

RESEARCH ARTICLE

Open Access



Anaerobic derivatives of mitochondria and peroxisomes in the free-living amoeba *Pelomyxa schiedti* revealed by single-cell genomics

Kristína Záhonová^{1,2*} , Sebastian Cristian Treitli¹ , Tien Le¹ , Ingrid Škodová-Sveráková^{2,3} , Pavla Hanousková⁴ , Ivan Čepička⁴ , Jan Tachezy¹  and Vladimír Hampel^{1*} 

Abstract

Background: Mitochondria and peroxisomes are the two organelles that are most affected during adaptation to microoxic or anoxic environments. Mitochondria are known to transform into anaerobic mitochondria, hydrogenosomes, mitosomes, and various transition stages in between, collectively called mitochondrion-related organelles (MROs), which vary in enzymatic capacity. Anaerobic peroxisomes were identified only recently, and their putatively most conserved function seems to be the metabolism of inositol. The group Archamoebae includes anaerobes bearing both anaerobic peroxisomes and MROs, specifically hydrogenosomes in free-living *Mastigamoeba balamuthi* and mitosomes in the human pathogen *Entamoeba histolytica*, while the organelles within the third lineage represented by *Pelomyxa* remain uncharacterized.

Results: We generated high-quality genome and transcriptome drafts from *Pelomyxa schiedti* using single-cell omics. These data provided clear evidence for anaerobic derivatives of mitochondria and peroxisomes in this species, and corresponding vesicles were tentatively identified in electron micrographs. In silico reconstructed MRO metabolism harbors respiratory complex II, electron-transferring flavoprotein, a partial TCA cycle running presumably in the reductive direction, pyruvate:ferredoxin oxidoreductase, [FeFe]-hydrogenases, a glycine cleavage system, a sulfate activation pathway, and an expanded set of NIF enzymes for iron-sulfur cluster assembly. When expressed in the heterologous system of yeast, some of these candidates localized into mitochondria, supporting their involvement in the MRO metabolism. The putative functions of *P. schiedti* peroxisomes could be pyridoxal 5'-phosphate biosynthesis, amino acid and carbohydrate metabolism, and hydrolase activities. Unexpectedly, out of 67 predicted peroxisomal enzymes, only four were also reported in *M. balamuthi*, namely peroxisomal processing peptidase, nudix hydrolase, inositol 2-dehydrogenase, and D-lactate dehydrogenase. Localizations in yeast corroborated peroxisomal functions of the latter two.

Conclusions: This study revealed the presence and partially annotated the function of anaerobic derivatives of mitochondria and peroxisomes in *P. schiedti* using single-cell genomics, localizations in yeast heterologous systems, and transmission electron microscopy. The MRO metabolism resembles that of *M. balamuthi* and most likely reflects

*Correspondence: kika.zahonova@gmail.com; vlada@natur.cuni.cz

¹ Department of Parasitology, Faculty of Science, Charles University, BIOCEV, Vestec, Czech Republic

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

the state in the common ancestor of Archamoebae. The peroxisomal metabolism is strikingly richer in *P. schiedti*. The presence of *myo*-inositol 2-dehydrogenase in the predicted peroxisomal proteome corroborates the situation in other Archamoebae, but future experimental evidence is needed to verify additional functions of this organelle.

Keywords: *Pelomyxa*, Mitochondrion-related organelle, Hydrogenosome, Anaerobic peroxisome, Anaerobiosis, FeS cluster assembly, Single-cell genomics

Background

Transition to life in low-oxygen environments requires significant modifications of cell biochemistry and organellar makeup. Several lineages of protists have undergone such transitions and exemplify partly convergent solutions [1–3]. Mitochondria and peroxisomes have been most significantly remodeled in this process, as they are the key places of oxygen-dependent metabolism and oxygen detoxification.

Mitochondria are double-membrane-bound organelles, which arose from the engulfment of members of a prokaryotic lineage related to alphaproteobacteria [2, 4–6]. Since then, they have diverged into a range of categories [1] and a plethora of transitional forms [7, 8], collectively designated as mitochondrion-related organelles (MROs), while only a single case of complete loss has been reported [9]. A substantial number of typical mitochondrial functionalities, such as oxidative phosphorylation; carbon, amino acid, and fatty acid metabolism; iron-sulfur (FeS) cluster assembly; homeostasis; and apoptosis, have been reduced to various extents in different MROs [10–12].

Peroxisomes are bound by a single membrane and characterized by a highly conserved set of proteins (peroxins) essential for their biogenesis [13, 14]. Their matrix content and consequently the repertoire of metabolic pathways is highly variable, reflecting a high degree of versatility in peroxisomal functions [15]. Most frequently, they possess oxidases reducing molecular oxygen to hydrogen peroxide (H₂O₂), and catalase for H₂O₂ detoxification. Not surprisingly, they are absent from most anaerobes, such as *Giardia* and *Trichomonas* [16]; however, anaerobic peroxisomes were recently reported from two Archamoebae, namely *Mastigamoeba balamuthi* and *Entamoeba histolytica* [17, 18].

Archamoebae represents a clade of microaerophilic protists nested within a broader group of predominantly aerobic amoebozoans [19, 20] represented, e.g., by *Dicystelium discoideum* (Eumycetozoa), known to bear a classical aerobic mitochondrion [21], or by their more distant amoebozoan relative *Acanthamoeba castellanii* (Centramoebida) with mitochondria potentially adapted to periods of anaerobiosis and exhibiting a highly complex proteome [12, 22]. Small to almost inconspicuous MROs have been characterized in two Archamoebae,

