Iron Hydrogenases and the Evolution of Anaerobic Eukaryotes

David S. Horner, Peter G. Foster, and T. Martin Embley

Department of Zoology, Natural History Museum, London, England

Hydrogenases, oxygen-sensitive enzymes that can make hydrogen gas, are key to the function of hydrogen-producing organelles (hydrogenosomes), which occur in anaerobic protozoa scattered throughout the eukaryotic tree. Hydrogenases also play a central role in the hydrogen and syntrophic hypotheses for eukaryogenesis. Here, we show that sequences related to iron-only hydrogenases ([Fe] hydrogenases) are more widely distributed among eukaryotes than reports of hydrogen production have suggested. Genes encoding small proteins which contain conserved structural features unique to [Fe] hydrogenases were identified on all well-surveyed aerobic eukaryote genomes. Longer sequences encoding [Fe] hydrogenases also occur in the anaerobic eukaryotes *Entamoeba histolytica* and *Spironucleus barkhanus,* both of which lack hydrogenosomes. We also identified a new [Fe] hydrogenase sequence from *Trichomonas vaginalis,* bringing the total of [Fe] hydrogenases reported for this organism to three, all of which may function within its hydrogenosomes. Phylogenetic analysis and hypothesis testing using likelihood ratio tests and parametric bootstrapping suggest that the [Fe] hydrogenases in anaerobic eukaryotes are not monophyletic. Iron-only hydrogenases from *Entamoeba*, *Spironucleus,* and *Trichomonas* are plausibly monophyletic, consistent with the hypothesis that a gene for [Fe] hydrogenase was already present on the genome of the common, perhaps also anaerobic, ancestor of these phylogenetically distinct eukaryotes. Trees where the [Fe] hydrogenase from the hydrogenosomal ciliate *Nyctotherus* was constrained to be monophyletic with the other eukaryote sequences were rejected using a likelihood ratio test of monophyly. In most analyses, the *Nyctotherus* sequence formed a sister group with a [Fe] hydrogenase on the genome of the eubacterium *Desulfovibrio vulgaris.* Thus, it is possible that *Nyctotherus* obtained its hydrogenosomal [Fe] hydrogenase from a different source from *Trichomonas* for its hydrogenosomes. We find no support for the hypothesis that components of the *Nyctotherus* [Fe] hydrogenase fusion protein derive from the mitochondrial respiratory chain.

Introduction

Reports of eukaryotic hydrogen production are limited to the plastids of some green algae (Happe and Naber 1993; Happe, Mosler, and Naber 1994) and to organelles found in certain protists which live in low-oxygen habitats (Müller 1993). These anaerobic, or, perhaps more accurately in some cases, microaerophilic, protists (for simplicity, we use the term ''anaerobic'' here throughout, but see Lloyd [1996] for a discussion of this point) include parabasalids such as *Trichomonas vaginalis,* some ciliates, and some chytrid fungi. These protists lack functional mitochondria but contain specialized redox organelles called hydrogenosomes which metabolize pyruvate using pyruvate : ferredoxin oxidoreductase (PFO), and produce hydrogen through the activity of hydrogenase (Steinbuchel and Müller 1986; Müller 1993). Hydrogenosome-containing protists do not form a monophyletic group but constitute separate lineages, some of which (e.g.,ciliates and chytrid fungi) are embedded among aerobic mitochondria-containing eukaryotes (Dore and Stahl 1991; Embley et al. 1995). Thus, hydrogenosomes in contemporary protists have been derived independently and it is this extraordinary capacity of eukaryotes to repeatedly evolve hydrogen-producing organelles which needs to be explained.

The issue of hydrogenosome evolution can be broken into two parts: one relating to the origin(s) of the organelle (here meaning the cell compartment) and one relating to the origin(s) of its distinctive biochemistry. As discussed previously (Müller 1993; Embley, Horner, and Hirt 1997), it is not necessary that all taxa have evolved hydrogenosomes in the same way, nor that organelle and biochemistry share a common origin. However, published data for hydrogenosomes in trichomonads, ciliates, and chytrid fungi are consistent with them sharing common ancestry with mitochondria (Benchimol, Durand, and Almeida 1997; Bradley et al. 1997; Embley, Horner, and Hirt 1997; van der Giezen et al. 1997*a*, 1997*b*, 1998; Akhmanova et al. 1998; Plümper, Bradley, and Johnson 1998). This immediately poses the question of the origins of the hydrogenosomal enzymes, particularly PFO and hydrogenase.

A number of alternative hypotheses for the origins of hydrogenosomal biochemistry have been proposed. For example, in the hydrogen hypothesis for the origin of eukaryotic cells, the common eubacterial ancestor of mitochondria and hydrogenosomes brought with it all of the components of ''hydrogenosomal'' metabolism, including PFO and hydrogenase (Martin and Müller 1998). The hydrogen hypothesis can thus be interpreted to infer a single common origin from an α-proteobacterium for both of these enzymes in different eukaryotes. In contrast, the syntrophic hypothesis posits an ancestral δ-proteobacterium as the hydrogen-evolving partner involved in eukaryogenesis (Moreira and Lopez-Garcia 1998). If the genes for these proteins were transferred to the host genome and retained after the initial selective pressure for hydrogen production was removed, then we could in principle test these inferences through phylogenetic analysis of PFO and hydrogenases from contem-

Abbreviations: HC, hydrogenase catalytic FeS cluster; PFO, pyruvate : ferredoxin oxidoreductase.

Key words: hydrogenase, hydrogenosomes, eukaryote origins, *Entamoeba, Spironucleus, Trichomonas.*

Address for correspondence and reprints: T. Martin Embley, Department of Zoology, Natural History Museum, London SW7 5BD, United Kingdom. E-mail: tme@nhm.ac.uk.

Mol. Biol. Evol. 17(11):1695–1709. 2000

porary eukaryotes (Embley and Martin 1998). Thus, in the simplest formulation of this scenario, each gene should trace its ancestry to the same single source related to either the α- or the δ-proteobacteria. Interestingly, recent analyses suggest that PFO from phylogenetically diverse anaerobic protists is monophyletic and was therefore likely present in the common ancestor of eukaryotes, having been retained in some lineages since then (Horner, Hirt, and Embley 1999). However, these same analyses were unable to infer a specific eubacterial sister group for eukaryotic PFO, and thus they provide no support for the predicted relationships of either the hydrogen or the syntrophic hypothesis for this particular enzyme.

Shortly after hydrogenosomes were discovered by Lindmark and Müller (1973), it was suggested that, because of biochemical similarities in hydrogen production, *Trichomonas* hydrogenosomes might have derived from strictly anaerobic endosymbiotic clostridia-like prokaryotes (Whatley, John, and Whatley 1979; Müller 1980). However, recent data are more consistent with the hydrogenosome organelle in *Trichomonas* sharing a common origin with mitochondria (Bui, Bradley, and Johnson 1996; Germot, Philippe, and Guyader 1996; Horner et al. 1996; Roger, Clark, and Doolittle 1996; Bradley et al. 1997; Plümper, Bradley, and Johnson 1998). Nonetheless, as mentioned above, a common origin with mitochondria need not preclude horizontal gene transfers from clostridia or other prokaryote(s) as a source of some hydrogenosome enzymes. Analyses of PFO do not provide support for a clostridial ancestry for eukaryote enzymes (Rosenthal et al. 1997; Horner, Hirt, and Embley 1999). However, two genes encoding iron-only ([Fe]) hydrogenases of about 50 kDa were recently reported for *T. vaginalis* (Bui and Johnson 1996), and these are similar to [Fe] hydrogenases found typically in obligate anaerobic prokaryotes including clostridia and sulphate-reducing bacteria. Interestingly, there is also a report in the literature describing a partially purified 64-kDa [Fe] hydrogenase from *T. vaginalis* (Payne, Chapman, and Cammack 1993).

The other eukaryote hydrogenase for which there is currently a complete gene sequence comes from *Nyctotherus ovalis,* an anaerobic ciliate which contains hydrogenosomes (Akhmanova et al. 1998). The gene encodes a putative fusion protein comprising a complete coding sequence for a 64-kDa [Fe] hydrogenase joined at its carboxyl terminus to two domains which share sequence similarity with the 24-kDa and 51-kDa subunits of mitochondrial complex I, respectively (Akhmanova et al. 1998). The addition of this carboxyl terminus may allow the fusion protein to use NAD(P) as an electron carrier (Akhmanova et al. 1998) rather than ferredoxins, as reported for the *Trichomonas* [Fe] hydrogenase(s) (Müller 1993). Akhmanova et al. (1998) listed a number of possible evolutionary origins for the *Nyctotherus* hydrogenase gene or its component parts, including lateral transfer from unknown prokaryotes and/or inheritance from the common ancestor of mitochondria and hydrogenosomes (Martin and Müller 1998). More recently, it has been suggested that the *Nyctotherus* hydrogenase was assembled from mitochondrial genes for complex I fused to a [Fe] hydrogenase gene (Andersson and Kurland 1999; Hackstein et al. 1999). However, neither the [Fe] hydrogenase part of the fusion protein nor its carboxyl terminus was subjected to phylogenetic analysis.

Hydrogenases are a prerequisite for hydrogenosome function in diverse anaerobic eukaryotes, and they also play a central role in recent hypotheses concerning the origins of the eukaryotic cell. Here, we present a complete coding sequence for a [Fe] hydrogenase for the diplomonad *Spironucleus barkhanus,* and a nearly complete sequence for a [Fe] hydrogenase from *Entamoeba histolytica.* In addition, we isolated almost the entire coding sequence for another [Fe] hydrogenase from *T. vaginalis.* Phylogenetic analysis, and hypothesis testing using likelihood ratio tests and parametric bootstrapping, suggest that the [Fe] hydrogenases in anaerobic eukaryotes are not monophyletic to the exclusion of bacterial sequences. Iron-only hydrogenases from *Entamoeba, Spironucleus,* and *Trichomonas* are plausibly monophyletic, consistent with the hypothesis that a gene for [Fe] hydrogenase was present on the genome of the common ancestor of these diverse eukaryotes. The [Fe] hydrogenase from the hydrogenosomal ciliate *Nyctotherus* may have an origin different from that of the other eukaryotic hydrogenases. We also report the occurrence of sequences on the genomes of aerobic eukaryotes which encode small proteins of unknown function, but which contain sequence features otherwise found only in [Fe] hydrogenases (see also Barton and Worman 1999). The implications of these findings for the evolution of eukaryotes and of hydrogenosomes, along with their relevance to the hydrogen and syntrophic hypotheses for eukaryogenesis, are discussed.

Materials and Methods

Strains, Culture, and Nucleic Acid Extractions

Spironucleus barkhanus (ATCC strain 50380) (Sterud, Mo, and Poppe 1997) was grown following the ATCC protocol. Cells of *E. histolytica* (strain HM-1: IMSS) were a gift of Dr Graham Clark (London School of Hygiene and Tropical Medicine). Cells of *T. vaginalis* (clone G3) were provided by Prof. Graham Coombs (University of Glasgow, Scotland). DNA was extracted from *S. barkhanus* and *T. vaginalis* using standard protocols. Messenger RNA was extracted from *E. histolytica, S. barkhanus,* and *T. vaginalis* using the Dynabead mRNA system (Dynal).

