Novel Hydrogenosomes in the Microaerophilic Jakobid *Stygiella incarcerata*

Michelle M. Leger,[1] Laura Eme,[1] Laura A. Hug,[‡,1] and Andrew J. Roger*[,1]

[1]Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, NS, Canada

[‡]Present address: Department of Biology, University of Waterloo, Waterloo, ON, Canada

***Corresponding author:** E-mail: andrew.roger@dal.ca.

Associate editor: Iñaki Ruiz-Trillo

Abstract

Mitochondrion-related organelles (MROs) have arisen independently in a wide range of anaerobic protist lineages. Only a few of these organelles and their functions have been investigated in detail, and most of what is known about MROs comes from studies of parasitic organisms such as the parabasalid *Trichomonas vaginalis*. Here, we describe the MRO of a free-living anaerobic jakobid excavate, *Stygiella incarcerata*. We report an RNAseq-based reconstruction of *S. incarcerata*'s MRO proteome, with an associated biochemical map of the pathways predicted to be present in this organelle. The pyruvate metabolism and oxidative stress response pathways are strikingly similar to those found in the MROs of other anaerobic protists, such as *Pygsuia* and *Trichomonas*. This elegant example of convergent evolution is suggestive of an anaerobic biochemical 'module' of prokaryotic origins that has been laterally transferred among eukaryotes, enabling them to adapt rapidly to anaerobiosis. We also identified genes corresponding to a variety of mitochondrial processes not found in *Trichomonas*, including intermembrane space components of the mitochondrial protein import apparatus, and enzymes involved in amino acid metabolism and cardiolipin biosynthesis. In this respect, the MROs of *S. incarcerata* more closely resemble those of the much more distantly related free-living organisms *Pygsuia biforma* and *Cantina marsupialis*, likely reflecting these organisms' shared lifestyle as free-living anaerobes.

Key words: **mitochondria, mitochondrion-related organelles, jakobids, excavates, anaerobic metabolism.**

Introduction

All known extant eukaryote lineages contain, or once contained, a mitochondrion of some description, originating from an alphaproteobacterial endosymbiont (Stairs et al. 2015). Highly derived mitochondrion-related organelles (MROs) are found in diverse anaerobic protist lineages (reviewed in Stairs et al. 2015; Makiuchi and Nozaki 2014), and these have been classified primarily based on their roles in ATP generation. Traditionally, MROs that produce energy anaerobically have been classified as hydrogenosomes (Lindmark and Muller 1973; Bui and Johnson 1996; Dyall et al. 2004; Hrdý et al. 2004), while more reduced organelles that lack a role in energy metabolism have been classified as mitosomes (Tovar et al. 1999; Jedelsky et al. 2011). However, a more recent classification scheme aims to encompass the broader spectrum of energy functions since discovered in MROs (Müller et al. 2012). In this scheme the first three classes have retained an electron transport chain. Class 1 mitochondria, typical mitochondria, are only capable of generating ATP through oxidative phosphorylation. Class 2 mitochondria are additionally capable of functioning anaerobically using alternative electron acceptors, such as fumarate, while class 3 mitochondria possess both an electron transport chain and a bacterial-like anaerobic ATP-generation pathway. Class 4 mitochondria, also known as hydrogenosomes, possess the anaerobic ATP-generation pathway, but have lost most components of the electron transport chain, such that they are no longer able to generate ATP through aerobic phosphorylation. The more highly reduced class 5 mitochondria, also known as mitosomes, lack any role in ATP generation and any of the associated enzymes. Based on the distribution of MROs within eukaryotic phyla (many of which contain predominantly aerobic mitochondriate taxa), it has been proposed that this adaptation to anaerobic conditions has occurred multiple times, including at least four independent events within ciliates alone (Embley et al. 1995). Nevertheless, striking biochemical similarities are apparent between these taxa in the form of enzymes shared between distantly-related anaerobic and microaerophilic eukaryotes (Andersson et al. 2003, 2006; Stairs et al. 2011; Müller et al. 2012; Tsaousis, Leger, et al. 2012; Noguchi et al. 2015).

Hydrogenosomes (class 4 mitochondria) were first described in trichomonads (Lindmark and Muller 1973), and have since been described in many anaerobic protistan lineages, including the excavates *Sawyeria marylandensis* (Barberà et al. 2010), *Psalteriomonas lanterna* (de Graaf et al. 2009), and *Spironucleus salmonicida* (Jerlstrom-Hultqvist et al. 2013); chytridiomycete fungi such as *Neocallimastix* spp. (Bowman et al. 1992); and some anaerobic ciliates (Yarlett et al. 1981; Akhmanova et al. 1998). They

contain a characteristic anaerobic ATP-generating pathway (Dyall et al. 2004; Hrdý et al. 2004, 2008; Tsaousis, Leger, et al. 2012), which relies on the presence of two key proteins rarely found within classical aerobic mitochondria: pyruvate:ferredoxin oxidoreductase (PFO) and [FeFe]-hydrogenase. PFO decarboxylates pyruvate, producing acetyl-CoA and carbon dioxide. In some anaerobic eukaryotes, such as *Blastocystis* sp., *Mastigamoeba balamuthi* or *Cryptosporidium parvum*, pyruvate to acetyl-CoA conversion is carried out by a pyruvate:NADP oxidoreductase (PNO) or a pyruvate:formate lyase (PFL) (Ctrnacta et al. 2006; Lantsman et al. 2008; Stairs et al. 2011). Coenzyme A is subsequently transferred from acetyl-CoA to succinate by an acetate:succinate CoA transferase (ASCT; Tielens et al. 2010), generating acetate and succinyl-CoA. The subsequent regeneration of succinate, catalyzed by the tricarboxylic acid cycle enzyme succinate thiokinase (STK), generates ATP through substrate-level phosphorylation. The electrons generated in the decarboxylation of pyruvate are transferred from PFO, via ferredoxin, to [FeFe]-hydrogenase, which catalyses the reductive formation of hydrogen gas from free protons. Three maturases, HydE, HydF, and HydG, are involved in the insertion of an iron sulfur (FeS) cluster into the active site of [FeFe]-hydrogenase (Posewitz et al. 2004; McGlynn et al. 2008; Pilet et al. 2009; Mulder et al. 2010). The 51- and 24-kDa subunits of mitochondrial electron transport chain Complex I are commonly found in anaerobic eukaryotes in the absence of other Complex I subunits, and are capable of transferring electrons to ferredoxin (Hrdý et al. 2004). The origin(s) of PFO and [FeFe]-hydrogenase and its maturases in diverse anaerobic eukaryotes are currently under debate. Phylogenetic analyses have provided some evidence for lateral transfers of these protein genes from either anaerobic bacteria or between eukaryotes, but these analyses suffer from lack of resolution (Vignais et al. 2001; Davidson et al. 2002; Horner et al. 2002; Hug et al. 2010; Leger et al. 2013; Stairs et al. 2014). Conversely, there is currently no strong phylogenetic link between eukaryotic homologs of these enzymes and alphaproteobacterial orthologs that would be suggestive of a mitochondrial origin.

With the advent of high-throughput sequencing, the proteomes of MROs from an increasing number of anaerobic lineages have been predicted, clarifying the totality of the functions of the organelles and expanding beyond energy metabolism. These studies have revealed a diversity of ancestrally mitochondrial and newly acquired biochemical pathways in MROs (Akhmanova et al. 1998; van Hoek et al. 2000; Boxma et al. 2005; Lantsman et al. 2008; Stechmann et al. 2008; Denoeud et al. 2011; Tsaousis et al. 2011; Nyvltova et al. 2013; Stairs et al. 2014; Nyvltova et al. 2015).

Here, we examine the MRO of the anaerobic protist *Stygiella incarcerata*. *S. incarcerata* (Panek et al. 2015) (originally *Jakoba incarcerata*; Bernard et al. 2000; then *Andalucia incarcerata*; Lara et al. 2006) is a member of the Jakobida (Lara et al. 2006), a group of flagellated protists with typical aerobic mitochondria. Jakobids are unique in having the most gene rich, bacterial-like mitochondrial genomes found to date (Lang et al. 1997; Burger et al. 2013; Kamikawa et al. 2014), bacterial-like RNA polymerases, and Shine–Dalgarno-like motifs upstream of translation sites (Lang et al. 1997). Stygiellidae represent a secondary derivation of the anaerobic lifestyle within the otherwise aerobic jakobids (Panek et al. 2015). The original identification by electron microscopy described uniformly-staining MROs lacking cristae (Simpson and Patterson 2001); however, the biochemical nature of the MROs has not been studied until now.

Stygiella incarcerata presents an opportunity to compare mitochondrial remodeling in a free-living organism with similar remodeling in the better-studied parasite lineages (Loftus et al. 2005; Aguilera et al. 2008; Jedelsky et al. 2011; Schneider et al. 2011; Jerlstrom-Hultqvist et al. 2013), specifically from a lineage containing mitochondria with particularly bacterial-like features.

Here, we reconstruct the MRO proteome of *S. incarcerata* based on RNASeq data, and we verify our predictions using immunolocalization of key enzymes.

Results and Discussion

Structural Characterization of the Organelle in *S. incarcerata*

Electron microscopy was used to examine the morphology of the organelle in greater detail than had previously been assessed. The organelle is uniform in staining density, with greater density than the surrounding cytosol, and lacking visible ribosomes (fig. 1). As previously described (Simpson and Patterson 2001), it is generally located close to the nucleus, anterior in the cell. It is roughly ovoid in shape, and 0.75–1 μm in length (fig. 1*A*). Examination of the organelle at a higher magnification confirmed the presence of a double membrane, consistent with its earlier designation as an endosymbiont-derived mitochondrion-like organelle (Simpson and Patterson 2001) (fig. 1*B*).

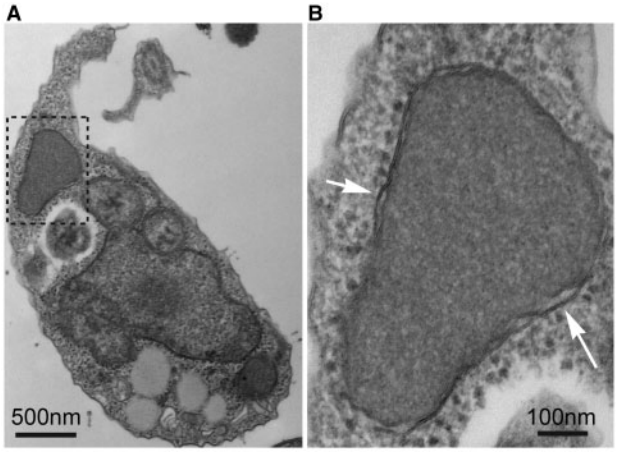

FIG. 1. Transmission electron micrographs of *S. incarcerata*, showing the MRO. (*A*) Whole *S. incarcerata* cell; the inset area shown in (*B*) is indicated with a dashed outline. (*B*) The MRO of *S. incarcerata*. White arrows indicate the double membrane.

Stygiella incarcerata MROs Likely Lack an Organellar Genome

To determine the degree of completeness of the transcriptome, we queried the transcripts using sequences present in a 159-protein dataset of conserved eukaryotic proteins (Brown et al. 2013). We identified *S. incarcerata* homologs of all but one of these proteins (nsf1-E; supplementary table 1, Supplementary Material online). This complement is comparable with the most complete complement present in Brown et al.'s dataset, that of *Homo sapiens*. We then searched for homologs of mitochondrially encoded genes of *S. incarcerata*'s close relative, *Andalucia godoyi* (Burger et al. 2013). Although *A. godoyi* has one of the largest mitochondrial gene complements known (72 protein-coding genes and 34 structural RNA genes; listed in supplementary table 2, Supplementary Material online), no *S. incarcerata* orthologs of any of these genes could be recovered from the transcriptome.

We made attempts at visualizing an organellar genome, if present, using pulsed-field gel electrophoresis. We observed only one candidate band, a transiently present band of ~9 kb that could not be cloned for sequencing, and that was not recognized by a 16S mitochondrial rDNA probe from the sister taxon *A. godoyi* on Southern blots (data not shown). Attempts were also made to isolate mitochondrial 16S ribosomal DNA sequences from *S. incarcerata* DNA extracts using PCR with eukaryote mitochondrial-specific primers. Only sequences with >95% identity to bacterial ribosomal DNAs in the GenBank nr database were retrieved. Positive controls using *A. godoyi* DNA successfully amplified mitochondrial 16S ribosomal DNA (data not shown). These findings are consistent with the apparent absence of mitochondrial genes, or any sequences encoding mitochondrial transcriptional or translational apparatus components. It therefore seems likely that *S. incarcerata* MROs, like those of *Trichomonas*, lack an organellar genome.

Identification of Organellar Proteins *In Silico*

CBOrg (see Materials and Methods) and comparative BLAST analyses identified 142 proteins of putative mitochondrial or hydrogenosomal origin. Of these, 89 sequences contain N-terminal targeting peptides as predicted by TargetP (supplementary table S3, Supplementary Material online). Given the identification of both α and β-subunits of the mitochondrial processing peptidase (MPP), it seems likely that typical targeting pathways are active within the organelle, and thus identification of N-terminal targeting peptides can be interpreted as further evidence for the localization of these proteins. Most of the proteins predicted to be localized to the MROs have N-terminal extensions relative to bacterial homologs; these extensions are generally predicted to be N-terminal targeting peptides by TargetP and Mitoprot. Exceptions are homologs of known membrane proteins, such as carrier proteins and some components of the mitochondrial protein import system, many of which are known to lack N-terminal targeting peptides (Becker et al. 2012). It is also reasonable to expect that a significant portion of the proteins actually targeted to the organelle contain cryptic internal targeting signals (Neupert and Herrmann 2007) and as such will not have been identified in this survey. Figure 2 shows a biochemical map of selected pathways predicted to be present in the MRO.

ATP Generation

Most enzymes and protein complexes typically present in, and associated with aerobic ATP generation in, classic mitochondria are all conspicuously absent from the transcriptome. These include subunits of complex I other than the 51 and 24 kDa subunits; components of the electron transport chain downstream of complex I [including Complex II subunits retained in the MROs of *Blastocystis* spp., *M. balamuthi*, *Pygsuia biforma*, and *Cantina marsupialis* (Gill et al. 2007; Lantsman et al. 2008; Stechmann et al. 2008; Stairs et al. 2014; Noguchi et al. 2015; Nyvltova et al. 2015)]; and subunits of the pyruvate dehydrogenase complex. In contrast, PFO, [FeFe]-hydrogenase and its maturases, ferredoxin, ASCT, STK, and the two complex I subunits were identified; these constitute the bulk of the typical hydrogenosomal ATP generation pathway described in *Trichomonas vaginalis* (Dyall et al. 2004; Hrdý et al. 2004, 2008; Carlton et al. 2007). Like *T. vaginalis*, *S. incarcerata* possesses multiple MRO-targeted homologs of both PFO and [FeFe]-hydrogenase; but both of its copies of ASCT belong to the subtype 1B family, which is distinct from the subtype 1C enzyme found in *T. vaginalis* (Tielens et al. 2010). The *S. incarcerata* transcriptome also encodes PFL and PNO, although these enzymes lack targeting peptides and are not predicted to localize to the MROs.

