Symbiotic Origin of a Novel Actin Gene in the Cryptophyte *Pyrenomonas helgolandii*

Thomas B. Stibitz, Patrick J. Keeling,† and Debashish Bhattacharya**

*Department of Biological Sciences, University of Iowa; and †Department of Botany, Canadian Institute for Advanced Research, University of British Columbia, Vancouver, British Columbia, Canada

Cryptophytes are photosynthetic protists that have acquired their plastids through the secondary symbiotic uptake of a red alga. A remarkable feature of cryptophytes is that they maintain a reduced form of the red algal nucleus, the nucleomorph, between the second and third plastid membranes (periplastidial compartment; PC). The nucleomorph is thought to be a transition state in the evolution of secondary plastids, with this genome ultimately being lost when photosynthesis comes under full control of the ''host'' nucleus (e.g., as in heterokonts, haptophytes, and euglenophytes). Genes presently found in the nucleomorph seem to be restricted to those involved in its own maintenance and to that of the plastid; other genes were lost as the endosymbiont was progressively reduced to its present state. Surprisingly, we found that the cryptophyte *Pyrenomonas helgolandii* possesses a novel type of actin gene that originated from the nucleomorph genome of the symbiont. Our results demonstrate for the first time that secondary symbionts can contribute genes to the host lineage which are unrelated to plastid function. These genes are akin to the products of gene duplication or lateral transfer and provide a source of evolutionary novelty that can significantly increase the genetic diversity of the host lineage. We postulate that this may be a common phenomenon in algae containing secondary plastids that has yet to be fully appreciated due to a dearth of evolutionary studies of nuclear genes in these taxa.

Introduction

Many algal plastids originated through secondary symbioses whereby a previously free-living alga (secondary symbiont) takes up permanent and heritable residence inside a nonphotosynthetic host eukaryote (Douglas et al. 1991; Bhattacharya and Medlin 1995, 1998; van de Peer et al. 1996; Delwiche and Palmer 1997; Zhang, Green, and Cavalier-Smith 1999; Oliveira and Bhattacharya 2000). In most instances, secondary plastids are recognizable only by the presence of extra plastid membranes (either three or four, as opposed to the two membranes surrounding primary plastids). However, in two lineages, the chlorarachniophytes and cryptophytes, the secondary symbiont retains a vestigial nucleus, or nucleomorph, which is located between the second and third plastid membranes (periplastidial compartment [PC]; Greenwood 1974; Hibberd and Norris 1984; McFadden, Gilson, and Hill 1994). Molecular characterization of nucleomorph chromosomes suggests that they exist primarily to facilitate photosynthesis and nucleomorph maintenance (Gilson and McFadden 1997; McFadden et al. 1997; Zauner et al. 2000). These genomes would presumably be lost (as they have been in other secondary endosymbionts) if all genes for plastid function that they encode were transferred to the host genome, as the host would then gain complete control of organellar expression (Zauner et al. 2000). Indeed, pressure to reduce the size of nucleomorph genomes is clearly reflected in large-scale gene loss from these chromosomes, limited intergenic regions, the existence of ''mini'' spliceosomal introns (in chlorarachniophytes), and overlapping open reading frames (e.g., Gilson and McFadden 1997; Fraunholz, Moerschel, and Maier 1998; Wastl et al. 1999; Zauner et al. 2000).

An issue of which nothing is presently known is the fate of nucleomorph genes that encode proteins with functions not required in the plastid or the PC. These sequences are presumably rapidly eliminated, but it is also conceivable that some may find their way to the host nuclear genome through gene transfer. This would be analogous to the process of endosymbiotic gene replacement that now appears to be common with prokaryotic endosymbiotic organelles such as the mitochondrion and the plastid (Keeling and Doolittle 1997; Martin and Herrmann 1998). Our study reports for the first time such a finding for the cryptophyte *Pyrenomonas helgolandii.* In this taxon, an actin gene from the red algal secondary symbiont has been transferred to the host nucleus, where it is expressed and apparently has assumed a novel function. This process is essentially a lateral gene transfer from one eukaryote to a distantly related eukaryote and provides a new source of genetic diversity that is potentially very important in eukaryotic evolution.

Materials and Methods

Isolation of Actin cDNAs and Genes

Total RNA was isolated from actively growing unialgal cultures of *P. helgolandii* (SAG B 28 87; Schlösser 1994), *Cryptomonas ovata* (SAG B 979-3), and *Guillardia theta* (CCMP 327; Andersen, Morton, and Sexton 1997) using either the RNeasy mini kit (Qiagen) or the Trizol reagent (BRL). Reverse transcriptase–polymerase chain reaction (RT-PCR) methods were used to isolate nearly complete (lacking approximately 40 nt at the 5′ end) actin cDNAs from these taxa using the Ac1 and Ac3 primers as previously described (Bhattacharya, Stickel, and Sogin 1991, 1993). Direct sequencing of

Abbreviations: ABP, actin-binding protein; PC, periplastidial compartment.

Key words: actin evolution, cryptophytes, symbiosis, gene transfer, nucleomorph, phylogeny.

Address for correspondence and reprints: Debashish Bhattacharya, Department of Biological Sciences, University of Iowa, 239 Biology Building, Iowa City, Iowa 52242.

