

Evidence for the Heterolobosea from Phylogenetic Analysis of Genes Encoding Glyceraldehyde-3-Phosphate Dehydrogenase

ANDREW J. ROGER,^{*1} MICHAEL W. SMITH,** RUSSELL F. DOOLITTLE*** and W. FORD DOOLITTLE*

^{*}Department of Biochemistry, Room 8C1, Sir Charles Tupper Medical Building, Dalhousie University, Halifax, Nova Scotia, B3H 4H7, Canada.

^{**}Science Applications International Corporation Frederick, National Cancer Institute, Research and Development Center, Post Office Box B, Frederick, Maryland 21702-1201, USA, and

^{***}Center for Molecular Genetics, University of California, San Diego, California 92093-0634, USA

ABSTRACT. The phylogenetic relationships between major slime mould groups and the identification of their unicellular relatives has been a subject of controversy for many years. Traditionally, it has been assumed that two slime mould groups, the acrasids and the dictyostelids were related by virtue of their cellular slime mould habit; a view still endorsed by at least one current classification scheme. However, a decade ago, on the basis of detailed ultrastructural resemblances, it was proposed that acrasids of the family Acrasidae were not relatives of other slime moulds but instead related to a group of mostly free-living unicellular amoebae, the Schizopyrenida. The class Heterolobosea was created to contain these organisms and has since figured in many discussions of protist evolution. We sought to test the validity of Heterolobosea by characterizing homologs of the highly conserved glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase (GAPDH) from an acrasid, *Acrasis rosea*; a dictyostelid, *Dictyostelium discoideum*; and the schizopyrenid *Naegleria andersoni*. Phylogenetic analysis of these and other GAPDH sequences, using maximum parsimony, neighbour-joining distance and maximum likelihood methods strongly supports the Heterolobosea hypothesis and discredits the concept of a cellular slime mould grouping. Moreover, all of our analyses place *Dictyostelium discoideum* as a relatively recently originating lineage, most closely related to the Metazoa, similar to other recently published phylogenies of protein-coding genes. However, GAPDH phylogenies do not show robust branching orders for most of the relationships between major groups. We propose that several of the incongruencies observed between GAPDH and other molecular phylogenies are artifacts resulting from substitutional saturation of this enzyme.

Supplementary key words. Amoebae, endosymbiosis, eukaryotic evolution, lateral gene transfer, Mycetozoa, protein evolution, protist systematics, slime mould.

SLIME moulds alternate between free-living unicellular and differentiated multicellular or multinucleate life cycle stages. Systematists have been debating their phylogenetic position and coherence for more than a century. De Bary [14] created the Mycetozoa as a taxon to accommodate the myxomycetes (plasmodial slime moulds) and cellular slime moulds. Since then, this heterogeneous group of amoebae have been variously treated, most often as fungi (see [64] for example). However, slime moulds resemble fungi only superficially [42] and are more correctly considered a protozoan group, a view reflected in most current classification schemes [9, 32, 42, 43, 47].

Among slime moulds, acrasids are little known organisms typified by small fruiting bodies with branched spore chains. They are often found growing on dead attached plant matter, bark, dung and soil [5, 43]. Initially, van Tiegham [62] included them with the dictyostelids in a mycetozoan order, the Acrasieae, because, in the development of a fruiting body, both groups form a cellular pseudoplasmodium by the aggregation of amoebae rather than the multinucleate plasmodium characteristic of the myxomycete slime moulds. However, unlike dictyostelids, acrasid amoebae display lobose pseudopodia, do not stream to aggregation centres in the formation of the pseudoplasmodium, and their fruiting bodies lack a cellulose stalk tube [5]. These distinctive properties led Olive to place acrasids in a separate class, the Acrasea, and to propose that they evolved independently of other eumycetozoans (true slime moulds), from unicellular flagellated soil amoebae with limax (cylindrical) morphology [43].

More recently, on the basis of ultrastructural similarities, Page and Blanton [45] have suggested that one family of acrasids, the Acrasidae, are specifically related to schizopyrenid amoebae, and that together these be included in a new sarcodine class, the Heterolobosea. The Schizopyrenida is an order of amoebae containing two families, the Vahlkampfiidae and the

Gruberellidae. Members of both these families resemble the Acrasidae in possessing limax morphology with eruptive hyaloplasmic lobose pseudopodia and a closed mitosis. In addition, these groups all share three strong ultrastructural characters: discoidal mitochondrial cristae, an envelope of endoplasmic reticulum often found close to or surrounding their mitochondria and the lack of a recognizable Golgi dictyosome (the Golgi stacks) [45].

The potential evolutionary importance of the Heterolobosea is increased by the fact that the available small subunit ribosomal RNA sequence data for vahlkampfiids show them to be one of the earliest branches of mitochondrion-bearing eukaryotes [9, 10, 12, 28]. This deep phylogenetic position has prompted their inclusion in the newly-erected protozoan phylum, the Percolozoa [10] containing organisms which possess discoidal mitochondrial cristae and primitively lack a Golgi dictyosome. Alternative phylogenetic hypotheses suggest that the Heterolobosea may be related to such groups as the Euglenozoa [47], nucleariid amoebae [47] and a jakobid flagellate [40], on the basis that these groups purportedly all possess discoidal mitochondrial cristae.

All of the above proposals take for granted the phylogenetic coherence of the Heterolobosea. However, for groups proposed solely on the basis of ultrastructural characters like the Heterolobosea, independent confirmation is needed and molecular phylogenetics has proved very useful in this regard. Small subunit ribosomal RNA phylogenetics has confirmed the monophyly of many groups previously proposed on ultrastructural grounds such as the alveolates [66], the Euglenozoa [56] and the Vahlkampfiidae [28]. Nevertheless, there are drawbacks to the use of small subunit ribosomal RNA in inferring phylogeny; extreme biases in base compositions in sequences may confound phylogenetic inference by generating artifactual topologies [22–24, 35]. By contrast, highly conserved protein-coding genes appear to be less sensitive to the effects of biased base composition [24, 25] and for this reason their development as phylogenetic markers is crucial to improve our understanding of organismal relationships.

In order to test the phylogenetic coherence of the Heterolobosea and their relationship to dictyostelid slime moulds and

¹ To whom correspondence should be addressed. Telephone: 902-494-3569; Fax: 902-494-1355; Email: aroger@is2.dal.ca

Note: Sequences described in this work have been deposited in the GenBank database with the accession numbers U55243, U55244, and U55245.

other protist groups, we chose to focus on the phylogeny of the gene encoding the glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Previous studies have shown that the GAPDH tree for eukaryotes differs markedly from those derived from small subunit ribosomal RNA comparisons [27, 38, 39, 51, 65, 67]. In addition, the discovery of eukaryotic-like GAPDH genes in both gamma-Proteobacteria [6, 15, 54] and cyanobacteria [39] has lent support to controversial hypotheses of lateral transfer of this enzyme between prokaryotes and eukaryotes [15, 26, 39, 54]. One of these hypotheses suggests that eukaryotes acquired their GAPDH gene from eubacteria [26], perhaps via an endosymbiotic event [27].

We have sequenced GAPDH homologs from a schizopyrenid, *Naegleria andersoni*, an acrasid, *Acrasis rosea*, and the dictyostelid, *Dictyostelium discoideum*. Our analyses yield the first molecular phylogenetic evidence for the Heterolobosea and against a cellular slime mould grouping. The phylogenetic positions of the Heterolobosea and *Dictyostelium* in the GAPDH tree have interesting implications for both organismal and GAPDH gene phylogeny.

MATERIALS AND METHODS

Organism culture and DNA extraction. *Naegleria andersoni* DNA (strain PPFMB-6) extracted from an axenic culture was kindly provided by S. Kilvington (Public Health Laboratory, Bath, England). *Dictyostelium discoideum* (strain AX4) DNA was provided by W. Loomis (University of California at San Diego, San Diego).

A culture of *Acrasis rosea* (strain T-235) was obtained from F. Spiegel (University of Arkansas, Fayetteville). *Acrasis* cells were maintained on 1.5% agar plates supplemented with 0.01% malt extract and 0.01% yeast extract with the yeast *Rhodotorula mucilaginosa* as a food source. Liquid media, consisting of 0.05% yeast extract, 0.25% proteose peptone, 0.1% glucose buffered with 3 mM K_2HPO_4 /16 mM KH_2PO_4 , was inoculated with a small block of solid media containing the yeast and amoebae and flasks were shaken slowly for 4–5 days at room temperature until significant *Acrasis* growth was observed under the light microscope. Cells were harvested by centrifugation, resuspended in lysis solution containing 200 mg/ml proteinase K and 1% SDS and were incubated at 50° C for 1 h. This treatment selectively disrupts the acrasid amoebae leaving the hard chitinous *Rhodotorula* cells intact. The yeast cells were subsequently removed by centrifugation and DNA was purified from the supernatant using the hexadecyltrimethyl-ammonium bromide (CTAB) extraction procedure described in [11].