In addition to the two MRO-targeted [FeFe]-hydrogenase homologs, the transcriptome encodes two further [FeFe]-hydrogenases that lack predicted N-terminal mitochondrial targeting peptides. Each of these is fused to a C-terminal CysJ similar to those PNOs and NADPH cytochrome p450 reductases. This type of [FeFe]-hydrogenase has previously been described in *T. vaginalis* (Carlton et al. 2007), and in the breviate *P. biforma* (Stairs et al. 2014); in both of these organisms, the CysJ-fused [FeF]-hydrogenases are also predicted to be cytosolic. We confirmed [FeFe]-hydrogenase localization within the MROs using immunogold labeling (fig. 3) with a heterologous polyclonal rat primary antibody raised against *Mastigamoeba* [FeFe]-hydrogenase peptide. Probing labeled cells with a secondary antibody conjugated to gold particles yielded a 7-fold higher density of gold particles in the MROs compared with the cytosol. We also observed cross-reaction of the antibody with engulfed bacteria, and with the nucleus (although staining in the nucleus was significantly lower than that in the MROs). The former likely results from cross-reaction with bacterial homologs, while the latter might be attributed to the presence of a transcript encoding Nuclear prelamin A Recognition Factor (NARF) protein, a distant homolog of [FeFe]-hydrogenase. While the antibody used recognizes all four [FeFe]-hydrogenases expressed in *S. incarcerata*—including the two that are predicted not to

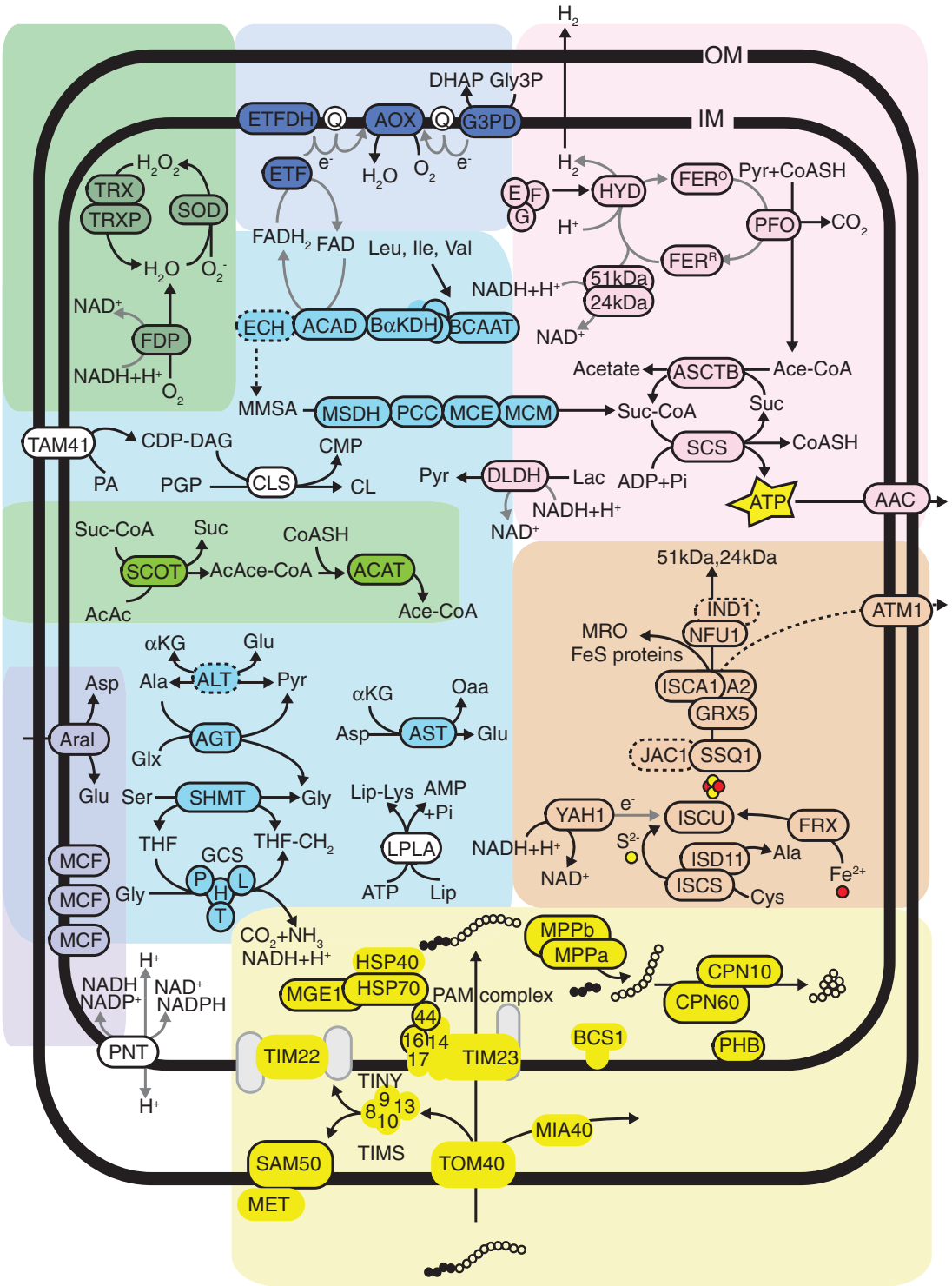

Fig. 2. A hypothetical biochemical map of selected pathways in the MRO of *S. incarcerata*. All proteins depicted are based on analyses of the EST and RNA-Seq data intended to identify mitochondrion or hydrogenosome-localized proteins. Circles enclosed by a solid outline indicate the protein contains a predicted N-terminal targeting peptide. Circles enclosed by a dashed outline indicate genes with an incomplete coding region at the 5′ end, and thus the presence/absence of a targeting peptide could not be evaluated. Circles with no outline indicate proteins with no predicted N-terminal targeting peptide. Grey circles indicate components of pathways or solute carriers which were not represented in the EST survey, but which are anticipated to be present in the organelle based on the presence of other components within complexes and pathways. Pathways and associated abbreviations of their enzyme components are depicted as follows: General: OM: outer mitochondrial membrane; IM: inner mitochondrial membrane; AMP: adenosine monophosphate; ADP: adenosine diphosphate; ATP: adenosine triphosphate; NAD: nicotinamide adenine dinucleotide; NADP: nicotinamide adenine dinucleotide phosphate; AcAc: acetoacetate; Ace-CoA: acetyl-CoA; AcAce-CoA: acetoacetyl-CoA; CoASH: Coenzyme A; MMSA: (S)-methylmalonate-semialdehyde; Suc: succinate; Suc-CoA: succinyl-CoA; Ala: alanine; Asp: aspartate; Glu: glutamate; Glx: glyoxylate; Gly: glycine; Ile: isoleucine; Leu: leucine; Lys: lysine; Ser: serine; Val: valine; Lac: lactate; Lip: lipoate;

localize to the MROs (supplementary fig. S1, Supplementary Material online)—the significantly higher degree of staining within the MROs suggests that the bulk of [FeFe]-hydrogenases is present there, confirming a likely role of this organelle in ATP production.

Evolutionary Origins of ATP Generation Enzymes

Eukaryote monophyly is recovered in phylogenies of PFO homologs, and in trees of all of the [FeFe]-hydrogenase maturases (fig. 4, supplementary figs. S2–S5, Supplementary Material online). Several major clades of [FeFe]-hydrogenases are recovered (fig. 5, supplementary figs. S6 and S7, Supplementary Material online), one of which is a long-branched clade that includes bacterial H_2-evolving periplasmic [FeFe]-hydrogenases (Vignais and Billoud 2007), as well as homologs from *Mastigamoeba*, *Entamoeba*, and *Trimastix*. Monophyly of all eukaryotic hydrogenases, both periplasmic-like and nonperiplasmic-like, was rejected in AU tests (table 1) (Hug et al. 2010; Leger et al. 2013). However, as in previous studies, monophyly of nonperiplasmic [FeFe]-hydrogenases of eukaryotes was not rejected (table 1, supplementary figs. S6 and S7, Supplementary Material online) (Leger et al. 2013). The origins of these anaerobic energy enzymes remain unclear, as there is no clear prokaryotic sister clade to any of these groups. In contrast to previous studies, monophyly of alphaproteobacteria and eukaryotes was not rejected. Nevertheless, the best constrained trees generated by RAxML for these hypotheses placed alphaproteobacteria branching from within eukaryotes (data not shown), and the reverse hypothesis was rejected in most cases. As in previous studies (Hug et al. 2010; Leger et al. 2013), the hypothetical grouping of alphaproteobacteria with eukaryotes was clearly rejected for PFO, whether sulfite reductases were included in the analyses or not.

The overall lack of resolution in these trees, and the lack of a clear prokaryotic sister group to eukaryotes that could point to these enzymes being either of mitochondrial origin, or laterally transferred from a specific prokaryotic group, is unfortunately typical of anaerobic energy generation enzyme phylogenies (Horner et al. 2000, 2002; Hug et al. 2010; Stairs et al. 2014). The prevalence of these enzymes among eukaryotes, and eukaryotic monophyly in some phylogenies, has been cited as evidence for their presence in the original mitochondrial endosymbiont (Horner et al. 1999, 2000, 2002; Müller et al. 2012). The suggestion of nonperiplasmic [FeFe]-hydrogenase eukaryotic monophyly, and the fact that alphaproteobacteria + eukaryote monophyly cannot be rejected for nonperiplasmic [FeFe]-hydrogenases, could be argued to be further evidence for this hypothesis. While these findings should not be ignored, they should, however, be treated with caution. If the monophyly of eukaryotes in anaerobic energy generation enzyme phylogenies were indeed a result of their presence in the protomitochondrial endosymbiont, then this must be weighed against the relative scarcity of [FeFe]-hydrogenases in alphaproteobacteria; the extreme scarcity of the three [FeFe]-hydrogenase maturases in the same group; and the rejection of an alphaproteobacterial affinity for PFO. Ancestral presence of modern anaerobic energy generation enzymes in the protomitochondrial endosymbiont requires massive loss from, or replacement of, these enzymes in alphaproteobacteria; or lateral transfer of at least some of them into the lineage giving rise to the endosymbiont shortly before the endosymbiotic event; and, certainly, large-scale subsequent loss in diverse eukaryote lineages. An alternative explanation involves lateral transfer of

Oaa: oxaloacetate; Pyr: pyruvate; Q: quinone; THF: tetrahydrofolate. Yellow circles: sulphur moiety of FeS clusters; red circles: iron moiety of FeS clusters; contiguous white circles: imported protein; contiguous black circles: mitochondrial targeting peptide of imported protein. Solid black arrows indicate biochemical pathways; dashed black arrows indicate poorly understood or unclear pathways; solid grey arrows indicate electron transfer. Pyruvate metabolism (pink): 24 kDa: 24 kDa subunit of Electron Transport Chain Complex I; 51 kDa: 51 kDa subunit of Electron Transport Chain Complex I; AAC: ADP/ATP translocator; ASCTB: acetate:succinate CoA-transferase, subtype 1B; E, F, G: [FeFe]-hydrogenase maturases HydE, HydF and HydG, respectively; FER^O: oxidized electron transport ferredoxin; HYD: FER^R: reduced electron transport ferredoxin; HYD: [FeFe]-hydrogenase; PFO: pyruvate:ferredoxin oxidoreductase; SCS: succinyl-CoA synthetase. Carriers (lilac): Aral: Aralar solute carrier; MCF: Mitochondrial Carrier Family protein. Ketone body degradation (light green): ACAT: acetyl-CoA C-acetyltransferase; SCOT: succinate:3-oxoacid CoA-transferase. ROS and oxygen stress (dark green): FDP: flavodiiron protein; SOD: superoxide dismutase; TRX: thioredoxin; TRXP: thioredoxin peroxidase. Mitochondrial protein import and folding (yellow): BCS: ubiquinol-Cytochrome c reductase Synthesis; CPN: Chaperonin; HSP: Heat Shock Protein; MET: Metaxin; MGE: Mitochondrial GrpE; MIA: Mitochondrial intermembrane space Import and Assembly; MPP: Mitochondrial Processing Peptidase; PHB: Prohibitin; SAM: Sorting and Assembly Machinery; TIM: Translocator of the Inner Membrane; TOM: Translocator of the Outer Membrane; 8, 9, 10, 13: Translocator of the Inner Membrane proteins Tim8, Tim9, Tim10, Tim13, respectively. Iron–sulfur cluster assembly (orange): ATM: ABC transporter, Mitochondrial; FRX: Frataxin; GRX: Glutaredoxin; IND: iron-sulfur protein required for NADH dehydrogenase; ISC: Iron-Sulfur Cluster; ISD: Iron-Sulfur protein biogenesis, Desulfurase-interacting; JAC: J-type Accessory Chaperone; NFU: NifU-like; SSQ: Stress Seventy subfamily Q; YAH: Yeast Adrenodoxin Homolog. Amino acid metabolism (light blue): AGT: alanine–glyoxylate aminotransferase; ALT: alanine aminotransferase; AST: aspartate aminotransferase; DLDH: D-lactate dehydrogenase; GCS: Glycine Cleavage System; H, L, P, T: H-, L-, P- and T-protein components of the glycine cleavage system, respectively; SHMT: serine hydroxymethyltransferase; ACAD: branched-chain acyl-CoA dehydrogenase; BαKDH: branched-chain α-ketoglutarate dehydrogenase complex; BCAAT: branched-chain amino acid aminotransferase; ECH: enoyl-CoA hydratase; MCE: methylmalonyl-CoA epimerase; MCM: methylmalonyl-CoA mutase; MSDH: methylmalonyl semialdehyde dehydrogenase; PCC: propionyl-CoA carboxylase. Electron transport (dark blue): AOX: alternative oxidase; ETF: electron transport flavoprotein; ETFDH: electron transport flavoprotein dehydrogenase; G3PD: glycerol-3-phosphate dehydrogenase. Other (white): CDP-DAG: cytidine diphosphate diacylglcerol; CMP: cytidine monophosphate; CL: cardiolipin; CLS: cardiolipin synthase; LPLA: lipoate protein ligase; PA: phosphatidic acid; PNT: pyridine nucleotide transhydrogenase; TAM: translocator assembly and maintenance.

