID:21315598; <http://dx.doi.org/10.1016/j.cub.2011.01.037>
- Price DC, Chan CX, Yoon HS, Yang EC, Qiu H, Weber AP, et al. *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. *Science* 2012; 335:843-7; PMID:22344442; <http://dx.doi.org/10.1126/science.1213561>
- Rodríguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Löffelhardt W, et al. Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr Biol* 2005; 15:1325-30; PMID:16051178; <http://dx.doi.org/10.1016/j.cub.2005.06.040>
- Archibald JM. The puzzle of plastid evolution. *Curr Biol* 2009; 19:R81-8; PMID:19174147; <http://dx.doi.org/10.1016/j.cub.2008.11.067>
- Keeling PJ. The endosymbiotic origin, diversification and fate of plastids. *Philos Trans R Soc Lond B Biol Sci* 2010; 365:729-48; PMID:20124341; <http://dx.doi.org/10.1098/rstb.2009.0103>
- Moore CE, Archibald JM. Nucleomorph genomes. *Annu Rev Genet* 2009; 43:251-64; PMID:19686079; <http://dx.doi.org/10.1146/annurev-genet-102108-134809>
- Baurain D, Brinkmann H, Petersen J, Rodríguez-Ezpeleta N, Stechmann A, Demoulin V, et al. Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles. *Mol Biol Evol* 2010; 27:1698-709; PMID:20194427; <http://dx.doi.org/10.1093/molbev/msq059>
- Burki F, Okamoto N, Pombert JF, Keeling PJ. The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. *Proc Biol Sci* 2012. In press. PMID:22298847; <http://dx.doi.org/10.1098/rspb.2011.2301>
- Parfrey LW, Grant J, Tekle YI, Lasek-Nesselquist E, Morrison HG, Sogin ML, et al. Broadly sampled multigene analyses yield a well-resolved eukaryotic tree of life. *Syst Biol* 2010; 59:518-33; PMID:20656852; <http://dx.doi.org/10.1093/sysbio/syq037>
- Nosenko T, Bhattacharya D. Horizontal gene transfer in chromalveolates. *BMC Evol Biol* 2007; 7:173; PMID:17894863; <http://dx.doi.org/10.1186/1471-2148-7-173>
- Archibald JM, Rogers MB, Toop M, Ishida K, Keeling PJ. 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* 2003; 100:7678-83; PMID:12777624; <http://dx.doi.org/10.1073/pnas.1230951100>
- Ishida K, Green BR. 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* 2002; 99:9294-9; PMID:12089328; <http://dx.doi.org/10.1073/pnas.142091799>
- Yoon HS, Hackett JD, Van Dolah FM, Nosenko T, Lidie KL, Bhattacharya D. Tertiary endosymbiosis driven genome evolution in dinoflagellate algae. *Mol Biol Evol* 2005; 22:1299-308; PMID:15746017; <http://dx.doi.org/10.1093/molbev/msi118>
- McFadden GI. Primary and secondary endosymbiosis and the origin of plastids. *J Phycol* 2001; 37:951-9; <http://dx.doi.org/10.1046/j.1529-8817.2001.01126.x>
- Moreira D, Le Guyader H, Philippe H. The origin of red algae and the evolution of chloroplasts. *Nature* 2000; 405:69-72; PMID:10811219; <http://dx.doi.org/10.1038/35011054>
- Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D. Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science* 2009; 324:1724-6; PMID:19556510; <http://dx.doi.org/10.1126/science.1172983>
- Cock JM, Sterck L, Rouzé P, Scornet D, Allen AE, Amoutzias G, et al. The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. *Nature* 2010; 465:617-21; PMID:20520714; <http://dx.doi.org/10.1038/nature09016>
- Chan CX, Reyes-Prieto A, Bhattacharya D. Red and green algal origin of diatom membrane transporters: insights into environmental adaptation and cell evolution. *PLoS One* 2011; 6:e29138; PMID:22195008; <http://dx.doi.org/10.1371/journal.pone.0029138>
- Krämer U, Talke IN, Hanikenne M. Transition metal transport. *FEBS Lett* 2007; 581:2263-72; PMID:17462635; <http://dx.doi.org/10.1016/j.febslet.2007.04.010>
- Dorrell RG, Smith AG. Do red and green make brown?: perspectives on plastid acquisitions within chromalveolates. *Eukaryot Cell* 2011; 10:856-68; PMID:21622904; <http://dx.doi.org/10.1128/EC.00326-10>
- Woehle C, Dagan T, Martin WF, Gould SB. Red and problematic green phylogenetic signals among thousands of nuclear genes from the photosynthetic and apicomplexa-related *Chromera velia*. *Genome Biol Evol* 2011; 3:1220-30; PMID:21965651; <http://dx.doi.org/10.1093/gbe/evr100>
