J. Eukaryot. Microbiol., 52(4), 2005 pp. 343–348

DOI: 10.1111/j.1550-7408.2005.00042x

A Transcriptional Fusion of Genes Encoding Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) and Enolase in Dinoflagellates

KIYOTAKA TAKISHITA,[a] NICOLA J. PATRON,[b] KEN-ICHIRO ISHIDA,[c] TADASHI MARUYAMA[a] and PATRICK J. KEELING[b]

[a]*Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Kanagawa 237-0061, Japan, and*
[b]*Canadian Institute for Advanced Research, Department of Botany, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada, and*
[c]*Division of Natural Science & Technology, Kanazawa University, Kanazawa, Ishikawa 920-1192, Japan*

ABSTRACT. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and enolase are enzymes essential for glycolysis and gluconeogenesis. Dinoflagellates possess several types of both GAPDH and enolase genes. Here, we identify a novel cytosolic GAPDH–enolase fusion protein in several dinoflagellate species. Phylogenetic analyses revealed that the GAPDH moiety of this fusion is weakly related to a cytosolic GAPDH previously reported in dinoflagellates, ciliates, and an apicomplexan. The enolase moiety has phylogenetic affinity with sequences from ciliates and apicomplexans, as expected for dinoflagellate genes. Furthermore, the enolase moiety has two insertions in a highly conserved region of the gene that are shared with ciliate and apicomplexan homologues, as well as with land plants, stramenopiles, haptophytes, and a chlorarachniophyte. Another glycolytic gene fusion in eukaryotes is the mitochondrion-targeted triose-phosphate isomerase (TPI) and GAPDH fusion in stramenopiles (i.e. diatoms and oomycetes). However, unlike the mitochondrial TPI–GAPDH fusion, the GAPDH–enolase fusion protein appears to exist in the same compartment as stand-alone homologues of each protein, and the metabolic reactions they catalyze in glycolysis and gluconeogenesis are not directly sequential. It is possible that the fusion is post-translationally processed to give separate GAPDH and enolase products, or that the fusion protein may function as a single bifunctional polypeptide in glycolysis, gluconeogenesis, or perhaps more likely in some previously unrecognized metabolic capacity.

Key Words. Alveolates, eukaryotes, evolution, gene fusion, insertion, phylogeny.

BOTH glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and enolase (2-phospho-D-glycerate hydrolase) are ubiquitous enzymes highly conserved at the amino acid sequence level. GAPDH catalyzes the reversible interconversion between glyceraldehyde-3-phosphate and 1,3-diphosphoglycerate and can be classified into two main types with respect to its function. One is a glycolytic/gluconeogenetic enzyme specific for NAD^+, and the other is involved in Calvin cycle reactions and can utilize both NAD^+ and $NADP^+$. Enolase is responsible for catalyzing the interconversion of 2-phospho-D-glycerate and phosphoenolpyruvate in glycolysis and gluconeogenesis, and plastid isoforms are also known.

Previous studies have demonstrated that dinoflagellates, a highly diverse group of photosynthetic and non-photosynthetic protists, possess several cytosolic and plastid-targeted GAPDH types, collectively recording to a complex evolutionary history including gene duplication, gene replacement, and lateral gene transfer (Fagan and Hastings 2002; Fagan, Hastings, and Morse 1998; Takishita, Ishida, and Maruyama 2003; Takishita, Ishida, and Maruyama 2004). For example, the molecular phylogeny of GAPDH suggests that a gene for the cytosolic GAPDH of an ancestor of chromalveolates (i.e. apicomplexans, cryptophytes, dinoflagellates, haptophytes, and heterokonts together with their non-photosynthetic relatives) was duplicated, and the product of one copy became targeted to the plastid, replacing the original GAPDH gene derived from the cyanobacterial endosymbiont (Fast et al. 2001; Harper and Keeling 2003). Recently, we have isolated in the dinoflagellate genus *Karenia*, one cytosolic type of GAPDH gene (*GapC4*), which branches within the chromalveolate cytosolic clade, but is phylogenetically distinct from other cytosolic GAPDH types so far discovered in dinoflagellates (Takishita et al. 2004). Remarkably, *GapC4* did not have a putative stop codon position in the expected region of the gene. Instead, the carboxy-terminus of *GapC4* shared a high degree of similarity with enolase, although the sequence was truncated and thus the significance of this similarity was not clear.

Enolase from dinoflagellates has itself been relatively well studied, in part on account of the presence of several highly conserved insertions in enolases from apicomplexans and ciliates. These insertions are shared with homologues from land plants, chlorarachniophytes, and some stramenopiles (Dzierszinski et al. 1999; Harper and Keeling 2004; Keeling and Palmer 2001; Read et al. 1994; Van der Straeten et al. 1991). The evolutionary history of these insertions has been a contentious issue, and led to the characterization of dinoflagellate homologues.