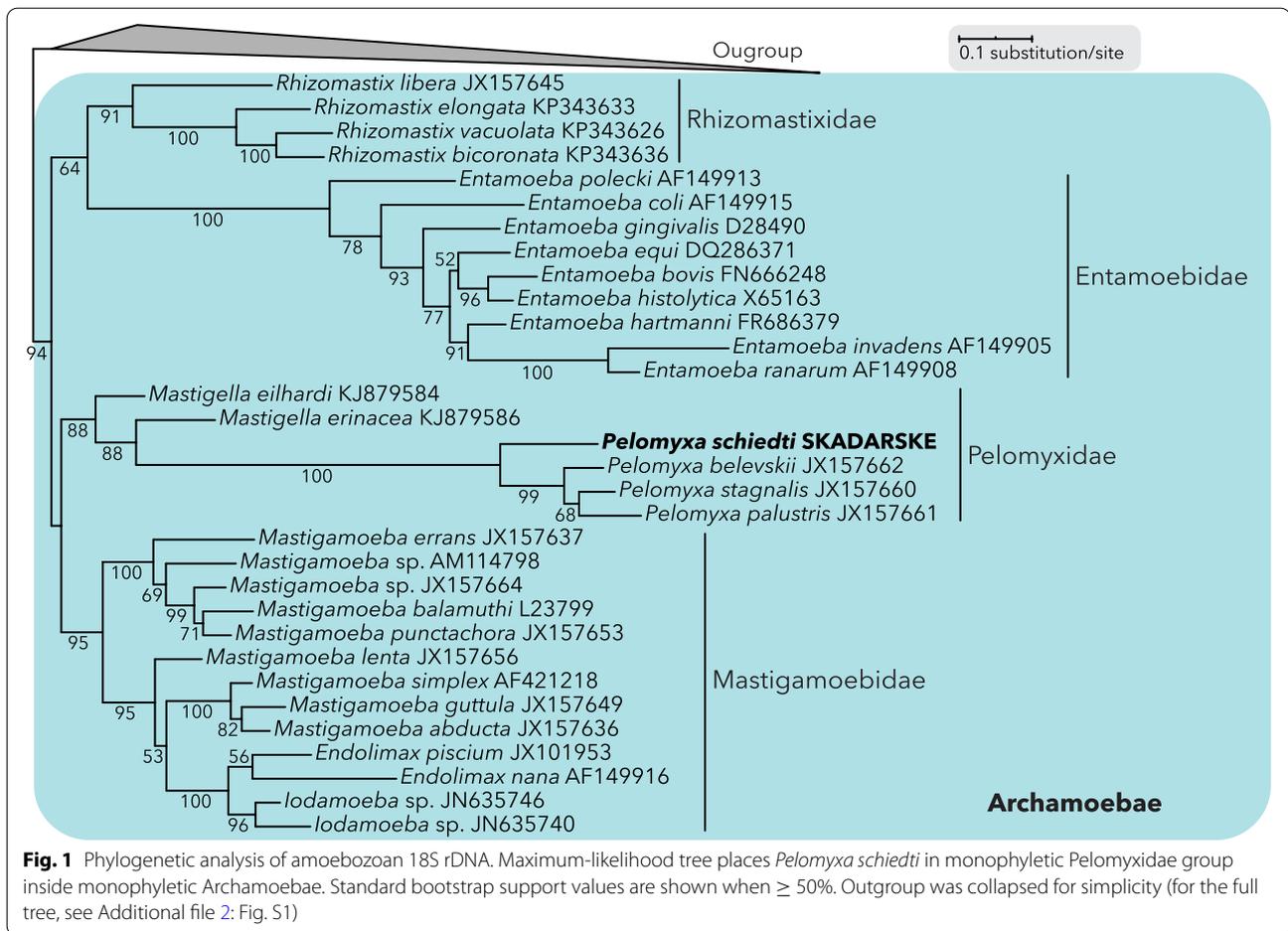
the parasitic *E. histolytica* and the free-living *M. balamuthi*. The only known function of the *E. histolytica* mitosome is production and export of activated sulfate—phosphoadenosine-5'-phosphosulfate (PAPS) [23]. The metabolic capacity of the *M. balamuthi* hydrogenosome is substantially broader, involving pyruvate and amino acid metabolism, ATP production, and FeS cluster assembly [24–26]. Another adaptation of both amoebae to the low oxygen environment is represented by anaerobic peroxisomes that lack catalase and enzymes of β -oxidation, and which are predicted to contain a diverse set of enzymes with only eight being common to both species, including *myo*-inositol 2-dehydrogenase, long-chain fatty acid-CoA ligase, and malate dehydrogenase [17, 18].

Pelomyxa is a free-living archamoeba that is distantly related to both *M. balamuthi* and *E. histolytica* [19], and so it represents a valuable point for tracing the evolution of anaerobic adaptations. *Pelomyxa schiedti* was previously isolated from freshwater sediments of lake Skadar in Albania [27]. It is characterized by a peripheral ring of chromatin granules in its nucleus and the presence of numerous prokaryotic endosymbionts in its cytosol [27]. There is a single report on MROs in the giant species *P. palustris* [28], but their metabolism is unknown. Using methods of single-cell -omics and electron microscopy, we bring clear evidence for the presence of both MROs and peroxisomes in its smaller cousin *P. schiedti* [27].

Results and discussion

General features of assemblies

Pelomyxa schiedti genome assembly generated from seven micromanipulated cells had a total length of 52.4 Mb and contained 5337 scaffolds with an N50 = 51,552 bp and 19,876 predicted proteins (Additional file 1: Table S1). We identified a single small subunit ribosomal RNA gene (18S rDNA). In our 18S phylogeny, *P. schiedti* was sister to other *Pelomyxa* species inside the Pelomyxidae clade (88% standard bootstrap support) within a robust clade (94% standard bootstrap) of Archamoebae (Fig. 1; Additional file 2: Fig. S1). The decontaminated transcriptome assembly of 76.6 Mb comprised 43,993 contigs. BUSCO was used to estimate the completeness of assemblies and to compare them to the *M. balamuthi* genome (Additional file 2: Fig. S2; Additional



file 1: Table S1). The transcriptome contained 83.2% complete and 2.0% fragmented BUSCO genes, while in the genome-derived proteome, the proportions were 81.9% and 3.6%, respectively. With 82.8% complete and 3.0% fragmented genes [29], the completeness of *M. balamuthi* data was comparable. Duplicated BUSCOs represented 36.0% and 8.6% in the transcriptome and genome assembly, respectively, reflecting a higher number of contigs, or the presence of isoforms, in the former. It should be noted that for non-model eukaryotes, which *Pelomyxa* certainly is, BUSCO completeness is not expected to reach 100%, because some orthologues might be absent and/or diverged beyond recognition. Altogether, our analyses showed considerably high completeness of both assemblies.