DNA amplification, cloning and sequencing. Degenerate oligonucleotide primers were designed against the conserved amino acid sequences NGFGRI and WYDNE found at the N- and C-termini of most GAPDH genes. These primers, designated GAPN and GAPC respectively, had the sequences GAGAGCTCRAYGGNTTYGGNMGNAT and GAGAGACTCWYTCRTRTCRTACCA. DNA amplification was performed using 100 ng of genomic DNA as template and 1 μ M final concentration of each primer in a 100 μ l reaction. All other reagents were present in standard concentrations. Cycling consisted of 92° C for 1 min (denaturation), 50° C for 1 min (annealing) and 72° C for 1 min (elongation) repeated 35 times with a final elongation step of 72° C for 10 min. Amplification products were run on an agarose gel and DNA was extracted from gel slices containing DNA of the appropriate molecular weight using the Prep-a-gene kit (BIO-RAD, Richmond, CA). Amplified fragments were cloned into the pCRII T-tailed vector (Invitrogen, San Diego, CA) and clones containing inserts were selected for further analysis. Plasmid DNA was extracted using the Nucleobond PC-20 kit (Macherey-Nagel, Düren, Germany).

Clones were sequenced using both manual and automated enzymatic sequencing techniques. Internal sequencing primers were synthesized in order to determine the complete sequence on both strands of the clones.

Alignment and phylogenetic analysis. Amino acid sequences of 86 GAPDH genes from eubacteria and eukaryotes were selected from GenBank release 91. These, along with the *Acrasis*, *Naegleria* and *Dictyostelium* sequences were aligned using the ClustalW program [61] using default settings. The alignment was then adjusted by eye and regions of equivocal alignment were identified and excluded from subsequent analysis. A preliminary analysis of this dataset was performed by estimating a distance matrix using the Dayhoff Accepted Point Mutation (PAM) correction (PROTDIST) and constructing a tree with the neighbour-joining algorithm (NEIGHBOR) (see citations below). From this tree, representatives of major groups (metazoans, plants, algae, protists, eubacteria and plastids) were selected to assemble a smaller dataset for subsequent analysis. For both parsimony and distance analysis, a final dataset consisting of 35 taxa and 350 positions was selected with ambiguously aligned regions removed. Insertions, deletions and missing sequence (at the ends of the alignment) were coded as missing data for sequences lacking the homologous regions.

Unweighted parsimony analysis was conducted using the PAUP 3.1.1 package [60]. Shortest trees were obtained by heuristic searches employing 50 random sequence addition replicates to avoid islands of local minima. Bootstrap analysis was performed using simple addition heuristic searches on 300 resamplings of the dataset.

Distance analyses were performed using the following programs of the PHYLIP package, version 3.57c [19]. A PAM-corrected distance matrix was obtained by using PROTDIST and from this a tree was calculated using the neighbour-joining method implemented in the NEIGHBOR program. Bootstrap analysis employed 300 resampled datasets generated by SEQBOOT and distance matrices and trees for each dataset were derived using the programs as described above. A majority-rule consensus bootstrap tree was then calculated by the CONSENSE program.

Due to the computationally intensive nature of maximum likelihood analysis, the use of the PROTML program (of the MOLPHY2.2 package [1]) required the analysis to be broken into two subanalyses. First, to place the root and determine the early branching order of the eukaryote GAPDH subtree, a dataset consisting of 17 sequences of eubacterial and eukaryote sequences was assembled. A semiconstrained tree was then developed based on groupings which were highly supported in both distance and parsimony analysis or for which there is strong reason to believe the sequences are closely related. The maximum likelihood tree was then obtained by exhaustive tree-searching using the semiconstrained tree and employing the Jones, Taylor and Thornton amino acid substitution matrix with transition probabilities adjusted for the amino acid frequencies observed in the dataset (the JTT-F option). To determine the precise position of *D. discoideum* within the crown of the eukaryote subtree, we chose the heterolobosean and euglenozoan glycosomal sequences as outgroups. Once again, exhaustive tree searching was performed on a semiconstrained tree (containing a total of 18 taxa) to find the maximum likelihood tree. Both maximum likelihood analyses required the removal of missing data shared by two or more taxa yielding a dataset consisting of 292 aligned positions. Bootstrap values for nodes were obtained by the resampling estimated log-likelihood (RELL) procedure employed in the PROTML program [1].

Nomenclature for genes and trees. The GAPDH tree can be broken up into two major subtrees: one that contains mostly,

but not exclusively, eubacterial sequences and another that contains mostly, and also not exclusively, eukaryotic sequences. We shall refer to these as the eubacterial and eukaryotic GAPDH subtrees respectively.

To minimize confusion, the following conventions will be observed to denote homologs of the GAPDH enzyme in organisms where multiple copies exist. As has been reported in [65], the kinetoplastids, such as *Trypanosoma brucei* and *Leishmania mexicana*, possess two phylogenetically distinct homologs of GAPDH. In kinetoplastids, one of these homologs is localized in an organelle, the glycosome, whereas the other is a cytosolic enzyme. For the more distantly related euglenozoans *Trypanoplasma borelli* and *Euglena gracilis*, only GAPDH homologs very similar to the kinetoplastid glycosomal GAPDH are thought to be present [27, 65]. We shall refer to the glycosomal versions of this enzyme as the gapG class of enzymes and the cytosolic ones as the gapC. Since *Euglena* lacks a glycosome and only possesses a single type of enzyme, its GAPDH homolog will be referred to as gapC. The entire grouping of glycosomal-related sequences (including the *Euglena* sequence) described in phylogenetic analyses will be referred to as the euglenozoan gapG clade.

The eubacteria *Escherichia coli*, *Anabaena variabilis* and *Synechocystis* sp. also possess multiple copies of GAPDH. In *Escherichia* these are denoted *gapA*, *gapB* and *gapC*. *Synechocystis* and *Anabaena* both have a *gap1* and *gap2* homologs while a third homolog, *gap3*, has only been isolated from *Anabaena* [39].

RESULTS

Sequence features. The use of the GAPN and the GAPC primers generated PCR products from *Acrasis*, *Naegleria* and *Dictyostelium* that were clearly homologous to GAPDH genes from other organisms and covered more than 90% of the predicted coding region. The inferred amino acid sequences from the *Naegleria andersoni* and *Acrasis rosea* genes were free of stop codons and aligned with a minimum of gaps to GAPDH sequences known from other organisms (Fig. 1), suggesting that neither gene is interrupted by spliceosomal introns. This is consistent with reports of a low density of introns in genomes of another vahlkampfiid species, *Naegleria gruberi* (there are only two introns currently known [49]). By contrast, the *Dictyostelium* sequence appears to be interrupted by a two introns displaying typical GT-AG spliceosomal intron boundaries. Their lengths are 90 bp and 85 bp and they map to codon 1, phase 1, and codon 65, phase 2, respectively. Although there is no direct evidence these are introns (there is no cDNA sequence), removal of the predicted intron sequences restores the reading frame and allows precise alignment of the amino acid sequence. Intron #1 appears to be unique to the *Dictyostelium* sequence whereas the position and phase of intron #2 correspond exactly to intron #17 described in [29], to date found exclusively in GAPDH genes of the metazoans, *Homo sapiens* and *Gallus gallus*. Although spliceosomal intron insertion is a controversial phenomenon [33, 34, 50], the unique presence of intron #1 in *Dictyostelium* suggests that it is a recent insertion into this lineage. The shared presence of intron #2 in *Dictyostelium* and the metazoans (likely closely related groups as described below and in [2, 31, 35, 36]) and the lack of this intron in all other known GAPDH genes suggest that it was also likely inserted, more anciently, into the Precambrian common ancestor of both groups. Conversely, the fact that several other introns shared between metazoans and plants [29] (probably an outgroup to a metazoan/*Dictyostelium* clade [31]) are not present in *Dictyostelium*, suggests that these introns were lost in the lineage leading to the dictyostelids.

Alignment features. Comparison of the three new sequences to other homologs (Fig. 1) reveals that they all possess the conserved catalytic cysteine and histidine residues (alignment positions 169 and 197 respectively in Fig. 1) [53] suggesting that they are functional GAPDH genes. Furthermore, they all possess a eukaryotic-like S-loop region suggesting they are most similar to the eukaryotic-like gapC class of GAPDH [21, 38].