Through a recent expressed sequence tag (EST) survey of the dinoflagellate *Heterocapsa triquetra*, three types of enolase (i.e. enolases 1, 2, and 3) have been discovered (Harper and Keeling 2004). Enolase 3 alone among them contains the insertions found in other alveolates. The presence of insertion-containing genes in all other alveolates examined so far suggests that the dinoflagellate insertion-containing enolase 3 is probably vertically inherited from an ancestor of alveolates, and that the insertion-lacking paralogues (i.e. enolases 1 and 2) originated from lateral gene transfer. This is supported in the case of enolase 2 by its close affinity to diatom enolases (Harper and Keeling 2004).

Here, we have characterized fusion proteins from *Karenia* and *Heterocapsa*, and analyze the origin of both the GAPDH and enolase moieties using phylogeny. Furthermore, the evolution and biochemistry of the fusion in this algal lineage are discussed.

MATERIALS AND METHODS

Culture strains. *Karenia mikimotoi* (Miyake & Kominami ex Oda) Hansen & Moestrup NIES680 and *Karenia brevis* (Davis) Hansen & Moestrup CCMP2229 were obtained from the National Institute for Environmental Studies (NIES) and the Provasoli-Guillard National Center for Culture of Marine Phytoplankton (CCMP), respectively. They were grown according to the protocols of the individual culture collections. *Heterocapsa triquetra* Stein CCMP449 was the basis for an EST project (http://amoebidia.bcm.umontreal.ca/public/pepdb/agrm.php) from which fusion proteins were identified.

Total RNA extraction. Cells were harvested from 100–150 ml of culture of each strain by centrifugation (5,000 *g*, 5 min, 4 °C). Total RNA was extracted using the Absolutely RNA™ RT-PCR Miniprep Kit (Stratagene, La Jolla, CA) after homogenizing the cell pellets with glass beads in lysis buffer provided with this kit.

Reverse transcription-polymerase chain reaction (RT-PCR) amplification. Synthesis of cDNA from total RNA of

Corresponding Author: K. Takishita—Telephone number: +81-46-867-9538; Fax number: +81-46-867-9525; e-mail: takishitak@jamstec.go.jp

K. mikimotoi and *K. brevis* and PCR amplification using the cDNA as a template were performed using SuperScript II RNase H-reverse transcriptase (Invitrogen, Carlsbad, CA) and HotStarTaq DNA polymerase (QIAGEN, Tokyo, Japan) according to the respective manufacturer's instructions. Enolase genes were PCR-amplified using either forward primers 5′-GCTTGCA TCGCTCTCAACGA-3′ (designed to anneal to the 3′ terminal region of the dinoflagellate *GapC4*) (Takishita, Ishida, and Maruyama 2004), 5′-AGCGGCAACCCGACNGTNGARGTNGA-3′ (enolase universal) or 5′-CCGGTCGACCGGNATHTAY-GARGC-3′ (enolase universal) with the reverse primer 5′-GCGCTCGCGRCANG GNGCNCCNGTYTT-3′ (enolase universal). The thermal cycle was completed under the following conditions: 35 cycles of 1 min at 94 °C, 1 min at 50 °C, and 2 min at 72 °C, with a final elongation step of 10 min at 72 °C. The amplified products were checked by 1.0% (w/v) agarose gel electrophoresis. All cDNAs encoding full-length copies of the fusion protein (GapC4-enolase 3) from *H. triquetra* were recovered from an existing EST project, and the clones were completely sequenced.

Cloning and sequencing. The PCR-amplified DNA fragments were cloned into the pCR2.1 vector of the TOPO TA Cloning Kit (Invitrogen) and sequenced with an ABI PRISMTM 3700 DNA Analyzer (PE Biosystems, Foster City, CA) using a BigDye™ Terminator Cycle Sequencing Ready Reaction Kit (PE Biosystems). The sequences were analyzed using GENETYX-MAC software version 8.0 (Software Development, Tokyo, Japan).