Cloning of Genes Encoding Putative [Fe] Hydrogenases from *E. histolytica, S. barkhanus,* and *T. vaginalis*

As part of an Expressed Sequence Tag (EST) project on the diplomonad *S. barkhanus,* we characterized a cDNA with a high level of identity to characterized [Fe] hydrogenases. This clone was used as a probe to isolate a genomic DNA clone from a lambda DASHII library (Horner, Hirt, and Embley 1999). The entire cod-

ing region for this open reading frame was sequenced directly using primer walking.

An *E. histolytica* EST partial sequence (AB002772) which showed similarity to the *S. barkhanus* hydrogenase was recovered from GenBank. We designed PCR primers EhisthydF (5′-ACATACAGTCACCGGACATGACCATAA-3′) and EhisthydR (5′-GGTGGAATTTTTGGTGGTTGTTTTACTA-3′) to amplify the entire 3′ and most of the 5′ region of the expressed sequence, respectively, in a rapid amplification of cDNA ends (RACE) experiment (Marathon System, Clontech).

Following up the report of a partially purified 64-kDa [Fe] hydrogenase from *T. vaginalis* (Payne, Chapman, and Cammack 1993), we designed primers to isolate the putative gene sequence. Published gene sequences for 64-kDa [Fe] hydrogenases from eubacteria were used to design PCR primers Tvhyd 45F (5′-TGYAGRGTYTGYCTYGTYGA-3′) and Tvhyd 340R (5′-CCARGCTGGRCANRWTGTRAACATTGG-3′), which incorporate known *T. vaginalis* codon usage bias (McInerney 1997). A fragment of approximately 900 bp was amplified from *T. vaginalis* DNA, and a RACE strategy was used to amplify the entire 3′ end and most of the 5′ part of the gene.

Identification of Database Sequences that Share Significant Sequence Similarity with [Fe] Hydrogenases

Fungal sequences related to [Fe] hydrogenases were recovered through BLASTP similarity searches of GenBank. Related ESTs from *Danio rerio, Drosophila melanogaster,* and *Homo sapiens* were identified through tBLASTn searches of the database of ESTs. Human genomic DNA encoding a related gene was identified though a BLAST search of the TIGR tentative human consensus database (http://www.ncbi.nlm.nih.gov/BLAST/thcblast.html). *Caenorhabditis elegans* genomic sequences were detected by tBLASTn searches at the Sanger Centre (www.sanger.ac.uk/Projects/C_elegans/blast_servers.html). *Arabidopsis thaliana* sequences were detected through tBLASTn searches at the *A. thaliana* database in Stanford, Calif. (genome-www2.stanford.edu/cgi-bin/AtDB/nph-blast2atdb). We also searched finished and unfinished prokaryote and eukaryote genome projects to identify [Fe] hydrogenase coding sequences, using prokaryote and eukaryote hydrogenase sequences as probes and the BLASTn, t-BLASTn, and iterative (PSI) BLAST programs of the BLAST2 package (Altschul et al. 1997).

Alignment of Inferred Protein Sequences and DNA Sequences

Inferred [Fe] hydrogenase amino acid sequences were aligned using CLUSTAL W (Thompson, Higgins, and Gibson 1994) and adjusted manually using GDE, version 2.2 (Maidak et al. 1996). DNA-coding regions were back-aligned to the protein alignment using PUTGAPS (J. O. McInerney, Natural History Museum, London). We excluded regions which could not be unambiguously aligned, leaving 270 aligned amino acid positions (18 sequences) for phylogenetic analyses. A second alignment accommodating the homologous sequences detected in human and yeast (102 aligned amino acids, 20 taxa) was also prepared but was considered too short for phylogenetic analysis. Alignments are available from D.S.H. (e-mail dsh@nhm.ac.uk).

Phylogenetic Analyses of [Fe] Hydrogenase Protein Sequences

To investigate potential saturation of [Fe] hydrogenase protein sequences, the numbers of inferred substitutions were calculated from a maximum-parsimony tree for each pairwise comparison and plotted against observed distances using the MUST package (Philippe 1993).

Maximum-likelihood (ML) analyses of protein alignments were performed with PROTML in MOLPHY, version 2.3 (Adachi and Hasegawa 1996), using the heuristic quick-add OTU search and local rearrangement (R-search) methods with the JTT-F amino acid replacement model. However, the ML models in PROTML make the assumption that all sites are equally free to vary, and when this assumption is violated, the wrong tree may be selected (Lockhart et al. 1996; Hirt et al. 1999). The [Fe] hydrogenase sequences contain sites which are of unvaried amino acid composition across the alignment and are thus potentially unable to vary (i.e., invariable). We therefore estimated the proportion of unvaried sites, inferred to be invariable (95%), using a two-rate model in PUZZLE (Strimmer and von Haeseler 1996) and removed them prior to phylogenetic analyses.

Phylogenetic Analyses of [Fe] Hydrogenase DNA Sequences

Potential mutational saturation of [Fe] hydrogenase DNA sequences and nucleotide base composition heterogeneities between sequences were investigated using plots of transitions versus transversions and ‘‘GC trees’’ as described previously (Lockhart et al. 1994; Lento et al. 1995; Charleston 1998; Hirt et al. 1999). Codon position 3 was potentially saturated for all comparisons and manifested large base composition variation between sequences (from 10% G+C for *Clostridium perfringens* to 94% G+C for *Desulfovibrio fructosovorans* cytochrome hydrogenase), so we excluded it from phylogenetic analyses. Nucleotide compositional heterogeneity at variable positions was observed for codon position 1 (from 44% G+C for *C. acetobutylicum* hydrogenase B to 68% G+C for *D. vulgaris* cytochrome hydrogenase) and position 2 (from 37% G+C for *Thermotoga maritima* B to 50% G+C for *D. vulgaris* cytochrome hydrogenase). In nucleotide composition ‘‘GC trees’’ (Lockhart et al. 1994) for codon positions 1 + 2, the eukaryote sequences were not monophyletic.

Because we observed base composition heterogeneity at codon positions 1 and 2, we used LogDet/Paralinear distances, a phylogenetic method in PAUP*, version 4.0b4a (Swofford 1998), which performs well in

simulations even when base composition varies across the tree (Lake 1994; Lockhart et al. 1994). However, the LogDet/Paralinear distances method does not incorporate a correction for site-by-site rate variation, and it assumes that all sites are free to vary. We therefore used PAUP*, version 4.0b4a, to estimate the fraction of sites free to vary across our alignment using a variable/invariable-sites ML model and restricted LogDet/Paralinear distances analyses to only these estimated variable positions (427 sites for codon positions 1 + 2). We also used an ML method to investigate relationships. For these analyses, we used a general time reversible (GTR) model (Yang 1994) with site rate heterogeneity modeled using a discrete four-category gamma distribution plus invariable sites (henceforth, GTR+Γ+I) (Yang 1996). Values for all model parameters were estimated using ML over the LogDet/Paralinear distances tree and then re-estimated over the resulting ML tree. The analysis was then repeated until no further improvement in likelihood was observed. Bootstrapping used PAUP*, version 4.0b4a, with either 100 (ML) or 1,000 (LogDet/Paralinear distances) replicates.

Investigation of Relationships Among Eukaryotic [Fe] Hydrogenases Using Parametric Bootstrapping and a Likelihood Ratio Test of Monophyly

We used a likelihood ratio test proposed by Huelsenbeck, Hillis, and Nielsen (1996) to examine relative support for monophyly of eukaryotic [Fe] hydrogenase DNA sequences. This test asks whether the difference in likelihood between the best trees estimated with and without the constraint of monophyly is significant. In the present case, the null hypothesis (H_0) is monophyly of eukaryotic enzymes and the alternative, less-constrained, hypothesis (H_1) is the best ML tree. The test statistic (δ) is $\delta = \ln L_1 - \ln L_0$, where L_1 is the likelihood of the best tree and $\ln L_0$ is the likelihood of the best monophyly tree. The null distribution of δ must be generated by simulation (parametric bootstrapping) under an appropriate model. Here, it is important to note that parametric bootstrapping may not perform well for approximating a distribution if the model of DNA substitution does not match the processes which generated the original sequences (Huelsenbeck, Hillis, and Nielsen 1996). Accordingly, we used a likelihood ratio test suggested by Goldman (1993) to first evaluate if the GTR+Γ+I model and the tree could have plausibly given rise to the sequences we observe (see *Results*). Custom software (available on request from P.G.F.) was used for all simulations and likelihood calculations.

For the monophyly test, the distribution of the test statistic δ was estimated by parametric bootstrapping, by repeatedly simulating sequences based on the best monophyly tree and model. With each simulated data set, two searches were made using PAUP*, version 4.0b4a, one with the eukaryote monophyly constraint and one without that constraint. Each search began with a parsimony search to obtain a tree upon which to optimize parameters for the GTR+Γ+I model. Using these parameters, a search using ML was made with the parsimony tree as a starting tree for tree bisection-reconnection (TBR) swapping. Parameters were reoptimized on the best tree from that search, and using these new parameters, a second round of swapping was done using subtree pruning-regrafting (SPR, with a reconnection limit of two nodes). After a final parameter optimization, the likelihood of the best tree from each search was used to calculate the simulation statistic.

Phylogenetic Analysis of the Carboxyl Terminus of the *Nyctotherus* Fusion Protein DNA and Inferred Protein Sequences

An amino acid alignment of mitochondrial complex I 51-kDa components and their prokaryote homologs (bacterial NADH dehydrogenase NuoF subunits and soluble hydrogenase components) was prepared as described above. Regions which could not be unambiguously aligned were excluded from phylogenetic analysis, as were inferred invariable sites. ML analysis was performed as described for a data set of 25 taxa and 145 variable positions.

For DNA analyses, we used the same strategy involving LogDet/Paralinear distances analyses and ML analyses as described above for the [Fe] hydrogenase sequences. For these analyses, we excluded codon position 3, leaving 370 codon 1 + 2 positions, of which 98 were constant across the alignment. ML analyses indicated that approximately 3% of the latter were potentially variable, so this proportion of observed constant sites was included in our analyses of variable positions (a total of 275 sequence positions).

Results

A 64-kDa [Fe] Hydrogenase from *T. vaginalis*

Two closely related [Fe] hydrogenase genes, TvhydA and TvhydB, have previously been characterized from *T. vaginalis* (Bui and Johnson 1996). Antisera raised against the protein of one of these was used to show the presence of a 50-kDa band in isolated hydrogenosomes of *T. vaginalis,* corresponding to the expected size of the products of the characterized genes (Bui and Johnson 1996). Here, we isolated a gene encoding a putative 64-kDa-like [Fe] hydrogenase from *T. vaginalis.* Conceptual translation of the new sequence revealed overall colinearity (fig. 1) with [Fe] hydrogenases from *Clostridium*, *D. fructosovorans*, *T. maritima*, and the [Fe] hydrogenase module of the fusion protein from the anaerobic ciliate *N. ovalis* (Akhmanova et al. 1998). The new *T. vaginalis* sequence is 47% identical to TvhydA in the overlapping region (450 amino acids) and, overall, shares 37% identity with *Clostridium pasteurianum* hydrogenase 1 and *D. fructosovorans* putative NADP reducing hydrogenase. With respect to the bacterial [Fe] hydrogenase sequences, the new *T. vaginalis* gene contains an N-terminal extension (lacking a putative initiator codon in our RACE clones) which is rich in alanine and serine and contains the 2-aa motif RN at positions 14–15. This sequence is similar to characterized *Trichomonas* hydrogenosome targeting motifs (table 1) (Plümper, Bradley, and Johnson 1998).