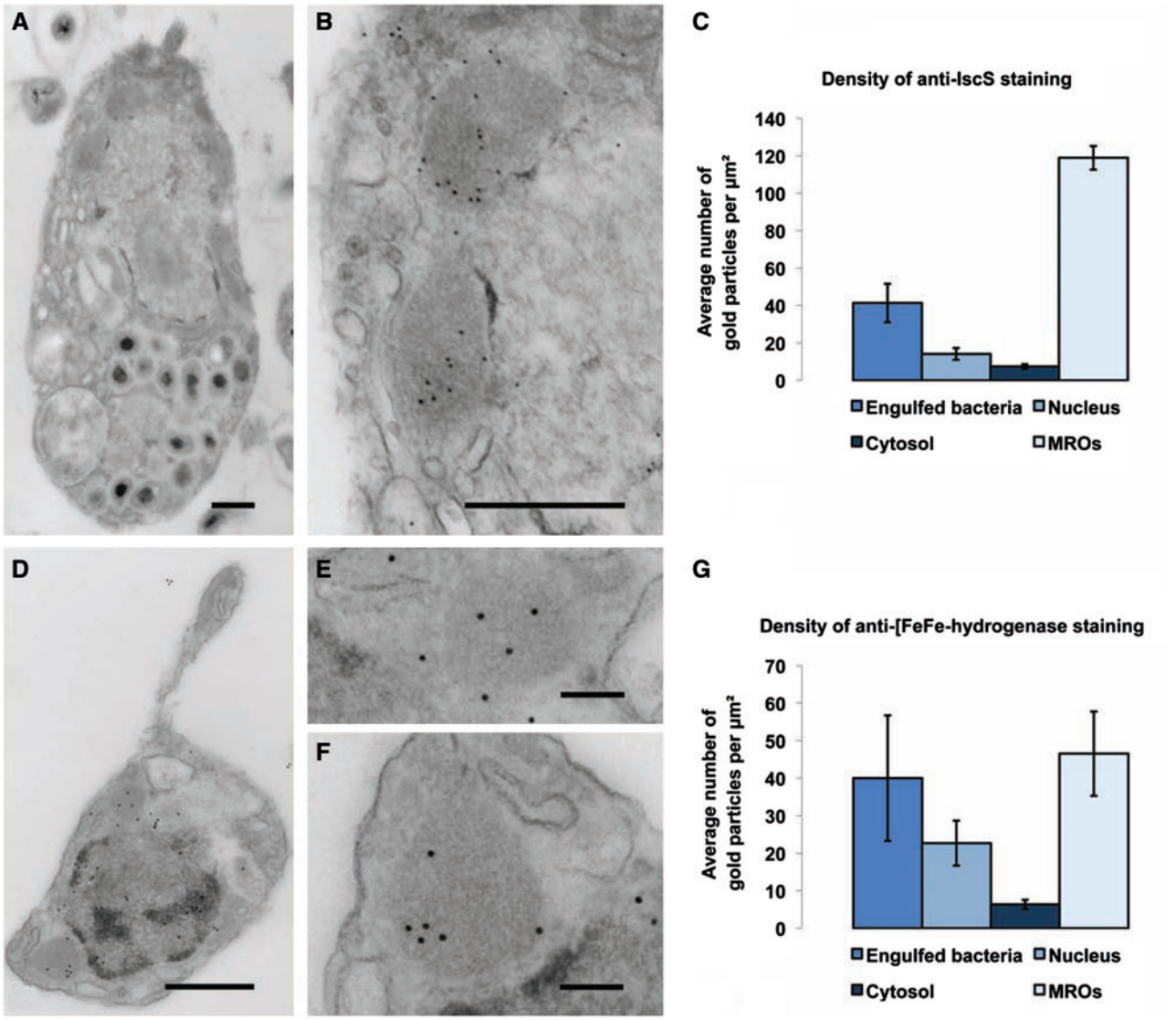

Fig. 3. Immunogold localization of IscS and [FeFe]-hydrogenase in *S. incarcerata* cells. (*A*) A representative whole cell fixed for immunogold staining; scale bar, 500 nm; this image has been cropped in order to show only the whole cell from which the insets shown were derived. (*B*) Magnified section from the cell pictured in (*A*), showing gold particles corresponding to anti-*Giardia* IscS localization; scale bar, 500 nm. (*C*) Mean density of immunogold labeling in engulfed bacteria, the cytosol, the nucleus, and MROs (seven cells), ± standard error of the mean. (*D*) Representative whole cell fixed for immunogold staining; scale bar, 500 nm; this image has been cropped in order to show only the whole cell from which the insets shown were derived. (*E*, *F*) Magnified sections from the cell pictured in (*A*), showing gold particles corresponding to anti-*Mastigamoeba* [FeFe]-hydrogenase localization; scale bar, 500 nm. (*G*) Mean density of immunogold labeling in engulfed bacteria, the cytosol, the nucleus, and MROs (13 cells), ± standard error of the mean. Brightness and contrast have been adjusted in each image to enhance visibility of the mitochondria and gold particles.

these genes between anaerobic eukaryotes, which presents an attractive explanation for an increasingly common pattern of genes shared only between anaerobic eukaryotes (Andersson et al. 2003, 2006; Stairset al. 2011, 2014; Tsaousis, Ollagnier de Choudens, et al. 2012). It should be noted that the latter hypothesis does not exclude the possibility of anaerobic energy generation enzymes originally being present in the protomitochondrial endosymbiont, but suggests that, if they were originally present, they have since been lost or replaced in extant eukaryotes (Tsaousis, Leger, et al. 2012).

In either case, these analyses highlight the importance of two factors. The first is the utility of gathering more data, shown by the increase in the representation of both eukaryotic and alphaproteobacterial taxa in our trees relative to even recent studies (Hug et al. 2010; Leger et al. 2013; Stairs et al. 2014); improved taxonomic sampling in the future may significantly alter or strengthen the inferences that can be made from these phylogenies. The second is that the periplasmic hydrogenases (and their eukaryotic homologs), although they may be functionally equivalent to the nonperiplasmic hydrogenases, may not be orthologs, and that caution should be used when grouping them together in analyses, as has been done in the past (Hug et al. 2010). AU tests performed on groupings of alphaproteobacteria and eukaryotes, for instance, will likely generate spurious results if periplasmic and nonperiplasmic hydrogenases are treated

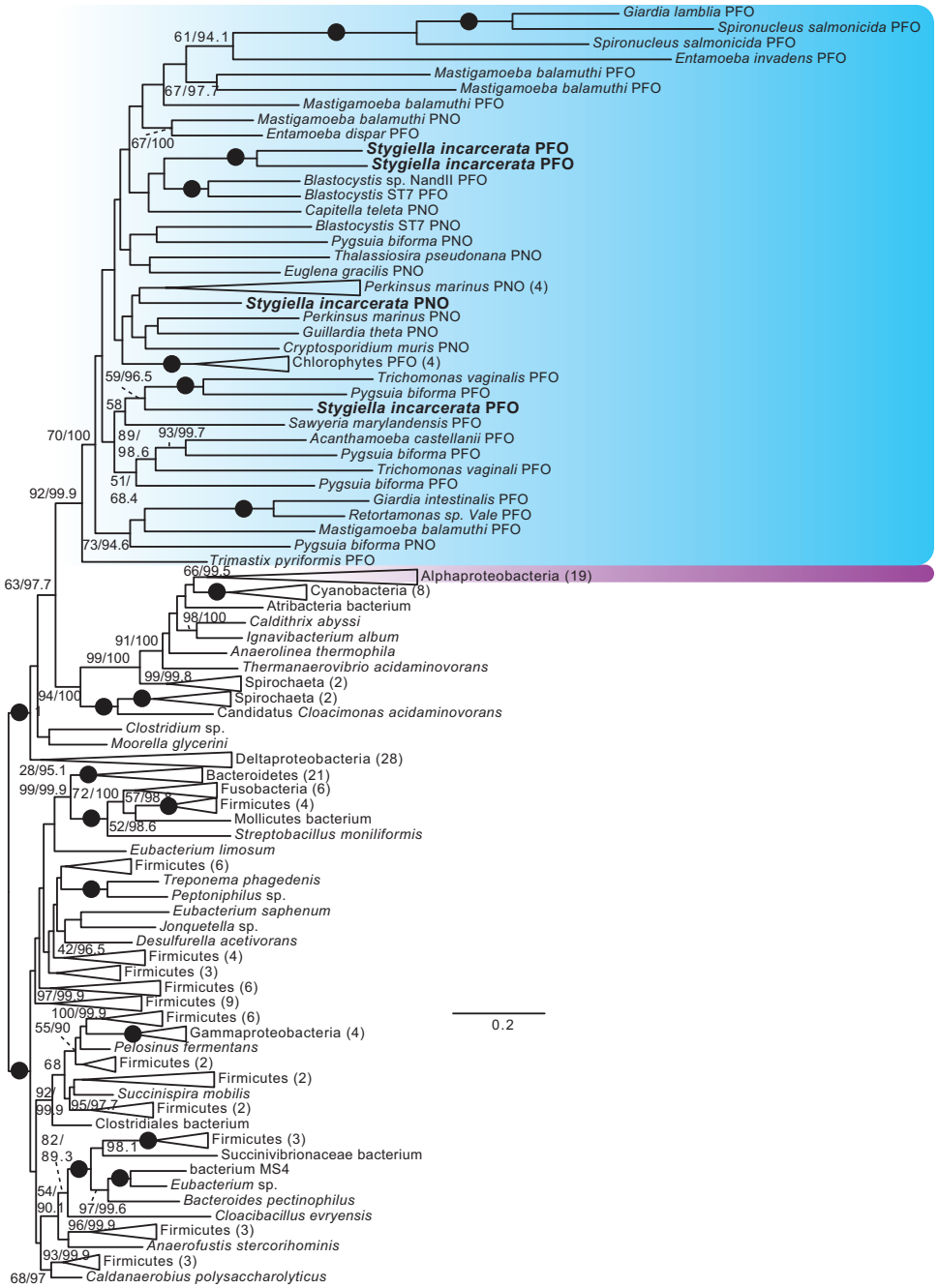

Fig. 4. Unrooted ML tree of PFO sequences. Phylogenetic analyses were performed on 212 sequences and 958 sites, using RAxML v.8.0.23 under the PROTGAMMALG4X model of amino acid substitution, and IQ-TREE v.1.3.11 under the C20+Gamma model of evolution. Bootstrap support values >50% (first number), and ultrafast bootstrap values >95% (second number), are shown next to relevant branches; both support values are shown even where only one is greater than the chosen cutoff. Branches with 100% bootstrap and ultrafast bootstrap support are indicated by black circles. Eukaryotes are shaded blue, and alphaproteobacteria magenta. Eukaryotic PFO and PNO sequences are indicated. For ultrafast bootstrap values, 95 was chosen as a cutoff value as per the recommendation of the IQ-TREE manual; for the ML tree under the C20+gamma model found using IQ-TREE, with full ultrafast bootstrap support values, see supplementary figs. S12–S24, Supplementary Material online.

as a single group, as these enzymes may have quite different origins in eukaryotes.

Iron–Sulfur Cluster Assembly

Iron–sulfur (Fe–S) cluster biosynthesis is known to be the only truly essential function of yeast mitochondria (Lill et al. 1999). The mitochondrial ISC Fe–S cluster assembly pathway is a crucial and conserved function of most MROs described to date (LaGier et al. 2003); it has been suggested that the need to provide an environment for this process explains the retention of the most reduced MROs (Embley et al. 2003; Tovar et al. 2003). Nevertheless, several examples of MROs are known in which the mitochondrial ISC system has been replaced by bacterial NIF systems [*Entamoeba histolytica*

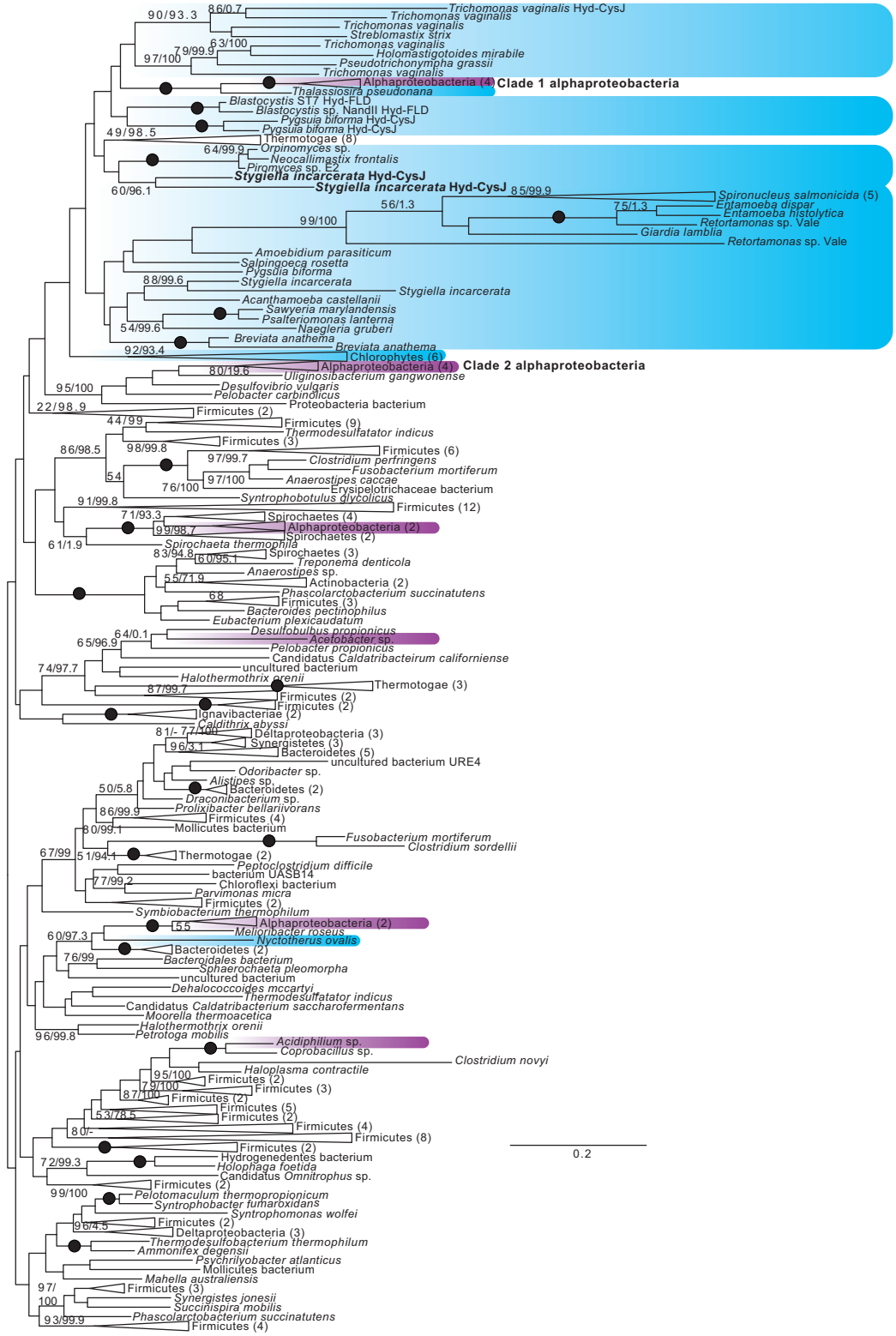

Fig. 5. Unrooted ML tree of nonperiplasmic-like [FeFe]-hydrogenase sequences. Phylogenetic analyses were performed on 238 sequences and 346 sites, using RAxML and IQ-TREE. Bootstrap and ultrafast bootstrap values are indicated as in figure 4. Eukaryotes are shaded blue, and alphaproteobacteria magenta. [FeFe]-hydrogenases with C-terminal CysJ or flavodoxin (FLD) domains are indicated.

(Ali et al. 2004; van der Giezenet al. 2004; Maralikova et al. 2010), *M. balamuthi* (Gill et al. 2007; Nyvltova et al. 2013)] or a SUF system [*P. biforma* (Stairs et al. 2014)].