Phylogenetic analyses. New sequences obtained in the present study were deposited in GenBank (Accession numbers AB195823-AB195834). By using CLUSTAL W version 1.8 (Thompson, Higgins, and Gibson 1994), the deduced amino acid sequences of GAPDH and enolase moieties of a fusion from *H. triquetra*, *K. mikimotoi*, and *K. brevis* were separately aligned with the corresponding sequences from a variety of eukaryotic groups. In the alignment of enolase, the sequences of *Oryza sativa* 2 and *Plasmodium yoelii* 2 (Harper and Keeling 2004) were not included, because of the strong likelihood that they are artifacts from genome projects and are actually prokaryotic sequences. The alignments were inspected by eye and manually edited. All ambiguous sites of the alignments were removed from the dataset for phylogenetic analyses. Consequently, we generated two alignment datasets: (1) 295 sites of 65 taxa (GAPDH) and (2) 326 sites of 82 taxa (enolase). The alignment data are available on request from the corresponding author. For the phylogenetic analyses of GAPDH and enolase sequences, the same methods were independently applied. Maximum-likelihood (ML) analyses were performed using PhyML (Guindon and Gascuel 2003). PhyML was performed using an input tree generated by BIONJ with the WAG model (Whelan and Goldman 2001) of amino acid substitution incorporating invariable sites and a discrete gamma distribution (eight categories) (WAG+I+Γ). The proportion of invariable sites, a discrete γ distribution, and base frequencies of the model were estimated from the dataset. PhyML bootstrap trees (500 replicates) were constructed using the same parameters as the individual ML trees. Bayesian trees were also reconstructed by using the program MrBayes 3.0 (Ronquist and Huelsenbeck 2003) under WAG+I+Γ. The Bayesian analysis was performed twice for each dataset to check for convergence. One cold and three heated Markov chain Monte Carlo (MCMC) chains with default-chain temperatures were run for 500,000 generations, sampling log-likelihoods (lnLs), and trees at 100-generation intervals (i.e. 5,000 lnLs and trees were saved during MCMC). The likelihood plot for GAPDH and enolase datasets suggested that MCMC reached the stationary phase after the first 40,000 and 50,000 generations, respectively. Thus, the remaining 4,600 and 4,500 trees of GAPDH and enolase were used to obtain clade probabilities and branch-length estimates, respectively.

RESULTS

A transcriptional fusion of GAPDH and enolase. In *K. mikimoti* and *K. brevis*, RT-PCR was performed by using a forward primer specific to the 3′ terminus of the coding region of *Karenia GapC4* and a reverse primer universal to the 3′ terminus of the coding region of enolase. The resulting sequences of the PCR-amplified fragments contained the insertion-containing enolase gene and were very similar to a previously reported gene for enolase 3 from *Heterocapsa* (Harper and Keeling 2004). In both species of *Karenia*, RT-PCR was also performed with a set of primers based on universally conserved regions, and multiple clones of PCR-amplified fragments were sequenced. Consequently, additional insertion-containing and insertion-lacking enolase sequences were found, and both types showed heterogeneity. Because the *Heterocapsa* enolase 3 gene was incomplete at the N-terminus, a full-length cDNA encoding *Heterocapsa* enolase 3 was sequenced. As expected, a GAPDH sequence was shown to be fused with enolase 3 at the 3′ terminus at the transcriptional level. The boundary between GAPDH and enolase moieties of the fusion was variable between the three species, so it is impossible to say unequivocally as to where one protein domain ends and the other begins (Fig. 1). Three amino acid residues at positions 32, 187, and 188 (residues numbered following Clermont et al. 1993)

```
                  H. triquetra (GapC4-enolase3)  ---SWYDNEWGYSNRLVDLACHMAVVDGVVPPPAKIVSIKAREIFDSRGNPTVEVD---
                  K. mikimotoi (GapC4-enolase3)  ---SWYDNEWGYSNRLVDLACHMACVDGIVPPAAKVVSIKAREIFDSRGNPTVEVD---
Dinoflagellates   K. brevis (GapC4-enolase3)     ---SWYDNEWGYSNRLVELASYMAIVDGVVPKPAKVVSIKAREIFDSRGNPTVEVD---
                  L. polyedra (GapC1-p)          ---MWYDNEWGYSCRVVRPDKAHGEG*
                  L. polyedra (GapC2)            ---SWYDNEWGYSNRLVDLAIYMAKKDG*
                  Symbiodinium (GapC3)           ---SWYDNEWGYSNRVVDLLMHMISA*
Apicomplexa       P. falciparum (GapC)           ---SWYDNEWGYSNRVLDLAVHITNN*
                  T. gondii (GapC)               ---SWYDNEWGYSNRLVELAHYMSVQDGA*
Streptophyte      O. sativa (GapC)               ---AWYDNEWGYSNRVIDLIRHMAKTQ*
Amoebae           M. balamuthi (GapC)            ---SWYDNEWGYSNRVVDLLLHSLSLH*
Dinoflagellates   H. triquetra (enolase1)                                          MAISKIHARQVFDSRGNPTVEVE---
                  H. triquetra (enolase2)                                           MIKSLFAREILDSRGNPTVEVD---
Apicomplexa       P. falciparum (enolase)                                        MAHVITRINAREILDSRGNPTVEVD---
                  T. gondii (enolase1)                                            MVAIKDITARQILDSRGNPTVEVD---
Streptophyte      O. sativa (enolase)                                            MAATIVSVKARQIFDSRGNPTVEVD---
Amoebae           M. balamuthi (enolase)                                          MSTIKSVFAREILDSRGNPTVEVD---
```