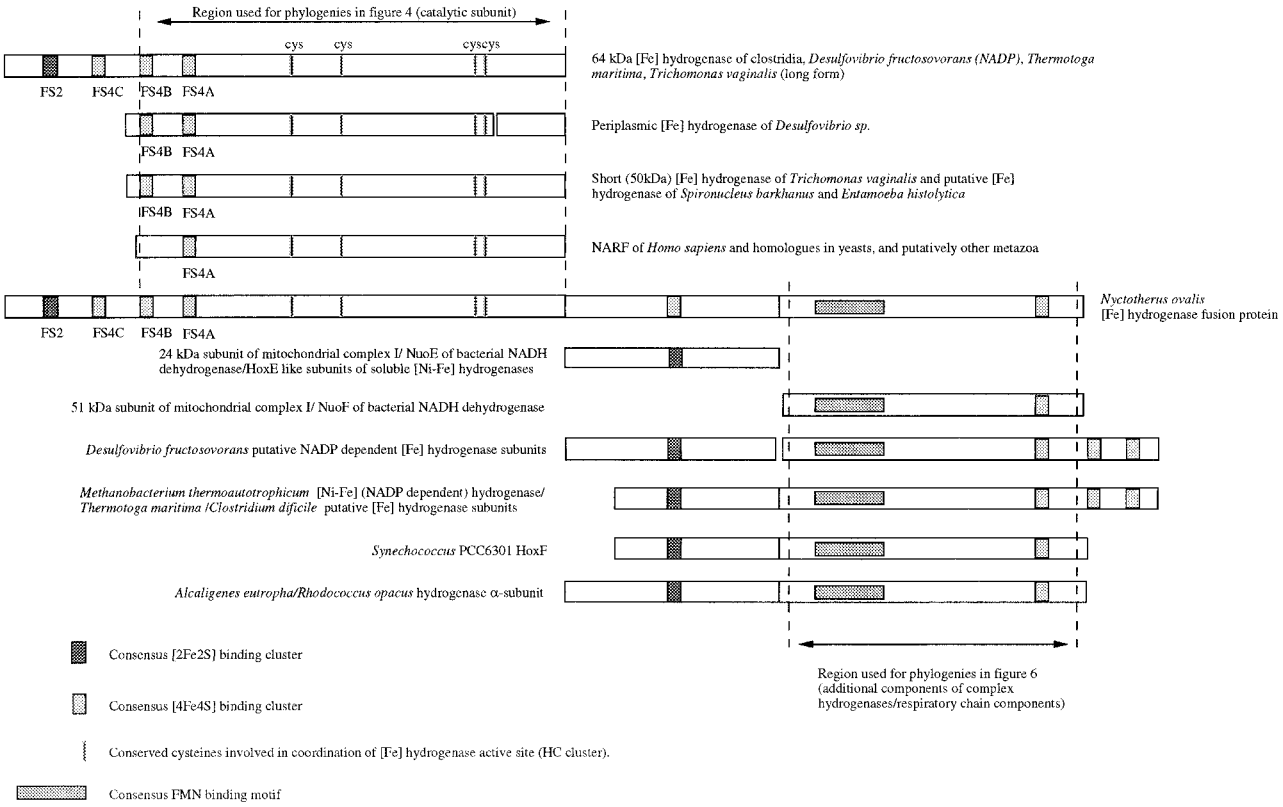

FIG. 1.—Schematic representation of [Fe] hydrogenase subunits.

Putative Cytosolic [Fe] Hydrogenases from *E. histolytica* and *S. barkhanus*

The *S. barkhanus* genomic clone encoded a putative [Fe] hydrogenase of 468 amino acids, which is similar in size to the TvhydA gene, with which it shares 32% identity at the amino acid level. Conceptual translation of the *S. barkhanus* sequence revealed several potential internal stop codons (TAA, TAG) which mainly corresponded to conserved glutamines in the alignment. This use of a noncanonical code has been reported previously for *S. barkhanus* (Keeling and Doolittle 1997).

We recovered a weak match from an EST survey of *E. histolytica* (Tanaka, Tanaka, and Mitsui 1997) to the new *Spironucleus* sequence. This clone (AB002772) had been designated as an unusual ferredoxin and contains conserved cysteines corresponding to potential [4Fe-4S] clusters. Alignment of the conceptually translated product (466 amino acids) from the almost-full-length sequence revealed the highest level of identity (35%) to the *S. barkhanus* [Fe] hydrogenase (fig. 1).

Sequences Related to [Fe] Hydrogenases on the Genomes of ''Higher'' Eukaryotes

Searches of GenBank using the *C. pasteurianum* hydrogenase 1 amino acid sequence as a probe revealed significant matches (S = 99–115, E = $5e^{-20}$–$5e^{-24}$, 23%–28% identity over a 300-aa overlap) on hypothetical open reading frames from *Saccharomyces cerevisiae* (YNYO_YEAST accession number P23503), *Shizosaccharomyces pombe* (SPCC1450.10c accession

Table 1
Comparison of N-Terminal Region of the Novel *Trichomonas vaginalis* [Fe] Hydrogenase with N-Terminal Presequences of a Selection of Hydrogenosomal Proteins

Hydrogenosomal Protein	Presequence															Start			
[Fe] hydrogenase 64 kDa[a] . .	?	A	S	T	G	I	N	S	T	A	N	I	L	R	N	I	T	V	T
Malic enzyme 1	M				L	T	S	S	V	S	V	P	V	R	N	I	C	R	A
Malic enzyme 2	M				L	T	S	V	S	Y	P	P	V	R	N	I	C	R	S
Adenylate kinase	M								L	S	G	V	S	R	N	A	A	R	T
SCS alpha subunit	M							L	A	G	D	F	S	R	N	L	K	Q	P
SCS beta subunit	M					L	S	A	S	S	N	F	A	R	N	F	N	I	L

NOTE.—Presequences are sequences encoded by the gene but absent from the N-terminus upon N-terminal sequencing of the mature protein isolated from hydrogenosomes. Start indicates the first four amino acids of the mature protein. SCS = succinyl CoA synthetase. The table was developed from Plümper, Bradley, and Johnson (1998).

[a] Putative presequences of the *T. vaginalis* [Fe] hydrogenase presented in this work, as suggested by comparison with bacterial hydrogenase sequences and empirically determined *T. vaginalis* hydrogenosome presequences.

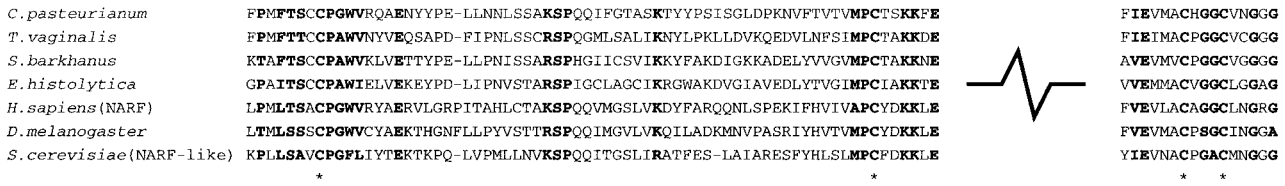

FIG. 2.—Alignment of putative HC cluster coordinating region for the crystallized *Clostridium pasteurianum* [Fe] hydrogenase 1 and the novel [Fe] hydrogenase homologs presented here. Conserved residues are shown in bold text. Cysteines implicated in HC coordination are each indicated by an asterisk.

number CAB40177), and *Kluveromyces lactis* (LET1 accession number X70373) and *H. sapiens* (*Narf* accession number AF128406). The encoded proteins are broadly colinear with the TvhydA gene product, as well as with the [Fe] hydrogenase sequences detected in *S. barkhanus* and *E. histolytica* (fig. 1). Like these proteins, the fungal and human gene products lack 120–140 amino acids with respect to the N-terminal region of many bacterial [Fe] hydrogenases, the *N. ovalis* [Fe] hydrogenase, and the new *T. vaginalis* [Fe] hydrogenase reported here. In reciprocated BLAST searches, these proteins match only themselves and [Fe] hydrogenases. We also recovered highly significant matches from *D. melanogaster* (EST accession number AA439709) and *D. rerio* cDNA (accession number AA495126). Closely related sequences were also detected on the genomes of *C. elegans* (clone Y54H5A) and *A. thaliana* (CAB10420). While none of these sequences represent complete open reading frames, the finding that most are expressed is consistent with the notion that they represent parts of functional genes.

To gain some insight into the possible functionality of our new sequences and those from the databases, we investigated the distribution of residues which are considered important for [Fe] hydrogenases based on the crystal structure of the *C. pasteurianum* [Fe] hydrogenase 1 and the *D. desulfuricans* periplasmic [Fe] hydrogenase (Peters et al. 1998; Nicolet et al. 1999). The new *T. vaginalis* [Fe] hydrogenase contains all of the [Fe-S] center coordinating residues, including those implicated in the coordination of FS2, FS4c, FS4b, FS4a, and the hydrogenase catalytic cluster (HC) (fig. 1), which have been demonstrated for the *C. pasteurianum* hydrogenase 1. The *S. barkhanus* and *E. histolytica* homologs, along with TvhydA and TvhydB and the dimeric hydrogenase of *Desulfovibrio* sp., share conserved cysteines implicated in FS4b, FS4a, and HC coordination, while the fungal, human, and *Drosophila* sequences specify the expected arrangement of cysteines for coordination of FS4a and the HC cluster, which is otherwise unique to [Fe] hydrogenases (figs. 1 and 2).

Phylogeny of [Fe] Hydrogenases

Since some of the analyzed sequences are very divergent, we investigated the level of potential saturation and other properties of the data prior to phylogenetic analysis. For the final alignment, pairwise identities at the amino acid level ranged from 35% (*C. pasteurianum* vs. *E. histolytica*) to 82% (*C. acetobutylicum* A vs. *C. perfringens*). Transition-versus-transversion plots for hydrogenase family DNA sequences revealed points resolved as distinct clusters (fig. 3*A*). Comparisons involving eukaryote sequences formed a cluster outside the range of distribution of most intrabacterial comparisons. Furthermore, the coefficient of regression for a linear distribution of all points was low ($r^2 = 0.580$). These observations suggest that there may be differences in the mode/rate of evolution of the different sequences and/or mutational saturation. To further investigate the latter possibility (Philippe 1993), we made ''saturation'' plots (Philippe and Adoutte 1998) of observed versus inferred substitutions (fig. 3*B*) for all pairwise comparisons of amino acid sequences (inferred substitutions were calculated from a maximum-parsimony tree using the TREEPLOT program in the MUST package (Philippe 1993). With the exception of some comparisons between closely related sequences, the ratio of inferred to observed substitutions was between 1.5 and 2.5, suggesting that mutational saturation may be a problem for some comparisons, particularly distant ones, within this data set. As these figures are derived from a maximum-parsimony tree, they likely represent a conservative estimate of the degree of mutational saturation for these sequences.