Mol. Biol. Evol. 17(11):1731–1738. 2000

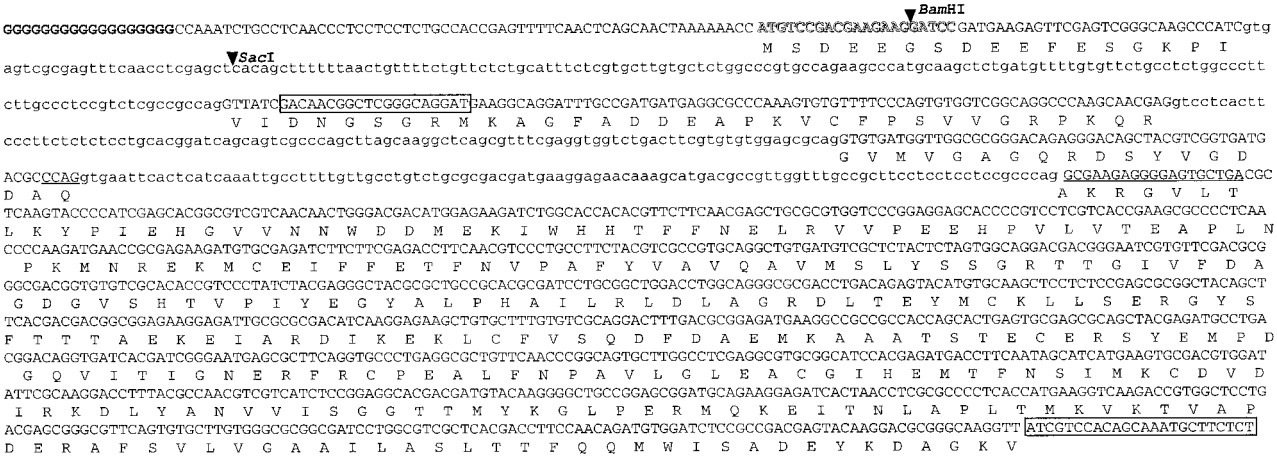

FIG. 1.—The divergent actin gene in *Pyrenomonas helgolandii.* This sequence begins with the anchor primer (shown in boldface) used for 5′ rapid amplification of cDNA ends (RACE). The region used to create the gene-specific 5′ RACE primers is underlined, whereas the Ac1 and Ac3 priming sites used for the RT/PCR experiments are boxed. The 5′-terminal genomic regions were isolated with the gene-specific primers and forward primers that recognize the 5′-terminus of the coding regions (shown in shadowed text). The inferred amino acid sequence of the red alga–like gene is shown, as are the restriction sites for *Bam*HI and *Sac*I. The intron regions are in lowercase.

these PCR products using a battery of conserved actin primers (Bhattacharya, Stickel, and Sogin 1991) with a fluorescent-labeled dideoxy terminator sequencing method on an ABI 373A sequencing device showed single dominant products in electropherograms except for *P. helgolandii.* The actin sequence for this alga contained many ambiguities. The *P. helgolandii* cDNA was, therefore, directionally cloned into the pBluescript vector (Stratagene), and four clones were further characterized. Three clones encoded identical products with high similarity to actins, and one was a divergent actin sequence. To determine the complete cDNA sequences of these two different *P. helgolandii* actin-coding regions, the 5′ rapid amplification of cDNA ends (RACE) procedure was done with total RNA and primers specific to each of the two actin genes according to the manufacturer's instructions (BRL; see fig. 1). Genomic PCRs were also performed with primers that were specific to the 5′-terminus to isolate the two actin genes. To determine a partial sequence of the actin genes in *Goniomonas truncata* (DNA kindly provided by G. I. McFadden, Melbourne, Australia), genomic DNA was used in PCR reactions with the actF2 (5′-GAGAAGATGACNCARATHATGTTYGA-3′) and actR1 (5′-GGCCTGGAARCAYTTNCGRTGNAC-3′) primers (as in Bhattacharya and Weber 1997). The PCR products were cloned, and six different actin sequences were determined. The GenBank accession numbers for actin sequences determined in this study are as follows: *P. helgolandii* highly divergent gene (AF284833), cryptophyte-like gene (AF284834); *C. ovata* (AF284836); *G. theta* (AF284835); *G. truncata* Ac1 (AF284837), Ac2 (AF284839), and Ac3 (AF284838).

Phylogenetic Analyses

Actin genic fragments were assembled using Sequencher, version 3.0 (Gene Codes Corporation), and SeqApp, version 1.9a169 (Gilbert 1992). The resulting sequences were aligned using the same programs, and the alignment was manually optimized. The final alignment was exported from Sequencher as a Nexus file for maximum-parsimony and neighbor-joining analyses using PAUP*, version 4.0b2 (Swofford 1999). First and second codon positions of actin genes (746 nt; 276 parsimony-informative sites) were used for the phylogenetic analyses. The maximum-parsimony method using 10 random additions in a heuristic tree search (tree bisection and reconnection) was used to infer the tree. Bootstrap analyses (2,000 replications) were performed with a weighted (rescaled consistency index over an interval of 1–1000) parsimony analysis using 10 random additions in a heuristic tree search. Bootstrap analyses (2,000 replications) were also performed using matrices calculated with the LogDet transformation and the neighbor-joining method. Quartet puzzling with the maximum-likelihood criterion (1,000 puzzling steps) was also used to estimate support for monophyletic groups. We also used protein distances (PROTDIST, Dayhoff distance matrix, PHYLIP, version 3.572 [Felsenstein 1997]) to build a neighbor-joining tree of actins.

Southern Blot Analyses

Genomic DNA was isolated from a *P. helgolandii* culture using the Plant DNeasy mini kit (Qiagen) and digested to completion with the *Bam*HI and *Sac*I restriction endonucleases under standard conditions (New England Biolabs). The digested DNA was transferred onto a nylon membrane (Schleicher and Schuell) and probed with partial cDNA fragments of the highly divergent and cryptophyte-like genes that had been isolated using gene-specific primers (underlined sequence in fig. 1) and 5′ RACE (see fragments in fig. 6*A*). Another Southern blot was probed with an 1,100-bp cDNA fragment that encoded the nearly complete cryptophyte-like actin gene (isolated using primers Ac1 and Ac3). The Southern hybridizations were performed using a nonradioactive method according to the manufacturer's instructions (Gene Images, Amersham). The hybridiza-

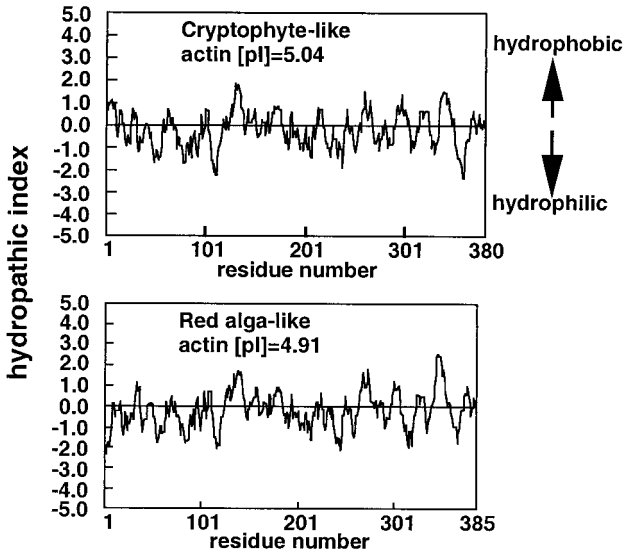

FIG. 2.—Hydrophobicity plots of the *Pyrenomonas helgolandii* actins. Note the highly hydrophilic N-terminus in the divergent protein.

