

Multiple Gene Phylogenies Support the Monophyly of Cryptomonad and Haptophyte Host Lineages

Nicola J. Patron,^{1,2} Yuji Inagaki,³
and Patrick J. Keeling^{1,*}

¹Department of Botany
University of British Columbia
3529-6270 University Boulevard
Vancouver, BC V6T 1Z4
Canada

²School of Botany
University of Melbourne
Parkville, Victoria 3010
Australia

³University of Tsukuba
Center for Computational Sciences
Institute of Biological Sciences
Tsukuba, Ibaraki 305-8577
Japan

Summary

Cryptomonad algae acquired their plastids by the secondary endosymbiotic uptake of a eukaryotic red alga. Several other algal lineages acquired plastids through such an event [1], but cryptomonads are distinguished by the retention of a relic red algal nucleus, the nucleomorph [2]. The nucleomorph (and its absence in other lineages) can reveal a great deal about the process and history of endosymbiosis, but only if we know the relationship between cryptomonads and other algae, and this has been controversial. Several recent analyses have suggested a relationship between plastids of cryptomonads and some or all other red alga-containing lineages [3–6], but we must also know whether host nuclear genes mirror this relationship to determine the number of endosymbiotic events, and this has not been demonstrated. We have carried out an expressed sequence tag (EST) survey of the cryptomonad *Guillardia theta*. Phylogenetic analyses of 102 orthologous nucleus-encoded proteins (18,425 amino acid alignment positions) show a robust sister-group relationship between cryptomonads and the haptophyte algae, which also have a red secondary plastid. This relationship demonstrates that loss of nucleomorphs must have taken place in haptophytes independently of any other red alga-containing lineages and that the ancestor of both already contained a red algal endosymbiont.

Results and Discussion

Nuclear Genes Significantly Support the Monophyly of Cryptomonads and Haptophytes

The cryptomonad endosymbiont retains two ancestral features that all other red algal secondary plastids have lost: the relic nucleus, or nucleomorph, and light-

harvesting accessory phycobiliproteins. These characteristics have singled out the cryptomonads and led to a number of hypotheses for their ancient origin, independent of other algal groups (e.g., [7–10]). Molecular phylogenetic analyses of nuclear genes have largely been congruent with this notion, because cryptomonads have not consistently been allied with any other algal lineage with any support, or indeed any lineage at all [11–15]. In contrast to this, recent analyses of plastid-encoded and nucleus-encoded plastid-targeted proteins have consistently suggested that cryptomonads are part of a large and diverse group that includes many or all other lineages that possess red algal-derived plastids, the so-called chromalveolates [8, 16–18]. Even in these data there is uncertainty, however, because the specific relationship between cryptomonads and other chromalveolates is generally unresolved or inconsistent. Plastid gene phylogenies have suggested a basal position [5, 6], whereas the shared presence of a laterally transferred bacterial *rp136* in their plastid genomes has suggested a relationship between cryptomonads and haptophyte algae [4]. Altogether, data from plastids are converging on the conclusion that the plastid components of cryptomonads are somehow related to the red algal-derived plastids found in three other lineages (alveolates, heterokonts, and haptophytes). This is an intriguing possibility, because this would suggest that the ancestral features found in cryptomonads had been retained in this group while being lost in its close relatives. However, until we determine whether the phylogenetic history of these plastids is congruent with that of their hosts, whether cryptomonad plastids arose directly from the same endosymbiosis as other algal plastids will remain unclear, as will the significance of the ancestral features retained by cryptomonads.

Large data sets consisting of multiple nuclear genes are increasingly being used to examine difficult, ancient relationships among eukaryotic lineages [19–22]. We sequenced 17,652 expressed sequence tags (ESTs) from the model cryptomonad, *Guillardia theta*, resulting in 6,267 unique clusters. Genes from this survey were added to alignments previously used to infer large-scale phylogenetic relationships among eukaryotes [20]. We also added data from two haptophyte EST surveys [23] and all other chromalveolate taxa for which sufficient sequence data were available.