There are three regions in the alignment where length heterogeneity in our sequences relative to other eukaryotic GAPDH homologs is observed and each is found between units of secondary structure of the protein (Fig. 1). Alignment positions 139–142 and 160–162 are regions where length heterogeneity in GAPDH genes in general is observed and no consistent pattern is displayed by our sequences. However, at position 286 (Fig. 1, box B) *Naegleria*, *Acrasis* and *Dictyostelium* all share a serine residue lacking in all other sequences. Conserved unique insertions or deletions in alignments can often be considered strong indicators of relationships [2]. However, a single amino acid insertion in this position is also found in GAPDH homologues from *Trichomonas vaginalis* (a glutamate residue), *Zymomonas mobilis* (a threonine) and *Anabaena variabilis gap3* (a glutamine). In addition, multiple amino acid insertions at this position are found in *Agaricus bisporus gap1* and *Haemophilus influenzae*. Since these sequences span extremely deep phylogenetic distances (Fig. 2, 3 and [27, 38, 39, 67]), these are likely polyphyletic insertion events indicating that this region of the protein tolerates length heterogeneity. Thus the shared serine displayed by our sequences cannot be considered strong evidence of their close relationship.

A potentially more stable character is found at alignment position 40 (Fig. 1, box A) where a single amino acid insertion is found only in eukaryotic-like gapC sequences and the eubacterial *Anabaena gap3* sequence. Interestingly, the eukaryotic-like GAPDH sequences *Anabaena gap1* and *Escherichia gapA* in addition to the *Trypanosoma* and *Leishmania* cytosolic (kinetoplastid gapC) GAPDH homologs all lack this insertion. Amongst eukaryotes, this position is occupied by a proline residue in all sequences except *Naegleria* and *Acrasis* (which have an isoleucine) and the euglenozoan gapG sequences (which typically possess a methionine residue). The deep divergence on trees (Fig. 2, 3) [39] between the eukaryotic-like gapC sequences and *Anabaena gap3* suggest that at least two independent insertion events likely occurred: one in the ancestor of *Anabaena gap3* and at least one in the ancestor of the eukaryotic sequences. The topology of the GAPDH tree within the eukaryotic gapC subtree (Fig. 2, 3) is not consistent with a single event of insertion at position 40; polyphyletic insertions or deletions must be postulated to obtain this consistency. However, the possibility exists that this topology is not correct in detail.

Phylogenetic analysis. Parsimony analysis of the aligned dataset yielded a set of 16 equally parsimonious trees with a length of 2434 steps from which a randomly selected one is shown in Fig. 2A. A neighbour-joining tree derived from PAM-corrected distance analysis of the dataset is depicted in Fig. 2B. The proportion of trees in the bootstrap analyses displaying a particular node are shown above branches. Since the bootstrap majority rule consensus trees for parsimony and distance analyses differed markedly from trees based on the original dataset, nodes of conflict are indicated in Fig. 2. Maximum likelihood trees inferred from two datasets, a eukaryotic/prokaryotic dataset and a eukaryotic crown dataset are shown in Fig. 3A & 3B respectively.

The phylogenetic coherence of the Heterolobosea. A *Naegleria/Acrasis* grouping is supported by trees generated using parsimony, distance and maximum likelihood methods (Fig. 2,

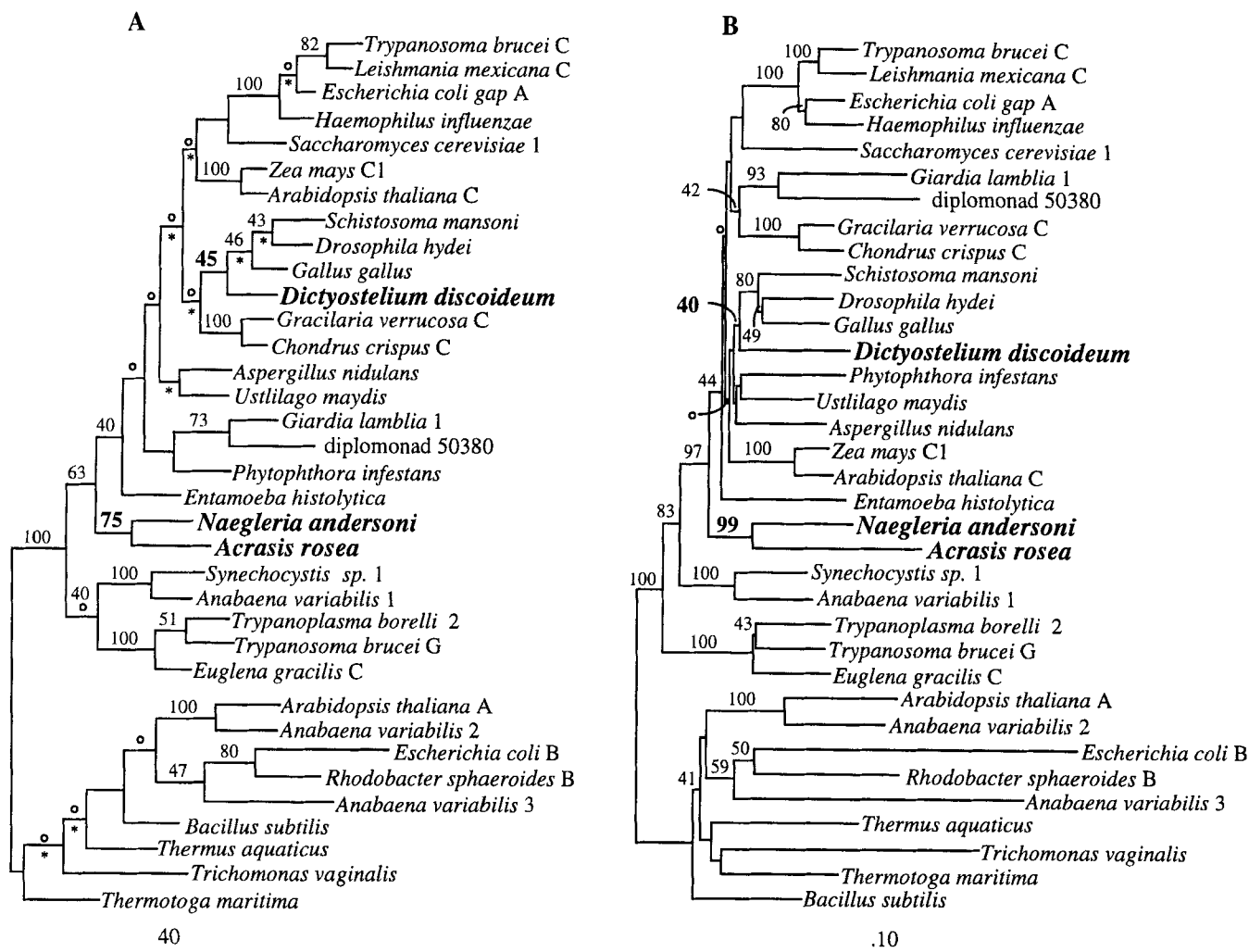


Fig. 2. Phylogenetic trees of glyceraldehyde-3-phosphate dehydrogenase inferred by parsimony and distance methods. Bootstrap proportions of greater than 40% are indicated above branches and nodes not present in the bootstrap majority rule tree are denoted by open circles (○) above the node. A. Unweighted maximum parsimony tree. This tree was randomly chosen from 16 equally parsimonious trees of total tree-length of 2434 steps. Branchlengths were inferred using the ACCTRAN method implemented in the PAUP 3.1.1 package [60]. Asterisks (*) under branches indicate that they were not present in all 16 shortest trees. The scale bar indicates the number of steps per unit branch length. B. Neighbour-joining tree constructed from PAM-corrected distances. PAM-corrected distances and the neighbour-joining tree was inferred using the PROTDIST and NEIGHBOR programs of the PHYLIP package (version 3.57) [19]. The scale bar indicates the numbers of substitutions per site for a unit branchlength.

3A). Bootstrap analysis suggested that this grouping was highly significant for both the distance and maximum likelihood trees (with bootstrap support of 99% and 95%, respectively). However, support for this node in parsimony analysis, although strong, is substantially lower (75%). Inspection of all groupings incompatible with a heterolobosean clade in the parsimony bootstrap analysis showed that all other groupings were found in less than 4% of the replicates. It is likely, therefore, that these other groupings are artifactual topologies where the long

branch of the *Acrasis* GAPDH is attracted to other long branch taxa, an artifact well known to occur in parsimony analysis [17].

Contrary to the hypothesis of a cellular slime mould grouping [37, 62], none of the phylogenetic methods obtained significant support for an *Acrasis/Dictyostelium* clade.