P. schiedti genes encompassed 149,016 introns (Additional file 1: Table S1), which accounts for a density of 7.43 introns/gene, almost twice as high as *M. balamuthi* (3.74). While protists' genomes have usually lower intron densities, several organisms in IntronDB [30] exhibit similar intron density as *Pelomyxa*, e.g., the choanoflagellates *Monosiga brevicolis* (6.53) and *Salpingoeca rosetta* (7.44),

the chromerid *Vitrella brassicaformis* (7.45), and the chlorarachniophyte *Bigelowiella natans* (7.85). The vast majority of introns (98.41%) contained canonical GT-AG boundaries, 1.59% possessed GC-AG boundaries, and one had an unusual GT-GG intron boundary (Additional file 1: Table S1). Similar frequencies of intron boundaries are observed in *M. balamuthi* (Additional file 1: Table S1 and [29]).

Putative MRO proteome

The major focus of this study was to test for the presence of, and then to characterize, the putative proteomes of the MRO and peroxisome of *P. schiedti*. We used a combined approach to search for proteins possibly involved in MRO metabolism and biogenesis by (i) retrieving homologues of MRO- or mitochondrion-targeted proteins of *E. histolytica*, *M. balamuthi*, and *A. castellanii*; (ii) predicting N-terminal mitochondrial targeting sequences (NTS) by four tools to be able to identify potentially novel pathways; and (iii) localizing selected candidate proteins in the heterologous system of *Saccharomyces cerevisiae*. The resulting predicted MRO proteome based