The final data set included 102 protein alignments amounting to 18,425 unambiguously aligned amino acid positions from 38 taxa. We also analyzed a 34-taxon alignment where the excavates, *Giardia intestinalis*, *Trichomonas vaginalis*, *Trypanosoma brucei*, and *T. cruzi*, were excluded, because these have been argued previously to disrupt the analyses because of their relatively fast-evolving nature [20]. Similarly, fast-evolving alignment positions have been shown to bias phylogenetic estimates under the linked model conditions [24, 25]. Accordingly, we also analyzed 34- and 38-taxon data sets where the 1,966 fastest evolving positions

*Correspondence: pkeeling@interchange.ubc.ca

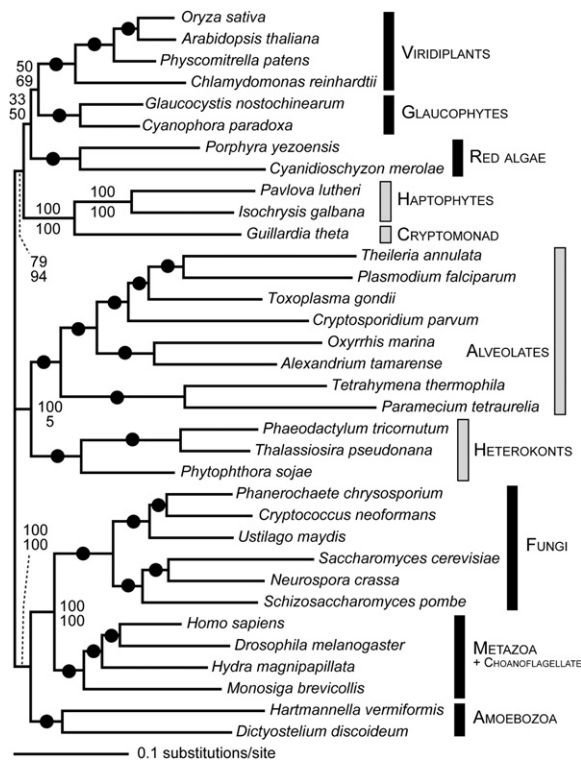


Figure 1. Linked Maximum-Likelihood Analyses of 16,459 Slow-Evolving Amino Acid Positions from 102 Proteins and 34 Taxa
Numbers at nodes are bootstrap values from the 34-taxon analysis (above branches) and the 38-taxon analysis including *Giardia*, *Trichomonas*, and two trypanosomes (below branches). Filled circles indicate bootstrap support of greater than 95% in both 34-taxon and 38-taxon analyses.

were excluded, leaving 16,459 slow-evolving positions (trees inferred from data including fast-evolving sites are shown in [Figure S1](#) in the [Supplemental Data](#) available online). Data were analyzed by linked (concatenated) and unlinked (separate) maximum likelihood (ML) models.

The linked ML tree of the 34-taxon data set ([Figure 1](#), support values above branches) recovered a number of relationships observed previously; monophyletic groups uniting heterokonts and alveolates, and opisthokonts and amoebzoa were robustly supported (BP = 100%). The union of glaucophytes, red algae, and viridiplants (Plantae) received poor support (BP = 50%), probably because of insufficient data [20]. The newly added data from cryptomonads and haptophytes is of particular interest, however, because these two groups formed a clade with robust support (BP = 100%). A monophyletic chromalveolates was not recovered, and instead the cryptomonad/haptophyte clade branched with moderate support as sister to the Plantae (BP = 79%). When excavates were included, linked ML analysis of slow-evolving sites revealed no impact on the relationship between cryptomonads and haptophytes ([Figure 1](#), support values below branches; [Figure S2](#)). In contrast, the relationship between heterokonts and alveolates was drastically changed, because excavates robustly grouped with alveolates (BP = 95%), a result not observed in analyses of larger data sets [20].