Fig. **1.** Amino acid sequence of the boundary between GapC4 and enolase 3 moieties of the fusion protein of *Karenia* and *Heterocapsa*, aligned with the sequences of stand-alone glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and enolase from dinoflagellates and other eukaryotes.

in the sequence of the GAPDH moiety of *Heterocapsa* were aspartic acid (D), glycine (G), and proline (P), respectively, conferring a strict specificity for NAD^+, and these characteristics are typically found in GAPDH involved in glycolysis or gluconeogenesis. In the enolase moiety from *Heterocapsa*, one in-frame stop codon was found, so this region of the gene was resequenced from a different strain of the same species (not shown), where it was found to be absent. This suggests that the stop codon is most likely an error, probably introduced during reverse transcription.

Origin of GAPDH and enolase moieties of the fusion protein. Large-scale phylogenies of GAPDH (not shown) were similar to published analyses (e.g. Fagan and Hastings 2002; Fagan, Hastings and Morse 1998; Fast et al. 2001; Harper and Keeling 2003; Takishita et al. 2003, 2004). To focus on the origin of the fusion protein, a phylogenetic tree of GAPDH restricted to the cytosolic type of various eukaryotes and nuclear-encoded and plastid-targeted type of chromalveolates was inferred (Fig. 2). The GAPDH moiety of the three fusion proteins branched together with strong support (82% bootstrap probability (BP) of ML and 100% posterior probability (PP) of Bayesian inference) and this branch was clustered with the dinoflagellate GapC2, cytosolic versions of *Toxoplasma gondii* (apicomplexan) and *Blepharisma intermedium* (ciliate) (62% BP and 100% PP).

A global phylogenetic tree of eukaryotic enolase (Fig. 3) was also similar to those seen in previous studies, in that several major eukaryotic groups were resolved, but overall deep-level relationships between groups were equivocal, and the phylogeny and distribution of insertions were not congruent (Bapteste and Philippe 2002; Hannaert et al. 2000; Harper and Keeling 2004; Keeling and Palmer 2001). The sequences of insertion-containing enolase from *K. mikimotoi* and *K. brevis* robustly clustered with the *Heterocapsa* enolase 3 (100% BP and 100% PP). These in turn grouped with the insertion-containing sequences from other alveolates (i.e. apicomplexans and ciliates), although the monophyly of alveolates was supported neither by BP nor by PP. The sequences of insertion-lacking enolase from *Karenia* species solidly constituted a monophyletic lineage with the *Heterocapsa* enolase 2 (100% BP and 100% PP). Sequences similar to the *Heterocapsa* enolase 1 gene were not found in *Karenia* species, although the number of clones sequenced was too small to draw any conclusions regarding its presence.

DISCUSSION

The complex evolution of dinoflagellate GAPDH and enolase. The GAPDH moiety (GapC4) from *Karenia* and *Heterocapsa* fusion proteins is weakly related to cytosolic GAPDH homologues from other alveolates, including dinoflagellate GapC2. This suggests that the gene evolved within the alveolate lineage and that perhaps one gene duplication took place. However, the lack of resolution within this group makes further conclusions impossible. Similarly, the dinoflagellate enolase 3 weakly branched with other insertion-containing sequences from alveolates, in contrast to previous phylogenetic analyses based on the truncated sequence alignment of enolase (Harper and Keeling 2004). Since all these genes also have the noted insertions, it is likely that the topology in Fig. 3 reflects the origin of the dinoflagellate gene and that the alveolate insertion-containing genes all evolved from a common ancestor. In contrast, the insertion-lacking enolase 2 sequences from *Karenia*, *Heterocapsa*, and *Amphidinium* are all closely related to diatom sequences, further supporting the hypothesis that this type of enolase is relatively ancient within dinoflagellates and was ultimately derived via lateral gene transfer from a diatom or a close relative of diatoms (Harper and Keeling 2004) (the *Amphidinium* enolase 2 was not included in the present phylogenetic tree due to its short sequence). In conclusion, therefore, phylogenetic evidence suggests that both the GAPDH and enolase moieties of the fusion protein were derived from exisiting genes in the dinoflagellate genome, and not from lateral gene transfer or some exogenous source. In the case of enolase, moreover, the fusion protein moiety is the only known ancestral dinoflagellate sequence, the free-standing genes all apparently being derived exogenously.