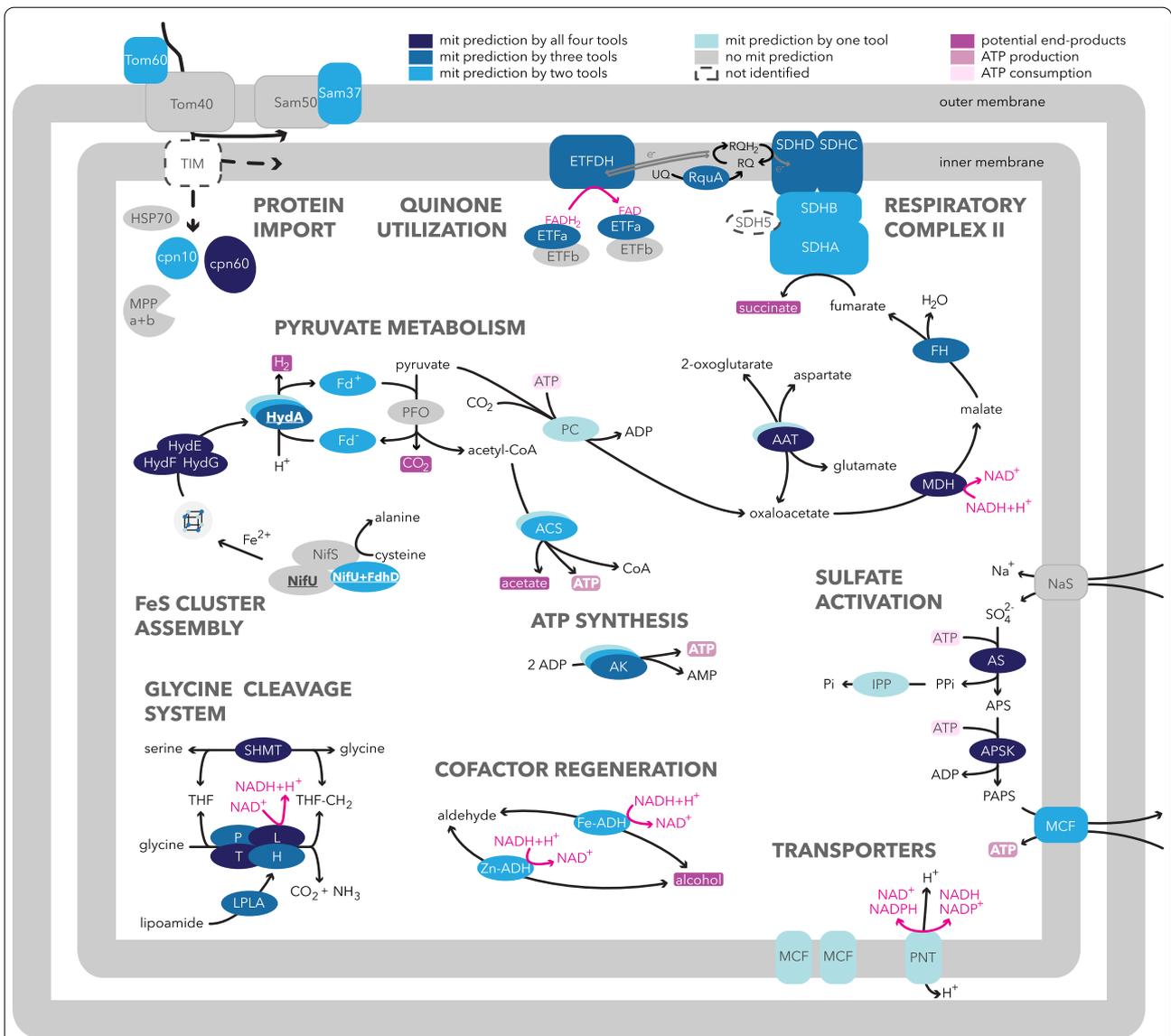


Fig. 2 Overview of the *Pelomyxa schiedti* MRO metabolism. Proteins were identified by BLAST or HMMER searches, and their intracellular localization was predicted by TargetP, PSORT II, MultiLoc2, and NommPred tools. Confidence of MRO localization is enhanced by shades of blue as explained in graphical legend above the scheme. Multiple copies of a protein are shown as overlapping ovals. Proteins that localized in *S. cerevisiae* mitochondria (Fig. 3) are in bold and underlined. AAT, aspartate aminotransferase; ACO, aconitase; ACS, acetyl-CoA synthetase; AK, adenylate kinase; APS, adenosine-5'-phosphosulfate; APSK, adenosine-5'-phosphosulfate kinase; AS, ATP sulfurylase; cpn10, chaperonin 10; cpn60, chaperonin 60; CoA, coenzyme A; ETFa, electron transferring flavoprotein subunit alpha; ETFb, electron transferring flavoprotein subunit beta; ETFDH, electron transferring flavoprotein dehydrogenase; Fe-ADH, iron-containing alcohol dehydrogenase; Fd, ferredoxin; FdhD, formate dehydrogenase accessory sulfurtransferase; FH, fumarase; H, glycine cleavage system H protein; HSP70, heat shock protein 70; HydA, [FeFe]-hydrogenase; HydE, hydrogenase maturase; HydF, hydrogenase maturase; HydG, hydrogenase maturase; IPP, inorganic pyrophosphatase; L, glycine cleavage system L protein; LPLA, lipoamide protein ligase; MCF, mitochondrial carrier family; MDH, malate dehydrogenase; MPP a+b, mitochondrial processing peptidase subunit alpha and beta; NaS, sodium/sulfate symporter; NifS, cysteine desulfurase; NifU, scaffold protein; P, glycine cleavage system P protein; PAPS, 3'-phosphoadenosine 5'-phosphosulfate; PAPS, 3'-phosphoadenosine 5'-phosphosulfate; PC, pyruvate carboxylase; PFO, pyruvate:ferredoxin oxidoreductase; PNT, pyridine nucleotide transhydrogenase; RQ, rodoquinone; RQH2, rholoquinol; RquA, RQ methyltransferase; SAM, sorting and assembly machinery; SDH5, succinate dehydrogenase assembly factor; SDH5, succinate dehydrogenase subunit A; SDHB, succinate dehydrogenase subunit B; SDHC, succinate dehydrogenase subunit C; SDHD, succinate dehydrogenase subunit D; SHMT, serine hydroxymethyltransferase; T, glycine cleavage system T protein; THF, tetrahydrofolate; THF-CH₂, N⁵,N¹⁰-methylene tetrahydrofolate; TOM/TIM, translocase of outer/inner membrane; UQ, ubiquinone; Zn-ADH, zinc-containing alcohol dehydrogenase

on three lines of evidence consisted of 46 proteins (Fig. 2; Additional file 1: Table S2) and putatively provides the functionalities described below.

Protein import machinery

Despite sensitive HMMER searching, we identified only four subunits of the outer membrane translocase (TOM) and the sorting and assembly machinery (SAM) complexes (Fig. 2). Tom40 (Additional file 1: Table S2) forms a pore through which proteins are translocated [31]. Tom60, a shuttle receptor of matrix and membrane proteins, was thought to be *Entamoeba*-specific [32], but its presence in *P. schiedti* (Additional file 1: Table S2) suggests that it may be a universal component of the TOM complex in all archamoebal species. The SAM complex is required for protein insertion into the outer membrane after its translocation via the TOM complex [31]. We found Sam50, which forms a central channel of the complex, and Sam37, which participates in protein insertion (Additional File 1: Table S2). All four proteins had corresponding domains predicted by InterProScan. However, many homologues of the canonical opisthokont subunits were missing (Additional file 1: Table S2), as were all parts of the translocase system of the inner membrane (TIM), and so the mechanism of protein import across this membrane remains unknown. This situation resembles that of other Archamoebae [24, 29, 33], suggesting that archamoebal translocons are in general either highly streamlined and/or contain highly divergent or lineage-specific subunits as reported from trichomonads or trypanosomes [34, 35].

Enzymes involved in proteolytic processing (matrix processing peptidase) and folding (chaperonins cpn10 and cpn60) were present in our predicted proteome. Although cpn10 is a strictly specific mitochondrial protein, which is undoubtedly localized in the MRO, heterologous expression in *S. cerevisiae* failed to localize this protein into its mitochondria (Fig. 3). This fact may be explained by the loss of positive charge in the NTS of some hydrogenosome and mitosome targeted proteins, which is crucial for protein import into yeast mitochondria [36]. At the same time, it suggests that observations from heterologous systems, particularly the absence of targeting, should be interpreted with caution.

HSP70 was detected in 14 copies; none of them confidently predicted to be mitochondrial (Additional file 1: Table S2). Phylogenetic analysis revealed a single putative mtHSP70 sequence (Pelo_10550) branching sister to *M. balamuthi* mtHSP70 within the mitochondrial clade (Additional file 2: Fig. S3). The other HSP70 paralogues fell into the ER or cytosolic clades, the latter being diversified in ten copies.

Although we have probably revealed only a fragment of the inventory needed for the protein import