The main components of the Fe–S cluster assembly pathway were identified in the transcriptomic survey of *S. incarcerata* (fig. 2, brown). Most notable are IscS, a cysteine desulfurase that generates sulfur destined for cluster assembly and the scaffold proteins IscU, IscA1 and IscA2 (Lill et al. 2012). Accessory proteins found include the mitochondrial transporter Atm1, as well as the iron donor frataxin, and

Table 1. Approximately Unbiased Tests of Alternate Topologies.

Hypothesis tested	AU test *P*-value
[FeFe]-hydrogenase and NARF-like (NARFL) proteins	
ML tree under the LG4X+Gamma model	0.744
ML tree under the C20+Gamma model	1e^{-09}[b]
Monophyly of eukaryotic hydrogenases and NARFL proteins; *Thalassiosira* and *Nyctotherus* unconstrained	2e^{-10}[b]
Monophyly of eukaryotic hydrogenases; NARFL proteins unconstrained; *Thalassiosira* and *Nyctotherus* unconstrained	0.001[b]
Monophyly of eukaryotic non-periplasmic-like hydrogenases; eukaryotic periplasmic-like hydrogenases and NARFL proteins unconstrained; *Thalassiosira* and *Nyctotherus* unconstrained	0.504
Monophyly of eukaryotic non-periplasmic-like hydrogenases and Clade 1 and 2 alphaproteobacteria; eukaryotic periplasmic-like hydrogenases and NARFL proteins unconstrained; all other alphaproteobacteria unconstrained; *Thalassiosira* and *Nyctotherus* unconstrained	0.266
Monophyly of eukaryotic non-periplasmic-like hydrogenases, and Clade 1 alphaproteobacteria basal to eukaryotic non-periplasmic-like hydrogenases; eukaryotic periplasmic-like hydrogenases and NARFL proteins unconstrained; all other alphaproteobacteria unconstrained; *Thalassiosira* and *Nyctotherus* unconstrained	0.459
[FeFe]-hydrogenase	
ML tree under the LG4X+Gamma model	0.704
ML tree under the C20+Gamma model	5e^{-08}[b]
Monophyly of eukaryotic hydrogenases; *Thalassiosira* and *Nyctotherus* unconstrained	0.002[b]
Monophyly of eukaryotic non-periplasmic-like hydrogenases; *Thalassiosira* and *Nyctotherus* and periplasmic hydrogenases unconstrained	0.555
Monophyly of eukaryotic non-periplasmic-like hydrogenases and Clade 1 and 2 alphaproteobacteria; *Thalassiosira* and *Nyctotherus* unconstrained; eukaryotic periplasmic hydrogenases unconstrained; all other alphaproteobacteria unconstrained	0.042[a]
Clade 1 alphaproteobacteria basal to eukaryotic non-periplasmic-like hydrogenases; *Thalassiosira* and *Nyctotherus* unconstrained; Clade 2 eukaryotes unconstrained; all other alphaproteobacteria unconstrained	0.504
[FeFe]-hydrogenase (non-periplasmic-like)	
ML tree under the LG4X+Gamma model	0.680
ML tree under the C20+Gamma model	0.016[a]
Eukaryote monophyly; *Thalassiosira* unconstrained	0.627
Monophyly of eukaryotes and Clade 1 and 2 alphaproteobacteria; *Thalassiosira* and *Nyctotherus* unconstrained; all other alphaproteobacteria unconstrained	0.406
Clade 1 alphaproteobacteria basal to eukaryotes; *Thalassiosira* and *Nyctotherus* unconstrained; all other alphaproteobacteria unconstrained	0.576
Clade 2 alphaproteobacteria basal to eukaryotes; *Thalassiosira* and *Nyctotherus* unconstrained; all other alphaproteobacteria unconstrained	0.362
PFO	
ML tree under the LG4X+Gamma model	1.000
ML tree under the C20+Gamma model	0.057
Monophyly of eukaryotes and alphaproteobacteria	1e^{-05}[b]
Alphaproteobacteria basal to eukaryotes	1e^{-05}[b]
PFO+sulfite reductase	
ML tree under the LG4X+Gamma model	0.999
ML tree under the C20+Gamma model	0.003[b]
Monophyly of eukaryotes and alphaproteobacteria	0.013[a]
Alphaproteobacteria basal to eukaryotes	0.035[a]
ASCT1B	
ML tree under the LG4X+Gamma model	0.734
ML tree under the C20+Gamma model	0.036[a]
Eukaryote monophyly	1e^{-33}[b]
Monophyly of Clade 1 eukaryotes and Clade 2 eukaryotes; other eukaryotes unconstrained	0.716
Monophyly of Clade 1, 2 and 3 eukaryotes; other eukaryotes unconstrained	0.601
Monophyly of Clade 2 and Clade 3 eukaryotes; other eukaryotes unconstrained	0.549
Monophyly of Clade 2 and Clade 4 eukaryotes; other eukaryotes unconstrained	8e^{-07}[b]
Monophyly of Clade 1, 2 and 4 eukaryotes; other eukaryotes unconstrained	7e^{-06}[b]
Monophyly of Clade 1 eukaryotes and alphaproteobacteria; other eukaryotes unconstrained	0.548
Monophyly of Clade 2 eukaryotes and alphaproteobacteria; other eukaryotes unconstrained	0.525
Flavodiiron protein	
ML tree under the LG4X+Gamma model	0.702
ML tree under the C20+Gamma model	1e^{-06}[b]
T. vaginalis sequence 2 constrained with *Mastigamoeba*, *Entamoeba*, *Sawyeria* and *Stygiella*	0.521
Eukaryote monophyly; *T. vaginalis* 2 unconstrained	0.001[b]

[a]Rejected with *P*-value <0.05.
[b]Rejected with *P*-value <0.01.

ferredoxin Yah1, which maintain a redox balance, completing the complement of Fe–S assembly proteins (for the complete set of proteins found in this pathway, see supplementary table S3, Supplementary Material online). We confirmed the localization of this typical mitochondrial pathway to the putative MROs using immunogold labeling (fig. 3), using a heterologous polyclonal primary antibody raised in rabbit against *Giardia* IscS (Tovar et al. 2003) and a secondary antibody conjugated to gold particles. As expected, we found a greater (16-fold higher) density of gold particles in the MROs compared with the cytosol; the antibody also cross-reacted with engulfed bacteria, consistent with the presence of IscS homologs in bacteria.

Protein Transport and Folding

Major components of the mitochondrial protein import machinery were represented in the transcriptome data (fig. 6). These include the Translocase of the Outer Membrane (TOM), the Translocase of the Inner Membrane 22 (TIM22), the Translocase of the Inner Membrane 23 (TIM23), and the Sorting and Assembly Machinery (SAM) complexes, as well as the ER-mitochondria encounter structure (ERMES). Several accessory subunits of these complexes that are present in *Saccharomyces cerevisiae* were not recovered, including components of the TOM and TIM22 complexes, and the Inner Membrane Proteases (Imp1/2). Less surprisingly, the mitochondrial oxidase assembly protein Oxa1 is also absent; this protein mediates the membrane insertion of mitochondrially encoded Cox1p and Cox3p (Hell et al. 2001), neither of which are found in *S. incarcerata*. All of these import proteins are also absent from *T. vaginalis* (Carlton et al. 2007). However, *S. incarcerata* has retained a larger mitochondrial protein import complement relative to *T. vaginalis*, including the intermembrane space chaperone proteins known as Tiny Tims—Tim8, Tim9, Tim10, and Tim13, as well as the mitochondrial intermembrane space import and assembly (MIA) protein Mia40 (fig. 6).

Many of the matrix proteins involved in folding the imported proteins are present (fig. 2, yellow, fig. 6). They include mtHsp70 and Cpn60, which are typical mitochondrial markers involved in protein transport and refolding, and both the α and β-subunits of the MPP, the enzyme that cleaves N-terminal targeting peptides after protein import into the organellar matrix.

As the majority of the proteins involved in protein import are highly divergent between taxa, *S. incarcerata* homologs of some proteins may simply not have been retrieved by our methods. However, mitochondrial protein import has been poorly studied in eukaryotes other than yeast and mammals, and recent surveys that include data from more diverse taxa (Eckers et al. 2012; Heinz and Lithgow 2013; Gawryluk et al. 2014; Stairs et al. 2014) suggest that the apparently 'reduced' complement found in *S. incarcerata* may in fact be more representative of eukaryotes overall. Despite uncertainly arising from the incomplete nature of transcriptome data, particularly in *Sawyeria* (Barberà et al. 2010) and *Psalteriomonas* (de Graaf et al. 2009), a contrast is apparent between the apparently smaller number of protein import machinery components in MROs of parasites such as diplomonads, *Entamoeba* and microsporidia (Dagley et al. 2009; Waller

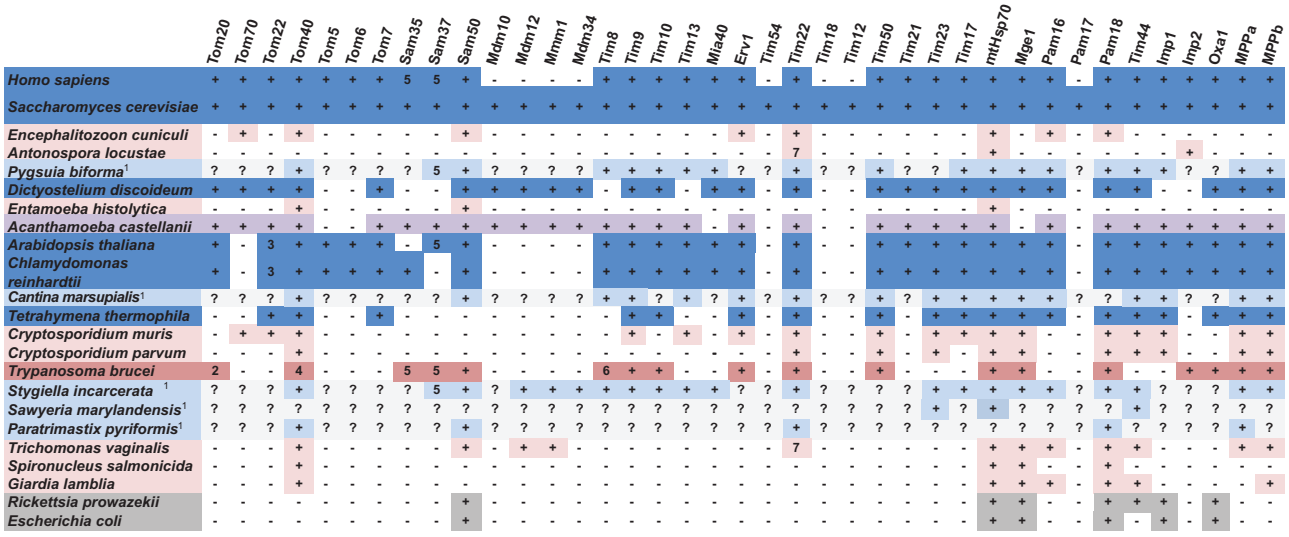

Fig. 6. Mitochondrial protein import proteins in various taxa. Dark blue: free-living mitochondrion; pale blue: free-living MRO; dark pink: parasite, mitochondrion; light pink, parasite, MRO; purple: facultative parasite, mitochondrion; dark grey: bacterium; pale grey: incomplete data. Data from Neupert (1997), Hoogenraad et al. (2002), Henriquez et al. (2005), Regoes et al. (2005), van der Laan et al. (2005); Williams and Keeling (2005), Dolezal et al. (2006), Figueroa-Martinez et al. (2008), Atteia et al. (2009), Dagley et al. (2009), Waller et al. (2009), Barberà et al. (2010), Dolezal et al. (2010), Jedelsky et al. (2011), Liu et al. (2011), Schneider et al. (2011), Eckers et al. (2012), Pusnik et al. (2012), Basu et al. (2013), Heinz and Lithgow (2013), Jerlstrom-Hultqvist et al. (2013), Wideman et al. (2013), Zubacova et al. (2013), Gawryluk et al. (2014), Murcha, Kmiec, et al. 2014; Murcha, Wang, et al. 2014; Stairs et al. (2014), Martincova et al. (2015), Murcha et al. (2015), Noguchi et al. (2015), Wojtkowska et al. (2015), and Buczek et al. (2016) this study. [1]Transcriptome data, incomplete, particularly in the cases of *S. marylandensis* and *Paratrimastix pyriformis*; [2]pATOM36; [3]Tom9; [4]pATOM; [5]Metaxin [6]Tim8/13; [7]Tim17/22/23. For accession numbers and additional proteins, see supplementary table S4, Supplementary Material online.

et al. 2009; Dolezal et al. 2010; Jedelsky et al. 2011; Schneider et al. 2011; Jerlstrom-Hultqvist et al. 2013), and the apparently more complex protein import machinery of typical mitochondria (fig. 6). This difference may reflect an overall reduction in the protein import machinery in parasites, or a more divergent complement of proteins that are not readily identified using methods based on sequence data from model organisms (Makiuchi et al. 2013; Martincova et al. 2015). As with the free-living organisms *Pygsuia* (Stairs et al. 2014) and the stramenopile *C. marsupialis* (Noguchi et al. 2015), the protein import machinery of *S. incarcerata* more closely resembles that of typical mitochondria. While more genomic data from organisms with MROs are needed to confirm it, this trend suggests that a more reduced or poorly conserved MRO protein import apparatus reflects a parasitic lifestyle, rather than necessarily accompanying the derivation of MROs from aerobic mitochondria.

The only other jakobid in which the mitochondrial protein import apparatus has been studied in some detail is *Reclinomonas americana*. This organism is notable in that, in addition to components of the canonical mitochondrial protein import pathway, it includes subunits of the bacterial SecY and twin-arginine transport complexes, which are involved in protein insertion into and secretion across the inner membrane (Tong et al. 2011). Subsequently, these proteins have been found to be mitochondrially encoded in other jakobids, including the sister-taxon to *S. incarcerata*, *A. godoyi* (Burger et al. 2013), and in the discobid *Tsukubamonas globosa* (Kamikawa et al. 2014). No nuclear-encoded homologs of these components have so far been found in eukaryotes, and it appears that they were lost from *S. incarcerata* along with the mitochondrial genome.

Amino Acid Metabolism

Genes encoding all components of the glycine cleavage system were identified from the transcripts, including the L, P, T, and H proteins and serine hydroxymethyl transferase (SHMT) (fig. 2, blue; fig. 7). Targeting peptides have been identified for all of these proteins, providing evidence that this pathway is active within the organellar compartment. Whereas only partial glycine cleavage systems are present in the MROs of the parasites *T. vaginalis* (Schneider et al. 2011) and *S. salmonicida* (Jerlstrom-Hultqvist et al. 2013), a complete glycine cleavage system is predicted to function in the MROs of *Trimastix pyriformis* (Zubacova et al. 2013), *M. balamuthi* (Nyvltova et al. 2015), and *P. biforma* (Stairs et al. 2014). The presence of a glycine cleavage system seems to be the rule, rather than the exception, in the MROs of free-living anaerobes.

All of the key enzymes in the initial stages of branched-chain amino acid degradation, as well as most downstream enzymes for the further breakdown of leucine, isoleucine, and valine were identified in *S. incarcerata* (fig. 2, blue), and have predicted targeting peptides. While enzymes carrying out the early steps of branched-chain amino acid degradation have been described in the *T. vaginalis* genome (Carlton et al. 2007), this pathway is not hypothesized to be present in the hydrogenosome (Carlton et al. 2007; Schneider et al. 2011), despite their mitochondrial localization being typical in aerobic organisms.