The general structure of the GAPDH tree. Our phylogenetic analyses (Fig. 2, 3) generate overall topologies of the GAPDH tree which are similar to other published analyses [27,

(*) below the alignment mark regions of ambiguous alignment which were excluded from the parsimony and distance analyses. Units of protein secondary structure, inferred from the *Bacillus stearothermophilus* crystal structure (PDB entry: 1GD1.pdb) [53], are shown above the alignment; alpha helices are indicated by (a) and beta-strands by (b). Two regions of the alignment are boxed where potentially informative single amino acid insertions and deletions exist. Box A indicates an insertion common to most eukaryotic nuclear sequences and found in a single eubacterial sequence while box B shows an identical insertion shared by *Acrasis*, *Naegleria* and *Dictyostelium*. Black circles (●) indicate residues involved in catalysis [53].

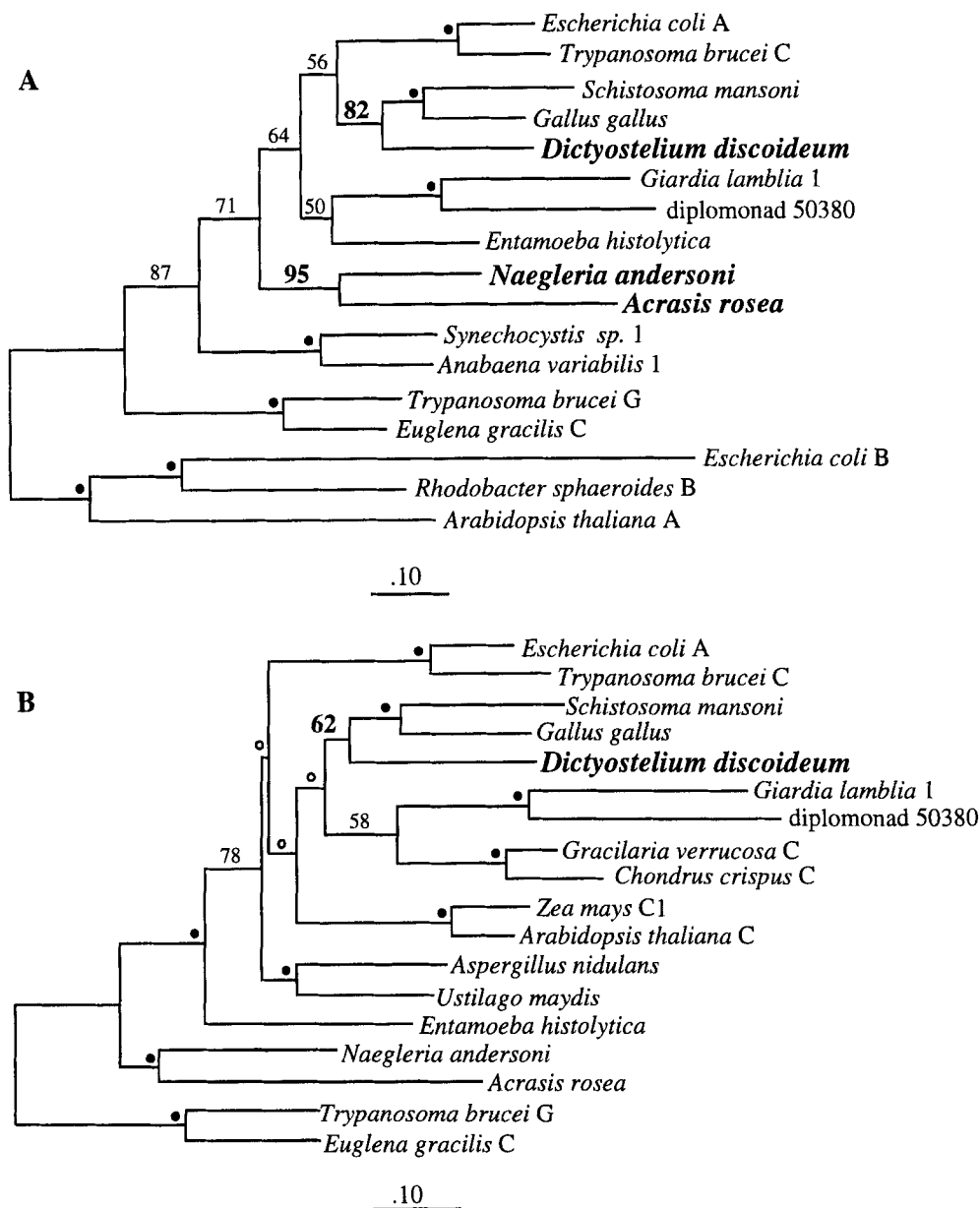


Fig. 3. Maximum likelihood trees derived from taxonomically reduced datasets. Each of these trees were inferred from a dataset consisting of 292 positions by exhaustive tree-searching on semiconstrained trees using the JTT-F model of amino acid replacement implemented in the PROTML program [1]. Nodes constrained during tree-searching are indicated by a black circles (●). Branches absent in the bootstrap majority-rule trees are indicated by open circles (○). Topological constraints were developed based on features of the tree supported strongly in distance and parsimony analysis or for groups where other strong evidence exists for monophyly (e.g. the Metazoa and the Fungi). Estimated bootstrap proportions are shown above nodes. **A.** The overall structure of the GAPDH tree. This tree had a log-likelihood = -7191.5. **B.** The location of *Dictyostelium* in the eukaryotic GAPDH tree. This tree had a log-likelihood = -6425.3

38, 39, 51, 65, 67] despite the differences in the methods employed and the taxonomic sampling. Although the primary focus of this study concerns the branching order within the eukaryote subtree, it should be noted that very different relationships in much of the eubacterial portion of tree are obtained in distance and parsimony analysis. Bootstrap support for most nodes is poor with the exception of the *Anabaena gap2/Arabidopsis gapA* grouping which is supported with a bootstrap value of 100% using both methods.

The deepest successive branches in the eukaryotic subtree in the distance and likelihood trees (Fig. 2B, 3A) are the euglenozoan gapG group followed by the cyanobacterial gap1 sequences. Support for the euglenozoan gapG group occupying the deepest branch is quite strong with bootstrap values exceeding 80% for both distance and maximum likelihood methods. By contrast, all 16 shortest trees in the parsimony analysis instead show that the euglenozoan gapG and the cyanobacterial gap1 sequences form a clade at the base of the eukaryote tree.

Interestingly, the parsimony bootstrap majority-rule consensus tree, which can be considered a bias-corrected estimate of the phylogeny [18], does not display this clade but instead shows the same topology as recovered by distance and maximum likelihood methods. However, the node placing the euglenozoan gapG sequences at the base of the eukaryote subtree is supported by a bootstrap value of 42% versus 40% for the cyanobacterial gap1/euglenozoan gapG clade suggesting that neither topology is significantly preferred over the other in parsimony analysis.

The node unifying the rest of the eukaryotic-like GAPDH sequences to the exclusion of both cyanobacterial gap1 and euglenozoan gapG sequences is highly supported by distance methods (with bootstrap support of 97%) and while only moderately by parsimony and likelihood (with bootstrap support of 63% and 71% respectively).

In all methods, the heteroloboseans are the next group to diverge on the GAPDH tree. Bootstrap values suggesting that

they are an outgroup to all remaining eukaryotic-like sequences are quite low for all methods (40%, 44% and 64% for parsimony, distance and maximum likelihood respectively). The heterolobosean sequences show no special affinity for any other group in the trees.

The overall structure of the eukaryotic subtree following the branching of the Heterolobosea is very poorly resolved. Several major groups, also described in previous analyses [27, 38, 39, 51, 65, 67], receive strong bootstrap support from both distance and parsimony methods. These include the land plants, the rhodophytes, the diplomonads and the kinetoplastid gapC/gamma-Proteobacterial clade. In sharp contrast, the monophyly of the Metazoa is only strongly supported by distance methods, receiving only 46% support in the parsimony analysis. Fungi are polyphyletic with the yeast *Saccharomyces cerevisiae* having no specific relationship to the two other fungi included in the analysis, *Aspergillus nidulans* and *Ustilago maydis*. The latter form only a weak clade in parsimony analysis (with bootstrap support less than 40%). In the distance tree, these fungi appear as an outgroup to a *Dictyostelium*/metazoan grouping but are paraphyletic, interrupted by the sequence from the oomycete *Phytophthora infestans*.

The relative branching order of well supported groups and all other sequences in the eukaryotic subtree is not consistent; many internal nodes are not shared by all of the 16 shortest trees in parsimony analysis (stars under nodes in Fig. 2A) or by the shortest trees and the bootstrap consensus tree in parsimony and distance analysis (open circles above nodes in Fig. 2), or between the trees generated by the the different methods (Fig. 2, 3). Moreover, very few of these internal nodes receive significant bootstrap support in any of the analyses.