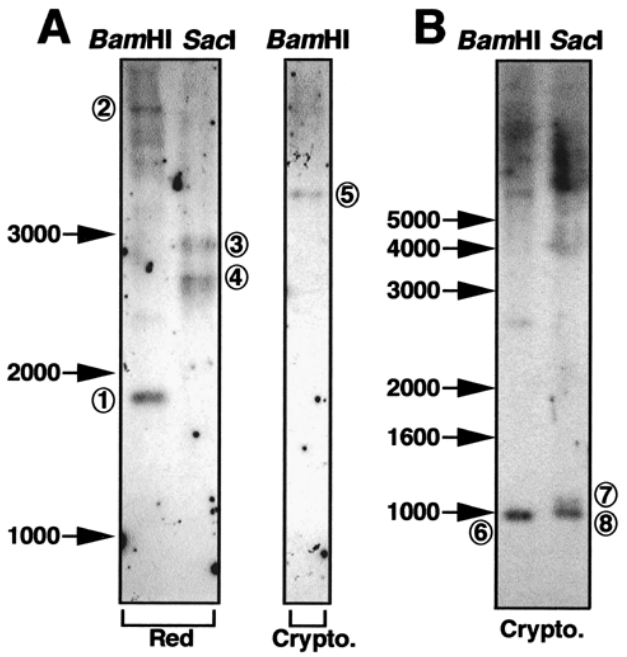

FIG. 3.—*A*, Southern blot analyses of *Pyrenomonas helgolandii* genomic DNA using red alga–like (Red) and cryptophyte-like (Crypto.) actin gene probes. Probe sequences encoded the distinct 5′-termini of these genes. *B*, Southern blot analyses of *P. helgolandii* genomic DNA using a cDNA that encoded most of the cryptophyte-like actin gene (1,100 nt) resulting from RT-PCR with the Ac1 and Ac3 primers. This probe includes 165 nt of sequence in common with the Crypto. probe. Cross-hybridizing bands are identified with numbers (see text for discussion). These data suggest that both types of actin genes are single-copy in *P. helgolandii.*

tion was done in 5 × SSC, 0.1% (w/v) SDS, and 5% (w/v) dextran sulphate at 60°C overnight. Filters were initially washed for 15 min at 60°C in 1 × SSC, and 0.1% SDS and then for 15 min at the same temperature in 0.1 × SSC and 0.1% SDS prior to detection. The same method was used to analyze a Southern blot of a pulsed-field gel containing DNA from the nuclear and nucleomorph compartments of *G. theta* (kindly provided by U.-G. Maier, Marburg, Germany). The divergent actin sequence was used as the probe in this experiment.

Results and Discussion

Characterization of Cryptophyte Actin Sequences

Single actin cDNA sequences were obtained from *G. theta* and *C. ovata,* and three closely related gene sequences were obtained from the plastid-lacking cryptophyte *G. truncata.* All of these sequences bore strong sequence similarity to one another, as expected from such closely related species, and are henceforth referred to as cryptophyte-like actins. However, sequence analysis of four different actin-encoding cDNAs isolated from *P. helgolandii* showed three identical sequences that also shared significant identity to other cryptophyte actins (96%–99% identical), but one actin was highly divergent (only 68%–70% identical to other cryptophyte actins; fig. 1). The provenance and expression of the divergent *P. helgolandii* gene was confirmed by RT-PCR using gene-specific primers from three independent aliquots of the *P. helgolandii* culture obtained from SAG (Schlösser 1994). Hydrophobicity plots (calculated with DNASIS-Mac, version 2.0) showed that the protein inferred from the divergent cDNA had a profile nearly identical to that of conventional *P. helgolandii* actin and was most certainly a functional protein (fig. 2).

Southern blot analysis with total genomic DNA from *P. helgolandii* showed that both highly divergent and cryptophyte-like actin genes exist as single copies (fig. 3*A* and *B*). The probes used for these analyses contained the distinct 5′-termini of the divergent and cryptophyte-like cDNAs that were isolated with 5′ RACE (see below). The appearance of a single, strongly hybridizing band in the *Bam*HI digest (marked as band 1 in fig. 3*A*), for example, is consistent with there being a single divergent actin gene in *P. helgolandii.* A second, more weakly hybridizing fragment of high molecular weight in the *Bam*HI digest (band 2) is likely explained by the binding of the probe to the 77 nucleotides that are upstream of the *Bam*HI site in the probe DNA. The two fragments resolved in the *Sac*I digest (bands 3 and 4) are explained by the presence of a *Sac*I site approximately 140 nt from the 5′-terminus of the divergent gene probe. There was also a single fragment resolved in the *Bam*HI digest when *P. helgolandii* genomic DNA was probed with the cryptophyte-like cDNA (band 5, fig. 3*A*). We did not perform this latter experiment with *Sac*I-digested DNA. To test the results of the Southern blot analyses, another genomic blot of *P. helgolandii* was probed with a cDNA that encoded most of the cryptophyte-like gene (1,100 nt) resulting from RT-PCR with the Ac1 and Ac3 primers (fig. 3*B*). This probe included 165 nt of sequence in common with the 5′-terminal probe. The existence of two *Bam*HI sites in the complete cryptophyte-like cDNA sequence (not shown) therefore explains the three bands that are resolved in this lane. The strongly hybridizing fragment of 990 nt (band 6) is the predicted size of the fragment that encodes the majority of the probe sequence in the absence of introns.

The two fragments resolved in the *Sac*I digest (bands 7 and 8) result from the existence of a single site for this restriction enzyme at position 280 in the complete cDNA sequence.