- Suzuki K, Miyagishima SY. Eukaryotic and eubacterial contributions to the establishment of plastid proteome estimated by large-scale phylogenetic analyses. *Mol Biol Evol* 2010; 27:581-90; PMID:19910386; <http://dx.doi.org/10.1093/molbev/msp273>
- Moustafa A, Bhattacharya D. PhyloSort: a user-friendly phylogenetic sorting tool and its application to estimating the cyanobacterial contribution to the nuclear genome of *Chlamydomonas*. *BMC Evol Biol* 2008; 8:6; PMID:18194581; <http://dx.doi.org/10.1186/1471-2148-8-6>
- Dagan T, Martin W. Seeing green and red in diatom genomes. *Science* 2009; 324:1651-2; PMID:19556490; <http://dx.doi.org/10.1126/science.1175765>
- Stiller JW. Experimental design and statistical rigor in phylogenomics of horizontal and endosymbiotic gene transfer. *BMC Evol Biol* 2011; 11:259; PMID:21923904; <http://dx.doi.org/10.1186/1471-2148-11-259>
- Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, et al. 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* 2002; 99:12246-51; PMID:12218172; <http://dx.doi.org/10.1073/pnas.182432999>
- Li S, Nosenko T, Hackett JD, Bhattacharya D. Phylogenomic analysis identifies red algal genes of endosymbiotic origin in the chromalveolates. *Mol Biol Evol* 2006; 23:663-74; PMID:16357039; <http://dx.doi.org/10.1093/molbev/msj075>
- Nowack EC, Vogel H, Groth M, Grossman AR, Melkonian M, Glöckner G. Endosymbiotic gene transfer and transcriptional regulation of transferred genes in *Paulinella chromatophora*. *Mol Biol Evol* 2011; 28:407-22; PMID:20702568; <http://dx.doi.org/10.1093/molbev/msq209>
- Reyes-Prieto A, Yoon HS, Moustafa A, Yang EC, Andersen RA, Boo SM, et al. Differential gene retention in plastids of common recent origin. *Mol Biol Evol* 2010; 27:1530-7; PMID:20123796; <http://dx.doi.org/10.1093/molbev/msq032>
- Andam CP, Gogarten JP. Biased gene transfer and its implications for the concept of lineage. *Biol Direct* 2011; 6:47; PMID:21943000; <http://dx.doi.org/10.1186/1745-6150-6-47>

33. Skippington E, Ragan MA. Lateral genetic transfer and the construction of genetic exchange communities. *FEMS Microbiol Rev* 2011; 35:707-35; PMID: 21223321; <http://dx.doi.org/10.1111/j.1574-6976.2010.00261.x>
34. Reyes-Prieto A, Bhattacharya D. Phylogeny of Calvin cycle enzymes supports Plantae monophyly. *Mol Phylogenet Evol* 2007; 45:384-91; PMID:17482838; <http://dx.doi.org/10.1016/j.ympev.2007.02.026>
35. Gross J, Bhattacharya D. Reevaluating the evolution of the Toc and Tic protein translocons. *Trends Plant Sci* 2009; 14:13-20; PMID:19042148; <http://dx.doi.org/10.1016/j.tplants.2008.10.003>
36. Linka N, Hurka H, Lang BF, Burger G, Winkler HH, Stamme C, et al. Phylogenetic relationships of non-mitochondrial nucleotide transport proteins in bacteria and eukaryotes. *Gene* 2003; 306:27-35; PMID: 12657464; [http://dx.doi.org/10.1016/S0378-1119\(03\)00429-3](http://dx.doi.org/10.1016/S0378-1119(03)00429-3)
37. Frommolt R, Werner S, Paulsen H, Goss R, Wilhelm C, Zauner S, et al. Ancient recruitment by chromists of green algal genes encoding enzymes for carotenoid biosynthesis. *Mol Biol Evol* 2008; 25:2653-67; PMID: 18799712; <http://dx.doi.org/10.1093/molbev/msn206>
38. Ball S, Colleoni C, Cenci U, Raj JN, Tirtiaux C. The evolution of glycogen and starch metabolism in eukaryotes gives molecular clues to understand the establishment of plastid endosymbiosis. *J Exp Bot* 2011; 62:1775-801; PMID:21220783; <http://dx.doi.org/10.1093/jxb/erq411>
39. Yoon HS, Price DC, Stepanauskas R, Rajah VD, Sieracki ME, Wilson WH, et al. Single-cell genomics reveals organismal interactions in uncultivated marine protists. *Science* 2011; 332:714-7; PMID:21551060; <http://dx.doi.org/10.1126/science.1203163>
40. Hackett JD, Yoon HS, Li S, Reyes-Prieto A, Rümmele SE, Bhattacharya D. Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of Rhizaria with chromalveolates. *Mol Biol Evol* 2007; 24:1702-13; PMID:17488740; <http://dx.doi.org/10.1093/molbev/msm089>