In addition to these proteins, we identified enzymes involved in alanine, arginine, glutamate, lysine, tryptophan, methionine, and tyrosine metabolism (fig. 7). These are largely absent from the reduced amino acid metabolism complement of *Trichomonas* MROs (Schneider et al. 2011) (and completely absent from the much more reduced proteomes of *Giardia* and *Entamoeba* MROs; Mi-ichi et al. 2009; Jedelsky et al. 2011). A number of these enzymes, including branched-chain amino acid degradation enzymes, are also present in the predicted MRO proteomes of *Pygsuia* (Stairs et al. 2014) and *Cantina* (Noguchi et al. 2015). This is consistent with the hypothesis that the MROs of free-living organisms retain more mitochondrial metabolic functions than those of parasites (Stairs et al. 2015). Nevertheless, *S. incarcerata* MROs possess a larger predicted complement of amino acid metabolic enzymes than *Pygsuia* MROs, notably those involved in valine, isoleucine and leucine degradation, and in glutamate metabolism (fig. 7). In some cases (e.g., tryptophanase, branched-chain amino acid aminotransferase, glutamate dehydrogenase, cysteine synthase), the genomes of *Spironucleus*, *Giardia*, and *Trichomonas* are known to encode amino acid metabolism enzymes homologous to those of prokaryotes, that are predicted to be active in the cytosol rather than in MROs, and that are predicted to have been acquired by lateral gene transfer (Andersson and Roger 2003; Andersson et al. 2003; Alsmark et al. 2013) (although most of these are not present in the detected or predicted MRO proteome of *Stygiella*; fig. 7). This suggests that the reduced MRO proteome of parasitic organisms may partly result from acquisition of homologs from prokaryotes and overall gene loss (a pattern commonly observed in parasites of animals; Leckenby and Hall 2015), rather than simple retargeting of ancestrally mitochondrial proteins to the cytosol.

Other Functions and Pathways

As previously reported (Leger et al. 2015), the *S. incarcerata* transcriptome encodes two homologs of the bacterial and mitochondrial division protein FtsZ, and homologs of three proteins, MinC, MinD and MinE, that, in bacteria, control the placement of FtsZ (Leger et al. 2015).

The phospholipid cardiolipin is a key component of both prokaryotic and mitochondrial inner membranes; in aerobic eukaryotes, it is synthesized in mitochondria. Cardiolipin synthesis proteins were until recently believed to be absent from MRO-bearing protists (Tian et al. 2012), but a bioinformatic survey of the *Pygsuia* MRO proteome uncovered the first evidence of a complete cardiolipin biosynthesis pathway in an MRO (Stairs et al. 2014). Our survey uncovered two enzymes in this pathway, a eukaryotic transferase-type cardiolipin synthesis and the CDP-diacylglycerol synthase Tam41 (Tamura et al. 2013). We found no evidence of phosphatidylglycerol phosphate (PGP) synthase or PGP phosphatase, which are involved in phosphatidylglycerol formation from CDP-diacylglycerol. In the absence of these enzymes, phosphatidylglycerol might be taken up from bacterial food sources; alternatively, homologs of these proteins present in *S.*

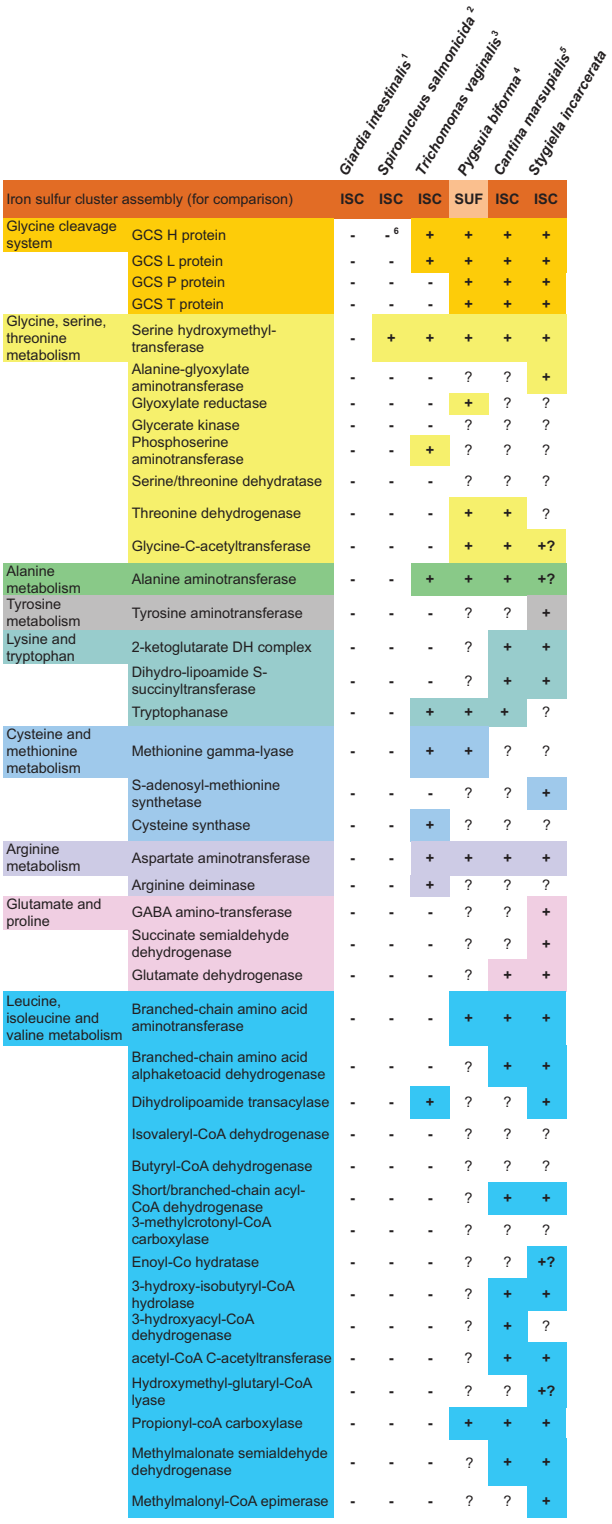

FIG. 7. Amino acid metabolism in MROs. [1]Source: (Jedelsky et al. 2011). [2]Source: (Jerlstrom-Hultqvist et al. 2013). [3]Source: (Schneider et al. 2011). [4]Source: (Stairs et al. 2014). [5]Source: (Noguchi et al. 2015). [6]Protein described in the literature as Glycine Cleavage System H protein-like, but which is not a reciprocal best BLAST hit to other eukaryotic H proteins. +: Predicted to be present in the MRO based on proteomic data or bioinformatic predictions. +? Incomplete sequence at the N-terminus, making localization prediction uncertain. Orange, FeS cluster assembly; Bright yellow, glycine cleavage system; pale yellow, glycine, serine and threonine metabolism; pink, alanine metabolism; brown, tyrosine metabolism; pale blue, lysine and tryptophan metabolism; green, arginine metabolism; purple, glutamate and proline metabolism; dark blue, leucine, isoleucine and valine (branched-chain amino acid) metabolism. For accession numbers, see supplementary table S5, Supplementary Material online.

incarcerata might simply not have been recovered in our transcriptome survey.

Recently, a novel, fused SufCB enzyme was identified in *Blastocystis* sp.—the first evidence for a Suf system enzyme in a eukaryote. While the ISC system is present in the MROs of *Blastocystis* sp., SufCB was expressed in the cytosol (Tsaousis, Ollagnier de Choudens, et al. 2012). Homologs of this enzyme were subsequently found to be encoded in the transcriptome of the anaerobic breviate *P. biforma*, which possesses both an MRO-targeted and a cytosolic copy (Stairs et al. 2014). Based on phylogenetic studies, the enzyme was hypothesized to have been laterally acquired from an archaeon related to the Methanomicrobiales, and subsequently transferred laterally from one eukaryote to the other (Tsaousis, Ollagnier de Choudens, et al. 2012; Stairs et al. 2014). We recovered a sequence encoding a homolog of this enzyme from our transcriptome data. The *S. incarcerata* enzyme displays the same fusion of the prokaryotic SufC and SufB proteins as those found in *Blastocystis* sp. and *P. biforma*; the enzyme lacks a targeting peptide, suggesting that it is cytosolic (supplementary table S3, Supplementary Material online). The genomic sequence of the enzyme that we amplified from *S. incarcerata* cultures does not include introns; however, it is also present in transcriptome data derived from the closely related jakobid *Velundella trypanoides* (Panek T and Cepicka I, personal communication), suggesting that it is not a prokaryotic contaminant. Because the fusion protein has only been identified in *Blastocystis*, *Pygsuia*, and *S. incarcerata*, we initially performed separate phylogenetic analyses of SufC and SufB in prokaryotes and eukaryotes (supplementary figs. S8 and S9, Supplementary Material online). The two analyses recovered broadly congruent results. We then performed analyses using concatenated SufC and SufB. In all three analyses, the three eukaryotes formed a well-supported clade, with Methanomicrobiales as a well-supported sister group (fig. 8, supplementary figs. S8 and S9, Supplementary Material online). Together with the fact that the fusion protein is found only in a small number of distantly-related eukaryotes (a stramenopile, an obazoan, and an excavate), this result is in line with those of previous studies positing an initial transfer from an archaeal source into a microbial eukaryote, followed by lateral transfer to other eukaryotes. In *Blastocystis*, SufCB is upregulated under oxygen stress, and it has been speculated that the enzyme provides a mechanism for the repair of oxygen-sensitive FeS proteins on exposure to oxygen (Tsaousis, Ollagnier de Choudens, et al. 2012). Because *Stygiella*, like *Blastocystis* and *Pygsuia*, inhabits anoxic environments, it seems likely that its SufCB has a similar function.

Like MROs of *Trichomonas* (Schneider et al. 2011), *Pygsuia* (Stairs et al. 2014), and *Cantina* (Noguchi et al. 2015), but unlike those of *Giardia* (Brown et al. 1995), *S. incarcerata*

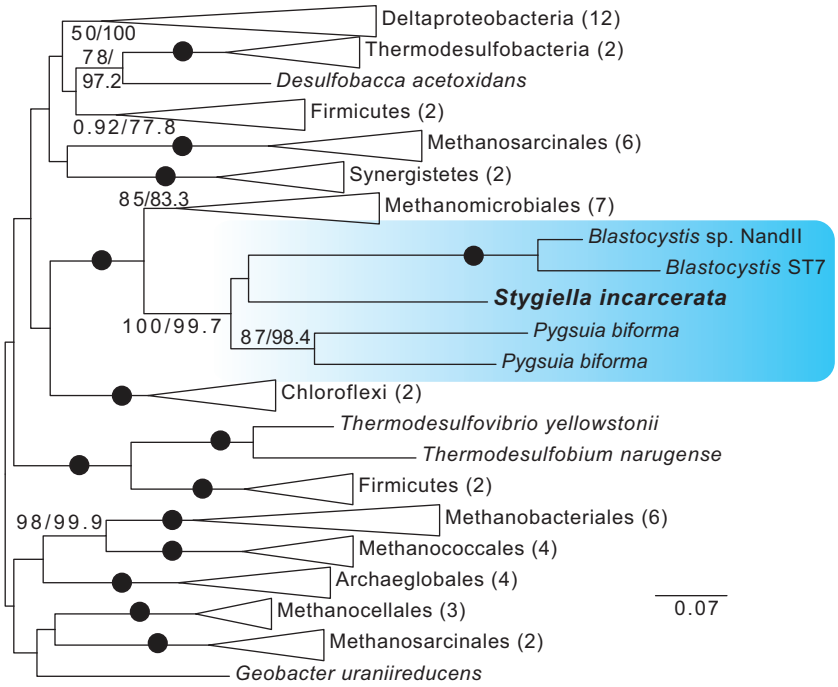

Fig. 8. Unrooted ML tree of concatenated SufC and SufB sequences. Phylogenetic analyses were performed on 63 sequences and 547 sites, using RAxML and IQ-TREE. Bootstrap and ultrafast bootstrap values are indicated as in figure 4. Eukaryotes are shaded in blue.

organelles have retained mitochondrial reactive oxygen stress response proteins (superoxide dismutase, thioredoxin, and thioredoxin peroxidase; fig. 2). In addition, *S. incarcerata* possesses homologs of a flavodiiron protein commonly found in anaerobic prokaryotes, as well as some anaerobic protists (Andersson et al. 2003); previously described in *Giardia* and *Trichomonas* as having a high affinity for O_2, this enzyme has been hypothesized to be involved in oxygen detoxification in eukaryotes (Di Matteo et al. 2008; Smutna et al. 2009; Vicente et al. 2009; Vicente et al. 2012). *S. incarcerata* possesses two paralogs of this protein that differ in their predicted localization and phylogenetic affinities. The first is predicted to be cytosolic, and groups together with *Entamoeba*, *Mastigamoeba*, and *Sawyeria* homologs. The second possesses a predicted mitochondrial targeting peptide, and forms a clade with *Pygsuia*, *Breviata*, *Giardia*, *Trichomonas*, and *Spironucleus*. Consistent with earlier results suggesting at least two eukaryotic acquisitions from prokaryotes (Andersson et al. 2003, 2006), AU tests reject the hypothesis of eukaryote monophyly for this flavodiiron protein (P-value <0.01) (table 1; supplementary fig. 10, Supplementary Material online). Lateral transfer of the protein between eukaryotes likely followed the initial acquisitions; it is particularly striking that *S. incarcerata* has acquired paralogs from two distinct sources. Interestingly, the MRO-targeted *S. incarcerata* protein is a fusion protein, possessing a C-terminal ferredoxin domain. This arrangement, which was confirmed by PCR amplification, appears to be unique to *S. incarcerata*. Flavodiiron proteins typically accept electrons from rubredoxins (which in turn accept electrons from NADH via a flavoprotein) (Di Matteo et al. 2008; Leger et al. 2015). Rubredoxins do not appear to be present in *S. incarcerata*'s MRO; the fused ferredoxin likely fulfills this role for the *S. incarcerata* organellar flavodiiron protein.

Conclusions

The availability of low-cost, high-coverage sequencing in recent years is making it possible to predict the proteomes of mitochondria and their related organelles on a much broader taxonomic and biochemical scale than was previously feasible. It is now possible to investigate the true diversity and flexibility of these organelles, rather than focusing mainly on energy generating functions in a limited number of taxa, as done previously. The examination of MROs in free-living anaerobes and microaerophiles is allowing us to separate adaptations to anoxia from byproducts of a parasitic lifestyle. The presence of a hydrogenosome-like organelle in a lineage closely related to organisms recognized for the unusual, primitive features of their mitochondria presents a particularly interesting opportunity. The complete absence of aerobic energy generation machinery and mitochondrial genome-encoded genes highlights the dramatic changes that can rapidly occur in the course of MRO derivation. Because Stygiellidae emerge from within jakobids, which possess classical mitochondria with genomes and electron transport chains, this absence provides evidence of the independent origins of MROs in *S. incarcerata*.