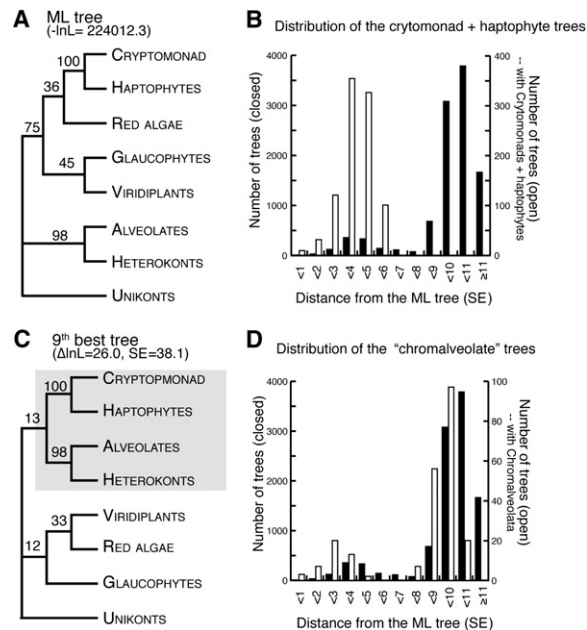


Figure 2. Unlinked Maximum-Likelihood Analyses of 8,359 Amino Acid Positions from 52 Proteins

(A) Unrooted ML tree with unscaled branch lengths and RELL bootstrap values.

(B) Distribution of 945 trees with cryptomonad/haptophyte monophyly. All cryptomonad-haptophyte trees (open bars) are less than 6 standard errors (SE) away from the ML tree. The distribution of the 10,395 test trees are indicated by closed bars. Note that the vast majority of the test trees are more than 8 SE away from the ML tree.

(C) Ninth-best tree (with chromalveolate monophyly). The details are same as (A).

(D) Distribution of 225 trees with chromalveolate monophyly. Comparing to the overall tree distribution (closed bars), no clear trend was found for the distribution of the chromalveolate trees (open bars).

Previously published studies consistently indicated that linked models, which treat a multigene data set as a single super matrix, cannot adequately describe data sets consisting of genes evolving with different tempos and modes, and this introduces artifacts to the tree reconstruction (e.g., [26–28]). We therefore conducted a multigene analysis under the unlinked model, accounting for gene-specific evolution. We calculated the log-likelihood (lnL) values for all the possible relationships among seven groups that received strong support in linked analyses: (1) unikonts, (2) heterokonts, (3) alveolates, (4) red algae, (5) glaucophytes, (6) viridiplants, (7) haptophytes, and the single species of cryptomonads. Each of the eight defined groups was therefore represented by at least one species in 52 single-gene data sets (8,359 positions). The same analysis was also performed excluding 663 fast-evolving positions, but the results were not significantly different (not shown). The 52-gene analysis recovered the monophyly of cryptomonads and haptophytes with strong support ([Figure 2A](#); RELL BP = 100%). Moreover, a close examination of the likelihood scores revealed that every tree including cryptomonad/haptophyte monophyly (945 trees in total) was distributed in the top 10% of the 10,395 test trees ([Figure 2B](#)), further supporting this relationship. The

approximate unbiased (AU) test also confirmed this conclusion, because every alternative tree in which cryptomonads and haptophytes were separated was rejected at the 1% level (see [Supplemental Data](#)).

Every analysis, both linked and unlinked, consistently and strongly indicate that cryptomonads and haptophytes share a common ancestor. This is significant because the ancestral features retained by the cryptomonad endosymbiont, in particular the nucleomorph and phycobiliproteins, must have been present in that ancestor. Therefore, regardless of whether or not the chromalveolate hypothesis is correct, these features must have been lost independently in the haptophyte and heterokont/alveolate lineages. In the case of the nucleomorph, its presence has led to speculation about why it has been retained in some lineages but lost in many others [2, 29]. The strongly supported relationship between cryptomonads and haptophytes (which certainly lack a nucleomorph) is the first definitive demonstration that nucleomorphs in two closely related lineages had opposite fates.