To investigate the relationships between the [Fe] hydrogenases, we constructed phylogenetic trees from protein and DNA data using the ML and LogDet/Paralinear distances (for DNA only) methods. As might be predicted from the ''saturation'' plot, there was little apparent signal from these data to resolve relationships between clades. Figure 4*A* shows the ML tree obtained for DNA, and figure 4*B* shows the LogDet/Paralinear distances DNA tree. The ML tree for proteins gave the same branching pattern as the DNA ML tree. The eukaryotic sequences did not form a monophyletic group in any of these analyses. Examination of the bootstrap partitions for DNA ML and LogDet/Paralinear distances revealed that monophyly of eukaryote sequences occurred in only about 2% of bootstrap trees. Thus, there is little support from this data set for monophyly of all eukaryotic [Fe] hydrogenases.

A relationship between the novel short form [Fe] hydrogenases from *Spironucleus* and *Entamoeba* was always strongly supported by our analyses. The long- and short-form *Trichomonas* [Fe] hydrogenases also formed a monophyletic group in most analyses. These four sequences together formed a monophyletic group in the ML tree (44% bootstrap support) but generally did not form a monophyletic group in analyses of LogDet/Paralinear distances (only 11% bootstrap support). In these

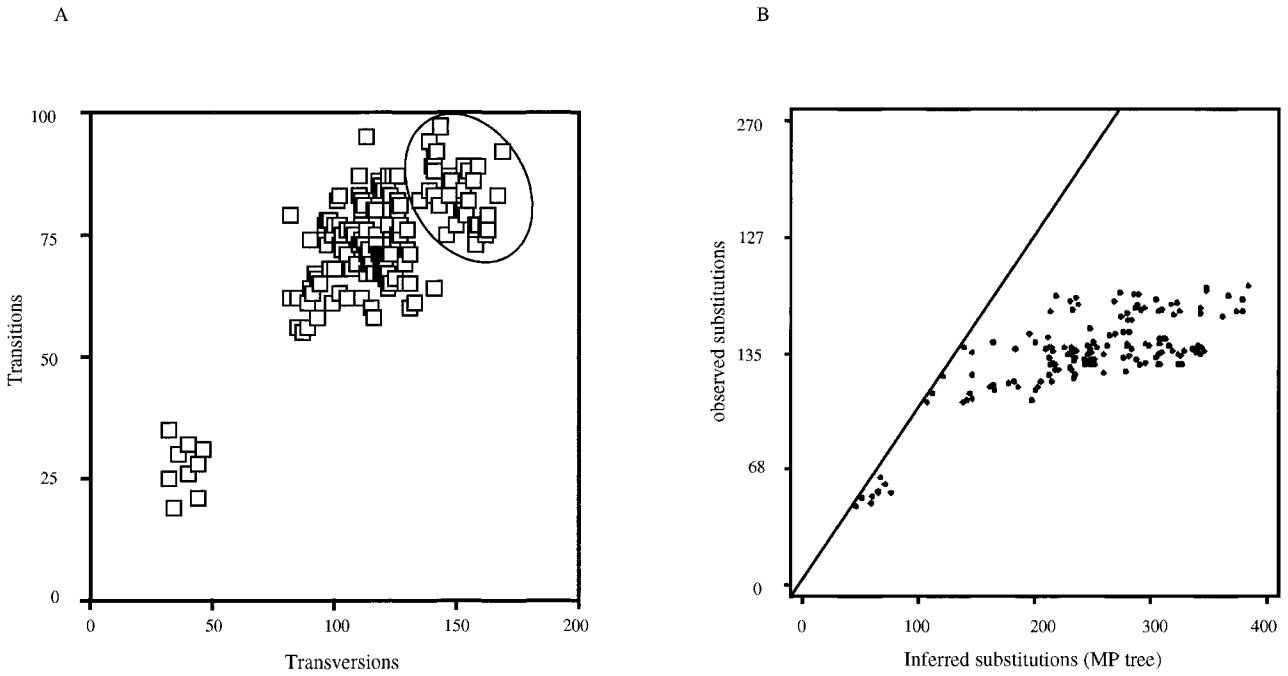

FIG. 3.—*A,* Plot of observed transitions versus transversions for DNA codon positions 1 and 2 (17 taxa, 540 characters). Extensive clustering of points is evident. Comparisons involving eukaryote taxa are circled and fall outside the range of comparisons for other taxa. *B,* Plot of inferred versus observed substitutions by pairwise comparison for the 17-taxon 229-site hydrogenase amino acid data set. Most comparisons fall to the right of the line of parity between observed and inferred substitutions, suggesting a significant degree of mutational saturation.

latter analyses, the two pairs of sequences formed weakly supported (23% and 28% by bootstrapping) sister groups with different prokaryote sequences (fig. 4*B*).

The favored position for the *Nyctotherus* sequence in all of our analyses was as the sister group to the *D. vulgaris* [Fe] hydrogenase gamma. In LogDet/Paralinear distances bootstrap analyses, 58% of trees contained this relationship, while it was found in 48% of DNA ML bootstrap trees. This compares with less than 2% of bootstrap trees, for either analysis, where the *Nyctotherus* sequence clustered with any of the other eukaryote sequences.

The slightly higher support obtained using LogDet/Paralinear distances for the relationship between the *Nyctotherus* [Fe] hydrogenase and the *D. vulgaris* [Fe] hydrogenase gamma may be due in part to the ability of this method to better deal with changing nucleotide compositions across the tree (Lake 1994; Lockhart et al. 1994). This can be illustrated by examining the other bootstrap partitions (data not shown) involving the *D. vulgaris* [Fe] hydrogenase gamma sequence. A relationship between this sequence and the *D. fructosovorans* NAD reducing hydrogenase was the next best supported partition for the *D. vulgaris* hydrogenase gamma sequence in both bootstrap analyses (ML, 31% of trees; LogDet/Paralinear distances, 20% of trees). However, comparison of base compositions revealed that these two sequences also have the highest G+C% compositions (61% and 54% GC at variable positions, respectively) of the sequences analyzed. Furthermore, in ''GC trees'' based solely on the nucleotide composition of variable sites, these two sequences clustered together. This might suggest that part of the reason why these two sequences come together in our phylogenetic analyses is because they share a common compositional bias to which LogDet/Paralinear distances are less susceptible and hence give lower support. In contrast, the nucleotide composition of the *Nyctotherus* sequence is much lower, at 44% G+C, and it does not cluster with the *D. vulgaris* hydrogenase gamma sequence in ''GC trees.'' Thus, positive support for this relationship in our phylogenetic analyses cannot be attributed to nucleotide composition bias, but the reduction in support in ML analyses might indeed be influenced by this problem.

Evaluation of the Relationships Between Eukaryotic [Fe] Hydrogenases Using Parametric Bootstrapping and a Likelihood Ratio Test of Monophyly

We used a likelihood ratio test (Goldman 1993) to evaluate whether the GTR+Γ+I model plus the best tree we found could have plausibly given rise to the observed data. The null hypothesis was that the data arose from the tree under the model. The alternative unconstrained hypothesis was that the data are not related by a tree or a model, i.e., assuming a multinomial distribution for the frequencies of site patterns. The test statistic was $L_1 - L_0$, where L_1 is the unconstrained likelihood and L_0 is the likelihood under the tree + model (Goldman 1993). When calculating the L_0 of simulated data, the branch lengths and model parameters were optimized. In order to calculate the unconstrained likelihood L_1, there can be no gaps or ambiguities in the alignment, and for this reason the two positions in the alignment which had gap characters were removed, leaving an alignment of 538 characters. Branch lengths

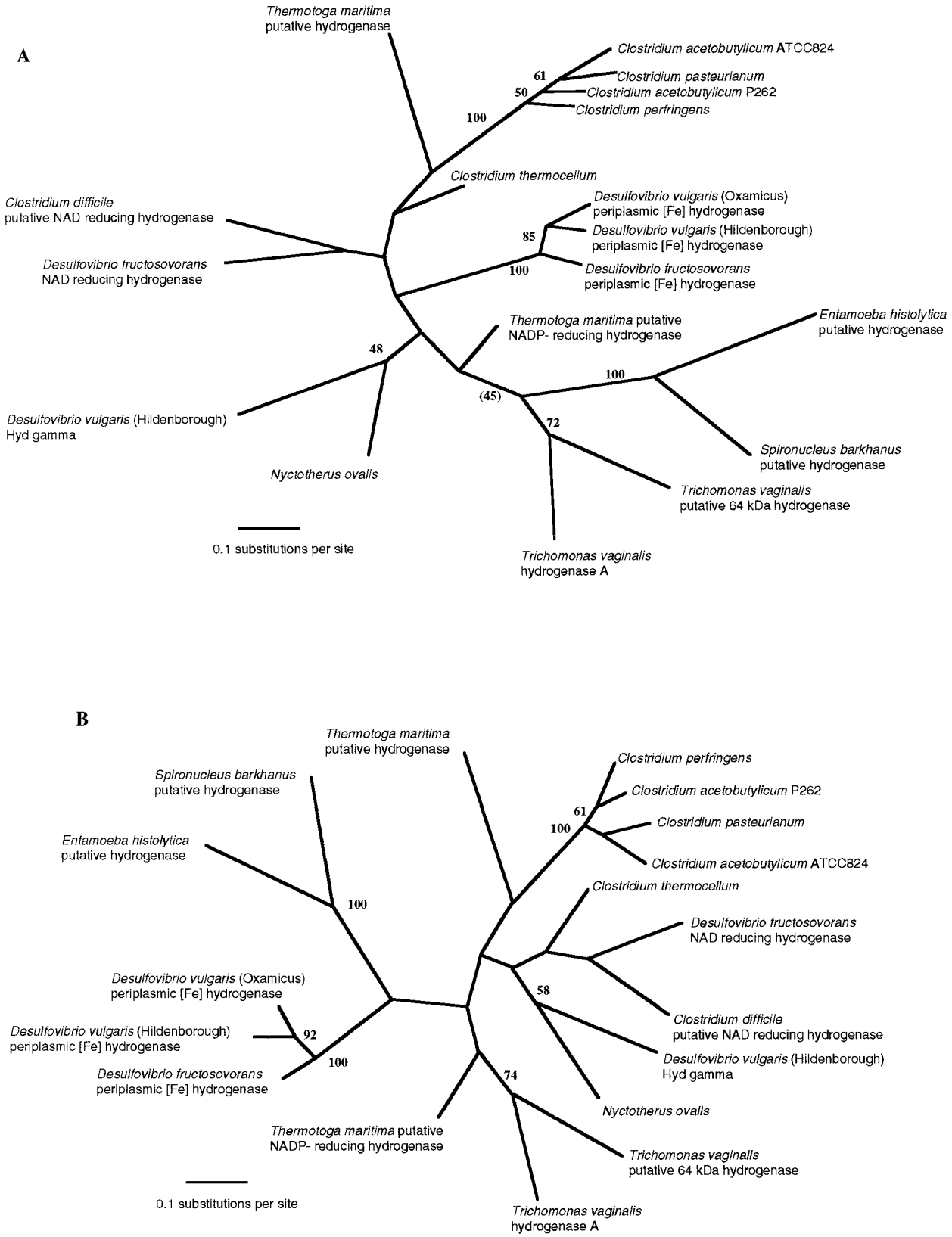

FIG. 4.—*A*, DNA maximum-likelihood tree with bootstrap support values (100 replicates). A general time reversible (GTR) model with site rate heterogeneity modeled using a discrete four-category gamma distribution (shape parameter 1.24) plus invariable sites (estimated proportion 0.11) was used. *B*, LogDet tree for variable DNA positions; bootstrap support values over 50% (1,000 replicates) are shown at relevant nodes.