A comparison of the MROs of *S. incarcerata* with those of other free-living and parasitic taxa reveals significant similarities as well as differences. Like the well-characterized MROs of *Trichomonas*, *S. incarcerata* MROs house an oxygen scavenging system for maintaining an anaerobic environment, and a anaerobic ATP generation pathway involving PFO, [FeFe]-

hydrogenase, [FeFe]-hydrogenase maturases, two Complex I subunits and ASCT, but lack additional electron transport chain complexes or subunits; the differences in ASCT subtypes between the two organisms suggest either some degree of convergent adaptation, or differential loss from a common ancestor possessing both subtypes 1B and 1C. A leitmotif in *S. incarcerata* and other anaerobic eukaryotes (Andersson et al. 2006, 2009; Stairset al. 2011, 2014) is the presence of enzymes shared with a highly restricted number of other eukaryotes with a similar lifestyle (and complete absence from the majority of eukaryotic genomes), that form small eukaryotic clades in trees (e.g., flavodiiron proteins, SufCB, squalene-tetrahymanol cyclase; Takishita et al. 2012). In light of the vast diversity of eukaryotes that have been sampled in recent years, and that still remain to be sampled, we cannot exclude the possibility that these enzymes were present in a distant common ancestor of these disparate eukaryotic anaerobes. Nevertheless, the parsimonious explanation is that this pattern, which is found in both MRO and cytosolic proteins, points to inter-eukaryote LGT as a mechanism that enables lineages to rapidly adapt to new habitats.

However, in its complement of typical mitochondrial enzymes, such as those involved in amino acid catabolism, cardiolipin synthesis and mitochondrial protein import, the MROs of *S. incarcerata* more closely resemble those of much more distantly related, free-living organisms, namely the breviate *P. biforma* and the stramenopile *C. marsupialis*. These similarities may result from a shorter amount of time elapsed since the process of MRO emergence in these taxa relative to diplomonads, but more likely reflect their shared lifestyle as free-living anaerobes; parasites may have less need of these catabolic pathways, as they can derive more metabolites from the host.

As a transcriptomic survey, this study is necessarily incomplete. Further examination of *S. incarcerata*'s genome, and, ultimately, direct characterization of the proteome, will clarify the role of this organelle within the cell, as well as its relationship to classical mitochondria. Such in-depth examinations will undoubtedly yield useful insights into the order and timing of protein and pathway loss in MROs. In coming years, it will be exciting to examine the mitochondrial proteomes of closely-related jakobids such as *A. godoyi* in order to draw inferences about the origins of *S. incarcerata* MRO proteins, particularly those involved in anaerobic energy generation and oxygen detoxification.

Materials and Methods

Cell Culture

Stygiella incarcerata MB1 had previously been isolated from Mahone Bay, Nova Scotia, Canada (Simpson et al. 2008). Cells were maintained at 21 °C in a medium composed of 50:50 Page's modified Neff's Amoeba Saline (AS; 2 mM Na_2HPO_4, 2 mM KH_2PO_4, 0.03 mM $MgSO_4{\cdot}7H_2O$, 0.05 mM $CaCl_2{\cdot}2H_2O$, 4.1 mM NaCl):Artificial seawater (ASW; 0.42 M NaCl, 9 mM KCl, 9.25 mM $CaCl_2{\cdot}2H_2O$, 23 mM $MgCl{\cdot}6H_2O$, 25.5 mM $MgSO_4{\cdot}7H_2O$, 2.14 mM $NaHCO_3$) plus 3% Luria-Bertani (LB) broth (10 g tryptone, 5 g yeast extract, 10 g NaCl in 1 l). Cultures were maintained in 50 ml Falcon tubes filled to ~45 ml.

Expressed Sequence Tag (EST) Library Creation

Total RNA was extracted from high-density cultures of *S. incarcerata* using TRIzol Reagent (Invitrogen) following the manufacturer's instructions for cells grown in suspension. A 2-µg aliquot of high-quality total RNA was sent to Amplicon Express (Pullman, USA) for EST library creation. RNA underwent two rounds of poly-A selection, followed by reverse-transcription to complementary DNA (cDNA). cDNA reads were size-selected, and cloned into pBluescript II SK+.

The completed EST library was sent to the National Research Council (NRC), Halifax location for Sanger sequencing. A total of 3,456 clones were sequenced and processed using Phred and Phrap (Ewing and Green 1998; Ewing et al. 1998).

The library was then amplified, nebulized, and subjected to 454 GS-FLX Titanium pyrosequencing at the McGill University Genome Quebec Innovation Centre, producing 11,905 reads.

A mixed Sanger/454 assembly was produced using Mira 2.9.39 (Chevreux et al. 1999) under default parameters with the following adjustments: job = *de novo*,est,accurate, sanger,454.

Illumina RNASeq

In order to expand transcript coverage and numbers, total RNA was extracted from high-density cultures of *S. incarcerata* as described above, and sent to Macrogen (Seoul, Korea) for two rounds of poly-A selection, followed by Illumina HiSeq2000 paired-end sequencing. 63.4 million raw reads were produced, and assembled *de novo* using Trinity (r2011-07-13) (Grabherr et al. 2011). Contigs were compared with the existing Sanger/454 transcriptome to verify the accuracy of the transcripts. These transcripts were used in all subsequent analyses.

Identification and Analysis of Putative Organellar Protein Genes

A conservative set of sequences encoding possible MRO proteins was identified using CBOrg, a comparative tool for predicting MRO proteomes based on *H. sapiens*, *Sa. cerevisiae*, *Tetrahymena thermophila*, and *T. vaginalis* proteomes (Gaston et al. 2009). The sequences in this set were re-examined using BLAST (Altschul et al. 1997) searches against the nonredundant nucleotide and protein databases in GenBank to confirm homology to known mitochondrial proteins of other eukaryotes, and to eliminate likely bacterial contaminants. Additionally, hidden Markov Model (HMM) searches were used to identify proteins involved in protein import and translocation, as these are often divergent. This was done using HMMER 3.0 (http://www.hmmer.org; last accessed 31 May 2016), with HMMs constructed from multiple sequence alignments of homologous sequences provided by Dr. Trevor Lithgow and aligned using MUSCLE v3.8.31 (Edgar 2004, 2010). Additional tBLASTn searches were

made against the transcriptome using sequences from a dataset of 159 conserved eukaryotic proteins (Brown et al. 2013), and mitochondrially-encoded *A. godoyi* genes (Burger et al. 2013) (for tRNA gene queries, BLASTn searches were performed). The presence or absence of a full 5′ end of the gene was determined through translation of the sequence and identification of a putative initiator methionine and an in-frame stop codon upstream. BLAST homology was used to confirm the presence of a complete 5′ end compared to other known sequences for the gene in question.

In Silico Detection of N-Terminal Targeting Peptides

For sequences with a full 5′ end of the coding sequence, the translated protein sequence was screened for the presence of a potential N-terminal targeting peptide using the internet-based targeting peptide prediction programs TargetP (Emanuelsson et al. 2000, 2007) and Mitoprot (Claros 1995). Targeting peptides were designated as probable if organelle localization probability was above 50%, and if an N-terminal extension was present relative to the closest bacterial homologs.

Extension of Incomplete Genes and Intron Searching

Total RNA was extracted as described above. From a total RNA extract, messenger RNA (mRNA) was enriched using the Poly(A) Purist mRNA purification kit (Ambion Inc, TX). cDNA was generated from mRNA using the GeneRacer kit (Invitrogen Corp., USA). Total DNA was extracted by centrifuging high-density cultures, resuspending the pellets in extraction buffer (1 M Tris–HCl pH 7.5, 5 M NaCl, 0.5 M EDTA, 20% sodium dodecyl sulfate (SDS), incubating at 50°C for 10 min., and centrifuging for 10 min. at 10,000 × g. The resulting supernatant was subjected to two rounds of phenol:chloroform:isoamyl alcohol (25:24:1) (Fluka, Germany) extraction, followed by a third extraction in chloroform:isoamyl alcohol (24:1), and precipitation through addition of 1/10th volume 3 M sodium acetate and 6/10th volume of isopropanol Rapid Amplification of cDNA Ends PCR was used to obtain missing sequence data from MinC, which was not completely represented in the transcriptomic data, and standard PCR was used to obtain full-length sequence data from specific genes of interest to verify sequence information, or to search for introns in genomic DNA (for HydE, HydF, HydG, all [FeFe]-hydrogenases, all PFOs, ASCT1B, acyl-CoA synthetase, Grx5, SufCB, and ferredoxin-fused flavodiiron protein). Gene-specific primers for genes of interest were designed based on sequence information from the *S. incarcerata* transcriptome.

Routine Transmission Electron Microscopy (TEM)

Cells were fixed with 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer with 5% sucrose. Post-fixation was performed in 1% osmium tetroxide and 0.25% uranyl acetate. Samples underwent stepwise dehydration in acetone, before being embedded in Epon/Araldite resin. 100-nm sections were cut using an LKB Huxley ultramicrotome with a Diatome diamond knife, and placed on 300-mesh copper grids. Sections were stained in 2% aqueous uranyl acetate and lead citrate, viewed using a JEOL JEM 1230 Transmission Electron Microscope at 80 kV, and photographed using a Hamamatsu ORCA-HR digital camera.

Western Blotting

IscS, [FeFe]-hydrogenases predicted to localize to the MROs ([FeFe]-hydrogenase 1 and [FeFe]-hydrogenase 4), and the N-terminal [FeFe]-hydrogenase domain of the predicted cytosolic CysJ-fused [FeFe]-hydrogenases ([FeFe]-hydrogenase 2 and [FeFe]-hydrogenase 3) were amplified by PCR from *S. incarcerata* cDNA using primers with N-terminal restriction sites (NdeI and XhoI for IscS, and *Xho*I and *Bam*HI for [FeFe]-hydrogenase 1). [FeFe]-hydrogenase 1, the N-terminal domains of [FeFe]-hydrogenase 2 and [FeFe]-hydrogenase 3, and IscS were cloned into pET-16b (Novagen) downstream of a 6×His tag. [FeFe]-hydrogenase 4 was cloned into pET-28a downstream of a 6×His tag. The constructs were expressed from Rosetta 2 (DE3) competent cells (Novagen). The recombinant protein was expressed in inclusion bodies, which were purified using BugBuster reagent (Novagen).

Antibodies were tested against inclusion bodies from Rosetta 2 cells expressing recombinant *S. incarcerata* IscS or [FeFe]-hydrogenases from the pET-16b or pET-28a vectors, or the empty pET-16b vector and pET-28a vectors (as negative controls). The antibodies used were a polyclonal anti-*Giardia* IscS antibody (Tovar et al. 2003) kindly provided by Dr. Jan Tachezy (Charles University in Prague, Czech Republic), and a custom polyclonal anti-[FeFe]-hydrogenase antibody previously raised against a *M. balamuthi* [FeFe]-hydrogenase peptide (CPFGAVMTRSFMLDVMRAMRDSRSAGSKVVAMVAPA VAGHMGNAPIWSICEALKRAGFDEALEVSIGADTTTENEAHE FEERFGEGAPKGKFSFMTTSCCPAYVACVRKHVPEIEDAVSHT RSPMHYTAKLAKERWPGCTTVFVGPCTAKLHEASIDEYTDFA ITVVEALSLLRGRGVALDNSQ; Abgent). A monoclonal antibody raised against the polyhistidine tag (anti-His; Thermo Scientific) was used as a positive control (data not shown). Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, blotted onto polyvinylidene difluoride membranes, and blocked overnight at 4°C in 5% milk. The blots were then incubated with primary antibody at a ratio of 1:500 (anti-IscS) or 1:250 (anti-[FeFe]-hydrogenase) in 1% milk for 1 h at room temperature, washed, and incubated with horseradish peroxidase-conjugated secondary anti-rabbit IgG antibodies (Sigma) at 1:50,000, for 1 h at room temp. The blots were then washed again, incubated with electrochemiluminescence western blotting reagents (Amersham), and photographed using a FluorChem E imaging system (Protein Simple).

Immuno-Electron Microscopy (Immuno-EM)

Cells were fixed in 4% paraformaldehyde and 0.5% glutaraldehyde in 0.1 sodium cacodylate buffer, rinsed in 0.1 M sodium cacodylate buffer, subjected to stepwise dehydration in ethanol, and embedded in LR White Resin. 100-nm sections were cut using an LKB Huxley ultramicrotome with a Diatome diamond knife, and placed on 200 mesh nickel grids. Sections were blocked in PBS containing 0.8% Bovine Serum Albumin and 0.01% Tween 20 for 30 min, then incubated with primary antibody (anti-*Giardia* IscS or

anti- *Mastigamoeba* [FeFe]-hydrogenase, as above) diluted in blocking solution for 3 h at room temperature or overnight at 4°C. The grids were then washed for 10 min each of four times in blocking solution before incubation with gold-conjugated anti-rabbit (IscS) or anti-rat ([FeFe]-hydrogenase) secondary antibody (Sigma or Electron Microscopy Sciences) diluted 1:200 in blocking solution. Labeled sections underwent one 10-min wash in blocking solution, three 10-min washes in PBS, and two 30-s washes in dH_2O. Following antibody labeling, grids were stained using 2% aqueous uranyl acetate and rinsed, stained with lead citrate, rinsed and air-dried. The sections were viewed using a JEOL JEM 1230 transmission electron microscope at 80 kV, and images were captured using a Hamamatsu ORCA-HR digital camera.

The areas of the nucleus, mitochondria, and cytosol of each cell cross section were measured using ImageJ for 8 cells (IscS localization) or 13 cells ([FeFe]-hydrogenase localization), and the gold particles in each part of the cell were counted by eye.

Phylogenetic Analyses

For each protein, all known eukaryotic and bacterial homologs in NCBI were collected and then multiple alignments were generated using PROBCONS v.1.2 (Do et al. 2005) or MAFFT-L-INSI v7.149b (Katoh et al. 2002, 2005; Katoh and Standley 2013), and trimmed of ambiguously aligned sites with BMGE 1.1 (Criscuolo and Gribaldo 2010) (-m BLOSUM30, all other parameters default). In the case of PFO, paralogous PNO sequences were included in alignments. Preliminary phylogenies were generated using FastTree (Price et al. 2009), and datasets were manually curated. Twenty independent Maximum Likelihood (ML) tree estimates and 200 bootstrap replicates were generated using RAxML v.8.0.23 (Stamatakis 2014) under the PROTGAMMALG4X (Le et al. 2012) model of amino acid substitution. Additional ML analyses using the C20 + Gamma model were carried out using IQ-TREE v.1.3.11 (Nguyen et al. 2015), with 1,000 ultrafast bootstrap replicates (Minh et al. 2013) (note that the IQ-TREE manual recommends treating only ultrafast bootstrap support values only over 95 as robust support; for the ML tree under the C20 + gamma model found using IQ-TREE, with full ultrafast bootstrap support values, see supplementary figs. S12–S24, Supplementary Material online). We tested whether specific phylogenetic hypotheses were rejected by the data using the Approximately Unbiased test implemented in CONSEL v.1.20 (Shimodaira and Hasegawa 2001). ML trees given specific constraints (i.e., corresponding to specific hypotheses) were generated using RAxML. The 200 trees from bootstrap replicates were included in analyses (as required by CONSEL), as were the ML trees generated by both RAxML and IQ-TREE. Site likelihoods were calculated under either the LG4X + Gamma model in RAxML. PFO and [FeFe]-hydrogenase phylogenies were performed both including and excluding paralogs (sulfite reductases, and periplasmic [FeFe]-hydrogenases and NARF-like proteins, respectively).