The position of *Dictyostelium*. One consistent feature in these analyses is the placement of the *Dictyostelium* sequence as a sister group to the metazoa. Once again, bootstrap support is low for both distance (40%) and parsimony analyses (45%) but it receives stronger support (82%) in the global maximum likelihood tree with reduced taxonomic representation (Fig. 3A). Such an effect is not surprising given that in this maximum likelihood analysis there were fewer taxa represented near the node of interest. This reduces the number of possible alternative suboptimal topologies and, on average, bootstrap support for any node will increase. In agreement with this, the eukaryotic GAPDH subtree maximum likelihood analysis (Fig. 3B), with several more eukaryotic sequences included, also shows the Metazoa/*Dictyostelium* grouping, but with reduced bootstrap support (62%).

DISCUSSION

The Heterolobosea and beyond. Our finding of a *Naegleria/Acrasis* grouping in the GAPDH tree strongly supports the proposal of a Schizopyrenida/Acrasidae grouping, the Heterolobosea. This suggests that the three ultrastructural characters described by Page and Blanton as defining the Heterolobosea are, taken together, good phylogenetic indicators [45].

It is unclear, however, if either of the classes Acrasidae or the Schizopyrenida are themselves monophyletic. In their study [45], Page and Blanton considered only one family of acrasids, the Acrasidae, containing the genera *Acrasis* and *Pocheina*. Three additional acrasid families, the Guttulinopsidae, the Copromyxidae and the Fonticulidae were excluded from that study because they did not possess all of the features defining the Heterolobosea and were suspected to be unrelated. Of the three, the Guttulinopsidae are most similar to the Acrasidae, possessing discoidal mitochondrial cristae and lacking a Golgi dictyosome. However, they differ by the absence of an envelope of rough endoplasmic reticulum surrounding their mitochondria.

The relationship of the Fonticulidae and the Copromyxidae to these groups is more doubtful; the former possess mitochondria with discoidal cristae but their amoebae are not of the eruptive limax type, while the latter have mitochondria with tubular cristae [5]. Our study only confirms the relationship of the family Acrasidae to the schizopyrenids and more developmental, ultrastructural and molecular data are necessary to determine the relationship of these other acrasid groups to the Heterolobosea.

The similarity of the flagellated species of the Acrasidae to some specific schizopyrenids suggests that the latter may be a paraphyletic group. The acrasid *Pocheina flagellata*, for example, has a flagellate stage which bears two equal length flagella [44], thus specifically resembling members of the vahlkampfiid genus *Naegleria*. Hinkle and colleagues [28] have recently shown that, within the Vahlkampfiidae, there is a deep phylogenetic gulf between *Naegleria* and three other genera, *Vahlkampfia*, *Tetramitus* and *Paratetramitus*. We suggest that, if the schizopyrenids are paraphyletic, then the Acrasidae may be specifically related to the *Naegleria* lineage. Further studies of the small subunit ribosomal RNA of acrasids will be useful in resolving this issue.

Several authors have suggested specific relatives of the Heterolobosea. In his description of the protozoan phylum Percolozoa, Cavalier-Smith [10] suggests that the Heterolobosea are related to members of the protist genera, *Percolomonas* (as first suggested by Fenchel and Patterson [20]) and *Stephanopogon* because both lack a Golgi dictyosome and possess discoidal mitochondrial cristae. Along similar lines, Patterson has argued that heteroloboseans are related to two other protist groups with discoidal mitochondrial cristae; nuclearioid amoebae and the Euglenozoa [47].

Of these organisms, only species of *Percolomonas* appear particularly similar to known heteroloboseans, resembling members of the vahlkampfiid genus *Tetramitus* but lacking an amoeboid stage [20]. Except for the Euglenozoa, no molecular data exist for any of these organisms and thus suggestions regarding their phylogenetic affinities must be considered only as tentative hypotheses.

By contrast, it is possible to test the validity of Patterson's euglenozoan/heterolobosean hypothetical grouping with existing molecular data. In our GAPDH trees, no affinity between the heterolobosean homologs and either the euglenozoan gapG or the kinetoplastid gapC enzymes is found. This coupled with the absence of a euglenozoan/heterolobosean grouping in trees based on small subunit ribosomal RNA [9, 56] and elongation factor 1 α (AJR., unpubl. data) can be regarded as evidence against Patterson's hypothesis. However, one molecular dataset does show such an affinity; trees based on α -tubulin [2] do show a *Naegleria*/euglenozoan clade, although the critical node is not strongly supported in bootstrap analysis. Clearly, the issue will only be decided when sequences of other phylogenetic marker molecules from heteroloboseans become available and a consistent pattern emerges.

Dictyostelium and the position of the Eumycetozoa. Suggestions [41, 62] that acrasids and dictyostelids are related, due to their common possession of a cellular slime mould habit, are discredited by the complete absence of an affinity of *Acrasis* for *Dictyostelium* in our GAPDH trees. If other molecular trees are congruent, then we suggest that the use of names based on the concept of an acrasid/dictyostelid grouping, such as the Acrasieae [14, 41, 62] or the Acrasiomycota [37, 64], be discontinued.

Several years ago, it was suggested that extremes in base composition have led to a consistent misplacement of *Dictyostelium* as branching prior to the radiation of the eukaryotic "crown" groups in small subunit ribosomal RNA trees [35].

Analyses of multiple protein-coding genes have instead shown that *Dictyostelium* consistently branches close to the metazoa and the fungi [2, 31, 36]. Our analysis also suggests that *Dictyostelium* is not a deeply branching eukaryotic group in the GAPDH tree, emerging as an immediate sister group to the Metazoa.

It is difficult to compare this result to those reported by others [2, 31, 35, 36] since, in our analysis, the different phylogenetic methods do not concur on the degree of monophyly nor the position of the fungal sequences relative to other eukaryotic GAPDH sequences. Nevertheless, taken literally, the trees in Fig. 2, 3 appear, superficially, to support the view that dictyostelids are more closely related to metazoans than either group are to fungi [36]. However, a recent maximum likelihood analysis of 19 protein datasets yielded greater overall support for a topology where *Dictyostelium* is an immediate outgroup to a metazoan/fungal clade [31], in agreement with the consistent placement of Fungi and Metazoa as sister groups in phylogenetic analyses of other molecular datasets [2, 9, 16, 63]. Moreover, the metazoa and fungi possess two strong synapomorphies lacking in *Dictyostelium*: a 12-amino acid insertion in their elongation factor 1 α homologues [2] and flattened mitochondrial cristae [9, 43]. We suggest, therefore, that the weight of evidence favours dictyostelids as an immediate outgroup to a metazoan/fungal clade. The polyphyly of the fungi as well as their lack of a strong affinity for the metazoan sequences in the GAPDH tree, indicates that the fungi are likely artifactually misplaced, perhaps as a result of an accelerated rate of evolution in the fungal enzymes.

The relationship of *Dictyostelium* to the plasmodial slime moulds is also a controversial issue. In nearly all small subunit ribosomal RNA trees published to date (for an exception see [9]), *Dictyostelium* and *Physarum polycephalum* are not sister taxa. This has led to the commonly-held notion that the dictyostelids and myxomycetes are unrelated [36, 52]. Curiously, however, the recent inclusion of a protostelid slime mould sequence into a partial small subunit ribosomal RNA dataset, appeared to move these species into a clade [59], consistent with the previously proposed slime mould group, the Eumycetozoa [43], an assemblage comprised of protostelids, dictyostelids and myxomycetes. In addition, this group was placed in the "crown" of the eukaryotic tree suggesting a more recent common ancestry with metazoans than previously observed. However, these trees, unlike those based on GAPDH and other protein data, group the Eumycetozoa specifically with the oomycetes and heterokonts rather than the animals and fungi. Interestingly, strong support for a dictyostelid/myxomycete clade is also found in trees of actin [4]. Thus it is clear that more data are needed; characterization of GAPDH and other protein coding genes from myxomycetes, protostelids and a wider variety of eukaryotes may be helpful in clarifying their phylogenetic position and testing the holophyly of the Eumycetozoa.

Reconciling GAPDH phylogeny with phylogeny of other molecules. When there are discrepancies between molecular trees in the absence of a known phylogeny, it is difficult to decide which, if any, molecules are correctly resolving true organismal relationships. One method of dealing with this problem is to look for congruence between different molecular and morphological phylogenies. If a single molecular tree conflicts with several others which are mutually congruent, then there are grounds to argue that that tree is not reflecting true organismal relationships.