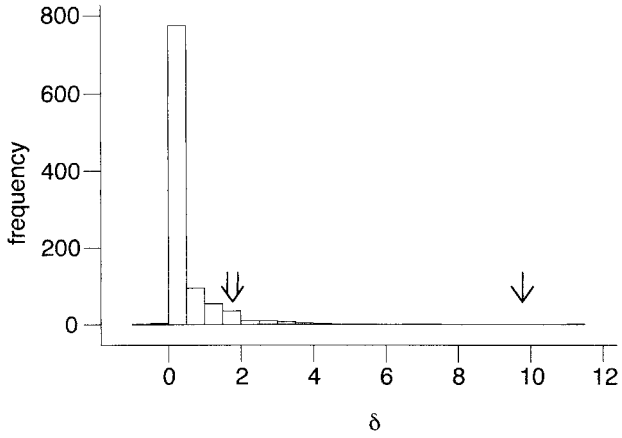

FIG. 5.—Monophyly statistic δ distribution from 1,000 simulations. Five hundred eighteen points were 0. The single arrow shows δ = 9.64 from the original data, and the double arrow shows the 95% interval of the simulation distribution. Monophyly is rejected. Four of the simulation analyses gave negative δ values. This suggests that the best tree constrained by monophyly was better than the best unconstrained tree. We interpret this impossible result as being a rare artifact of the search strategy employed. Had a more thorough, but more computationally expensive, search been done, it would be expected that such results would not occur. Indeed, in preliminary analyses using computationally less expensive search strategies, negative δ values were more of a problem.

and GTR+Γ+I model parameters were reoptimized with the new alignment of 538 characters. Data sets were first simulated on the ML tree + model, and the test statistic was calculated. The original data had a statistic of 4,083.39, compared with a critical 95% value estimated through simulation of 4,232.99. Therefore, there was no evidence to reject the null hypothesis, as the original data set is a plausible realization of the tree + model.

The likelihood ratio test of monophyly (Huelsenbeck, Hillis, and Nielsen 1996) was used to test whether we could reject the null hypothesis of monophyly of eukaryotic [Fe] hydrogenases. The test statistic (δ) was $\delta = \ln L_1 - \ln L_0$, and the null distribution of δ was generated by parametric bootstrapping using the GTR+Γ+I model and the best eukaryote monophyly tree. If the likelihood ratio statistic δ calculated from the real data is greater than the 95% confidence interval determined through simulation, the null hypothesis that the group in question is monophyletic can be rejected (Huelsenbeck, Hillis, and Nielsen 1996). Figure 5 shows the distribution of δ for 1,000 simulations. The observed δ for the original data was 9.64 and thus was well outside of the critical 95% value determined through parametric bootstrapping. Thus, monophyly of eukaryote [Fe] hydrogenase DNA sequences can be rejected using this test.

Does the Carboxyl Terminus of the *N. ovalis* [Fe] Hydrogenase Fusion Protein Comprise Components Derived from the Mitochondrial Respiratory Chain?

The N-terminal 600 amino acids of the *N. ovalis* fusion protein encode an [Fe] hydrogenase, but the carboxyl terminus of about 600 amino acids shares sequence similarity with NAD^+-utilizing components of mitochondrial complex I (24-kDa and 51-kDa components), bacterial NADH dehydrogenases (NuoE and NuoF), and some bacterial NADH–dependent [Ni-Fe] hydrogenases (various nomenclature) (fig. 1). Based on these similarities, it has been suggested that the *N. ovalis* [Fe] fusion protein contains components derived from the mitochondrial respiratory (complex I modules) chain (Andersson and Kurland 1999; Hackstein et al. 1999). This generates the testable prediction that the components of the carboxyl terminus of the *Nyctotherus* fusion protein should form a monophyletic group with mitochondrial complex I components.

The 24-kDa component of complex I/NuoE family of genes is poorly conserved and could not be reliably aligned for phylogenetic analysis. However, the 51-kDa component of mitochondrial complex I, bacterial NuoF gene products, and bacterial [Ni-Fe] hydrogenase components are highly conserved and easy to align. The 51-kDa component of mitochondrial complex I formed a strongly supported monophyletic group with the NuoF sequences from α-proteobacteria (fig. 6). This topology is consistent with other data which suggest that the mitochondrion endosymbiont was a member of the α-proteobacteria (Yang et al. 1985; Andersson et al. 1998). However, the 51-kDa-like domain of the *N. ovalis* fusion protein was never part of this clade in our analyses. Thus, there is no support from our analyses for the hypothesis that the carboxyl terminus of the *N. ovalis* fusion protein is derived from the mitochondrial respiratory chain. In LogDet/Paralinear distances (fig. 6) and protein ML trees, the 51-kDa-like domain of the *N. ovalis* fusion protein clustered with the carboxyl terminus components of NAD(H)-dependent [Ni-Fe] hydrogenases isolated from the β-proteobacterium *Alcaligenes eutrophus* and the Gram-positive actinomycete *Rhodococcus opacus.* In DNA ML analyses, the 51-kDa-like domain of the *N. ovalis* fusion protein clustered with the γ-proteobacteria *Escherichia coli* and *Salmonella typhimurium* in the ML tree ($\ln L = 5{,}413.6$). However, this tree was no better, considering the estimated error, than the best ML tree we found ($\ln L = 5{,}414.6 \pm 6.3$), which clustered it with *A. eutrophus* and *R. opacus.*

Discussion

All of the previously described eukaryote [Fe] hydrogenases appear to be localized within organelles (Happe, Mosler, and Naber 1994; Bui and Johnson 1996). The two novel [Fe] hydrogenase gene sequences from *E. histolytica* and *S. barkhanus* may thus be the first cytosolic eukaryote [Fe] hydrogenases to be discovered. The *Spironucleus* sequence is complete, and it shows no obvious motif at its N-terminus which might target it to an organelle. Furthermore, *Spironucleus* and its better-known relative *Giardia* are reported to lack typical eukaryotic organelles such as mitochondria and peroxisomes (e.g., Cavalier-Smith 1993). *Giardia lamblia* (Roger et al. 1998) and *S. barkhanus* (unpublished data) do contain a mitochondrial chaperonin 60, suggesting that they once had mitochondria, but there is no

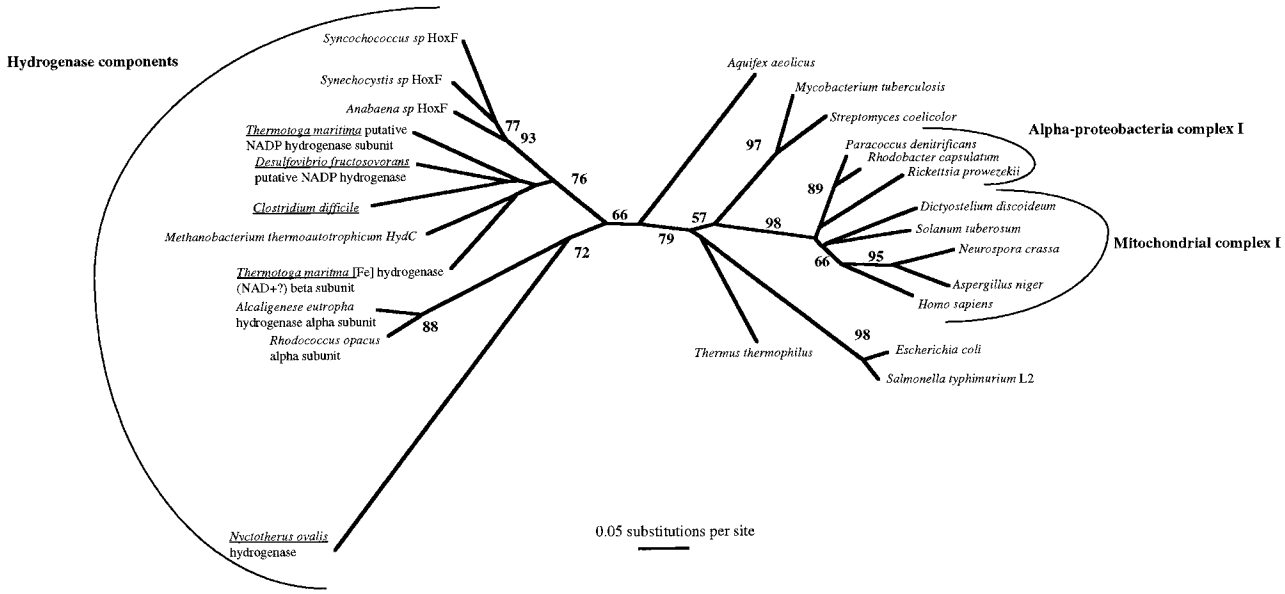

FIG. 6.—LogDet tree for variable DNA positions of respiratory chain complex I 51-kDa components (NuoF) and homologs in complex [Ni-Fe]- and [Fe]-only hydrogenases. Components implicated in the function of [Fe]-only hydrogenases are underlined. Bootstrap support values over 50% (1,000 replicates) are shown at relevant nodes.

evidence from antibody localization experiments in *Giardia* that this protein is now targeted to a particular membranous compartment (Soltys and Gupta 1994; Roger et al. 1998).