Availability of Supporting Data

The datasets supporting the results of this article are available in the GenBank repository, Accession Numbers SRR2566811 (raw reads) and KT984512–KT984652 and KP258196-KP258200 (coding sequences reported in this manuscript).

Supplementary Material

Supplementary figures S1–S24 and tables S1–S5 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We would like to thank Drs. John Archibald, Ford Doolittle, Thomas Richards, Alastair Simpson and Anastasios Tsaousis for helpful discussion, Dr. Trevor Lithgow for providing the protein sequences used to construct protein import HMMs, Dr. Jan Tachezy for kindly providing antibodies, Greg McCluskey for providing Rosetta 2 cells, and Mary Ann Trevors for help in preparing the electron microscopy samples.

This work was supported by the Canadian Institutes of Health Research (CIHR) (MOP-142349 to A.J.R.); the Nova Scotia Health Research Foundation (NSHRF) (Regional Partnerships Program FRN 62809 to A.J.R., Scotia Scholarship 2012-8781 to M.M.L.); the National Research Fund, Luxembourg (FNR) (PHD-08-001); the Natural Sciences and Engineering Research Council of Canada (PGS-M to L.A.H.); a Killam Pre-Doctoral Fellowship to L.A.H.; and the Tula Foundation (Centre for Comparative Genomics and Evolutionary Bioinformatics postdoctoral fellowship to L.E.).

References

Aguilera P, Barry T, Tovar J. 2008. *Entamoeba histolytica* mitosomes: organelles in search of a function. *Exp Parasitol*. 118:10–16.

Akhmanova A, Voncken F, van Alen T, van Hoek A, Boxma B, Vogels G, Veenhuis M, Hackstein JH. 1998. A hydrogenosome with a genome. *Nature* 396:527–528.

Ali V, Shigeta Y, Tokumoto U, Takahashi Y, Nozaki T. 2004. An intestinal parasitic protist, *Entamoeba histolytica*, possesses a non-redundant nitrogen fixation-like system for iron-sulfur cluster assembly under anaerobic conditions. *J Biol Chem*. 279:16863–16874.

Alsmark C, Foster PG, Sicheritz-Ponten T, Nakjang S, Martin Embley T, Hirt RP. 2013. Patterns of prokaryotic lateral gene transfers affecting parasitic microbial eukaryotes. *Genome Biol*. 14:R19.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 25:3389–3402.

Andersson JO. 2009. Horizontal gene transfer between microbial eukaryotes. *Methods Mol Biol*. 532:473–487.

Andersson J, Hirt R, Foster P, Roger A. 2006. Evolution of four gene families with patchy phylogenetic distributions: influx of genes into protist genomes. *BMC Evol Biol* 6:27.

Andersson JO, Roger AJ. 2003. Evolution of glutamate dehydrogenase genes: evidence for lateral gene transfer within and between prokaryotes and eukaryotes. *BMC Evol Biol*. 3:14.

Andersson JO, Sjogren AM, Davis LA, Embley TM, Roger AJ. 2003. Phylogenetic analyses of diplomonad genes reveal frequent lateral gene transfers affecting eukaryotes. *Curr Biol*. 13:94–104.

Atteia A, Adrait A, Brugiere S, Tardif M, van Lis R, Deusch O, Dagan T, Kuhn L, Gontero B, Martin W, et al. 2009. A proteomic survey of *Chlamydomonas reinhardtii* mitochondria sheds new light on the

metabolic plasticity of the organelle and on the nature of the alphaproteobacterial mitochondrial ancestor. *Mol Biol Evol*. 26:1533–1548.
Barberà MJ, Ruiz-Trillo I, Tufts JY, Bery A, Silberman JD, Roger AJ. 2010. *Sawyeria marylandensis* (Heterolobosea) has a hydrogenosome with novel metabolic properties. *Eukaryot Cell* 9:1913–1924.
Basu S, Fey P, Pandit Y, Dodson R, Kibbe WA, Chisholm RL. 2013. DictyBase 2013: integrating multiple Dictyostelid species. *Nucleic Acids Res*. 41:D676–D683.
Becker T, Bottinger L, Pfanner N. 2012. Mitochondrial protein import: from transport pathways to an integrated network. *Trends Biochem Sci*. 37:85–91.
Bernard C, Simpson AGB, Patterson DJ. 2000. Some free-living flagellates (protista) from anoxic habitats. *Ophelia* 52:113–142.
Bowman BH, Taylor JW, Brownlee AG, Lee J, Lu SD, White TJ. 1992. Molecular evolution of the fungi: relationship of the Basidiomycetes, Ascomycetes, and Chytridiomycetes. *Mol Biol Evol*. 9:285–296.
Boxma B, de Graaf RM, van der Staay GW, van Alen TA, Ricard G, Gabaldón T, van Hoek AH, Moon-van der Staay SY, Koopman WJ, van Hellemond JJ, et al. 2005. An anaerobic mitochondrion that produces hydrogen. *Nature* 434:74–79.
Brown DM, Upcroft JA, Upcroft P. 1995. Free radical detoxification in Giardia duodenalis. *Mol Biochem Parasitol*. 72:47–56.
Brown MW, Sharpe SC, Silberman JD, Heiss AA, Lang BF, Simpson AG, Roger AJ. 2013. Phylogenomics demonstrates that breviate flagellates are related to opisthokonts and apusomonads. *Proc Biol Sci*. 280:20131755.
Buczek D, Wojtkowska M, Suzuki Y, Sonobe S, Nishigami Y, Antoniewicz M, Kmita H, Makalowski W. 2016. Protein import complexes in the mitochondrial outer membrane of Amoebozoa representatives. *BMC Genomics* 17:99.
Bui E, Johnson P. 1996. Identification and characterization of [Fe]-hydrogenases in the hydrogenosome of *Trichomonas vaginalis*. *Mol Biochem Parasitol*. 76:305–310.
Burger G, Gray MW, Forget L, Lang BF. 2013. Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. Genome. *Biol Evol*. 5:418–438.
Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, Zhao Q, Wortman JR, Bidwell SL, Alsmark CM, Besteiro S, et al. 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science* 315:207–212.
Chevreux B, Wetter T, Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information. *Comput Sci Biol: Proc German Conf Bioinfo*. 99:45–56.
Claros MG. 1995. MitoProt, a Macintosh application for studying mitochondrial proteins. *Comput Appl Biosci*. 11:441–447.
Criscuolo A, Gribaldo S. 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol Biol*. 10:210.
Ctrnacta V, Ault JG, Stejskal F, Keithly JS. 2006. Localization of pyruvate:NADP+ oxidoreductase in sporozoites of *Cryptosporidium parvum*. *J Eukaryot Microbiol*. 53:225–231.
Dagley MJ, Doležal P, Likic VA, Smid O, Purcell AW, Buchanan SK, Tachezy J, Lithgow T. 2009. The protein import channel in the outer mitosomal membrane of *Giardia intestinalis*. *Mol Biol Evol*. 26:1941–1947.
Davidson EA, van der Giezen M, Horner DS, Embley TM, Howe CJ. 2002. An [Fe] hydrogenase from the anaerobic hydrogenosome-containing fungus *Neocallimastix frontalis* L2. *Gene* 296:45–52.
de Graaf RM, Duarte I, van Alen TA, Kuiper JW, Schotanus K, Rosenberg J, Huynen MA, Hackstein JH. 2009. The hydrogenosomes of *Psalteriomonas lanterna*. *BMC Evol Biol*. 9:287.
Denoeud F, Roussel M, Noel B, et al. 2011. Genome sequence of the stramenopile Blastocystis, a human anaerobic parasite. *Genome Biol*. 12:R29.
Di Matteo A, Scandurra FM, Testa F, Forte E, Sarti P, Brunori M, Giuffre A. 2008. The O2-scavenging flavodiiron protein in the human parasite *Giardia intestinalis*. *J Biol Chem*. 283:4061–4068.
Do CB, Mahabhashyam MS, Brudno M, Batzoglou S. 2005. ProbCons: probabilistic consistency-based multiple sequence alignment. *Genome Res*. 15:330–340.
Doležal P, Dagley MJ, Kono M, Wolynec P, Likic VA, Foo JH, Sedinova M, Tachezy J, Bachmann A, Bruchhaus I, et al. 2010. The essentials of protein import in the degenerate mitochondrion of *Entamoeba histolytica*. *PLoS Pathog*. 6:e1000812.
Doležal P, Likic V, Tachezy J, Lithgow T. 2006. Evolution of the molecular machines for protein import into mitochondria. *Science* 313:314–318.
Dyall SD, Yan W, Delgadillo-Correa MG, Lunceford A, Loo JA, Clarke CF, Johnson PJ. 2004. Non-mitochondrial complex I proteins in a hydrogenosomal oxidoreductase complex. *Nature* 431:1103–1107.
Eckers E, Cyrklaff M, Simpson L, Deponte M. 2012. Mitochondrial protein import pathways are functionally conserved among eukaryotes despite compositional diversity of the import machineries. *Biol Chem*. 393:513–524.
Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 32:1792–1797.
Edgar RC. 2010. Quality measures for protein alignment benchmarks. *Nucleic Acids Res*. 38:2145–2153.
Emanuelsson O, Brunak S, von Heijne G, Nielsen H. 2007. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc*. 2:953–971.
Emanuelsson O, Nielsen H, Brunak S, von Heijne G. 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol*. 300:1005–1016.
Embley TM, Finlay BJ, Dyal PL, Hirt RP, Wilkinson M, Williams AG. 1995. Multiple origins of anaerobic ciliates with hydrogenosomes within the radiation of aerobic ciliates. *Proc Biol Sci*. 262:87–93.
Embley TM, van der Giezen M, Horner DS, Dyal PL, Foster P. 2003. Mitochondria and hydrogenosomes are two forms of the same fundamental organelle. *Philos Trans R Soc Lond B Biol Sci*. 358:191–201.
Ewing B, Green P. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res*. 8:186–194.
Ewing B, Hillier L, Wendl MC, Green P. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res*. 8:175–185.
Figueroa-Martinez FS, Funes LG. Franzen D, Gonzalez-Halphen 2008. Reconstructing the mitochondrial protein import machinery of *Chlamydomonas reinhardtii*. *Genetics* 179:149–155.
Gaston D, TsaousisAD, RogerAJ. 2009. Predicting proteomes of mitochondria and related organelles from genomic and expressed sequence tag data. *Methods Enzymol*. 457:21–47.
Gawryluk RM, ChisholmKA, PintoDM, GrayMW. 2014. Compositional complexity of the mitochondrial proteome of a unicellular eukaryote (*Acanthamoeba castellanii*, supergroup Amoebozoa) rivals that of animals, fungi, and plants. *J Proteomics* 109:400–416.
Gill EE, Diaz-Trivino S, BarberaMJ, SilbermanJD, Stechmann A, Gaston D, Tamas I, Roger AJ. 2007. Novel mitochondrion-related organelles in the anaerobic amoeba *Mastigamoeba balamuthi*. *Mol Microbiol*. 66:1306–1320.
Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 29:644–652.
Heinz E, Lithgow T. 2013. Back to basics: a revealing secondary reduction of the mitochondrial protein import pathway in diverse intracellular parasites. *Biochim Biophys Acta* 1833:295–303.
Hell K, Neupert W, Stuart RA. 2001. Oxa1p acts as a general membrane insertion machinery for proteins encoded by mitochondrial DNA. *Embo J*. 20:1281–1288.
Henriquez FL, RichardsTA, Roberts F, McLeod R, Roberts CW. 2005. The unusual mitochondrial compartment of *Cryptosporidium parvum*. *Trends Parasitol*. 21:68–74.
Hoogenraad NJ, Ward LA, Ryan MT. 2002. Import and assembly of proteins into mitochondria of mammalian cells. *Biochim Biophys Acta* 1592:97–105.
Horner DS, Foster PG, Embley TM. 2000. Iron hydrogenases and the evolution of anaerobic eukaryotes. *Mol Biol Evol*. 17:1695–1709.

Horner DS, Heil B, Happe T, Embley TM. 2002. Iron hydrogenases–ancient enzymes in modern eukaryotes. *Trends Biochem Sci*. 27:148–153.

Horner DS, Hirt RP, Embley TM. 1999. A single eubacterial origin of eukaryotic pyruvate: ferredoxin oxidoreductase genes: implications for the evolution of anaerobic eukaryotes. *Mol Biol Evol*. 16:1280–1291.

Hrdý I, Hirt RP, Dolezal P, Bardonova L, Foster PG, Tachezy J, Embley TM. 2004. *Trichomonas* hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. *Nature* 432:618–622.

Hrdý I, Tachezy J, Müller M. 2008. Metabolism of trichomonad hydrogenosomes. In: Tachezy J, editor. Hydrogenosomes and mitosomes: mitochondria of anaerobic eukaryotes. New York: Springer. p. 113–146.

Hug LA, Stechmann A, Roger AJ. 2010. Phylogenetic distributions and histories of proteins involved in anaerobic pyruvate metabolism in eukaryotes. *Mol Biol Evol*. 27:311–324.

Jedelský PL, Doležal P, Rada P, Pyrih J, Šmid O, Hrdý I, Sedinova M, Marcincikova M, Voleman L, Perry AJ, et al. 2011. The minimal proteome in the reduced mitochondrion of the parasitic protist *Giardia intestinalis*. *PLoS One* 6:e17285.

Jerström-Hultqvist J, Einarsson E, Xu F, Hjort K, Ek B, Steinhauf D, Hultenby K, Bergquist J, Andersson JO, Svard SG. 2013. Hydrogenosomes in the diplomonad *Spironucleus salmonicida*. *Nat Commun*. 4:2493.

Kamikawa R, Kolisko M, Nishimura Y, Yabuki A, Brown MW, Ishikawa SA, Ishida K, Roger AJ, Hashimoto T, Inagaki Y. 2014. Gene content evolution in Discobid mitochondria deduced from the phylogenetic position and complete mitochondrial genome of *Tsukubamonas globosa*. *Genome Biol Evol*. 6:306–315.

Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res*. 33:511–518.

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 30:3059–3066.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30:772–780.

LaGier MJ, Tachezy J, Stejskal F, Kutisova K, Keithly JS. 2003. Mitochondrial-type iron-sulfur cluster biosynthesis genes (IscS and IscU) in the apicomplexan *Cryptosporidium parvum*. *Microbiology* 149:3519–3530.

Lang BF, Burger G, O'Kelly CJ, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Gray MW. 1997. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* 387:493–497.

Lantsman Y, Tan KS, Morada M, Yarlett N. 2008. Biochemical characterization of a mitochondrial-like organelle from *Blastocystis* sp. subtype 7. *Microbiology* 154:2757–2766.

Lara E, Chatzinotas A, Simpson AG. 2006. Andalucia (n. gen.)—the deepest branch within jakobids (Jakobida; Excavata), based on morphological and molecular study of a new flagellate from soil. *J Eukaryot Microbiol*. 53:112–120.