Origin and function of a novel fusion protein. It has been suggested that gene or transcriptional fusion leading to the formation of multidomain proteins is a major route of protein evolution and provides a mechanism for the physical association of different catalytic domains or catalytic and regulatory structures (Jensen 1996). Fusions frequently involve genes coding for proteins that function in a concerted manner, such as enzymes catalyzing sequential steps within a metabolic pathway (Yanai, Wolf, and Koonin 2002). The selective advantage of domain fusion in these cases probably lies in the increased efficiency of coupling of the corresponding biochemical reactions (Marcotte et al. 1999). For example, in diatoms and oomycetes, the glycolytic enzymes triosephosphate isomerase (TPI) and GAPDH are expressed as a fusion protein, which, at least in diatoms, is imported into mitochondria prior to its assembly into a tetrameric bifunctional enzyme complex (Liaud et al. 2000). Such cases, and the present fusion protein, show that genes of related function do fuse on occasion, but how? The answer to this is not obvious, because if the order of genes in the genome is random, then the likelihood of two functionally related genes falling side by side is very low. Interestingly, this is particularly so in dinoflagellates, where genome size has expanded dramatically so that many species have genome sizes well in excess of humans (e.g. Allen et al. 1975; Rizzo, Jones, and Ray 1982). It may be that fusions are very common, but those involving functionally related proteins are more likely to be useful and therefore retained, the vast majority being lost. It is important to note that this does not necessarily mean proteins we think of as acting in the same pathway, but rather enzymes with similar substrates or products.

However the GAPDH–enolase fusion originated, its presence begs the question of what it is presently doing in dinoflagellates. Its continued maintenance in these two distantly related species suggests it has some function, but what? It is impossible to make firm conclusions with the available data, but there are two differences between this fusion and the TPI–GAPDH fusion that may provide some ideas. First, unlike TPI–GAPDH, the metabolic reactions catalyzed by GAPDH and enolase in glycolysis and gluconeogenesis are not sequential: there are two steps catalyzed by phosphoglycerate kinase and phosphoglyceromutase between them. Indeed, in the hypothetical structure of the complex of glycolytic enzymes, GAPDH and enolase do not even interact with each other (Kurganov, Sugrobova, and Mil'man 1985). The other difference of note is that the TPI–GAPDH fusion protein is targeted to the mitochondrion where it is sequestered from the stand-alone TPI and GAPDH proteins thought to act in glycolysis and gluconeogenesis. However, it is possible that the GAPDH–enolase fusion is cytosolic and is expressed alongside stand-alone proteins, suggesting that it has an alternate function. Intriguingly, it is known that both GAPDH and enolase promote immunoglobulin production, which is irrelevant to their enzymatic function in glycolysis or gluconeogenesis in humans (Sugahara et al. 1991, 1992). GAPDH isoforms in other species are also known to recognize erythrose-4-phosphate rather than glyceraldehydes-3-phosphate (Zhao et al. 1995). With the functional diversity of these enzymes probably underestimated in general, as has been found in other proteins (Schwender et al. 2004), and the presence of multiple types of both GAPDH and enolase in dinoflagellates suggesting further functional diversity specific to