Because there were no algal contaminants in the *P. helgolandii* culture (confirmed with extensive light microscopic analyses), we hypothesized that the divergent gene may have originated from the secondary symbiont of this taxon. To test this hypothesis, we first reconstructed the phylogeny of the cryptophyte actin genes. Phylogenetic analyses were done with the actin-encoding DNA and amino acid sequences of *P. helgolandii,* conventional actin sequences that we determined from other cryptomonads (*C. ovata, G. theta,* and the aplastidial *G. truncata*), and existing sequences from a variety of eukaryotes. The results of the phylogenetic analyses of the first and second codon positions of actin genes are summarized in figure 4. Three independent methods (maximum parsimony, LogDet transformation, quartet puzzling) were used to build the trees (and calculate bootstrap values) using two different data sets. In the first data set, all of the taxa shown in figure 4 were used, whereas in the second, the bootstrap analyses were recalculated after the exclusion of the highly divergent *P. helgolandii, Chondrus crispus,* and *G. truncata* Ac2 and Ac3 actin sequences. All phylogenetic analyses supported the independent origins of the divergent *P. helgolandii* actin and other cryptophyte actin genes, in all cases but LogDet showing a relationship between the divergent *P. helgolandii* actin gene and those of the red algae (e.g., puzzle value = 78%; fig. 4). The parsimony analysis also provided bootstrap support (96%) for a close relationship between the divergent *P. helgolandii* actin and that of *C. crispus.* These results are consistent with the hypothesis that the *P. helgolandii* actin gene originated from the secondary symbiont, and this gene is henceforth referred to as the red alga–like gene of *P. helgolandii.*

To test the proposed symbiotic origin of the divergent actin gene in *P. helgolandii,* we determined the number of amino acid changes that unambiguously supported the monophyly of the red alga–like and *C. crispus* actins. To do this, the tree shown in figure 4 was used with an actin protein alignment to ''map'' amino acid changes on the topology (using MacClade, version 3.07; Maddison and Maddison 1997). This analysis showed 13 unambiguous changes on the branch that supported a sister group relationship between the red alga–like and *C. crispus* actins. In comparison, there are 4 and 13 changes that support the monophyly of the fungi and the highly derived haptophyte algae, respectively. This result supports a red algal origin of this *P. helgolandii* actin gene. In support of this view, phylogenetic analysis of the more conserved actin protein sequences using a distance method and the neighbor-joining method also showed a sister group relationship between the red alga–like and *C. crispus* sequences (results not shown). The position of these taxa within the crown radiation varied, however, with the use of different outgroups. The destabilizing effect (see Felsenstein 1978; Hendy and Penny 1989; Stiller, Duffield, and Hall 1998) of the highly divergent actin sequences is particularly apparent when they are excluded from the bootstrap analyses in the reduced data set. In these consensus trees, there is a significant increase in bootstrap support for a number of groups, including the cryptomonads, the red algae, and the green algae/land plants (see fig. 4).

Evidence that the *P. helgolandii* Red Alga–like Actin Is Encoded in the Host Genome and Functions in the Host Cytoplasm

The presence of a red alga–like actin gene in *P. helgolandii* could be explained in three ways. First, the sequence is encoded on the *P. helgolandii* nucleomorph genome, and its product is a cytoskeletal element in the PC. Recently, α-, β-, and γ-tubulin genes have been found in the nucleomorph genome of the cryptomonad *G. theta* (Keeling et al. 1999). Therefore, even though electron microscopic studies fail to detect any cytoskeletal elements such as actin, tubulin, or intermediate filaments in the PC of cryptomonads or chlorarachniophytes (Gillott and Gibbs 1980; McKerracher and Gibbs 1982; Morrall and Greenwood 1982; Ludwig and Gibbs 1989), at least tubulin-based cytoskeletal structures must still exist in the endosymbiont (Keeling et al. 1999). Second, the gene has been transferred from the endosymbiont genome to the host nuclear genome of *P. helgolandii* and is posttranslationally imported back to the endosymbiont cytoplasm much as plastid proteins are targeted. Third, the gene has been transferred from the endosymbiont genome to the host nuclear genome and now functions as a novel actin in the host cytoplasm.

The first alternative was excluded by determining whether the red alga–like gene was encoded in the nuclear or the nucleomorph genome. A heterologous Southern blot analysis was done with total DNA from the well-studied cryptomonad *G. theta.* The DNA in the different genetic compartments of *G. theta* was separated by pulsed-field gel electrophoresis (Hofmann et al. 1994), and the resulting blot was probed with the *P. helgolandii* red alga–like gene. This analysis showed a strong hybridization of the red alga–like gene probe with the nuclear genome, but no hybridization with the three nucleomorph chromosomes (fig. 5). This suggests that no actin-like sequence is encoded by the *G. theta* nucleomorph genome, and it provides circumstantial evidence that the red alga–like gene is more likely encoded in the *P. helgolandii* host nucleus. However, RT-PCR analyses with *G. theta* total RNA and a 5′ primer specific to the 5′-terminus of the *P. helgolandii* red alga–like gene failed to amplify an ortholog of the red alga–like actin gene in *G. theta.* If the red alga–like actin gene in *P. helgolandii* is still being encoded in the nucleomorph genome, then the hybridization results from *G. theta* must be due to either differential transfer of nucleomorph genes in *P. helgolandii* and *G. theta,* the high divergence of the 5′-terminus of this gene, or the loss of this gene in *G. theta.* Pulsed-field gel analysis of total genomic DNA from *P. helgolandii* is needed to conclusively distinguish between these hypotheses.