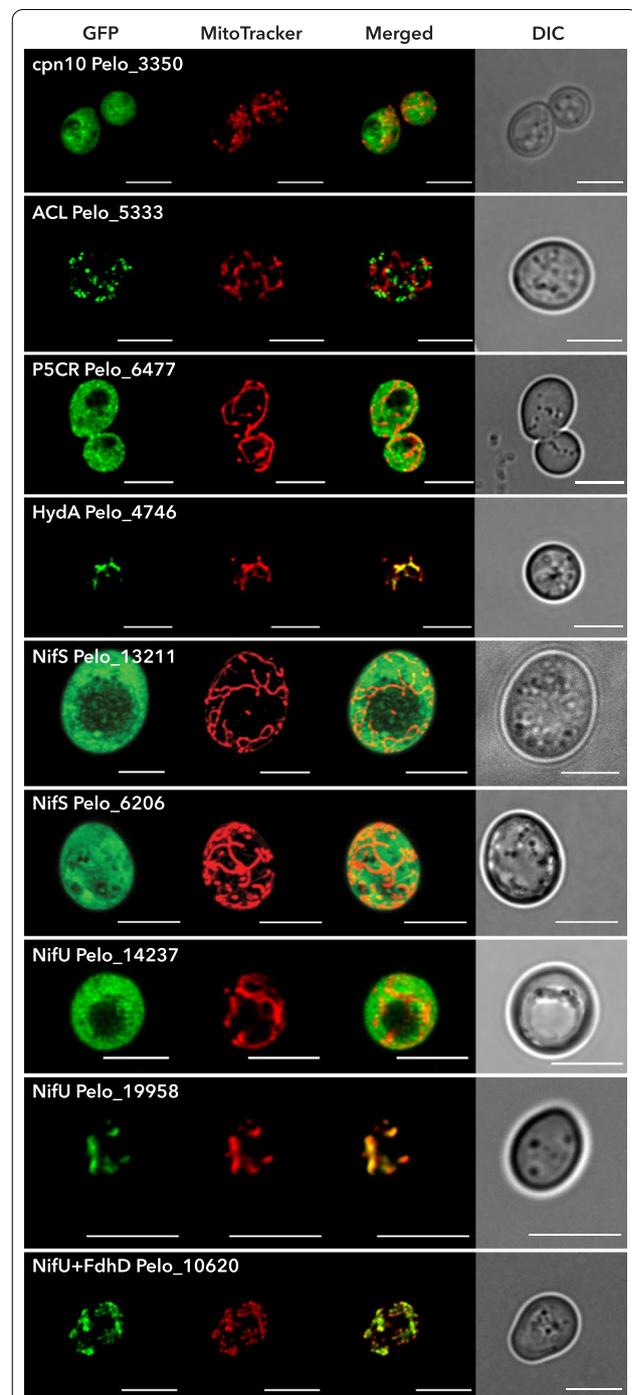
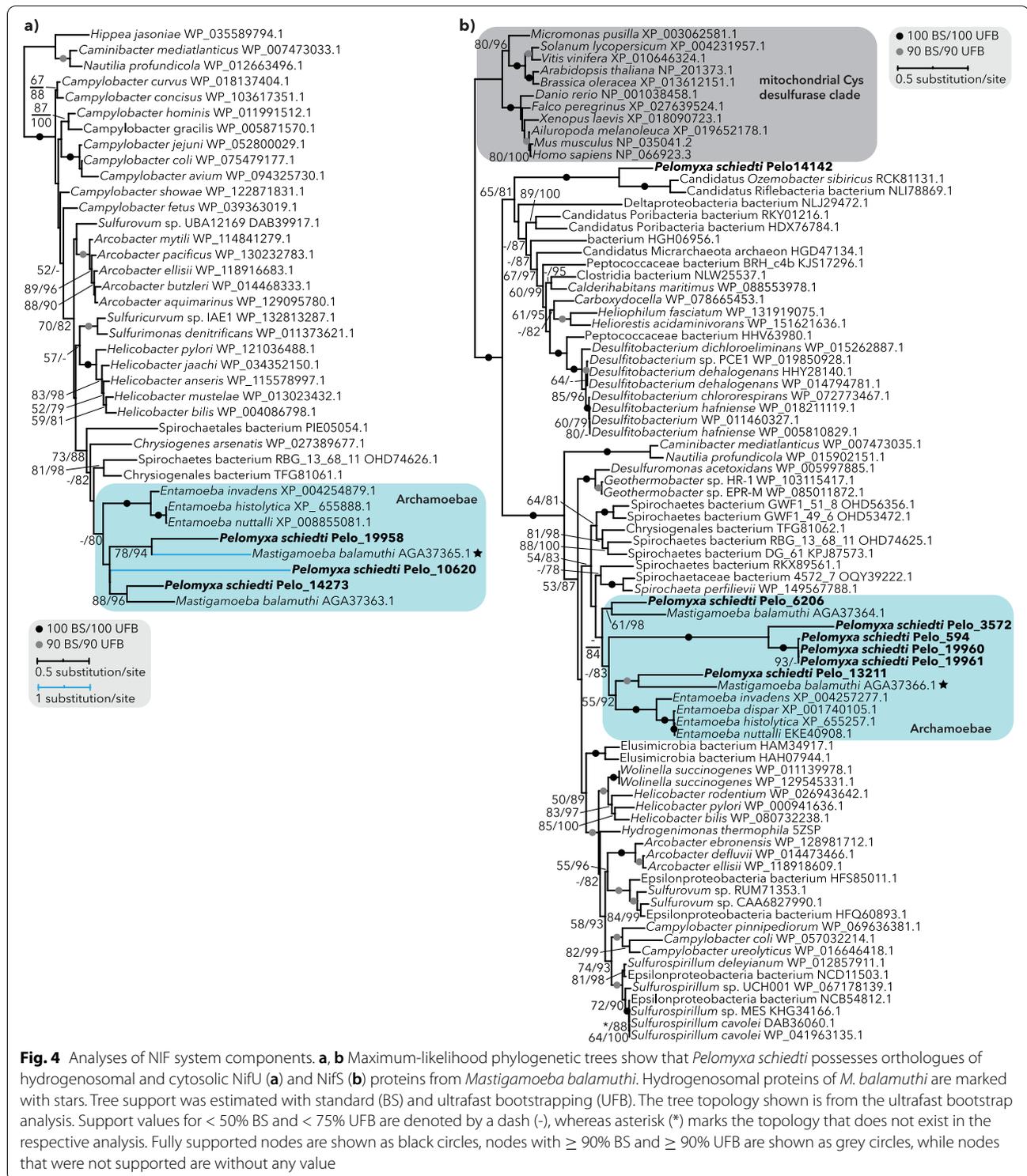


Fig. 3 Heterologous localization of candidate MRO proteins of *Pelomyxa schiedti*. C-terminally GFP-tagged proteins were expressed in *Saccharomyces cerevisiae*. Mitochondria were stained with MitoTracker. Note that gene sequences were amplified from cDNA using specific primers (Additional file 1: Table S4), except for P5CR that was synthesized in vitro. ACL, ATP-citrate lyase; cpn10, chaperonin 10; DIC, differential interference contrast; FdhD, formate dehydrogenase accessory sulfurtransferase; HydA, [FeFe]-hydrogenase; NifS, cysteine desulfurase; NifU, scaffold protein; P5CR, pyrroline-5-carboxylate reductase. Scale bar, 5 μ m



into the *P. schiedtii* MRO, the presence of mitochondrial hallmarks—Tom40, Sam50, mtHSP70, cpn60, and cpn10—conclusively shows that the MRO is truly present.