The question of nucleomorph fate is even more interesting in light of the recent discoveries of two other lineages that are closely related to the cryptomonads: katablephorids and picobiliphytes [30–32]. Katablephorids are heterotrophic predators that possess neither plastid nor nucleomorph [33], although in at least one case they do take up semipermanent green algal endosymbionts [33]. They are now known to be specifically related to cryptomonads [30, 31], so if the red algal plastid and its nucleus are ancestral cryptomonads and haptophytes, then they must also have been present in the ancestor of katablepharids. The same is true for the nonphotosynthetic cryptomonad, *Goniomonas*, where no evidence of a plastid or nucleomorph has been found [34]. Picobiliphytes, on the other hand, are a group of common, but until recently overlooked, marine picoplankton that possess plastids and appear to possess phycobiliproteins and a DNA-contained structure that may be a nucleomorph [31]. In molecular trees, they branch with the cryptomonad/katablepharid clade, but only with very weak support. Given the suite of characteristics so far reported for these little-understood cells, it is most likely that they are more closely related to cryptomonads than to haptophytes, but an intriguing possibility yet to be rejected is that picobiliphytes are the ancestor of both cryptomonads and haptophytes.

A Single Red-Algal Endosymbiosis?

The failure of a large collection of nuclear genes to recover a monophyletic chromalveolate clade may appear to challenge the hypothesis that all secondary red algal plastids derive from a single endosymbiosis. Data from plastid-targeted proteins glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and fructose-1,6-bisphosphate aldolase (FBA) suggest a monophyletic origin of chromalveolate plastids [3, 13, 16], although they cannot confirm a single endosymbiotic event. The separation of the cryptomonad/haptophyte clade from the heterokont/alveolate clade is also inconsistent with phylogenies of plastid-encoded genes, recent analyses of which support a grouping of cryptomonads, haptophytes, and heterokonts [5, 35], and to a lesser extent also peridinin-containing dinoflagellates [6, 36].

Once again, to definitively reconstruct the endosymbiotic history of these plastids, it is necessary to determine whether plastid and host relationships are congruent or not. Accordingly, we examined the nuclear gene data more closely to determine whether our data conclusively rejects chromalveolate monophyly. In the unlinked phylogeny inferred from the 52-gene data set, neither chromalveolate monophyly nor Plantae monophyly was recovered in the best, unrooted tree ([Figure 2A](#)). However, the ninth best tree, which was worse by only 26.0 InL units, possessed both chromalveolate and Plantae groups ([Figure 2C](#)). There is no practical difference between these two trees, because the difference in InL is within standard error (38.1) and the AU test failed to reject this alternative topology (p value of 0.223; [Supplemental Data](#)). Indeed, the distribution of the trees supporting chromalveolate monophyly among the 10,395 test trees shows no clear trend distinguishing them from a random sample ([Figure 2D](#)), strikingly unlike the trees that include cryptomonad/haptophyte monophyly ([Figure 2B](#)). This distribution suggests that there are an insufficient number of phylogenetically informative positions in the 52-gene data set to either confirm or reject chromalveolate monophyly with confidence.

We therefore performed a second unlinked analysis where monophyly of both cryptomonads/haptophytes and Plantae were assumed, which allowed additional single-gene data sets including either haptophytes or cryptomonads and those including either glaucophytes, red algae, or viridiplants to be included (resulting in 95 genes and 16,745 positions). Fast-evolving positions were retained because their exclusion had no impact in the first unlinked analysis (see above). In the 95-gene analysis, all 15 topologies possible for five groups, (1) unikonts, (2) cryptomonads/haptophytes, (3) alveolates, (4) heterokonts, and (5) Plantae, were evaluated. In the best tree, the cryptomonad/haptophyte clade was separate from an alveolate/heterokont clade, but in the second-best tree, chromalveolates were monophyletic and the InL was only 16.6 units lower. Similarly, the AU tests on all 15 topologies failed to reject this tree ($p = 0.318$; [Supplemental Data](#)). Including excavates did not alter the results ([Supplemental Data](#)).

Conclusions

Both linked and unlinked analyses of large protein data sets consistently and strongly support a close relationship between the host (nuclear) components of cryptomonads and haptophytes. This conclusion is also supported by data from their plastid genomes [4], and together they dispel two prevailing theories about cryptomonads: they are not the earliest-diverging lineage chromalveolates, and their plastids did not originate independently of other red algal-derived plastids. One implication of this is that the ancestral features of the cryptomonad endosymbiont, in particular the presence of the relic nucleomorph and the light-harvesting accessory phycobiliproteins, must have been lost repeatedly in other algae with secondary plastids of red algal origin: once in haptophytes and at least once in the other lineages. A number of other characters that differ in these groups (e.g., carbohydrate storage products and mitochondrial structure) will also be of interest when the tree is fully resolved and a detailed analysis of their

evolution can be performed. A second implication is that there was no inherent characteristic of the red algal nucleus that led inevitably to its retention, because the same nucleus was lost in haptophytes, katablepharids, and *Goniomonas*.