Further evidence for a host nuclear localization of the *P. helgolandii* red alga-like actin came from char-

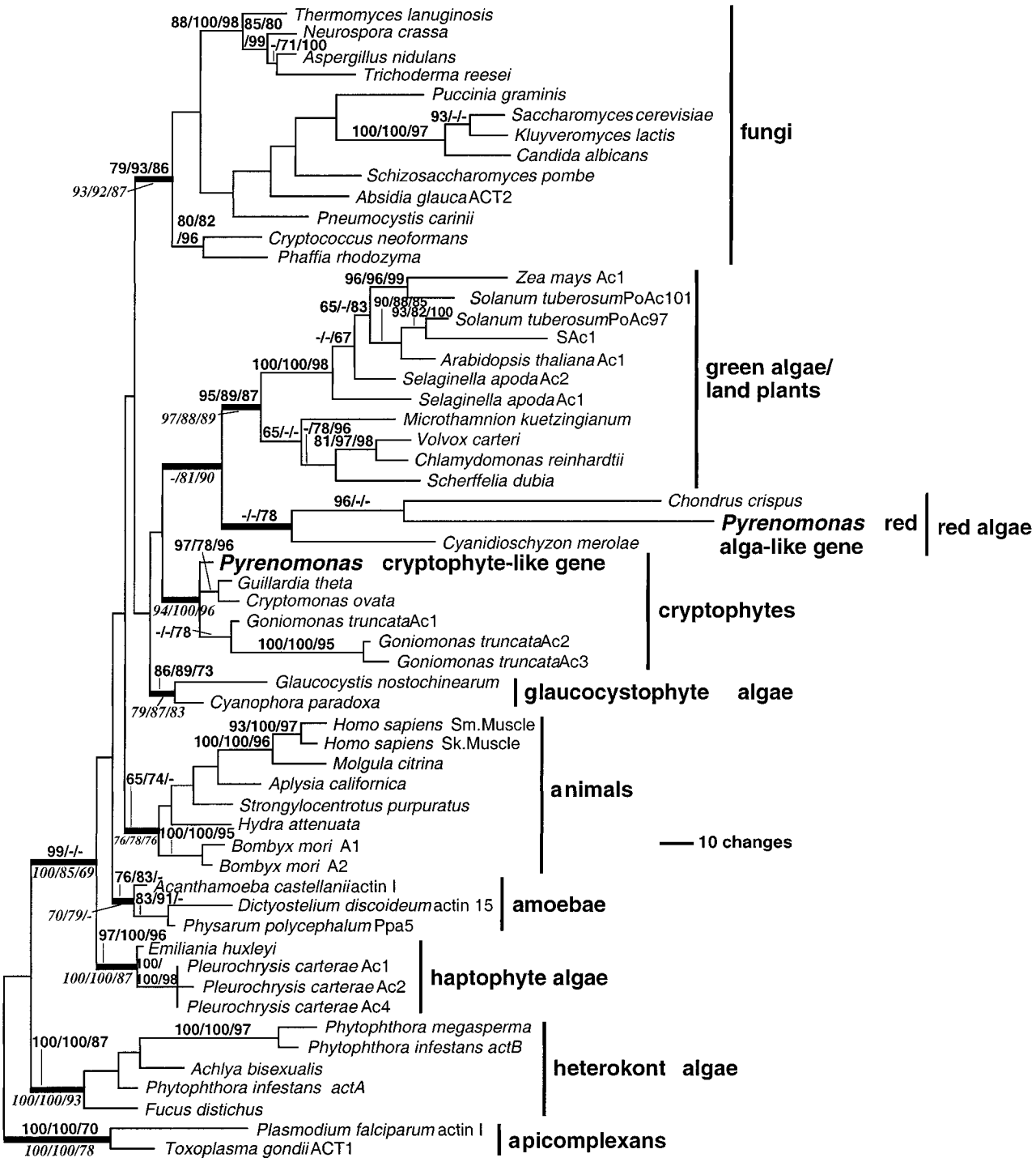

FIG. 4.—Phylogenetic analyses of the first and second codon positions (746 nt) of actin genes. This tree was inferred with a maximum-parsimony method and is the strict consensus of nine equally parsimonious phylogenies of 1,680 steps (consistency index = 0.37, rescaled consistency index = 0.22) and has been outgroup-rooted with the apicomplexan actin sequences. The alternative trees differed only with respect to local shuffling of the *Pleurochrysis carterae* and *Pyrenomonas helgolandii* cryptophyte-like actin genes. The bootstrap values (2,000 replications) shown above the nodes on the left result from a weighted (rescaled consistency index over an interval of 1–1000) parsimony analysis using 10 random additions in a heuristic tree search. The values shown in the middle result from a bootstrap analysis (2,000 replications) using LogDet matrices and trees built with neighbor-joining. The rightmost values are puzzle values (1,000 steps) calculated with the quartet puzzling method. The bootstrap and puzzle values below the branches (in italics) were inferred as described above after exclusion of the *P. helgolandii* red alga–like, *Chondrus crispus*, and *Goniomonas truncata* Ac2 and Ac3 actin sequences. These values are only shown for branches (relatively thick lines) that unite the different eukaryotic lineages. Only bootstrap/puzzle values ≥60% are shown.

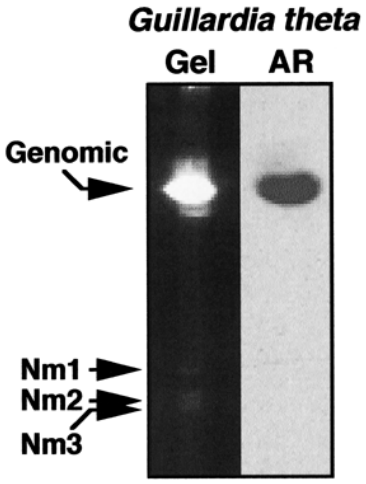

FIG. 5.—Southern blot analysis of DNA from the nuclear and nucleomorph compartments of *Guillardia theta* using a red alga–like actin gene probe. This pattern of hybridization shows that the nucleomorph chromosomes in *G. theta* do not encode an actin gene.