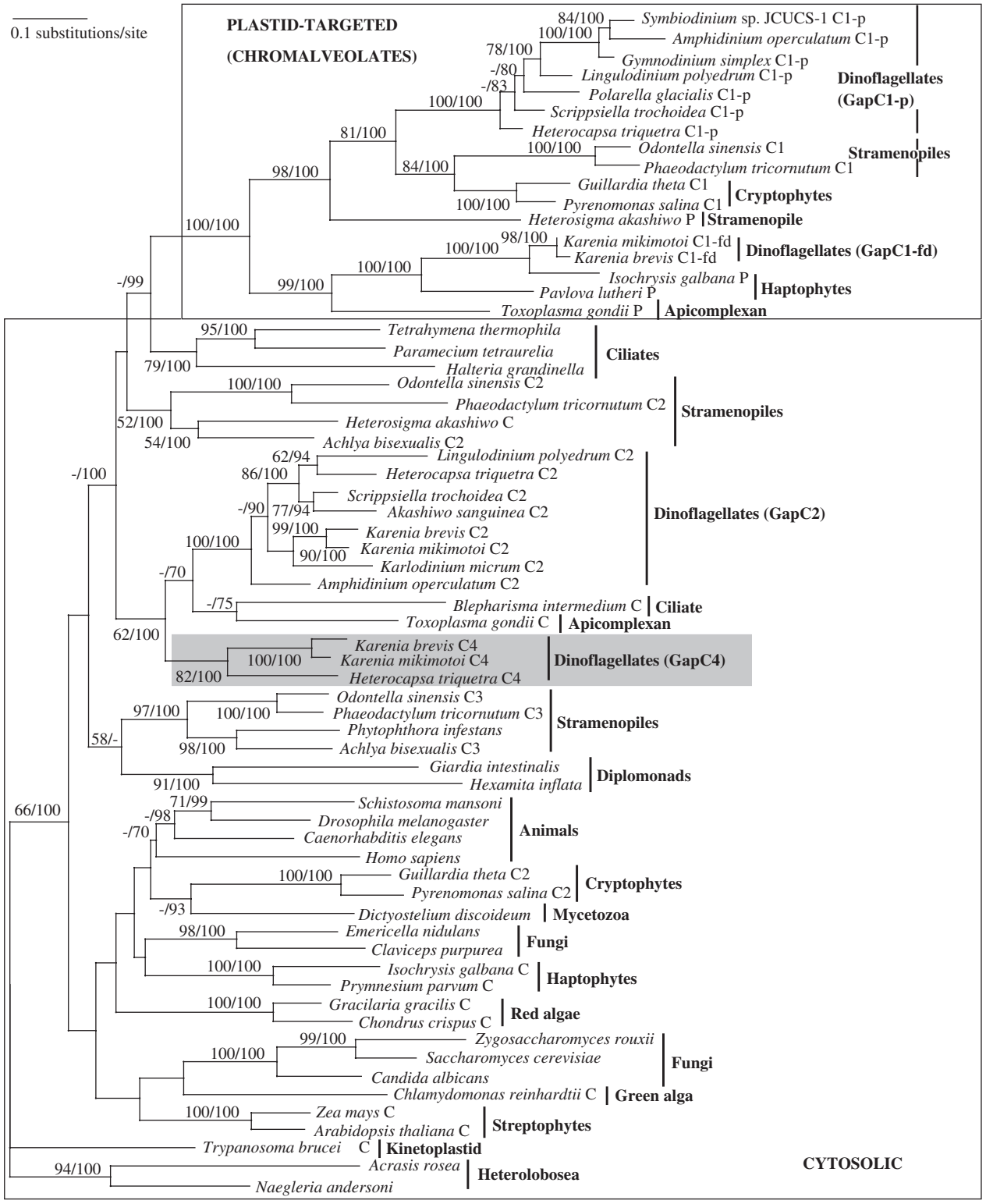

Fig. **2.** Protein maximum likelihood phylogeny (PhyML) of glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Bootstrap probabilities of PhyML and posterior probabilities of MrBayes are shown for nodes with support over 50% (dashes represent support lower than 50%). The dinoflagellate GapC4 sequences fused with the insertion-containing enolase (enolase 3) are highlighted with a shaded box and major groups are labeled to the right.