The *Entamoeba* sequence is from cDNA, and it is incomplete at its 5′ terminus, so we do not know whether it carries a targeting sequence or not. *E. histolytica* has recently been shown to contain an organelle, called a mitosome (Tovar, Fischer, and Clark 1999) or crypton (Mai et al. 1999), which likely shares common ancestry with the mitochondrion, since it contains a mitochondrial chaperonin 60, and it uses a mitochondrial-like cleavable N-terminal targeting motif in protein import (Clark and Roger 1995; Roger et al. 1998; Mai et al. 1999; Tovar, Fischer, and Clark 1999). Trichomonad hydrogenosomes use a similar import mechanism (Plümper, Bradley, and Johnson 1998), and there are strong similarities between the sequences of some trichomonad targeting motifs and those for mitosome/crypton proteins (Mai et al. 1999; Tovar, Fischer, and Clark 1999). Because of these similarities, it has been suggested that *Entamoeba* might once have contained a hydrogenosome-like organelle (Tovar, Fischer, and Clark 1999). Further work is required to investigate whether the novel [Fe] hydrogenase we have detected is targeted to the mitosome, but it may be relevant that in *E. histolytica*, PFO lacks an obvious targeting sequence and is apparently active in the cytoplasm (Takeuchi, Weinbach, and Diamond 1975; Reeves et al. 1977). In trichomonads, PFO is a key enzyme of the hydrogenosome organelle releasing reducing equivalents which are transferred to protons with the formation of hydrogen by hydrogenase (Müller 1993).

We also found evidence of short sequences (hereafter termed ''*Narf*-like,'' after the human homolog) related to [Fe] hydrogenases on all of the well-surveyed eukaryotic genomes (see also Barton and Worman 1999). The organisms comprise the fungi *K. lactis*, *S. cerevisiae*, and *S. pombe*, the metazoa *C. elegans*, *D. rerio*, *D. melanogaster*, and *H. sapiens*, and the plant *A. thaliana*. These sequences do not contain all of the residues which are thought to be important for the functioning of the *C. pasteurianum* [Fe] hydrogenase 1 (Peters et al. 1998), and they also lack one of the iron sulfur clusters (FS4B) which is found on the short-form [Fe] hydrogenases found in *Entamoeba*, *Spironucleus*, and *Trichomonas*. They do contain the expected arrangement of cysteines for coordination of FS4a and also for the HC cluster, which is otherwise unique to [Fe] hydrogenases (Peters et al. 1998). The paucity of complete *Narf*-like genes and their short length confounded our attempts to make phylogenetic inferences from these proteins. However, based on their similarities in structure, we speculate that the *Narf*-like genes are descended from an ancestral [Fe] hydrogenase which had lost some structural features, perhaps reducing its sensitivity to inactivation by oxygen, while retaining others.

It is clearly important to try and establish what the sequences related to [Fe] hydrogenases, which are found otherwise only in anaerobic prokaryotes and eukaryotes, are doing in such a diverse range of aerobic eukaryotes. The [Fe] hydrogenases from prokaryotes and eukaryotes which have been investigated are very oxygen-sensitive. For example, the [Fe] hydrogenase in the green alga *Scenedesmus* is only active after a period of anaerobic adaptation, and once exposed to aerobic conditions, hydrogen production rapidly stops (Urbig, Schulz, and Senger 1993; Appel and Schulz 1998). It is thus unlikely, given also the lack of reports of hydrogen production, that the new sequences encode proteins which are functioning as classic hydrogenases. The *H. sapiens* gene has already been the subject of investigation and apparently

encodes a protein which interacts with prenylated prelamin A (Barton and Worman 1999). The encoded protein was shown to localize to the nucleus of human cells and to interact with the C-terminal region of prenylated prelamin A, which is processed to become mature lamin A. The protein has thus been named nuclear prelamin A recognition factor (Narf), but its precise function is still unknown. In *S. cerevisiae,* deletion of the yeast *Narf*-like sequence has been shown to be lethal in a haploid background (Munich Information Centre for Protein Sequences—http://www.mips.biochem.mpg.de/). Interestingly, *S. cerevisiae* does not encode a homolog of lamin A, suggesting that the yeast Narf-like protein may not have the same function in yeast as in humans.

Iron sulfur proteins occur in all organisms and assume a diversity of functions. Iron sulfur clusters in these proteins may facilitate electron transfer, contribute to catalytic function, or help to maintain structural integrity (Roualt and Klausner 1996). Other [Fe-S] center–containing proteins have been implicated in biosensing of oxidative stress and redox-dependent gene regulation in prokaryotes and, potentially, eukaryotes (Zheng and Storz 2000). While some of these proteins do not contain [Fe-S] centers identical to those predicted to be present in the Narf-like proteins investigated here and elsewhere (Barton and Worman 1999), the function of these sensory/regulatory molecules is believed to be dependent on the exquisite redox sensitivity of diverse types of [Fe-S] clusters (Zheng and Storz 2000). It is known that redox/nitric oxide sensing and processing/cleavage of nuclear lamins are key events in programmed cell death (apoptosis) of higher animal cells (Kluck et al. 1997). It therefore seems reasonable to speculate that the features of eukaryotic Narf-like proteins might allow them to participate as sensory/regulatory proteins in such processes.

Additional studies are also required to investigate the functions of the proteins encoded by the *E. histolytica* and *S. barkhanus* [Fe] hydrogenase gene sequences. These sequences closely resemble in structure the two short-form [Fe] hydrogenases which are targeted to hydrogenosomes in *T. vaginalis* (Bui and Johnson 1996). However, it is still not known if the short-form *Trichomonas* hydrogenases are actually responsible for hydrogen production from *Trichomonas* hydrogenosomes. For example, Payne, Chapman, and Cammack (1993) found hydrogenase activity purified from *T. vaginalis* to be associated with a polypeptide of approximately 64 kDa in size, which is larger than predicted (ca. 50 kDa) for the published sequences. In the present investigation, we isolated an almost complete coding sequence for a [Fe] hydrogenase of approximately the same size as the polypeptide discovered by Payne, Chapman, and Cammack (1993). The new *T. vaginalis* [Fe] hydrogenase contains all of the key residues known to be important for the function of the *C. pasteurianum* hydrogenase 1 (Peters et al. 1998). The new *T. vaginalis* gene apparently contains an N-terminal extension which is similar to characterized *Trichomonas* hydrogenosome-targeting motifs (Plümper, Bradley, and Johnson 1998), suggesting that the encoded protein is also targeted to the hydrogenosome. Thus, the currently available data suggest that there may be at least three [Fe] hydrogenases, possibly doing different things, within the hydrogenosomes of *T. vaginalis.* Moreover, phylogenetic analyses suggest that the observed structural diversity of *Trichomonas* [Fe] hydrogenases arose within the *Trichomonas* lineage.

The data presented here demonstrate that sequences related to [Fe] hydrogenases are more widely distributed on the genomes of eukaryotes than previously thought. According to published hypotheses, this distribution might result from vertical inheritance of ancestral genes via speciation (orthology), from one or more horizontal transfers (xenology), or from a mixture of both (Whatley, John, and Whatley 1979; Müller 1980; Embley, Horner, and Hirt 1997; Akhmanova et al. 1998; Martin and Müller 1998). An additional complication when trying to interpret patterns of gene history is the potential of confusing orthologs with paralogs arising via gene duplication. It is already clear from available data that multiple [Fe] hydrogenases exist on some prokaryotic and eukaryotic genomes (fig. 4). Limited and/or skewed sampling of taxa or genomes may further compound difficulties of interpretation and limit inferences.

ML analyses support monophyly of *Entamoeba, Spironucleus,* and *Trichomonas* [Fe] hydrogenases, with the *Entamoeba* and *Spironucleus* sequences forming a strongly supported sister group in all analyses. Current ideas about species relationships between microbial eukaryotes are in a state of flux, with the relationships between parasitic anaerobes such as the taxa investigated here being considered particularly problematic (Embley and Hirt 1998; Hirt et al. 1999; Roger 1999). However, there are no published data to suggest that *Entamoeba, Spironucleus,* and *Trichomonas* form a monophyletic group to the exclusion of other eukaryotes. Thus, the hypothesis that at least one [Fe] hydrogenase gene was present on the genome of a common ancestor shared by these species and other eukaryotes is not excluded by the data.

Monophyly of all of the eukaryote sequences was never observed in unconstrained analyses and, indeed, was rarely recovered among nonparametric bootstrap replicates. Testing the robustness of phylogenetic hypotheses can be problematic, because the statistical basis of some tests is poorly understood (Huelsenbeck and Rannala 1997). For example, the most common likelihood-based statistical test for comparing two topologies is probably the Kishino-Hasegawa test (Kishino and Hasegawa 1989), but the validity of this test has recently been questioned for its most common usage, that of testing the ML tree against another topology (Shimodaira and Hasegawa 1999; Goldman, Anderson, and Rodrigo 2000). Here, we used parametric bootstrapping and a likelihood ratio test of monophyly (Huelsenbeck, Hillis, and Nielsen 1996) to investigate if the [Fe] hydrogenase data could reject trees where eukaryote sequences were constrained to be monophyletic. We made no other assumptions regarding the internal relationships among eukaryote sequences relying on our search strategy to find the optimal arrangement for each data set. Likeli-

hood ratio tests coupled with parametric bootstrapping are among the most powerful approaches for hypothesis testing (Huelsenbeck and Rannala 1997), especially when it can be shown, as here, that the model and tree which are used for simulations also fit the original data. The likelihood ratio test of monophyly strongly rejected the hypothesis that eukaryote [Fe] hydrogenase sequences formed a monophyletic group.

In most analyses, the [Fe] hydrogenase module of the *N. ovalis* fusion protein formed a sister group relationship with the *D. vulgaris* [Fe] hydrogenase gamma sequence (Stokkermans et al. 1989), albeit with rather low support from nonparametric bootstrapping. One possible explanation of this topology is that horizontal gene transfer has occurred between the ancestors of these two organisms. An alternative hypothesis is one of hidden paralogy whereby the *Nyctotherus* and *D. vulgaris* sequences are paralogous to the other eukaryote and prokaryote sequences in the tree. To discriminate between these possibilities, we need more data than we have here. To make a strong case for horizontal transfer, sequences should ideally be available for a good taxonomic sample of organisms, and the tree relating the sequences should correspond to a conventional phylogeny—except that one or more sequences appears in a radically unexpected position (Smith, Feng, and Doolittle 1992) and preferably nested with strong support in an otherwise accepted monophyletic group.

It has been suggested that the *N. ovalis* [Fe] fusion protein contains components derived from the mitochondrial respiratory (complex I modules) chain (Andersson and Kurland 1999; Hackstein et al. 1999). Our analyses provide no support for this hypothesis. The 51-kDa component of mitochondrial complex I formed a strongly supported monophyletic group with the NuoF sequences from α-proteobacteria, in agreement with other data for mitochondrial origins (Gray et al. 1998; Andersson and Kurland 1999). The 51-kDa subunit–like component of the *Nyctotherus* fusion protein was never part of this clade, but its relationships to other proteins were not strongly resolved. In most analyses, the *Nyctotherus* sequence clustered with the 51-kDa-like component of the NAD(H)-dependent [Ni-Fe] hydrogenases isolated from the β-proteobacterium *A. eutrophus* and the Gram-positive actinomycete *R. opacus.* This result is interesting because, like the C-terminal domain of the *N. ovalis* fusion protein, these subunits are part of a fusion comprising 24-kDa/NuoE–like components and 51-kDa/NuoF–like components. The overall homologies between components of respiratory chain complex I and hydrogenase components has previously been noted (Pilkington et al. 1991; Albracht, Mariette, and de Jong 1997).