acteristics of the genomic fragment of this gene. We amplified partial 5′-terminal genomic fragments using gene-specific primers (see fig. 1) for the red alga–like and cryptophyte-like actin genes and the remaining coding region of the cryptophyte-like gene with the Ac1 and Ac3 primers to test for the presence/absence of introns. All cryptophyte-like genomic PCR fragments were of the size predicted from the cDNA sequence, showing that they do not encode spliceosomal introns throughout their entire coding region. The red alga–like gene, on the other hand, contained three introns in the 5′-terminus. These introns were at codon positions 11-0 (i.e., between the 10th and 11th amino acid codons), 44-1, and 62-0 (relative to the skeletal muscle α-actin of vertebrates; Weber and Kabsch 1994; Bhattacharya and Weber 1997) and had the typical GT-AG intron splice junctions (Rymond et al. 1990; Lopez and Séraphin 1999). Interestingly, the 44-1 actin intron is shared with *Cyanophora paradoxa* and is one nucleotide removed from the only intron (44-0) found in the single-copy actin gene of the red alga *C. crispus* (Bouget et al. 1995). These introns may be homologous and could have originated through intron ''sliding'' (reviewed in Stoltzfus et al. 1997). The 11-0 intron in the red alga–like actin gene is also found in the fungus *Phaffia rhodozyma*, whereas the 62-0 intron is novel (based on the intron catalog in Bhattacharya and Weber 1997). The presence of introns in the red alga–like gene is in and of itself evidence for a host nuclear location for this gene, because cryptophyte nucleomorph genes have not yet been found to contain spliceosomal introns (Zauner et al. 2000), so the probability of finding three in this one gene is remote. Moreover, nucleomorph genes are also known to be exceedingly AT-rich, but the *P. helgolandii* red alga–like actin is actually slightly GC-rich (40% AT) and very similar to the *P. helgolandii* cryptophyte-like actin (41% AT). Taken together, there is no evidence that this gene is encoded by the nucleomorph genome, and there is considerable, albeit circumstantial, evidence that it is encoded by the host nuclear genome.

To address the possibility that the red alga–like actin is posttranslationally targeted to the endosymbiont, we used 5′ RACE procedures to complete the 5′-terminal sequences of the red alga–like and cryptophyte-like actin cDNAs. Proteins targeted to the secondary endosymbiotic plastid of cryptophytes are first directed to the endomembrane system using an amino-terminal signal peptide (McFadden 1999; Wastl and Maier 2000). It is likely that any protein targeted to the endosymbiont cytosol will initially use the same protein trafficking pathway, so if this actin is targeted to the endosymbiont, it is predicted to encode a signal peptide. The prediction of an ''extra'' sequence at the 5′-terminus of the red alga–like gene was supported with the finding of a 5′ RACE product that was of a relatively larger size than the cryptophyte-like cDNA (fig. 6*A*). However, sequence analyses of these cloned PCR products provided an unexpected result. The extra sequence at the 5′-terminus of the red alga–like gene did not share any properties with cryptophyte signal sequences (e.g., Durnford et al. 1999; McFadden 1999; Wastl et al. 1999, 2000; Wastl and Maier 2000) but, rather, was a highly hydrophilic, novel domain not yet described for actins or any other protein (fig. 6*B*). Although the N-terminus is the most variable region of actins, this domain never shows an additional sequence, as is apparent in the red alga–like actin (e.g., see alignment in Sheterline, Clayton, and Sparrow 1995). This is significant because the variable N-terminus is thought to facilitate the binding of actin to a variety of different actin-binding proteins (ABPs; Her-

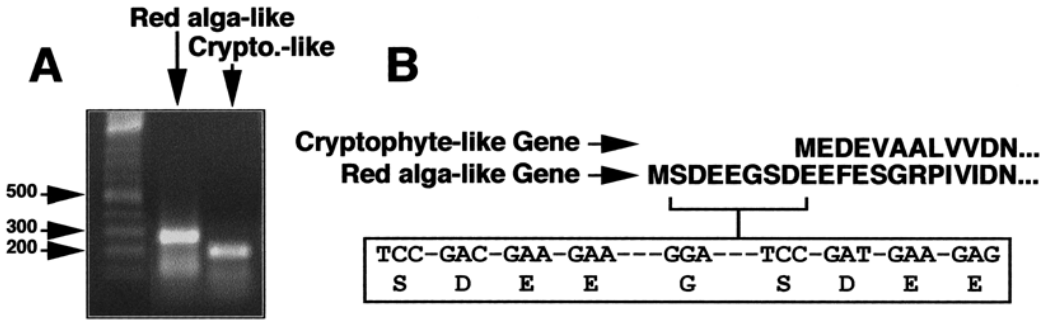

FIG. 6.—*A,* Results of 5′ rapid amplification of cDNA ends (RACE) procedures with the red alga–like and cryptophyte-like actin cDNAs. The anchor primer and the gene-specific reverse primers were used to isolate these sequences. *B,* Comparison of the N-termini of the red alga–like and cryptophyte-like actins.

man 1993; Bhattacharya and Ehlting 1995; Sheterline, Clayton, and Sparrow 1995). The different ABPs confer the diversity of functions ascribed to this otherwise highly conserved protein. With this in mind, we postulate that the red alga–like actin has evolved a new actin-related function in the host cell through the evolution of a novel, divergent N-terminal domain. Analysis of the amino acid sequence in the domain suggests that it originated through the duplication of an SDEE amino acid motif (fig. 6*B*) separated by a glycine (G) residue. That the serine (S) residues are encoded by the same codon (TCC; out of six possible triplets) in both repeats is consistent with a recent origin of this sequence through an internal duplication. The preponderance of charged amino acids such as aspartic acid (D) and glutamic acid (E) suggests that the novel, likely α-helical, domain is completely exposed in the red alga–like actin tertiary structure and interacts with another protein (Branden and Tooze 1991). Although we propose a host function for the red alga–like actin, it is also possible but unlikely that the N-terminal extension may facilitate its interaction with an as yet undescribed protein to allow transfer into the PC. Since no endosymbiont-cytosol–targeted proteins have been characterized yet, nothing is presently known about the transfer of proteins through the second membrane into the PC in four-membraned plastids (McFadden 1999; Wastl and Maier 2000).