All these conclusions are equally true whether or not cryptomonads and haptophytes are part of the larger chromalveolate group. The 102 genes analyzed here did not support or reject this hypothesis, which is not unexpected because the addition of taxa has been shown to reduce support for previously robust lineages whereas the addition of more proteins has been shown to increase support regardless of the topology [37, 38]. Testing the monophyly of the chromalveolate host lineage will require more data, but the present results significantly alter any predictions about what might be found in such an analysis. Taking plastid data together with the results reported here, we suggest that additional analyses are most likely to show that chromalveolates are monophyletic but split into two major subgroups: cryptomonads/haptophytes and heterokonts/alveolates.

Experimental Procedures

Construction of cDNA Libraries, Sequencing, and Multiple Alignments

A description of the cDNA library construction and sequencing of *G. theta* (CCMP 327) can be found with the [Supplemental Data](#). The nuclear gene data set is based on an available alignment [20]. Data from *G. theta*, *I. galbana*, *P. lutheri*, *O. marina*, and *H. vermiformis* were added from TBestDB, and potential paralogs were identified and removed as described [20, 39]. Only positions containing unambiguously aligned characters were included. Details of representation of each taxon and each protein are given in the [Supplemental Data](#).

Phylogenetic Analyses

The 34-taxon and 38-taxon data sets were first subjected to site-by-site rate estimation. The positions assigned to the fastest-evolving category were omitted from the alignments, and 16,459 positions were used for linked ML phylogenetic analyses. Further details are given in the [Supplemental Data](#).

For the unlinked analyses of the 52-gene data set, the internal branching patterns in each group that was strongly supported to be monophyletic in the linked analyses were constrained in advance, and the 10,395 resulting tree topologies were exhaustively searched. The ML analyses under the unlinked model conditions were also conducted on the 89-gene and 95-gene data sets as described above. 945 and 15 tree topologies were assessed in the 89-gene and 95-gene analyses, respectively. Further details are given in the [Supplemental Data](#).

Supplemental Data

Two figures, five tables, and Experimental Procedures are available at <http://www.current-biology.com/cgi/content/full/17/10/887/DC1/>.

Acknowledgments

This work was supported by the Protist EST Program of Genome Canada/Genome Atlantic and a grant from the Natural Sciences and Engineering Research Council of Canada to P.J.K. Y.I. was supported in part by grants from the Japan Society for the Promotion of Science (No. 18570214). We thank N. Rodríguez-Ezpeleta, F. Lang, and H. Philippe for providing alignments, A. de Koning for assistance with updating them, and TBestDB. P.J.K. is a Fellow of the Canadian Institute for Advanced Research and a Senior Investigator of the Michael Smith Foundation for Health Research.