In addition to those mentioned in the preceding section, there are several major discrepancies between phylogenies of eukaryotes inferred by GAPDH and those of other molecules. For instance, as several authors have pointed out [27, 38], *Giar-*

dia lamblia (and now several other diplomonads [51]) and *Entamoeba histolytica*, both amitochondrial protists, do not branch deeply in the eukaryotic GAPDH tree, incongruent with their deep position in phylogenies of small subunit ribosomal RNA [9, 56], elongation factor 1 α [2, 24, 25] and the A-subunit of RNA polymerase II [30]. We suggest that the lack of strong bootstrap support for internal nodes connecting these with other eukaryotic groups in the GAPDH tree and the major differences between parsimony, distance and maximum likelihood-derived topologies is consistent with a lack of resolution at deep phylogenetic distances caused by substitutional saturation in this enzyme. If this is true, then there is no significant conflict with other datasets. Consistent with this view, strong support was obtained for clades of recently diverged organisms such as the monocots and the dicots, represented by *Arabidopsis* and *Zea*, the two red algae *Chondrus* and *Gracilaria* and the kinetoplastids *Trypanosoma* and *Leishmania*. By contrast, poorer resolution is found for all of the Precambrian branches; those connecting the groups described above with most of the other groups found in the eukaryotic subtree.

As has been previously reported [51, 55, 67] and mentioned above, the ascomycete yeast *Saccharomyces*, fails to group with the filamentous ascomycete *Aspergillus* and the basidiomycete *Ustilago*, despite a wealth of morphological and molecular evidence that these are members of a holophyletic fungal group [3, 8]. This case of incongruence appears more problematic since the two ascomycetes, *Saccharomyces* and *Aspergillus*, likely shared a common fungal branch for at least 200 million years (my) prior to their splitting roughly 320 million years ago (mya) [3]. Since the much earlier divergence (540 mya), and shorter shared common history (probably lasting less than 100 million years [48]) of the metazoan groups is resolved by GAPDH, this suggests that the globally high rate of sequence evolution of this enzyme cannot alone be responsible for the anomaly. This coupled with the aberrant placement of Fungi relative to the Metazoa (discussed above), suggests that fungal GAPDH sequences, especially *Saccharomyces* and the other ascomycete yeasts [51], may have suffered a change in their pattern of substitution or increase in their rate of evolution which has obscured their true phylogenetic relationships.

The origin(s) of the cytosolic GAPDH genes of eukaryotes. Perhaps the two most peculiar features of GAPDH phylogeny are the presence of two phylogenetically distinct homologues of the enzyme found in the kinetoplastids (gapG and gapC) and the grouping of gamma-Proteobacterial (represented by *Escherichia gapA* and the *Haemophilus* homologs) [6, 15, 54] and cyanobacterial gap1 homologs [39] firmly within the eukaryotic-like GAPDH subtree. For these anomalies two alternative explanations have been offered.

R. F. Doolittle et al. [15] suggested that an event of lateral gene transfer occurred whereby a eukaryotic GAPDH gene (presumably from the kinetoplastid gapC lineage) was transferred to the gamma-Proteobacterial lineage. The subsequent discovery of a eukaryotic-like GAPDH gene in cyanobacteria (the cyanobacterial gap1 enzymes), not specifically related to the gamma-Proteobacterial version, makes it necessary to propose that two such events occurred independently.

A second explanation, offered by Martin and colleagues [39], accounts for these anomalies by suggesting that the gene encoding the GAPDH enzyme underwent a series of gene-duplications early in eubacterial evolution, giving rise to several paralogous gene families. Subsequently, GAPDH homologs from lineages of these eubacterial paralogous subtrees were laterally transferred to the eukaryotic nucleus several times; once, relatively recently from the gamma-Proteobacteria to kinetoplastids (giving rise to the kinetoplastid gapC enzyme) and once, or

possibly several times, anciently in a common ancestor of extant eukaryotes perhaps in the context of the endosymbiotic origin of mitochondria [27, 39].

It is difficult to decide between these two hypotheses as they both require multiple lateral transfer events. However, it is possible to discern the directionality of the transfer for the case of the gamma-Proteobacterial/kinetoplastid gapC clade. The fact that two successive outgroups to the kinetoplastids, *Trypanoplasma* and *Euglena*, lack the gapC enzyme is most parsimoniously explained if the transfer occurred from the gamma-Proteobacteria to the kinetoplastids after the splitting of these groups [27]. If one accepts this argument, then it seems simplest to assume that the cyanobacterial gap1 is also an enzyme native to these bacteria and that the ancestral GAPDH of the eukaryotic gapC clade resided in a eubacterium, in agreement with Martin et al. [39].

However, the inclusion of our heterolobosean sequences in the GAPDH tree makes this scenario more complex. If the topologies of our GAPDH trees (Fig. 2, 3) are taken literally, then many independent transfer events to the eukaryotic nucleus must be postulated (Fig. 4A). If we instead presume that most of the internal structure of the eukaryotic GAPDH tree is unresolved (as demonstrated by the incongruency of trees inferred by different methods and the low bootstrap values for most of the internal nodes) and that aberrant rates and patterns of amino acid substitution in the euglenozoan gapG lineages are artifactually placing them deeply in the eukaryotic subtree (consistent with the large number of unique insertions present in these enzymes), then the hypothetical tree shown in Fig. 4B may be closer to the true gene phylogeny.

This topology requires that only two lateral transfers to the eukaryotic nucleus took place. One transfer occurred anciently to the ancestor of all extant eukaryotes giving rise to most eukaryotic cytosolic enzymes. A second, more recent transfer from the gamma-Proteobacterial lineage to the kinetoplastid lineage (after *Trypanoplasma* and *Euglena* diverged) then took place giving rise to the kinetoplastid gapC enzyme [65]. As circumstantial evidence for this suggestion, we find that the hypothetical topology shown in Fig. 4B also reduces the number of independent insertion events necessary at position 40 in the alignment (Fig. 1, box A) to two; one occurring on the branch leading to the eukaryotic gapC enzymes and one on the branch leading to *Anabaena gap3*.

If the endosymbiotic origin scenario for nuclear GAPDH is correct, two major problems still remain. Firstly, entamoebids and diplomonads, both amitochondrial protist groups, possess GAPDH genes derived from mitochondrial endosymbiosis. Secondly, the GAPDH gene found in the protist *Trichomonas vaginalis* falls firmly within the eubacterial subtree [38]. These facts can be rationalized by considering the growing body of evidence that these amitochondrial protists may have secondarily lost their mitochondria [13, 58] (AJR, unpubl. data). If this is the case, then the ancestral mitochondrial endosymbiont's genome may have possessed several of the anciently-duplicated GAPDH homologs, and diplomonads, entamoebids and most other eukaryotes could have retained one copy that gave rise to the typical eukaryotic gapC enzyme, while trichomonads retained another version belonging to a different eubacterial paralogous gene family.

However, these speculative hypotheses will only be tenable when evidence for a mitochondrial origin for these enzymes, in the form of alpha-Proteobacterial GAPDH homologs of both the eukaryotic and trichomonad enzymes, is forthcoming. The recent finding of a typical eukaryotic/eubacterial-like GAPDH gene in the archaeobacterium *Haloarcula vallismortis* [7], in contrast to the highly divergent version found in many other