Origin of the Red Alga–like Actin Gene: A New Source of Molecular Diversity

Taken together, our data provide strong evidence for the existence of a divergent, novel form of actin in *P. helgolandii* that originated from the nucleus (i.e., nucleomorph) of its red algal secondary symbiont. This gene was apparently transferred to the host genome, where it integrated, acquired regulatory sequences to drive its expression, and assumed some role in its new environment. While such a series of events may be complex, it should not be so surprising, because the host nucleus is known to contain a considerable number of genes encoding plastid-localized proteins, and these genes have all been transferred in much the same way (Gilson and McFadden 1997; McFadden 1999; Blanchard and Lynch 2000). Moreover, it is well established that the eukaryotic hosts of prokaryotic endosymbiotic organelles (mitochondria and plastids) have also acquired genes from these endosymbionts—a specialized type of lateral transfer called ''endosymbiotic gene replacement'' (e.g., Keeling and Doolittle 1997; Martin and Herrmann 1998). The case of the red algal actin in *P. helgolandii* differs primarily in that a eukaryotic gene has been transferred to another, albeit distantly related, eukaryotic genome.

This actin lateral transfer has one other important characteristic of general significance. Unlike many endosymbiosis-associated gene transfers, in this case the red alga–like actin did not simply replace the cryptophyte actin, but seems to have been integrated into the cryptophyte cytosol alongside the original protein. The relatively divergent sequence of the red alga–like gene and its evolution of a novel 5′-terminus are consistent with this scenario. Our results show, therefore, that a host can reap the benefits of duplicate coding regions harbored in the symbiont genome in an entirely novel way—as a source of instant sequence diversity that has evolved under subtly, or even substantially, different selective pressures. Transferred genes with nonplastid functions have likely gone unnoticed until now because of the limited number of host nuclear genes that have been characterized from these alga. The red alga–like actin in *P. helgolandii* may be the tip of an iceberg of endosymbiont-derived coding regions that turn up in the host genome of cryptophytes and other algae with secondary endosymbionts.

Acknowledgments

D.B. acknowledges generous support from the College of Liberal Arts (University of Iowa) and a grant from the Carver Foundation. P.J.K. is a Scholar of the Canadian Institute for Advanced Research and is supported by a grant from the Natural Sciences and Engineering Research Council of Canada (227301-00). We thank U. Maier and C. Hofmann for providing the Southern blot of the pulsed-field gel analysis of *G. theta*, G. I. McFadden for providing *G. truncata* DNA, and J. D. Palmer, in whose laboratory some of this work was done.

LITERATURE CITED

ANDERSEN, R. A., S. L. MORTON, and J. P. SEXTON. 1997. Provasoli-Guillard National Center for Culture of Marine Phytoplankton: 1997 list of strains. J. Phycol. **33**(Suppl.): 1–75.

BHATTACHARYA, D., and J. EHLTING. 1995. Actin coding regions: gene family evolution and use as a phylogenetic marker. Arch. Protistenkd. **145**:155–164.

BHATTACHARYA, D., and L. MEDLIN. 1995. The phylogeny of plastids: a review based on comparisons of small subunit ribosomal RNA coding regions. J. Phycol. **31**:489–498.

———. 1998. Algal phylogeny and the origin of land plants. Plant Physiol. **116**:9–15.

BHATTACHARYA, D., S. K. STICKEL, and M. L. SOGIN. 1991. Molecular phylogenetic analysis of actin genic regions from *Achlya bisexualis* (Oomycota) and *Costaria costata* (Chromophyta). J. Mol. Evol. **33**:525–536.

———. 1993. Isolation and molecular phylogenetic analysis of actin coding regions from the prymnesiophyte alga, *Emiliania huxleyi*, using reverse transcriptase and PCR methods. Mol. Biol. Evol. **10**:689–703.

BHATTACHARYA, D., and K. WEBER. 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.

BLANCHARD, J. L., and M. LYNCH. 2000. Organellar genes: why do they end up in the nucleus? Trends Genet.**16**:315–320.

BOUGET, F. Y., C. KERBOURC'H, M. F. LIAUD, S. LOISEAUX DE GOER, R. S. QUATRANO, R. CERFF, and B. KLOAREG. 1995. Structural features and phylogeny of the actin gene of *Chondrus crispus* (Gigartinales, Rhodophyta). Curr. Genet. **28**:164–172.

BRANDEN, C., and J. TOOZE. 1991. Introduction to protein structure. Garland, New York.

DELWICHE, C. F., and J. D. PALMER. 1997. The origin of plastids and their spread via secondary symbiosis. Pp. 53–86 *in*

D. BHATTACHARYA, ed. Origins of algae and their plastids, Springer, Vienna.

DOUGLAS, S. E., C. A. MURPHY, D. F. SPENCER, and M. W. GRAY. 1991. Molecular evidence that cryptomonad algae are chimaeras of two phylogenetically distinct unicellular eukaryotes. Nature **350**:148–151.

DURNFORD, D. G., J. A. DEANE, S. TAN, G. I. MCFADDEN, E. GANTT, and B. R. GREEN. 1999. A phylogenetic assessment of the eukaryotic light-harvesting antenna proteins, with implications for plastid evolution. J. Mol. Evol. **48**:59–68.

FELSENSTEIN, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. **25**: 401–410.

———. 1997. PHYLIP manual. Version 3.572 P-PC. Department of Genetics, University of Washington, Seattle.

FRAUNHOLZ, M. J., E. MOERSCHEL, and U.-G. MAIER. 1998. The chloroplast division protein FtsZ is encoded by a nucleomorph gene in cryptomonads. Mol. Gen. Genet. **260**: 207–211.

GILBERT, D. 1992. SeqApp, A biosequence editor and analysis application, V1.9. Indiana University, Bloomington, Ind.

GILLOTT, M. A., and S. P. GIBBS. 1980. The cryptomonad nucleomorph: its ultrastructure and evolutionary significance. J. Phycol. **16**:558–568.

GILSON, P. R., and G. I. MCFADDEN. 1997. Good things in small packages: the tiny genomes of chlorarachniophyte endosymbionts. Bioessays **19**:167–173.

GREENWOOD, A. D. 1974. The Cryptophyta in relation to phylogeny and photosynthesis. Pp. 566–567 *in* J. V. SANDERS and D. J. GOODCHILD, eds. Electron microscopy 1974. Australian Academy of Science, Canberra, Australia.

HENDY, M. D., and D. PENNY. 1989. A framework for the quantitative study of evolutionary trees. Syst. Zool. **38**:297–309.

HERMAN, I. M. 1993. Actin isoforms. Curr. Opin. Cell Biol. **5**: 48–55.