Received: February 28, 2007

Revised: March 18, 2007

Accepted: March 28, 2007

Published online: April 26, 2007

References

- Falkowski, P.G., Katz, M.E., Knoll, A.H., Quigg, A., Raven, J.A., Schofield, O., and Taylor, F.J. (2004). The evolution of modern eukaryotic phytoplankton. *Science* 305, 354–360.
- Douglas, S., Zauner, S., Fraunholz, M., Beaton, M., Penny, S., Deng, L.T., Wu, X., Reith, M., Cavalier-Smith, T., and Maier, U.G. (2001). The highly reduced genome of an enslaved algal nucleus. *Nature* 410, 1091–1096.
- Patron, N.J., Rogers, M.B., and Keeling, P.J. (2004). Gene replacement of fructose-1,6-bisphosphate aldolase supports the hypothesis of a single photosynthetic ancestor of chromalveolates. *Euk. Cell* 3, 1169–1175.
- Rice, D.W., and Palmer, J.D. (2006). An exceptional horizontal gene transfer in plastids: gene replacement by a distant bacterial paralog and evidence that haptophyte and cryptophyte plastids are sisters. *BMC Biol.* 4, 31.
- Yoon, H.S., Hackett, J.D., Pinto, G., and Bhattacharya, D. (2002). The single, ancient origin of chromist plastids. *Proc. Natl. Acad. Sci. USA* 99, 15507–15512.
- Bachvaroff, T.R., Sanchez Puerta, M.V., and Delwiche, C.F. (2005). Chlorophyll c-containing plastid relationships based on analyses of a multigene data set with all four chromalveolate lineages. *Mol. Biol. Evol.* 22, 1772–1782.
- Bodyl, A. (2005). Do plastid-related characteristics support the chromalveolate hypothesis? *J. Phycol.* 41, 712.
- Cavalier-Smith, T. (1998). A revised six-kingdom system of life. *Biol. Rev. Camb. Philos. Soc.* 73, 203–266.
- Dougherty, E.C., and Allen, M.B. (1960). Is pigmentation a clue to protistan phylogeny? In *Comparative Biochemistry of Photo-reactive Systems*, M.B. Allen, ed. (New York: Academic Press), pp. 129–144.
- Douglas, S.E., Murphy, C.A., Spencer, D.F., and Gray, M.W. (1991). Cryptomonad algae are evolutionary chimaeras of two phylogenetically distinct unicellular eukaryotes. *Nature* 350, 148–151.
- Bhattacharya, D., and Weber, K. (1997). The actin gene of the glaucocystophyte *Cyanophora paradoxa*: analysis of the coding region and introns, and an actin phylogeny of eukaryotes. *Curr. Genet.* 31, 439–446.
- Cavalier-Smith, T., Allsopp, M.T., and Chao, E.E. (1994). Chimeric conundra: are nucleomorphs and chromists monophyletic or polyphyletic? *Proc. Natl. Acad. Sci. USA* 91, 11368–11372.
- Harper, J.T., and Keeling, P.J. (2003). Nucleus-encoded, plastid-targeted glyceraldehyde-3-phosphate dehydrogenase (GAPDH) indicates a single origin for chromalveolate plastids. *Mol. Biol. Evol.* 20, 1730–1735.
- Keeling, P.J., Deane, J.A., Hink-Schauer, C., Douglas, S.E., Maier, U.G., and McFadden, G.I. (1999). The secondary endosymbiont of the cryptomonad *Guillardia theta* contains alpha-, beta-, and gamma-tubulin genes. *Mol. Biol. Evol.* 16, 1308–1313.
- Liaud, M.F., Brandt, U., Scherzinger, M., and Cerff, R. (1997). Evolutionary origin of cryptomonad microalgae: two novel chloroplast/cytosol-specific GAPDH genes as potential markers of ancestral endosymbiont and host cell components. *J. Mol. Evol.* 44 (Suppl 1), S28–S37.
- Fast, N.M., Kissinger, J.C., Roos, D.S., and Keeling, P.J. (2001). Nuclear-encoded, plastid-targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids. *Mol. Biol. Evol.* 18, 418–426.
- Cavalier-Smith, T. (2003). Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimaeras (meta-algae). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 109–133.
- Palmer, J.D. (2004). The symbiotic birth and spread of plastids: how many times and whodunit? *J. Phycol.* 39, 4–11.
- Parfrey, L.W., Barbero, E., Lasser, E., Dunthorn, M., Bhattacharya, D., Patterson, D.J., and Katz, L.A. (2006). Evaluating