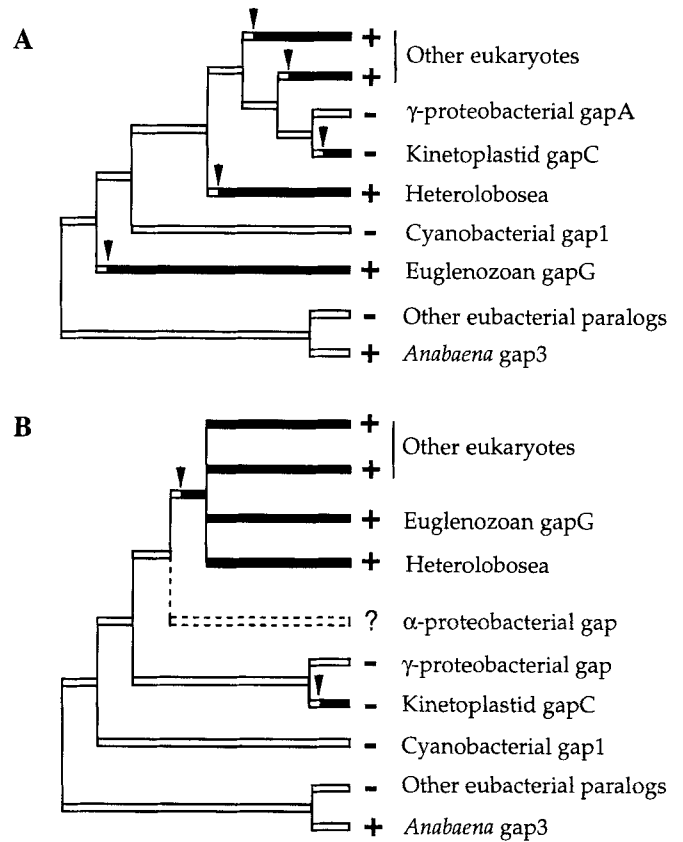


Fig. 4. Two possible scenarios of transfer of GAPDH to the eukaryotic nucleus from a prokaryotic source superimposed on the gene tree. Blackened branches indicate that the gene is residing in a eukaryote and white branches indicate its presence in a prokaryote. Plus signs (+) indicate the presence of an insertion at position 40 in the alignment (Fig. 1, box A) in that lineage and, conversely, minus signs (-) denote the lack of this insertion. Arrows (▼) indicate events of gene transfer from prokaryote to eukaryote. **A.** A topology of the GAPDH tree similar to those recovered by all of the phylogenetic methods. **B.** A hypothetical topology which minimizes the number of transfer and insertion events necessary to explain the data. If this topology is correct, then the trees depicted in Fig. 2, 3 misplace the euglenozoan gapG lineage too deeply in the eukaryotic-like subtree and fail to place the kinetoplastid/gamma-Proteobacterial gapC clade as an outgroup to all other eukaryotic sequences. If the GAPDH of the eukaryotic cytosol is derived from the mitochondrial endosymbiont, we propose that a hypothetical eukaryotic-like alpha-Proteobacterial homolog, indicated by the dashed branch, will be found.

archaeobacteria [26], suggests that we may have to reconsider the possibility that the nuclear GAPDH is not of endosymbiotic origin but instead evolved by direct filiation from the common ancestor of eukaryotes and archaeobacteria. If this view were correct, eukaryotic-like cyanobacterial and gamma-Proteobacterial GAPDH genes have originated by several lateral transfers from eukaryote to prokaryote in agreement with the hypothesis advocated by R. F. Doolittle et al. [15, 54].

ACKNOWLEDGMENTS

We are grateful to F. W. Spiegel for providing the *Acrasis* culture, advice on cultivation and interesting discussions regarding mycetozoa relationships. We are also thankful of DNA provided by S. Kilvington and W. F. Loomis. Thanks are due to M. Müller, H. Philippe and W. Martin for illuminating discussions about GAPDH phylogeny and to T. Cavalier-Smith for

interesting discussions about protistan relationships. For critical reading of early versions of the manuscript, we thank S. L. Baldauf, D. R. Edgell, N. M. Fast, J. M. Logsdon, C. G. Clark and J. Silberman. This work was supported by a grant, MT4467, from the Medical Research Council of Canada awarded to WFD and an Natural Sciences and Engineering Research Council 1967 Science and Engineering Scholarship awarded to AJR.

LITERATURE CITED

1. Adachi, J. & Hasegawa, M. 1992. MOLPHY: Programs for Molecular Phylogenetics, Vers. 2.2. Institute of Statistical Mathematics, Tokyo, Japan.
2. Baldauf, S. L. & Palmer, J. D. 1993. Animals and fungi are each other's closest relatives: congruent evidence from multiple proteins. *Proc. Nat. Acad. Sci. USA*, **90**:11558–11562.
3. Berbee, M. L. 1993. Dating the evolutionary radiations of the true fungi. *Can. J. Bot.*, **71**:1114–1127.
4. Bhattacharya, D. & Ehling, J. 1995. Actin coding regions: gene family evolution and use as a phylogenetic marker. *Arch. Protistenkd.*, **145**:155–164.
5. Blanton, R. L. 1989. Phylum Acrasea. In: Margulis, L., Corliss, J. O., Melkonian, M. & Chapman, D. J. (ed.), *Handbook of Protozoa*. Jones & Bartlett, Boston, Massachusetts. Pp. 75–87.
6. Brantant, G. & Brantant, C. 1985. Nucleotide sequence of the *Escherichia coli* gap gene. Different evolutionary behavior of the NAD⁺-binding domain and of the catalytic domain of D-glyceraldehyde-3-phosphate dehydrogenase. *Eur. J. Biochem.*, **150**:61–66.
7. Brinkmann, H. & Martin, W. 1996. Plant chloroplast and cytosolic 3-phosphoglycerate kinases: a case of endosymbiotic gene replacement. *Plant Mol. Biol.*, **30**:65–75.
8. Bruns, T. D., Vilgalys, R., Barns, S. M., Gonzalez, D., Hibbett, D. S., Lane, D. J., Simon, L., Stickele, S., Szaro, T. M., Weisburg, W. G. & Sogin, M. L. 1992. Evolutionary relationships within the fungi: analyses of nuclear small subunit rRNA sequences. *Mol. Phylogen. Evol.*, **1**:231–241.
9. Cavalier-Smith, T. 1993a. Kingdom Protozoa and its 18 phyla. *Microbiol. Rev.*, **57**:953–994.
10. Cavalier-Smith, T. 1993b. Percolozoa and the symbiotic origin of the metakaryote cell. In: Ishikawa, H., Ishida, M. & Sato, S. (ed.), *Entocytobiology V*, Tübingen University Press, Tübingen, Germany. Pp. 399–406.
11. Clark, C. G. 1992. DNA purification from polysaccharide-rich cells. In: Lee, J. J. & Soldo, A. (ed.), *Protocols in Protozoology*. Allen Press, Lawrence, Kansas. **1**:D3.1–D3.2.
12. Clark, C. G. & Cross, G. A. 1988. Small-subunit ribosomal RNA sequence from *Naegleria gruberi* supports the polyphyletic origin of amoebae. *Mol. Biol. Evol.*, **5**:512–518.
13. Clark, C. G. & Roger, A. J. 1995. Direct evidence for secondary loss of mitochondria in *Entamoeba histolytica*. *Proc. Nat. Acad. Sci. USA*, **92**:6518–6521.
14. de Bary, A. 1887. *Comparative Morphology and Biology of the Fungi, Mycetozoa and Bacteria*. Clarendon Press, Oxford.
15. Doolittle, R. F., Feng, D. F., Anderson, K. L. & Alberro, M. R. 1990. A naturally occurring horizontal gene transfer from a eukaryote to a prokaryote. *J. Mol. Evol.*, **31**:383–388.
16. Doolittle, R. F., Feng, D. F., Tsang, S., Cho, G. & Little, E. 1996. Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science*, **271**:470–477.
17. Felsenstein, J. 1978. Cases in which parsimony and compatibility methods will be positively misleading. *Syst. Zool.*, **27**:401–410.
18. Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**:783–791.
19. Felsenstein, J. 1996. PHYLIP (Phylogeny Inference Package), Vers. 3.57c. Department of Genetics, University of Washington, Seattle.
20. Fenchel, T. & Patterson, D. J. 1986. *Percolomonas cosmopolitus* (Ruinen) n. gen., a new type of filter feeding flagellate from marine plankton. *J. Mar. Biol. Ass. U. K.*, **66**:465–482.
21. Fothergill-Gilmore, L. A. & Michels, P. A. 1993. Evolution of glycolysis. *Prog. Biophys. Mol. Biol.*, **59**:105–235.
22. Galtier, N. & Gouy, M. 1995. Inferring phylogenies from DNA sequences of unequal base compositions. *Proc. Nat. Acad. Sci. USA*, **92**:11317–11321.
23. Hasegawa, M. & Hashimoto, T. 1993. Ribosomal RNA trees misleading? *Nature*, **361**:23.
24. Hasegawa, M., Hashimoto, T., Adachi, J., Iwabe, N. & Miyata, T. 1993. Early branchings in the evolution of eukaryotes: ancient divergence of *Entamoeba* that lacks mitochondria revealed by protein sequence data. *J. Mol. Evol.*, **36**:380–388.
25. Hashimoto, T., Nakamura, Y., Nakamura, F., Shirakura, T., Adachi, J., Goto, N., Okamoto, K. & Hasegawa, M. 1994. Protein phylogeny gives a robust estimation for early divergences of eukaryotes: phylogenetic place of a mitochondria-lacking protozoan, *Giardia lamblia*. *Mol. Biol. Evol.*, **11**:65–71.
26. Hensel, R., Zwickl, P., Fabry, S., Lang, J. & Palm, P. 1989. Sequence comparison of glyceraldehyde-3-phosphate dehydrogenases from the three kingdoms: evolutionary implication. *Can. J. Microbiol.*, **35**:81–85.
27. Henze, K., Badr, A., Wettern, M., Cerff, R. & Martin, W. 1995. A nuclear gene of eubacterial origin in *Euglena gracilis* reflects cryptic endosymbioses during protist evolution. *Proc. Nat. Acad. Sci. USA*, **92**:9122–9126.
28. Hinkle, G. & Sogin, M. L. 1993. The evolution of the Vahlkampfiidae as deduced from 16S-like ribosomal RNA analysis. *J. Euk. Microbiol.*, **40**:599–603.
29. Kersanach, R., Brinkmann, H., Liaud, M. F., Zhang, D. X., Martin, W. & Cerff, R. 1994. Five identical intron positions in ancient duplicated genes of eubacterial origin. *Nature*, **367**:387–389.
30. Klenk, H.-P., Zillig, W., Lanzendorfer, M., Grampp, M. & Palm, P. 1995. Location of protist lineages in a phylogenetic tree inferred from sequences of DNA-dependent RNA polymerases. *Arch. Protistenkd.*, **145**:221–230.
31. Kuma, K., Nikoh, N., Iwabe, N. & Miyata, T. 1995. Phylogenetic position of *Dictyostelium* inferred from multiple protein data sets. *J. Mol. Evol.*, **41**:238–246.
32. Levine, N. D., Corliss, J. O., Cox, F. E. G., Deroux, G., Grain, J., Honigberg, B. M., Leedale, G. F., Loeblich, A. R., III, Lom, J., Lynn, D., Merinfeld, F. G., Page, F. C., Poljansky, G., Sprague, V., Vavra, J. & Wallace, F. G. 1980. A newly revised classification of the Protozoa. *J. Protozool.*, **27**:37–58.
33. Logsdon, J., Jr., Tyshenko, M. G., Dixon, C., D., Jafari, J., Walker, V. K. & Palmer, J. D. 1995. Seven newly discovered intron positions in the triose-phosphate isomerase gene: evidence for the introns-late theory. *Proc. Nat. Acad. Sci. USA*, **92**:8507–8511.
34. Logsdon, J. M., Jr. & Palmer, J. D. 1994. Origin of introns—early or late? *Nature*, **369**:526.
35. Loomis, W. F. & Smith, D. W. 1990. Molecular phylogeny of *Dictyostelium discoideum* by protein sequence comparison. *Proc. Nat. Acad. Sci. USA*, **87**:9093–9097.
36. Loomis, W. F. & Smith, D. W. 1995. Consensus phylogeny of *Dictyostelium*. *Experientia*, **51**:1110–1115.
37. Margulis, L. & Schwartz, K. V. 1988. *Five kingdoms: an illustrated guide to the phyla of life on earth*. W. H. Freeman, New York.
38. Markós, A., Miretsky, A. & Müller, M. 1993. A glyceraldehyde-3-phosphate dehydrogenase with eubacterial features in the amitochondriate eukaryote, *Trichomonas vaginalis*. *J. Mol. Evol.*, **37**:631–643.
39. Martin, W., Brinkmann, H., Savonna, C. & Cerff, R. 1993. Evidence for a chimeric nature of nuclear genomes: eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. *Proc. Nat. Acad. Sci. USA*, **90**:8692–8696.
40. O'Kelly, C. J. 1993. The jakobid flagellates: structural features of *Jakoba*, *Reclinomonas* and *Histiona* and implications for the early diversification of eukaryotes. *J. Euk. Microbiol.*, **40**:627–636.
41. Olive, E. W. 1902. Monograph of the Acrasieae. *Proc. Boston Soc. Nat. Hist.*, **30**:451–513.
42. Olive, L. S. 1970. The Mycetozoa: a revised classification. *Bot. Rev.*, **36**:59–87.
43. Olive, L. S. 1975. *The Mycetozoa*. Academic Press, New York.
44. Olive, L. S., Stoianovitch, C. & Bennet, W. E. 1983. Descriptions of acrasid cellular slime molds: *Pocheina rosea* and a new species, *Pocheina flagellata*. *Mycologia*, **75**:1019–1029.
45. Page, F. C. & Blanton, R. L. 1985. The Heterolobosea (Sarcodina: Rhizopoda), a new class uniting the Schizopyrenida and the Acrasidae (Acrasida). *Protistologica*, **21**:121–132.