Tricarboxylic acid cycle and electron transport chain
Pelomyxa schiedti encodes four enzymes of the tricarboxylic acid (TCA) cycle possessing NTS (Additional file 1: Table S2) and catalyzing consecutive reactions. Instead

of citrate synthase, the common TCA enzyme, we identified an ATP citrate lyase (ACL) homologous to human ACL (EC: 2.3.3.8), which harbors citryl-CoA synthetase alpha and citryl-CoA lyase domains, and thus may provide oxaloacetate for the reductive direction of the TCA. However, ACL was only weakly predicted to localize in the MRO (i.e., by a single tool) and, when heterologously expressed in yeast, ACL localized into vesicles that did not overlap with stained mitochondria (Fig. 3). Thus, ACL was excluded from the list of putative MRO proteins. Other identified proteins, fumarate hydratase (fumarase/FH), malate dehydrogenase (MDH), and four subunits of the succinate dehydrogenase complex (SDH/complex II/CII), were robustly predicted to localize in the MRO. These proteins are common for both oxidative and reductive TCA, and the absence of the CII subunit SDH5/SDHAF involved in the flavination of the SDHA subunit [37] is likely common for Archamoebae as it is absent also from *M. balamuthi* [25].

Homologues of *A. castellanii* respiratory complexes were not identified, except for the aforementioned CII and a quinone-dependent electron-transferring flavoprotein (ETF; Additional file 1: Table S2). Both soluble subunits, alpha (ETFa) and beta (ETFb), and the membrane-bound ETF dehydrogenase (ETFDH), are present but only ETFDH and ETFa contain a recognizable NTS. It has been proposed in *M. balamuthi* that electrons may be transferred in an unknown direction between ETF and rhodoquinone (RQ), a quinone molecule with a lower electron potential than ubiquinone [3, 38]. RQ is synthesized in *M. balamuthi* by a hydrogenosomal methyltransferase dubbed RquA [38], which was also detected in *P. schiedti* (Additional file 1: Table S2). RQ presence allows the delivery of electrons to CII that could function as fumarate reductase [39] producing succinate, the putative end product of the partial reversed TCA in both archamoebal species [25], which may be secreted as in *Trypanosoma* [40].

Pyruvate and ATP metabolism

Pyruvate is oxidatively decarboxylated in aerobic mitochondria to acetyl-coenzyme A (CoA) by the pyruvate dehydrogenase (PDH) complex. In most MROs, PDH is substituted by pyruvate:ferredoxin oxidoreductase

(PFO), pyruvate:NADP⁺ oxidoreductase (PNO), or pyruvate formate lyase (PFL) [2]. We identified six copies of PFO and one copy of PNO in the *P. schiedti* genome, all without a predicted NTS (Additional file 1: Table S2). However, one of the *P. schiedti* PFOs was sister to one of the putatively hydrogenosomal PFOs in *M. balamuthi* [25] (Additional file 2: Fig. S4) and hence might represent a PFO homologue that operates in the *P. schiedti* MRO. Another pyruvate-metabolizing enzyme predicted to function in the MRO is pyruvate carboxylase (PC; Additional file 1: Table S2), which catalyzes carboxylation of pyruvate to oxaloacetate [41], a substrate of MDH. In *M. balamuthi*, pyruvate may be produced by the activity of NAD⁺-dependent D-lactate dehydrogenases (D-LDH), one of which localizes in the hydrogenosome and the other in the peroxisome [17, 25]. However, *P. schiedti* bears only one homologue of D-LDH that was predicted to localize in its peroxisomes (Additional file 1: Table S3). Since we have no evidence for any pyruvate-producing enzyme in MRO, we hypothesize that pyruvate is most likely imported from the cytosol.

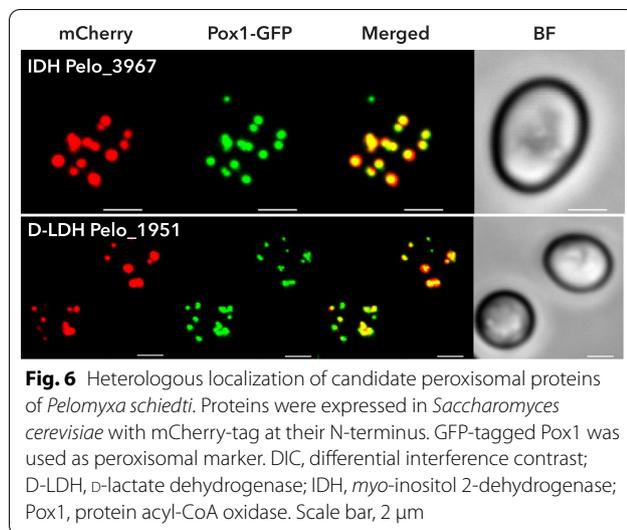
Two ATP-synthesizing enzymes are putatively present. Acetyl-CoA synthetase (ACS), an enzyme converting acetyl-CoA to acetate, CoA, and ATP, was found in eight copies, four of which possessed a putative NTS. ATP may also be formed by an adenylate kinase (AK) catalyzing the interconversion of adenine nucleotides. Three of the six AKs that we identified are putatively localized in the MRO (Additional file 1: Table S2). In this respect, the situation resembles that of the *M. balamuthi* hydrogenosome, in which both enzymes produce ATP [25]. A third putative source of ATP is the antiport against PAPS produced via the sulfate activation pathway (see below).