HIBBERD, D. J., and R. E. NORRIS. 1984. Cytology and ultrastructure of *Chlorarachnion reptans* (Chlorarachniophyta divisio nova, Chlorarachniophyceae classis nova). J. Phycol. **20**:310–330.

HOFMANN, C. J., S. A. RENSING, M. M. HAUBER, W. F. MARTIN, S. B. MULLER, J. COUCH, G. I. MCFADDEN, G. L. IGLOI, and U.-G. MAIER. 1994. The smallest known eukaryotic genomes encode a protein gene: towards an understanding of nucleomorph functions. Mol. Gen. Genet. **243**:600–604.

KEELING, P. J., J. A. DEANE, C. HINK-SCHAUER, S. E. DOUGLAS, U.-G. MAIER, and G. I. MCFADDEN. 1999. The secondary endosymbiont of the cryptomonad *Guillardia theta* contains alpha-, beta-, and gamma-tubulin genes. Mol. Biol. Evol. **16**:1308–1313.

KEELING, P. J., and W. F. DOOLITTLE. 1997. Evidence that eukaryotic triosephosphate isomerase is of alpha-proteobacterial origin. Proc. Natl. Acad. Sci. USA **94**:1270–1275.

LOPEZ, P. J., and B. SÉRAPHIN. 1999. Genomic-scale quantitative analysis of yeast pre-mRNA splicing: implications for splice-site recognition. RNA **5**:1135–1137.

LUDWIG, M., and S. GIBBS. 1989. Evidence that nucleomorphs of *Chlorarachnion reptans* (Chlorarachniophyceae) are vestigial nuclei: morphology, division and DNA-DAPI fluorescence. J. Phycol. **25**:385–394.

MCFADDEN, G. I. 1999. Plastids and protein targeting. J. Eukaryot. Mirobiol. **46**:339–346.

MCFADDEN, G. I., P. R. GILSON, S. E. DOUGLAS, T. CAVALIER-SMITH, C. J. HOFMANN, and U.-G. MAIER. 1997. Bonsai genomics: sequencing the smallest eukaryotic genomes. Trends Genet. **13**:46–49.

MCFADDEN, G. I., P. R. GILSON, and D. R. A. HILL. 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.

MCKERRACHER, L., and S. P. GIBBS. 1982. Cell and nucleomorph division in the alga *Cryptomonas*. Can. J. Bot. **60**: 2440–2452.

MADDISON, W. P., and D. R. MADDISON. 1997. MacClade. Version 3.07. Sinauer, Sunderland, Mass.

MARTIN, W., and R. G. HERRMANN. 1998. Gene transfer from organelles to the nucleus: how much, what happens, and why? Plant Physiol. **118**:9–17.

MORRALL, S., and A. D. GREENWOOD. 1982. Ultrastructure of nucleomorph division in species of Cryptophyceae and its evolutionary implications. J. Cell Sci. **54**:311–328.

OLIVEIRA, M., and D. BHATTACHARYA. 2000. Phylogeny of the Bangiophycidae (Rhodophyta) and the secondary endosymbiotic origin of algal plastids. Am. J. Bot. **87**:482–492.

RYMOND, B. C., C. PIKIELNY, B. SERAPHIN, P. LEGRAIN, and M. ROSBASH. 1990. Measurement and analysis of yeast premRNA sequence contribution to splicing efficiency. Methods Enzymol. **181**:122–147.

SCHLÖSSER, U. G. 1994. SAG—Sammlung von Algenkulturen at the University of Göttingen. Bot. Acta **107**:111–186.

SHETERLINE, P., J. CLAYTON, and J. C. SPARROW. 1995. Actin. Protein Profile **2**:27–28.

STILLER, J. W., E. C. S. DUFFIELD, and B. D. HALL. 1998. Amitochondriate amoebae and the evolution of DNA-dependent RNA polymerase II. Proc. Natl. Acad. Sci. USA **95**:11769–11774.

STOLTZFUS, A., J. M. LOGSDON JR., J. D. PALMER, and W. F. DOOLITTLE. 1997. Intron ''sliding'' and the diversity of intron positions. Proc. Natl. Acad. Sci. USA **94**:10739–10744.

SWOFFORD, D. L. 1999. PAUP*: phylogenetic analysis using parsimony. Version 4.0b2. Smithsonian Institution, Washington, D.C.

VAN DE PEER, Y., S. A. RENSING, U.-G. MAIER, and R. DE WACHTER. 1996. Substitution rate calibration of small subunit rRNA identifies chlorarachniophyte endosymbionts as remnants of green algae. Proc. Natl. Acad. Sci. USA **93**: 7732–7736.

WASTL, J., E. C. DUIN, L. IUZZOLINO, W. DORNER, T. LINK, S. HOFFMANN, H. STICHT, H. DAU, K. LINGELBACH, and U.-G. MAIER. 2000. Eukaryotically encoded and chloroplast located rubredoxin is associated with photosystem II. J. Biol. Chem. (in press).

WASTL, J., M. FRAUNHOLZ, S. ZAUNER, S. DOUGLAS, and U.-G. MAIER. 1999. Ancient gene duplication and differential gene flow in plastid lineages: the GroEL/Cpn60 example. J. Mol. Evol. **48**:112–117.

WASTL, J., and U.-G. MAIER. 2000. Transport of proteins into cryptomonad complex plastids. J. Biol. Chem. **275**:23194–23198.

WEBER, K., and E. KABSCH. 1994. Intron positions in actin genes seem unrelated to the secondary structure of the protein. EMBO J. **13**:1280–1286.

ZAUNER, S., M. FRAUNHOLZ, J. WASTL, S. PENNY, M. BEATON, T. CAVALIER-SMITH, U. G. MAIER, and S. DOUGLAS. 2000. Chloroplast protein and centrosomal genes, a tRNA intron, and odd telomeres in an unusually compact eukaryotic genome, the cryptomonad nucleomorph. Proc. Natl. Acad. Sci. USA **97**:200–205.

ZHANG, Z., B. R. GREEN, and T. CAVALIER-SMITH. 1999. Single gene circles in dinoflagellate chloroplast genomes. Nature **400**:155–159.

WILLIAM MARTIN, reviewing editor

Accepted August 9, 2000