46. Palmer, J. D. & Logsdon, J. M., Jr. 1991. The recent origins of introns. *Curr. Opin. Genet. Dev.*, **1**:470–477.
47. Patterson, D. J. 1994. Protozoa: evolution and systematics. In: Hausmann, K. & Hülsmann, N. (ed.), *Progress in Protozoology*. Gustav Fischer, Stuttgart, Germany. Pp. 1–14.
48. Philippe, H., Chenuil, A. & Adoutte, A. 1994. Can the Cambrian explosion be inferred through molecular phylogeny? *Development* **1994**(suppl.):15–25.
49. Remillard, S. P., Lai, E. Y., Levy, Y. Y. & Fulton, C. 1995. A calcineurin-B-encoding gene expressed during differentiation of the amoeboid flagellate *Naegleria gruberi* contains two introns. *Gene*, **154**: 39–45.
50. Roger, A. J., Keeling, P. J. & Doolittle, W. F. 1994. Introns, the broken transposons. *Soc. Gen. Physiol. Ser.*, **49**:27–37.
51. Rozario, C., Morin, L., Roger, A. J., Smith, M. W. & Müller, M. 1996. Primary structure and phylogenetic relationships of glyceraldehyde-3-phosphate dehydrogenase genes of free-living and parasitic diplomonad flagellates. *J. Euk. Microbiol.*, **43**:330–340.
52. Schlegel, M. 1994. Molecular phylogeny of eukaryotes. *Trends Ecol. Evol.*, **9**:330–335.
53. Skarzynski, T., Moody, P. C. & Wonacott, A. J. 1987. Structure of holo-glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* at 1.8 ° resolution. *J. Mol. Biol.*, **193**:171–187.
54. Smith, M. W., Feng, D. F. & Doolittle, R. F. 1992. Evolution by acquisition: the case for horizontal gene transfers. *Trends Biochem. Sci.*, **17**:489–493.
55. Smith, T. L. 1989. Disparate evolution of yeasts and filamentous fungi indicated by phylogenetic analysis of glyceraldehyde-3-phosphate dehydrogenase genes. *Proc. Nat. Acad. Sci. USA*, **86**:7063–7066.
56. Sogin, M. L. 1991. Early evolution and the origin of eukaryotes. *Curr. Opin. Gen. Devel.*, **1**:457–463.
57. Sogin, M. L. & Gunderson, J. H. 1987. Structural diversity of eukaryotic small subunit ribosomal RNAs. Evolutionary implications. *Ann. N. Y. Acad. Sci.*, **503**:125–139.
58. Soltys, B. J. & Gupta, R. S. 1994. Presence and cellular distribution of a 60-kDa protein related to mitochondrial hsp60 in *Giardia lamblia*. *J. Parasitol.*, **80**:580–590.
59. Spiegel, F. W., Lee, S. B. & Rusk, S. A. 1995. Eumycetozoans and molecular systematics. *Can. J. Bot.*, **73**(suppl. 1):S738–S746.
60. Swofford, D. L. 1993. PAUP: Phylogenetic Analysis Using Parsimony, Vers. 3.1.1. Illinois Natural History Survey, Champaign, Illinois.
61. 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. *Nucl. Acids Res.*, **22**:4673–4680.
62. van Tieghem, M. P. 1880. Sur quelques myxomycètes à plasmode agrégé. *Bull. Soc. Bot. Fr.*, **27**:317–322.
63. Wainright, P. O., Hinkle, G., Sogin, M. L. & Stickel, S. K. 1993. Monophyletic origins of the metazoa: an evolutionary link with fungi. *Science*, **260**:340–342.
64. Whittaker, R. H. 1969. New concepts of kingdoms of organisms. *Science*, **163**:150–160.
65. Wiemer, E. A., Hannaert, V., van den Ijssel, P. R., Van Roy, J., Opperdoes, F. R. & Michels, P. A. 1995. Molecular analysis of glyceraldehyde-3-phosphate dehydrogenase in *Trypanoplasma borelli*: an evolutionary scenario of subcellular compartmentation in kinetoplastida. *J. Mol. Evol.*, **40**:443–454.
66. Wolters, J. 1991. The troublesome parasites—molecular and morphological evidence that Apicomplexa belong to the dinoflagellate-ciliate clade. *Biosystems*, **25**:75–83.
67. Zhou, Y. & Ragan, M. A. 1995. The nuclear gene and cDNAs encoding cytosolic glyceraldehyde-3-phosphate dehydrogenase from the marine red alga *Gracilaria verrucosa*: cloning, characterization and phylogenetic analysis. *Curr. Genet.*, **28**:324–332.

Received 4-19-96, 7-5-96; accepted 7-8-96.

10th International Congress of Protozoology (ICOP-10)

July 21–25, 1997

Sydney, Australia

For information, contact:

Prof. D. J. Patterson

School of Biological Sciences AO8

University of Sydney

NSW 2006 Australia

FAX: 612 351 4119

Tel: 612 351 2438

email: djp@bio.usyd.edu.au