Conclusions

The extraordinary capacity of eukaryotes to repeatedly evolve organelles capable of producing hydrogen (hydrogenosomes) is a fascinating biological puzzle. Some have hypothesized that the genes encoding the key enzymes PFO and hydrogenase were acquired only once in eukaryotic evolution and have subsequently been retained (Embley, Horner, and Hirt 1997; Martin and Müller 1998). Under this scenario, the enzymes for hydrogen production might be readily available for recruitment into the compartment destined to become a hydrogenosome, with the mitochondrion being favored by the available data (for review, see Embley, Horner, and Hirt 1997; Martin and Müller 1998; Hackstein et al. 1999). Recent analyses have demonstrated that eukaryotic cytosolic and hydrogenosomal PFO from diverse eukaryotes, including *Entamoeba, Giardia, Spironucleus,* and *Trichomonas,* do form a monophyletic group, which is consistent with the ancestral acquisition and retention hypothesis for this particular enzyme (Horner, Hirt, and Embley 1999). In the present investigation, we focused on the origins of hydrogenase, the canonical enzyme for hydrogenosomes (Müller 1993). Our analyses suggest, albeit with relatively weak support, that cytosolic and hydrogenosomal [Fe] hydrogenases from *Entamoeba, Spironucleus,* and *Trichomonas* are also monophyletic. Thus, the available data and analyses are consistent with the hypothesis that the common, perhaps also anaerobic, ancestor of these species and, potentially, other eukaryotes contained genes for both PFO and [Fe] hydrogenase.

Two recently published hypotheses have postulated that eukaryotic cells originated through a symbiosis between two prokaryotes, and each posits that interspecies hydrogen transfer was the key to binding one partner to the other. In the hydrogen hypothesis, the prokaryotes comprise an anaerobic hydrogen-consuming archaebacterium (the host) and a hydrogen-producing facultative anaerobic α-proteobacterium (the symbiont), with the α-proteobacterium subsequently becoming the mitochondrion (Martin and Müller 1998). The syntrophic hypothesis posits that eukaryotes arose from a primary symbiosis between an anaerobic hydrogen-consuming methanogenic archaebacterium which became the nucleus and an anaerobic hydrogen-producing δ-proteobacterium which became the host, with mitochondria originating from a later symbiosis involving a facultatively anaerobic methane-consuming α- proteobacterium (Moreira and Lopez-Garcia 1998). While they differ in key features, both hypotheses can be interpreted to predict that PFO and hydrogenase were present early in eukaryotic evolution, a prediction for which our present and previous analyses provide some support. However, our analyses failed to identify a eubacterial sister group for either enzyme and thus cannot discriminate between the hydrogen and syntrophic hypotheses regarding the identity of a potential eubacterial donor for eukaryotic enzymes. This situation may improve with more sampling—particularly of α- and δ-proteobacterial genes. However, unless the trees for PFO and [Fe] hydrogenase can be rooted, it may remain difficult to identify a potential source for ancestral enzymes.

Trees where the [Fe] hydrogenase from the hydrogenosomal ciliate *N. ovalis* was constrained to be monophyletic with the other eukaryote sequences were rejected by our analyses. The favored position for the *Nyctotherus* sequence was as the sister group to a [Fe] hy-

drogenase on the genome of the eubacterium *D. vulgaris.* Thus, as previously suggested (Akhmanova et al. 1998) it seems possible that *Nyctotherus* has obtained its hydrogenosomal [Fe] hydrogenase from a different source (either from a different paralog or via horizontal transfer) from *Trichomonas* for its hydrogenosomes.

Acknowledgments

We would like to thank Mark van der Giezen and Robert Hirt for useful discussion and comments on the manuscript. We are indebted to Nick Goldman for allowing us access, prior to publication, to his manuscript on likelihood-based tests of topologies. The EST survey on *Spironucleus* is in collaboration with Mark Ragan (Canadian Institute of Advanced Research). We utilized data generated by the Sanger Centre (www.sanger.ac.uk) and the Genome Therapeutics Corporation (www.cric.com/genomesequences). D.S.H. is supported by a fellowship awarded by the Natural History Museum. The sequences reported here have been deposited in GenBank (accession numbers AF262400, AF262401, and AF262402).

LITERATURE CITED

ADACHI, J., and M. HASEGAWA. 1996. MOLPHY version 2.3: programs for molecular phylogenetics based on maximum likelihood. Comput. Sci. Monogr. **28**:1–150.

AKHMANOVA, A., F. VONCKEN, T. VAN ALEN, A. VAN HOEK, B. BOXMA, G. VOGELS, M. VEENHUIS, and J. H. HACKSTEIN. 1998. A hydrogenosome with a genome. Nature **396**: 527–528.

ALBRACHT, S. P. J., A. MARIETTE, and P. DE JONG. 1997. Bovine heart NADH:ubiquinone oxidoreductase is a monomer with 8 Fe-S clusters and 2 FMN groups. Biochim. Biophys. Acta **1318**:92–106.

ALTSCHUL, S. F., T. L. MADDEN, A. A. SCHAFFER, J. ZHANG, Z. ZHANG, W. MILLER, and D. J. LIPMAN. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**:3389–3402.

ANDERSSON, S. G. E., and C. G. KURLAND. 1999. Origins of mitochondria and hydrogenosomes. Curr. Opin. Microbiol. **2**:535–541.

ANDERSSON, S. G. E., A. ZOMORODIPOUR, J. O. ANDERSSON, T. SICHERITZ-PONTÉN, U. C. M. ALSMARK, R. M. PODOWSKI, A. K. NÄSLUND, A.-S. ERIKSSON, H. H. WINKLER, and C. G. KURLAND. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature **396**:133–140.

APPEL, J., and R. SCHULZ. 1998. Hydrogen metabolism in organisms with oxygenic photosynthesis: hydrogenases as important regulatory devices for a proper redox poising? J. Photochem. Photobiol. B **47**:1–11.

BARTON, R. M., and H. J. WORMAN. 1999. Prenylated prelamin A interacts with Narf, a novel nuclear protein. J. Biol. Chem. **274**:30008–30018.

BENCHIMOL, M., R. DURAND, and J. C. A. ALMEIDA. 1997. A double membrane surrounds the hydrogenosomes of the anaerobic fungus *Neocallimastix frontalis.* FEMS Microbiol. Lett. **154**:277–282.

BRADLEY, P. J., C. J. LAHTI, E. PLÜMPER, and P. L. JOHNSON. 1997. Targeting and translocation of proteins into the hydrogenosome of the protist *Trichomonas*: similarities with mitochondrial protein import. EMBO J. **16**:3484–3493.

BUI, E. T. N., P. J. BRADLEY, and P. J. JOHNSON. 1996. A common evolutionary origin for mitochondria and hydrogenosomes. Proc. Natl. Acad. Sci. USA **93**:9651–9656.

BUI, E. T., and P. J. JOHNSON. 1996. Identification and characterization of [Fe]-hydrogenases in the hydrogenosome of *Trichomonas vaginalis.* Mol. Biochem. Parasitol. **76**:305–310.

CAVALIER-SMITH, T. 1993. Kingdom Protozoa and its 18 phyla. Microbiol. Rev. **57**:953–994.

CHARLESTON, M. 1998. Spectrum—a tool for spectral analysis of phylogenetic data. Bioinformatics **14**:98–99.

CLARK, C. G., and A. J. ROGER. 1995. Direct evidence for secondary loss of mitochondria in *Entamoeba histolytica.* Proc. Natl. Acad. Sci. USA **92**:6518–6521.

DORE, J., and D. A. STAHL. 1991. Phylogeny of anaerobic rumen chytridiomycetes inferred from small subunit ribosomal RNA sequence comparisons. Can. J. Bot. **69**:1964–1971.

EMBLEY, T. M., B. J. FINLAY, P. L. DYAL, R. P. HIRT, M. WILKINSON, and A. G. WILLIAMS. 1995. Multiple origins of anaerobic ciliates with hydrogenosomes within the radiation of aerobic ciliates. Proc. R. Soc. Lond. B Biol. Sci. **262**: 87–93.

EMBLEY, T. M., and R. P. HIRT. 1998. Early branching eukaryotes? Curr. Opin. Genet. Dev. **8**:624–629.

EMBLEY, T. M., D. S. HORNER, and R. P. HIRT. 1997. Anaerobic eukaryote evolution: hydrogenosomes as biochemically modified mitochondria? Trends Ecol. Evol. **12**:437–441.

EMBLEY, T. M., and W. MARTIN. 1998. Molecular evolution—a hydrogen-producing mitochondrion. Nature **396**:517–519.

GERMOT, A., H. PHILIPPE, and H. L. GUYADER. 1996. Presence of a mitochondrial type 70kDa heat shock protein in *Trichomonas vaginalis* suggests a very early mitochondrial endosymbiosis in eukaryotes. Proc. Natl. Acad. Sci. USA **93**: 14614–14617.

GOLDMAN, N. 1993. Statistical tests of models of DNA substitution. J. Mol. Evol. **36**:182–198.

GOLDMAN, N., J. P. ANDERSON, and A. G. RODRIGO. 2000. Likelihood-based tests of topologies in phylogenetics. Syst. Biol. (in press).

GRAY, M. W., B. F. LANG, R. CEDERGREN et al. (15 co-authors). 1998. Genome structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res. **26**:865–878.

HACKSTEIN, J. H. P., A. AKHMANOVA, B. BOXMA, H. R. HARHANGI, and G. J. VONCKEN. 1999. Hydrogenosomes: eukaryotic adaptations to anaerobic environments. Trends Microbiol. **7**:441–447.

HAPPE, T., B. MOSLER, and J. D. NABER. 1994. Induction, localization and metal content of hydrogenase in the green-alga *Chlamydomonas reinhardtii.* Eur. J. Biochem. **222**: 769–774.

HAPPE, T., and J. D. NABER. 1993. Isolation, characterization and n-terminal amino-acid-sequence of hydrogenase from the green-alga *Chlamydomonas reinhardtii.* Eur. J. Biochem. **214**:475–481.

HIRT, R. P., J. M. LOGSDON, B. HEALY, M. W. DOREY, W. F. DOOLITTLE, and T. M. EMBLEY. 1999. Microsporidia are related to fungi: evidence from the largest subunit of RNA polymerase II and other proteins. Proc. Natl. Acad. Sci. USA **96**:580–585.

HORNER, D. S., R. P. HIRT, and T. M. EMBLEY. 1999. A single eubacterial origin of eukaryotic pyruvate : ferredoxin oxidoreductase genes: Implications for the evolution of anaerobic eukaryotes. Mol. Biol. Evol. **16**:1280–1292.

HORNER, D. S., R. P. HIRT, S. KILVINGTON, D. LLOYD, and T. M. EMBLEY. 1996. Molecular data suggest an early acquisition of the mitochondrion endosymbiont. Proc. R. Soc. Lond. Biol. Sci. B Biol. Sci. **263**:1053–1059.

HUELSENBECK, J. P., D. M. HILLIS, and R. NIELSEN. 1996. A likelihood-ratio test of monophyly. Syst. Biol. **45**:546–558.