- support for the current classification of eukaryotic diversity. *PLoS Genet.* 2, e220. 10.1371/journal.pgen.0020220.
20. Rodriguez-Ezpeleta, N., Brinkmann, H., Burey, S.C., Roure, B., Burger, G., Löffelhardt, W., Bohnert, H.J., Philippe, H., and Lang, B.F. (2005). Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr. Biol.* 15, 1325–1330.
 21. Jeffroy, O., Brinkmann, H., Delsuc, F., and Philippe, H. (2006). Phylogenomics: the beginning of incongruence? *Trends Genet.* 22, 225–231.
 22. Rokas, A., Williams, B.L., King, N., and Carroll, S.B. (2003). Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* 425, 798–804.
 23. Patron, N.J., Waller, R.F., and Keeling, P.J. (2006). A tertiary plastid uses genes from two endosymbionts. *J. Mol. Biol.* 357, 1373–1382.
 24. Delsuc, F., Brinkmann, H., and Philippe, H. (2005). Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet.* 6, 361–375.
 25. Sanderson, M.J. (2002). Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* 19, 101–109.
 26. Pupko, T., Huchon, D., Cao, Y., Okada, N., and Hasegawa, M. (2002). Combining multiple data sets in a likelihood analysis: which models are the best? *Mol. Biol. Evol.* 19, 2294–2307.
 27. Simpson, A.G., Inagaki, Y., and Roger, A.J. (2006). Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of “primitive” eukaryotes. *Mol. Biol. Evol.* 23, 615–625.
 28. Takishita, K., Inagaki, Y., Tsuchiya, M., Sakaguchi, M., and Maruyama, T. (2005). A close relationship between Cercozoa and Foraminifera supported by phylogenetic analyses based on combined amino acid sequences of three cytoskeletal proteins (actin, α -tubulin, and β -tubulin). *Gene* 362, 153–160.
 29. Cavalier-Smith, T. (2006). The tiny enslaved genome of a rhizarian alga. *Proc. Natl. Acad. Sci. USA* 103, 9379–9380.
 30. Okamoto, N., and Inouye, I. (2005). The katablepharids are a distant sister group of the Cryptophyta: a proposal for Katablepharidophyta divisio nova/ Kathablepharida phylum novum based on SSU rDNA and beta-tubulin phylogeny. *Protist* 156, 163–179.
 31. Kim, E., Simpson, A.G., and Graham, L.E. (2006). Evolutionary relationships of apusomonads inferred from taxon-rich analyses of 6 nuclear encoded genes. *Mol. Biol. Evol.* 23, 2455–2466.
 32. Not, F., Valentin, K., Romari, K., Lovejoy, C., Massana, R., Tobe, K., Vaulot, D., and Medlin, L.K. (2007). Picobiliphytes: a marine picoplanktonic algal group with unknown affinities to other eukaryotes. *Science* 315, 253–255.
 33. Okamoto, N., and Inouye, I. (2006). *Hatena arenicola* gen. et sp. nov., a katablepharid undergoing probable plastid acquisition. *Protist* 157, 401–419.
 34. McFadden, G.I., Gilson, P.R., and Hill, D.R.A. (1994). *Goniomonas*: rRNA sequences indicate that this phagotrophic flagellate is a close relative of the host component of cryptomonads. *Eur. J. Phycol.* 29, 29–32.
 35. Rogers, M.B., Gilson, P.R., Su, V., McFadden, G.I., and Keeling, P.J. (2007). The complete chloroplast genome of the chlorarachniophyte *Bigelowiella natans*: evidence for independent origins of chlorarachniophyte and euglenid secondary endosymbionts. *Mol. Biol. Evol.* 24, 54–62.
 36. Waller, R.F., Patron, N.J., and Keeling, P.J. (2006). Phylogenetic history of plastid-targeted proteins in the peridinin-containing dinoflagellate *Heterocapsa triquetra*. *Int. J. Syst. Evol. Microbiol.* 56, 1439–1447.
 37. Hillis, D.M., Pollock, D.D., McGuire, J.A., and Zwickl, D.J. (2003). Is sparse taxon sampling a problem for phylogenetic inference? *Syst. Biol.* 52, 124–126.
 38. Rokas, A., and Carroll, S.B. (2005). More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. *Mol. Biol. Evol.* 22, 1337–1344.
 39. Philippe, H., Snell, E.A., Baptiste, E., Lopez, P., Holland, P.W., and Casane, D. (2004). Phylogenomics of eukaryotes: impact of missing data on large alignments. *Mol. Biol. Evol.* 21, 1740–1752.

Accession Numbers

The GenBank accession numbers for the EST sequences reported in this paper are EG715451–EG729649.