Amino acid metabolism

The glycine cleavage system (GCS) is at least partially retained in many MROs [42]. It consists of four enzymes (H-, L-, T-, and P-protein) and methylates tetrahydrofolate (THF) while decomposing glycine into CO₂ and ammonia. THF methylation is also provided by serine hydroxymethyltransferase (SHMT) [43]. We identified all GCS enzymes and SHMT in *P. schiedti*, all with predicted NTS (Fig. 2; Additional file 1: Table S2). L-protein

(See figure on next page.)

Fig. 5 Overview of the *Pelomyxa schiedti* peroxisomal metabolism. Peroxins were identified by BLAST searches. Peroxisomal matrix proteins were predicted by searching for peroxisomal targeting signals (PTS; Additional file 1: Table S3). Proteins that localized in *S. cerevisiae* peroxisomes (Fig. 6) are in bold and underlined. AAT, aspartate aminotransferase; ADH, alcohol dehydrogenase; GDH, glutamate dehydrogenase; GPI, glucose-6-phosphate isomerase; IDH, *myo*-inositol 2-dehydrogenase; LDH, D-lactate dehydrogenase; MPI, mannose-6-phosphate isomerase; Nudt, nudix hydrolase; P5CR, pyrroline-5-carboxylate reductase; PCL, phenylacetate-CoA ligase; PMP, peroxisomal membrane protein; PNPO, pyridoxamine 5'-phosphate oxidase; PPP, peroxisomal processing peptidase; PTS, peroxisomal targeting signal; PSPH, phosphoserine phosphatase; TDH, tartrate dehydrogenase. Numbers indicate peroxins



was present in two copies with only one bearing an NTS, similarly to *M. balamuthi*, in which the function of the cytosolic copy is unknown [25]. Lipoamide protein ligase (LPLA), necessary for lipoamide attachment to GCSH, was present with an NTS. The resulting N⁵,N¹⁰-methylene tetrahydrofolate (CH₂-THF) is an intermediate in one-carbon metabolism and a cofactor for the synthesis of pyrimidines and methionine in both mitochondria and the cytosol. Two cytosolic enzymes requiring this cofactor, B12-dependent methionine synthase, and THF dehydrogenase/cyclohydrolase, were detected (Additional file 1: Table S2). Glycine can be produced in mitochondria from threonine by threonine dehydrogenase (TDH) and alpha-amino-beta-ketobutyrate CoA ligase (AKL) [44] but both proteins lack a recognizable NTS in *P. schiedti* (Additional file 1: Table S2). Consistently, TDH activity was measured only in the cytosolic fraction of *M. balamuthi* [25]. It is highly probable that this pathway operates in the cytosol of *P. schiedti* and that glycine is imported into the MRO.

To our surprise, pyrroline-5-carboxylate reductase (P5CR) was predicted to be mitochondrion-targeted by one predictor (Additional file 1: Table S2). This enzyme could potentially be involved in the proline degradation pathway, which operates in the mitochondria of the distantly related amoebozoan *A. castellanii* [22]. As P5CR represents the only part of the pathway with localization signal whatsoever (Additional file 1: Table S2), we rather tested its localization in the *S. cerevisiae* heterologous system. There, the enzyme was clearly localized in the cytosol (Fig. 3), diminishing the probability of localizing the proline degradation pathway in the *P. schiedti* MRO. Moreover, further examination of the P5CR sequence revealed the presence of a peroxisomal targeting signal

(Additional file 1: Table S3), which was masked in localization experiments by a C-terminal tag, and thus precluded its peroxisomal targeting in yeast. For these reasons, we included P5CR among putative peroxisomal proteins.

Cofactor regeneration

NADH produced by GCS would be reoxidized in most mitochondria by NADH dehydrogenases in the electron transport chain [45]. Since the electron transport chain is absent in *P. schiedti*, we explored other pathways for cofactor regeneration. One possibility is fermentation of aldehydes to alcohols by alcohol dehydrogenases [46] putatively targeted to the MRO (Additional file 1: Table S2). Another option is the reductive partial TCA cycle, running from oxaloacetate to succinate, consuming not only NADH but also electrons from ETFDH via CII [47]. Oxaloacetate may be produced from pyruvate by PC with ATP consumption or by the action of aspartate aminotransferase (Additional file 1: Table S2); however, the source of 2-oxoglutarate and aspartate and the fate of glutamate is unclear due to the absence of a glutamate-aspartate antiporter (Additional file 1: Table S2).

Pyridine nucleotide transhydrogenase (PNT) is predicted to localize to the MRO by a single predictor (Additional file 1: Table S2). PNT usually localizes in the inner mitochondrial membrane and pumps protons while transferring electrons between NADH and NADPH [48]. PNT is present in *M. balamuthi* and *E. histolytica* [24, 49]; however, in *E. histolytica*, it was shown to localize outside mitochondria [50], which calls into question its MRO localization in other Archamoebae.

The *P. schiedti* MRO contains an ETF complex and a set of [FeFe]-hydrogenases, two additional electron sinks. ETF and ETFDH proteins are known to use electrons from the oxidation of fatty acids, a pathway that is absent in the *P. schiedti* MRO. [FeFe]-hydrogenases uptake electrons from reduced ferredoxins and produce molecular hydrogen. Three of the six detected hydrogenases bear putative NTS, and indeed, the one with the strongest prediction (Pelo_4746) was targeted to *S. cerevisiae* mitochondria (Fig. 3). Hydrogenases contain catalytic H clusters, and their maturation is dependent on maturases HydE, HydF, and HydG [51], which are all present and contain NTS (Additional file 1: Table S2). Meanwhile, reduced ferredoxin could originate from pyruvate oxidation.