HUELSENBECK, J. P., and B. RANNALA. 1997. Phylogenetic methods come of age: testing hypotheses in an evolutionary context. Science **276**:227–232.

KEELING, P. J., and W. F. DOOLITTLE. 1997. Widespread and ancient distribution of a noncanonical genetic code in diplomonads. Mol. Biol. Evol. **14**:895–901.

KISHINO, H., and M. HASEGAWA. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. **29**:170–179.

KLUCK, R. M., S. J. MARTIN, B. M. HOFFMAN, J. S. ZHOU, D. R. GREEN, and D. D. NEWMEYER. 1997. Cytochrome c activation of CPP32-like proteolysis plays a critical role in a *Xenopus* cell-free apoptosis. EMBO J. **16**:4639–4649.

LAKE, J. A. 1994. Reconstructing evolutionary trees from DNA and protein sequences: paralinear distances. Proc. Natl. Acad. Sci. USA **91**:1455–1459.

LENTO, G. M., R. E. HICKSON, G. K. CHAMBERS, and D. PENNY. 1995. Use of spectral analysis to test hypotheses on the origins of pinnipeds. Mol. Biol. Evol. **12**:28–52.

LINDMARK, D. G., and M. MÜLLER. 1973. Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate, *Tritrichomonas foetus*, and its role in pyruvate metabolism. J. Biol. Chem. **248**:7724–7728.

LLOYD, D. 1996. Obligate anaerobe or not? Nature **381**:121.

LOCKHART, P. J., A. W. D. LARKUM, M. A. STEEL, P. J. WADDEL, and D. PENNY. 1996. Evolution of chlorophyll and bacteriochlorophyll: the problem of invariant sites in sequence analysis. Proc. Natl. Acad. Sci. USA **93**:1930–1934.

LOCKHART, P. J., M. A. STEEL, M. D. HENDY, and D. PENNY. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. **11**:605–612.

MCINERNEY, J. O. 1997. Codon usage patterns in *Trichomonas vaginalis.* Eur. J. Protistol. **33**:266–273.

MAI, Z., S. GHOSH, M. FRISARDI, B. ROSENTHAL, R. ROGERS, and J. SAMUELSON. 1999. Hsp60 is targeted to a cryptic mitochondrion-derived organelle (''crypton'') in the microaerophilic protozoan parasite *Entamoeba histolytica.* Mol. Cell. Biol. **19**:2198–2205.

MAIDAK, B. L., G. J. OLSEN, N. LARSEN, R. OVERBEEK, M. J. MCCAUGHEY, and C. R. WOESE. 1996. The Ribosomal Database Project (RDP). Nucleic Acids Res. **24**:82–85.

MARTIN, W., and M. MÜLLER. 1998. The hydrogen hypothesis for the first eukaryote. Nature **392**:37–41.

MOREIRA, D., and P. LOPEZ-GARCIA. 1998. Symbiosis between methanogenic archaea and δ-proteobacteria as the origin of eukaryotes: the syntrophic hypothesis. J. Mol. Evol. **47**: 517–530.

MÜLLER, M. 1980. The hydrogenosome. Pp. 127–142 *in* G. W. GOODAY, D. LLOYD, and A. P. J. TRINCI, eds. The eukaryotic microbial cell. Cambridge University Press, Cambridge, England.

———. 1993. The hydrogenosome. J. Gen. Microbiol. **139**: 2879–2889.

NICOLET, Y., C. PIRAS, P. LEGRAND, C. E. HATCHIKIAN, and J. C. FONTECILLA-CAMPS. 1999. *Desulfovibrio desulfuricans* iron hydrogenase: the structure shows unusual coordination to an active site Fe binuclear cluster. Structure **7**:13–23.

PAYNE, M. J., A. CHAPMAN, and R. CAMMACK. 1993. Evidence for an [Fe]-type hydrogenase in the parasitic protozoan *Trichomonas vaginalis.* FEBS Lett. **317**:101–104.

PETERS, J. W., W. N. LANZILOTTA, B. J. LEMON, and L. C. SEEFELDT. 1998. X-ray crystal structure of the Fe-only hydrogenase (CpI) from *Clostridium pasteurianum* to 1.8 angstrom resolution. Science **282**:1853–1858.

PHILIPPE, H. 1993. MUST, a computer package of management utilities for sequences and trees. Nucleic Acids Res. **21**: 5264–5272.

PHILIPPE, H., and A. ADOUTTE. 1998. The molecular phylogeny of eukaryota: solid facts and uncertainties. Pp. 25–56 *in* G. H. COOMBS, K. VICKERMANN, M. A. SLEIGH, and A. WARREN, eds. Evolutionary relationships among protozoa. Kluwer Academic Publishers, London.

PILKINGTON, S. J., J. M. SKEHEL, R. B. GENNIS, and J. E. WALKER. 1991. Relationship between mitochondrial NADH-ubiquinone reductase and a bacterial NAD-reducing hydrogenase. Biochemistry **30**:2166–2175.

PLÜMPER, E., P. J. BRADLEY, and P. J. JOHNSON. 1998. Implications of protein import on the origin of hydrogenosomes. Protist **149**:303–311.

REEVES, R. E., L. G. WARREN, B. SUSKIND, and H. S. LO. 1977. An energy conserving pyruvate to acetate pathway in *Entamoeba histolytica.* J. Biol. Chem. **252**:726–731.

ROGER, A. J. 1999. Reconstructing early events in eukaryotic evolution. Am. Nat. **154**:S146–S163.

ROGER, A. J., C. G. CLARK, and W. F. DOOLITTLE. 1996. A possible mitochondrial gene in the early branching amitochondriate protist *Trichomonas vaginalis.* Proc. Natl. Acad. Sci. USA **93**:14618–14622.

ROGER, A. J., S. G. SVARD, J. TOVAR, C. G. CLARK, M. W. SMITH, F. D. GILLIN, and M. L. SOGIN. 1998. A mitochondrial-like chaperonin 60 gene in *Giardia lamblia*: evidence that diplomonads once harbored an endosymbiont related to the progenitor of mitochondria. Proc. Natl. Acad. Sci. USA **95**:229–234.

ROSENTHAL, B., M. ZHIMING, D. CAPLIVSKI, S. GHOSH, H. DE LA VEGA, T. GRAF, and J. SAMUELSON. 1997. Evidence for the bacterial origin of genes encoding fermentation enzymes of the amitochondriate protozoan parasite *Entamoeba histolytica.* J. Bacteriol. **179**:3736–3745.

ROUALT, T. A., and R. D. KLAUSNER. 1996. Iron-sulfur clusters as biosensors of oxidants and iron. Trends Biochem. **21**: 174–177.

SHIMODAIRA, H., and M. HASEGAWA. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. **16**:1114–1116.

SMITH, M. W., D.-F. FENG, and R. F. DOOLITTLE. 1992. Evolution by acquisition: the case for horizontal gene transfers. Trends Ecol. Evol. **17**:489–493.

SOLTYS, B. J., and R. S. GUPTA. 1994. Presence and cellular distribution of a 60KDA protein related to mitochondrial HSP-60 in *Giardia lamblia.* J. Parasitol. **80**:580–590.

STEINBUCHEL, A., and M. MÜLLER. 1986. Anaerobic pyruvate metabolism of *Tritrichomonas foetus* and *Trichomonas vaginalis* hydrogenosomes. Mol. Biochem. Parasitol. **20**:57–65.

STERUD, E., T. A. MO, and T. T. POPPE. 1997. Ultrastructure of *Spironucleus barkhanus* n. sp. (Diplomonadida: Hexamitidae) from grayling *Thymallus thymallus* (L.) and Atlantic salmon *Salmo salar* L. (Salmonidae). J. Eukaryot. Microbiol. **44**:399–407.

STOKKERMANS, J., W. VANDONGEN, A. KAAN, W. VANDENBERG, and C. VEEGER. 1989. Hyd-gamma, a gene from *Desulfovibrio vulgaris* (Hildenborough) encodes a polypeptide homologous to the periplasmic hydrogenase. FEMS Microbiol. Lett. **58**:217–222.

STRIMMER, K., and A. VON HAESELER. 1996. Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. Mol. Biol. Evol. **13**:964–969.

SWOFFORD, D. L. 1998. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sinauer, Sunderland, Mass.

TAKEUCHI, T., E. C. WEINBACH, and L. S. DIAMOND. 1975. Pyruvate oxidase (CoA acetylating) in *Entamoeba histolytica*. Biochem. Biophys. Res. Commun. **65**:591–596.

TANAKA, T., M. TANAKA, and Y. MITSUI. 1997. Analysis of expressed sequence tags (ESTs) of the parasitic protozoa *Entamoeba histolytica*. Biochem. Biophys. Res. Commun. **236**:611–615.

THOMPSON, J. D., D. G. HIGGINS, and T. J. GIBSON. 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.

TOVAR, J., A. FISCHER, and C. G. CLARK. 1999. The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. Mol. Microbiol. **32**: 1013–1021.

URBIG, T., R. SCHULZ, and H. SENGER. 1993. Inactivation and reactivation of the hydrogenases of the green algae *Scenedesmus obliquus* and *Chlamydomonas reinhardtii*. Z. Naturforsch. **48**:41–45.

VAN DER GIEZEN, M., J. A. K. W. KIEL, K. A. SJOLLEMA, and R. A. PRINS. 1998. The hydrogenosomal Malic enzyme is targeted to mitochondria of the methylotrophic yeast *Hansenula polymorpha*. Curr. Genet. **33**:131–135.

VAN DER GIEZEN, M., K. B. RECHINGER, I. SVENDSEN, R. DURAND, R. P. HIRT, M. FEVRE, T. M. EMBLEY, and R. A. PRINS. 1997. A mitochondrial-like targeting signal on the hydrogenosomal malic enzyme from the anaerobic fungus *Neocallimastix frontalis*: support for the hypothesis that hydrogenosomes are modified mitochondria. Mol. Microbiol. **23**:11–21.

VAN DER GIEZEN, M., K. A. SJOLLEMA, R. R. E. ARTZ, W. ALKEMA, and R. A. PRINS. 1997. Hydrogenosomes in the anaerobic fungus *Neocallimastix frontalis* have a double membrane but lack an associated organelle genome. FEBS Lett. **408**:147–150.

WHATLEY, J. M., P. JOHN, and F. R. WHATLEY. 1979. From extracellular to intracellular: the establishment of mitochondria and chloroplasts. Proc. R. Soc. Lond. B Biol. Sci. **204**: 165–187.

YANG, D., Y. OYAIZU, H. OYAIZU, G. J. OLSEN, and C. R. WOESE. 1985. Mitochondrial origins. Proc. Natl. Acad. Sci. USA **82**:4443–4447.

YANG, Z. 1994. Estimating the pattern of nucleotide substitution. J. Mol. Evol. **39**:105–111.

———. 1996. Among-site rate variation and its impact on phylogenetic analyses. Trends Ecol. Evol. **11**:367–372.

ZHENG, M., and G. STORZ. 2000. Redox sensing by prokaryotic transcription factors. Biochem. Pharmacol. **59**:1–6.

WILLIAM MARTIN, reviewing editor

Accepted July 10, 2000