Le SQ, Dang CC, Gascuel O. 2012. Modeling protein evolution with several amino acid replacement matrices depending on site rates. *Mol Biol Evol*. 29:2921–2936.

Leckenby A, Hall N. 2015. Genomic changes during evolution of animal parasitism in eukaryotes. *Curr Opin Genet Dev*. 35:86–92.

Leger MM, Gawryluk RM, Gray MW, Roger AJ. 2013. Evidence for a hydrogenosomal-type anaerobic ATP generation pathway in *Acanthamoeba castellanii*. *PLoS One* 8:e69532.

Leger MM, Petru M, Zarsky V, Eme L, Vlcek C, Harding T, Lang BF, Elias M, Doležal P, Roger AJ. 2015. An ancestral bacterial division system is widespread in eukaryotic mitochondria. *Proc Natl Acad Sci U S A*. 112:10239–10246.

Lill R, Diekert K, Kaut A, Lange H, Pelzer W, Prohl C, Kispal G. 1999. The essential role of mitochondria in the biogenesis of cellular iron-sulfur proteins. *Biol Chem*. 380:1157–1166.

Lill R, Hoffmann B, Molik S, Pierik AJ, Rietzschel N, Stehling O, Uzarska MA, Webert H, Wilbrecht C, Muhlenhoff U. 2012. The role of mitochondria in cellular iron-sulfur protein biogenesis and iron metabolism. *Biochim Biophys Acta* 1823:1491–1508.

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

Liu Z, Li X, Zhao P, Gui J, Zheng W, Zhang Y. 2011. Tracing the evolution of the mitochondrial protein import machinery. *Comput Biol Chem*. 35:336–340.

Loftus B, Anderson I, Davies R, Alsmark CM, Samuelson J, Amedeo P, Roncaglia P, Berriman M, Hirt RP, Mann BJ, et al. 2005. The genome of the protist parasite *Entamoeba histolytica*. *Nature* 433:865–868.

Makiuchi T, Mi-ichi F, Nakada-Tsukui K, Nozaki T. 2013. Novel TPR-containing subunit of TOM complex functions as cytosolic receptor for Entamoeba mitosomal transport. *Sci Rep*. 3:1129.

Makiuchi T, Nozaki T. 2014. Highly divergent mitochondrion-related organelles in anaerobic parasitic protozoa. *Biochimie* 100:3–17.

Maralikova B, Ali V, Nakada-Tsukui K, Nozaki T, van der Giezen M, Henze K, Tovar J. 2010. Bacterial-type oxygen detoxification and iron-sulfur cluster assembly in amoebal relict mitochondria. *Cell Microbiol*. 12:331–342.

Martincova E, Voleman L, Pyrih J, Zarsky V, Vondrackova P, Kolisko M, Tachezy J, Dolezal P. 2015. Probing the biology of *Giardia intestinalis* mitosomes using in vivo enzymatic tagging. *Mol Cell Biol*. 35:2864–2874.

McGlynn SE, Shepard EM, Winslow MA, Naumov AV, Duschene KS, Posewitz MC, Broderick WE, Broderick JB, Peters JW. 2008. HydF as a scaffold protein in [FeFe] hydrogenase H-cluster biosynthesis. *FEBS Lett*. 582:2183–2187.

Mi-ichi F, Abu Yousuf M, Nakada-Tsukui K, Nozaki T. 2009. Mitosomes in *Entamoeba histolytica* contain a sulfate activation pathway. *Proc Natl Acad Sci U S A*. 106:21731–21736.

Minh BQ, Nguyen MA, von Haeseler A. 2013. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol*. 30:1188–1195.

Mulder DW, Boyd ES, Sarma R, Lange RK, Endrizzi JA, Broderick JB, Peters JW. 2010. Stepwise [FeFe]-hydrogenase H-cluster assembly revealed in the structure of HydA(DeltaEFG). *Nature* 465:248–251.

Müller M, Mentel M, van Hellemond JJ, Henze K, Woehle C, Gould SB, Yu RY, van der Giezen M, Tielens AG, Martin WF. 2012. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol Mol Biol Rev*. 76:444–495.

Murcha MW, Kmiec B, Kubiszewski-Jakubiak S, Teixeira PF, Glaser E, Whelan J. 2014. Protein import into plant mitochondria: signals, machinery, processing, and regulation. *J Exp Bot*. 65:6301–6335.

Murcha MW, Narsai R, Devenish J, Kubiszewski-Jakubiak S, Whelan J. 2015. MPIC: a mitochondrial protein import components database for plant and non-plant species. *Plant Cell Physiol*. 56:e10.

Murcha MW, Wang Y, Narsai R, Whelan J. 2014. The plant mitochondrial protein import apparatus - the differences make it interesting. *Biochim Biophys Acta* 1840:1233–1245.

Neupert W. 1997. Protein import into mitochondria. *Annu Rev Biochem*. 66:863–917.

Neupert W, Herrmann JM. 2007. Translocation of proteins into mitochondria. *Annu Rev Biochem*. 76:723–749.

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 32:268–274.

Noguchi F, Shimamura S, Nakayama T, Yazaki E, Yabuki A, Hashimoto T, Inagaki Y, Fujikura K, Takishita K. 2015. Metabolic capacity of mitochondrion-related organelles in the free-living anaerobic stramenopile *Cantina marsupialis*. *Protist* 166:534–550.

Nývltová E, Stairs CW, Hrdy I, Ridl J, Mach J, Paces J, Roger AJ, Tachezy J. 2015. Lateral gene transfer and gene duplication played a key role in the evolution of *Mastigamoeba balamuthi* hydrogenosomes. *Mol Biol Evol*. 32:1039–1055.

Nývltová E, Sutak R, Harant K, Sedinova M, Hrdy I, Paces J, Vlcek C, Tachezy J. 2013. NIF-type iron-sulfur cluster assembly system is duplicated and distributed in the mitochondria and cytosol of *Mastigamoeba balamuthi*. *Proc Natl Acad Sci U S A*. 110:7371–7376.

Panek T, Taborsky P, Pachiadaki MG, Hroudova M, Vlcek C, Edgcomb VP, Cepicka I. 2015. Combined Culture-Based and Culture-Independent Approaches Provide Insights into Diversity of Jakobids, an Extremely Plesiomorphic Eukaryotic Lineage. *Front Microbiol*. 6:1288.

Pilet E, Nicolet Y, Mathevon C, Douki T, Fontecilla-Camps JC, Fontecave M. 2009. The role of the maturase HydG in [FeFe]-hydrogenase active site synthesis and assembly. *FEBS Lett*. 583:506–511.

Posewitz MC, King P, Smolinski SL, Zhang L, Seibert M, Ghirardi M. 2004. Discovery of two novel radical S-adenosylmethionine proteins required for the assembly of an active [Fe] hydrogenase. *J Biol Chem*. 279:25711–25720.

Price MN, Dehal PS, Arkin AP. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol*. 26:1641–1650.

Pusnik M, Mani J, Schmidt O, Niemann M, Oeljeklaus S, Schnarwiler F, Warscheid B, Lithgow T, Meisinger C, Schneider A. 2012. An essential novel component of the noncanonical mitochondrial outer membrane protein import system of trypanosomatids. *Mol Biol Cell*. 23:3420–3428.

Regoes A, Zourmpanou D, Leon-Avila G, van der Giezen M, Tovar J, Hehl AB. 2005. Protein import, replication, and inheritance of a vestigial mitochondrion. *J Biol Chem*. 280:30557–30563.

Schneider RE, Brown MT, Shiflett AM, Dyall SD, Hayes RD, Xie Y, Loo JA, Johnson PJ. 2011. The *Trichomonas vaginalis* hydrogenosome proteome is highly reduced relative to mitochondria, yet complex compared with mitosomes. *Int J Parasitol*. 41:1421–1434.

Shimodaira H, Hasegawa M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17:1246–1247.

Simpson AG, Patterson DJ. 2001. On core jakobids and excavate taxa: the ultrastructure of *Jakoba incarcerata*. *J Eukaryot Microbiol*. 48:480–492.

Simpson AG, Perley TA, Lara E. 2008. Lateral transfer of the gene for a widely used marker, alpha-tubulin, indicated by a multi-protein study of the phylogenetic position of *Andalucia* (Excavata). *Mol Phylogenet Evol*. 47:366–377.

Smutna T, Goncalves VL, Saraiva LM, Tachezy J, Teixeira M, Hrdy I. 2009. Flavodiiron protein from *Trichomonas vaginalis* hydrogenosomes: the terminal oxygen reductase. *Eukaryot Cell* 8:47–55.

Stairs CW, Eme L, Brown MW, Mutsaers C, Susko E, Dellaire G, Soanes DM, van der Giezen M, Roger AJ. 2014. A SUF Fe-S cluster biogenesis system in the mitochondrion-related organelles of the anaerobic protist *Pygsuia*. *Curr Biol*. 24:1176–1186.

Stairs CW, Leger MM, Roger AJ. 2015. Diversity and origins of anaerobic metabolism in mitochondria and related organelles. *Philos Trans R Soc Lond B Biol Sci*. 370(1678):20140326.

Stairs CW, Roger AJ, Hampl V. 2011. Eukaryotic pyruvate formate lyase and its activating enzyme were acquired laterally from a firmicute. *Mol Biol Evol*. 28:2087–2099.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313.

Stechmann A, Hamblin K, Perez-Brocal V, Gaston D, Richmond GS, van der Giezen M, Clark CG, Roger AJ. 2008. Organelles in *Blastocystis* that blur the distinction between mitochondria and hydrogenosomes. *Curr Biol*. 18:580–585.

Takishita K, Chikaraishi Y, Leger MM, Kim E, Yabuki A, Ohkouchi N, Roger AJ. 2012. Lateral transfer of tetrahymanol-synthesizing genes has allowed multiple diverse eukaryote lineages to independently adapt to environments without oxygen. *Biol Direct* 7:5.

Tamura Y, Harada Y, Nishikawa S, Yamano K, Kamiya M, Shiota T, Kuroda T, Kuge O, Sesaki H, Imai K, et al. 2013. Tam41 Is a CDP-diacylglycerol synthase required for cardiolipin biosynthesis in mitochondria. *Cell Metabolism* 17:709–718.

Tian HF, Feng JM, Wen JF. 2012. The evolution of cardiolipin biosynthesis and maturation pathways and its implications for the evolution of eukaryotes. *BMC Evol Biol*. 12:32.

Tielens AG, van Grinsven KW, Henze K, van Hellemond JJ, Martin W. 2010. Acetate formation in the energy metabolism of parasitic helminths and protists. *Int J Parasitol*. 40:387–397.

Tong J, Doležal P, Selkrig J, Crawford S, Simpson AG, Noinaj N, Buchanan SK, Gabriel K, Lithgow T. 2011. Ancestral and derived protein import pathways in the mitochondrion of *Reclinomonas americana*. *Mol Biol Evol*. 28:1581–1591.

Tovar J, Fischer A, Clark CG. 1999. The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. *Mol Microbiol*. 32:1013–1021.

Tovar J, Leon-Avila G, Sanchez LB, Sutak R, Tachezy J, van der Giezen M, Hernandez M, Muller M, Lucocq JM. 2003. Mitochondrial remnant organelles of *Giardia* function in iron-sulphur protein maturation. *Nature* 426:172–176.

Tsaousis AD, Gaston D, Stechmann A, Walker PB, Lithgow T, Roger AJ. 2011. A functional Tom70 in the human parasite *Blastocystis* sp.: implications for the evolution of the mitochondrial import apparatus. *Mol Biol Evol*. 28:781–791.

Tsaousis AD, Leger MM, Stairs CAW, Roger AJ. 2012. The biochemical adaptations of mitochondrion-related organelles of parasitic and free-living microbial eukaryotes to low oxygen environments. In: Altenbach AV, Bernhard JM, Seckbach J, editors. Anoxia: evidence for eukaryote survival and paleontological strategies. Dordrecht: Springer. p. 51–81.

Tsaousis AD, Ollagnier de Choudens S, Gentekaki E, Long S, Gaston D, Stechmann A, Vinella D, Py B, Fontecave M, Barras F, et al. 2012. Evolution of Fe/S cluster biogenesis in the anaerobic parasite *Blastocystis*. *Proc Natl Acad Sci U S A*. 109:10426–10431.

van der Giezen M, Cox S, Tovar J. 2004. The iron-sulfur cluster assembly genes iscS and iscU of *Entamoeba histolytica* were acquired by horizontal gene transfer. *BMC Evol Biol*. 4:7.

van der Laan M, Chacinska A, Lind M, Perschil I, Sickmann A, Meyer HE, Guiard B, Meisinger C, Pfanner N, Rehling P. 2005. Pam17 is required for architecture and translocation activity of the mitochondrial protein import motor. *Mol Cell Biol*. 25:7449–7458.

van Hoek AH, Akhmanova AS, Huynen MA, Hackstein JH. 2000. A mitochondrial ancestry of the hydrogenosomes of *Nyctotherus ovalis*. *Mol Biol Evol*. 17:202–206.

Vicente JB, Testa F, Mastronicola D, Forte E, Sarti P, Teixeira M, Giuffre A. 2009. Redox properties of the oxygen-detoxifying flavodiiron protein from the human parasite *Giardia intestinalis*. *Arch Biochem Biophys*. 488:9–13.

Vicente JB, Tran V, Pinto L, Teixeira M, Singh U. 2012. A detoxifying oxygen reductase in the anaerobic protozoan *Entamoeba histolytica*. *Eukaryot Cell* 11:1112–1118.

Vignais PM, Billoud B. 2007. Occurrence, classification, and biological function of hydrogenases: an overview. *Chem Rev*. 107:4206–4272.

Vignais PM, Billoud B, Meyer J. 2001. Classification and phylogeny of hydrogenases. *FEMS Microbiol Rev*. 25:455–501.

Waller RF, Jabbour C, Chan NC, Celik N, Likic VA, Mulhern TD, Lithgow T. 2009. Evidence of a reduced and modified mitochondrial protein import apparatus in microsporidian mitosomes. *Eukaryot Cell* 8:19–26.

Wideman JG, Gawryluk RM, Gray MW, Dacks JB. 2013. The ancient and widespread nature of the ER-mitochondria encounter structure. *Mol Biol Evol*. 30:2044–2049.

Williams BA, Keeling PJ, 2005. Microsporidian mitochondrial proteins: expression in *Antonospora locustae* spores and identification of genes coding for two further proteins. *J Eukaryot Microbiol*. 52:271–276.

Wojtkowska M, Buczek D, Stobienia O, Karachitos A, Antoniewicz M, Slocinska M, Makalowski W, Kmita H. 2015. The TOM complex of Amoebozoans: the cases of the amoeba *Acanthamoeba castellanii* and the slime mold *Dictyostelium discoideum*. *Protist* 166:349–362.

Yarlett N, Hann AC, Lloyd D, Williams A. 1981. Hydrogenosomes in the rumen protozoon Dasytricha ruminantium Schuberg. *Biochem J*. 200:365–372.

Zubacova Z, Novak L, Bublikova J, Vacek V, Fousek J, Ridl J, Tachezy J, Doležal P, Vlcek C, Hampl V. 2013. The mitochondrion-like organelle of *Trimastix pyriformis* contains the complete glycine cleavage system. *PLoS One* 8:e55417.