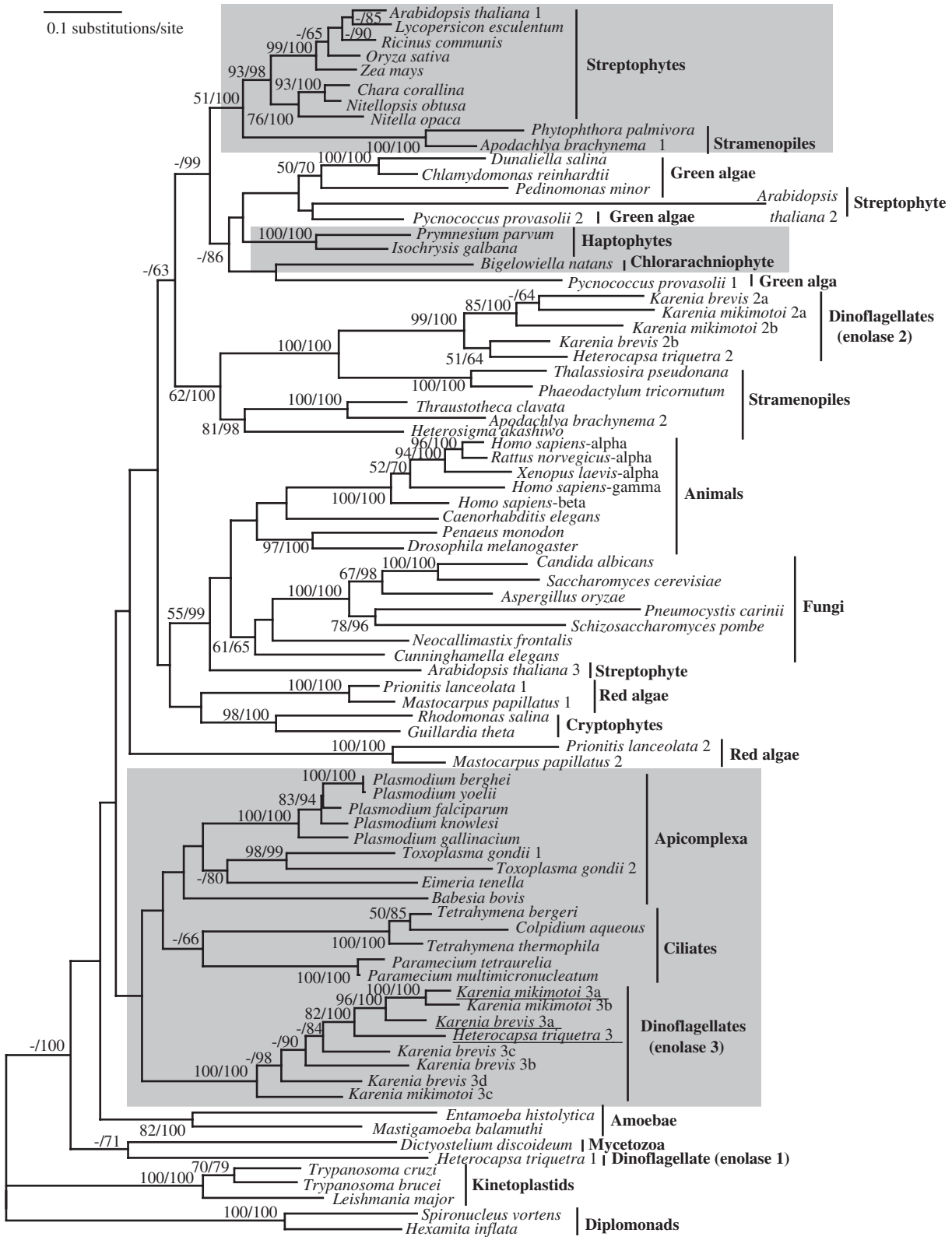

Fig. **3.** Protein maximum likelihood phylogeny (PhyML) of enolase. Bootstrap probabilities of PhyML and posterior probabilities of MrBayes are shown for nodes with support over 50% (dashes represent support lower than 50%). Insertion-containing enolase genes are highlighted with shaded boxes and major eukaryotic groups are labeled to the right. The dinoflagellate enolase 3 sequences confirmed to be fused with GapC4 are underlined.

dinoflagellates, unambiguously determining the role of this unusual fusion will likely require physiological experiments at the protein level.

ACKNOWLEDGMENTS

We thank Dr. J. D. Reimer (JAMSTEC) for critical comments and Mr. M. Kawato (JAMSTEC) for technical assistance. This work was performed as a part of the project ''Study for Understanding of Function and Structure of the Marine Ecosystems in the Earth Systems'' conducted at JAMSTEC. *Heterocapsa* EST sequencing is part of the Protist EST Program supported by Genome Canada and Genome Atlantic. PJK is a Scholar of the CIAR and a New Investigator of the MSFHR and CIHR.

LITERATURE CITED

Allen, J. R., Roberts, M., Loeblich, A. R. III & Klotz, L. C. 1975. Characterization of the DNA from the dinoflagellate *Crypthecodinium cohnii* and implications for nuclear organization. *Cell*, **6**:161–169.

Bapteste, E. & Philippe, H. 2002. The potential value of indels as phylogenetic markers: position of trichomonads as a case study. *Mol. Biol. Evol.*, **19**:972–977.

Clermont, S., Corbier, C., Mely, Y., Gerard, D., Wonacott, A. & Branlant, G. 1993. Determinants of coenzyme specificity in glyceraldehyde-3-phosphate dehydrogenase: role of the acidic residue in the fingerprint region of the nucleotide binding fold. *Biochemistry*, **32**:10178–10184.

Dzierszinski, F., Popescu, O., Toursel, C., Slomianny, C., Yahiaoui, B. & Tomavo, S. 1999. The protozoan parasite *Toxoplasma gondii* expresses two functional plant-like glycolytic enzymes. Implications for evolutionary origin of apicomplexans. *J. Biol. Chem.*, **274**:24888–24895.

Fagan, T. F. & Hastings, W. J. 2002. Phylogenetic analysis indicates multiple origins of chloroplast glyceraldehyde-3-phosphate dehydrogenase genes in dinoflagellates. *Mol. Biol. Evol.*, **19**:1203–1207.

Fagan, T., Hastings, W. J. & Morse, D. 1998. The phylogeny of glyceraldehyde-3-phosphate dehydrogenase indicates lateral gene transfer from cryptomonads to dinoflagellates. *J. Mol. Evol.*, **47**:633–639.

Fast, N. M., Kissinger, J. C., Roos, D. S. & 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.

Guindon, S. & Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.*, **52**: 696–704.

Hannaert, V., Brinkmann, H., Nowitzki, U., Lee, J. A., Albert, M. A., Sensen, C. W., Gaasterland, T., Müller, M., Michels, P. & Martin, W. 2000. Enolase from *Trypanosoma brucei*, from the amitochondriate protist *Mastigamoeba balamuthi*, and from the chloroplast and cytosol of *Euglena gracilis*: pieces in the evolutionary puzzle of the eukaryotic glycolytic pathway. *Mol. Biol. Evol.*, **17**:989–1000.

Harper, J. T. & 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.

Harper, J. T. & Keeling, P. J. 2004. Lateral gene transfer and the complex distribution of insertions in eukaryotic enolase. *Gene*, **340**:227–235.

Jensen, R. A. 1996. Evolution of metabolic pathways in enteric bacteria. *In*: Neidhard, F. C., Curtiss III, R., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., & Umbarger, H. E. (ed.), *Escherichia coli* and *Salmonella typhimurium*, Cellular and Molecular Biology. ASM Press, Washington, DC. p. 2649–2662.

Keeling, P. J. & Palmer, J. D. 2001. Lateral transfer at the gene and subgenic levels in the evolution of eukaryotic enolase. *Proc. Natl. Acad. Sci. USA*, **98**:10745–10750.

Kurganov, B. I., Sugrobova, N. P. & Mil'man, L. S. 1985. Supramolecular organization of glycolytic enzymes. *J. Theor. Biol.*, **116**:509–526.

Liaud, M. F., Lichtle, C., Apt, K., Martin, W. & Cerff, R. 2000. Compartment-specific isoforms of TPI and GAPDH are imported into diatom mitochondria as a fusion protein: evidence in favor of a mitochondrial origin of the eukaryotic glycolytic pathway. *Mol. Biol. Evol.*, **17**:213–223.

Marcotte, E. M., Pellegrini, M., Ng, H. L., Rice, D. W., Yeates, T. O. & Eisenberg, D. 1999. Detecting protein function and protein-protein interactions from genome sequences. *Science*, **285**:751–753.

Read, M., Hicks, K. E., Sims, P. F. & Hyde, J. E. 1994. Molecular characterisation of the enolase gene from the human malaria parasite *Plasmodium falciparum*. Evidence for ancestry within a photosynthetic lineage. *Eur. J. Biochem.*, **220**:513–520.

Rizzo, P. J., Jones, M. & Ray, S. M. 1982. Isolation and properties of isolated nuclei from the Florida red-tide dinoflagellate *Gymnodinium breve* (Davis). *J. Protozool.*, **29**:217–222.

Ronquist, F. & Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**:1572–1574.

Schwender, J., Goffman, F., Ohlrogge, J. B. & Shachar-Hill, Y. 2004. Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. *Nature*, **432**:779–782.

Sugahara, T., Nakajima, H., Shirahata, S. & Murakami, H. 1992. Purification and characterization of immunoglobulin production stimulating factor-II beta derived from Namalwa cells. *Cytotechnology*, **10**: 137–146.

Sugahara, T., Shirahata, S., Akiyoshi, K., Isobe, T., Okuyama, T. & Murakami, H. 1991. Immunoglobulin production stimulating factor-II alpha (IPSF-II alpha) is glyceraldehyde-3-phosphate dehydrogenase like protein. *Cytotechnology*, **6**:115–120.

Takishita, K., Ishida, K. & Maruyama, T. 2003. An enigmatic GAPDH gene in the symbiotic dinoflagellate genus *Symbiodinium* and its related species (the order Suessiales): possible lateral gene transfer between two eukaryotic algae, dinoflagellate and euglenophyte. *Protist*, **154**:443–454.

Takishita, K., Ishida, K. & Maruyama, T. 2004. Phylogeny of nuclear-encoded plastid-targeted GAPDH gene supports separate origins for the peridinin- and the fucoxanthin derivative-containing plastids of dinoflagellates. *Protist*, **155**:447–458.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence-weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**:4673–4680.

Van der Straeten, D., Rodrigues-Pousada, R. A., Goodman, H. M. & Van Montagu, M. 1991. Plant enolase: gene structure, expression, and evolution. *Plant Cell*, **3**:719–735.

Whelan, S. & Goldman, N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.*, **18**:691–699.

Yanai, I., Wolf, Y. I. & Koonin, E. V. 2002. Evolution of gene fusions: horizontal transfer versus independent events. *Genome Biol.*, **3**:1–13.

Zhao, G., Pease, A. J., Bharani, N. & Winkler, M. E. 1995. Biochemical characterization of *gapB*-encoded erythrose 4-phosphate dehydrogenase of *Escherichia coli* K-12 and its possible role in pyridoxal 59-phosphate biosynthesis. *J. Bacteriol.*, **177**:2804–2812.

Received: 12/16/04, 03/31/05; accepted: 04/23/05