Iron-sulfur cluster assembly

Mitochondria usually house the iron-sulfur (FeS) cluster assembly (ISC) pathway inherited from the alphaproteobacterial ancestor and serve for the maturation of both mitochondrial and cytosolic FeS proteins [52]. Some

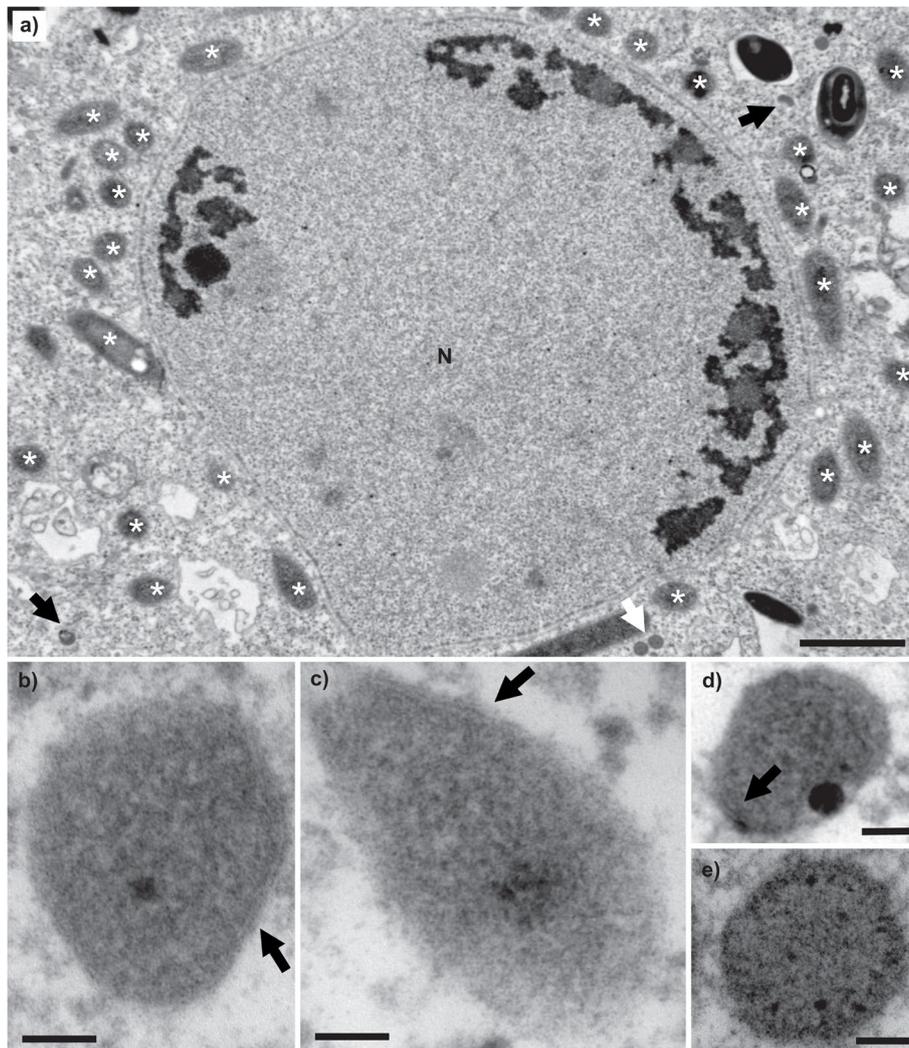


Fig. 7 Transmission electron micrograph of *Pelomyxa schiedti*, ultra-thin sections. **a** Nuclear area. N, nucleus; black arrow, putative mitochondrion-related organelle; white arrow, small dense body (putative peroxisome); asterisk (*), prokaryotic endosymbiont. **b–d** High magnification of putative mitochondrion-related organelle. Specimens were fixed with 0.1 M cacodylate (**b, c**) or PHEM buffer (**d**). Black arrow, bounding double membrane. **e** High magnification of small dense body (putative peroxisome). Scale bars, 400 nm for **a** and 50 nm for **b–e**

organisms, including Archamoebae, have replaced it with another pathway via horizontal gene transfer [26, 53]. *M. balamuthi* bears two copies of the nitrogen fixation (NIF) system, both comprising NifS and NifU proteins. While one pair of NIFs operates in the cytosol, the other localizes in the hydrogenosome [26]. In *E. histolytica*, only the cytosolic copy has been retained [25].

In the *P. schiedti* MRO, hydrogenases and their maturases HydE and HydF, SDH, ferredoxin, and PFO are putative clients for the NIF system and the situation with the gene inventory of the NIF system is by far the most complex among the investigated Archamoebae. We identified seven NifS and three NifU proteins, of which only

one NifU (Pelo_10620) contained predicted NTS (Additional file 1: Table S2). Interestingly, this protein consists of a NifU N-terminal domain fused to a formate dehydrogenase accessory sulfurtransferase (FdhD) C-terminal domain (Additional file 2: Fig. S5a). In *Escherichia coli*, FdhD transfers sulfur from IscS to formate dehydrogenase (FdhF) and is thus essential for the FdhF activity [54]. *P. schiedti* indeed encodes a FdhF homologue but without an NTS (Additional file 1: Table S2). In the NifU phylogeny (Fig. 4a), the NifU domain of the fusion protein formed a long branch within a moderately supported (80% ultrafast bootstrap) clade of all archamoebal NifUs. The other *P. schiedti* NifU sequences branched sister to

hydrogenosomal and cytosolic *M. balamuthi* homologues. All three *P. schiedti* NifU sequences contained conserved cysteine residues (Additional file 2: Fig. S5b) necessary for their function [55].

None of the seven NifS proteins was predicted to localize in the MRO (Additional file 1: Table S2). Two sequences were identical but incomplete at their C-termini and could not be completed by read mapping or PCR amplification. In our phylogenetic analysis (Fig. 4b), Pelo_13211 and Pelo_6206 branched sister to the hydrogenosomal and cytosolic *M. balamuthi* sequences, respectively. Pelo_14142 was sister to Candidati Riflebacteria species, and the remaining four sequences formed a long branch nested within the Archamoebae clade. All amino acid residues required for functioning [56] were present in both of the *P. schiedti* sequences that were sister to *M. balamuthi* (Additional file 2: Fig. S5c).

This phylogenetic pattern offers an elegant hypothesis in which NifS Pelo_6206 and NifU Pelo_14273 act in the cytosol, while NifS Pelo_13211 and NifU Pelo_19958 act in the MRO. The remaining NifS copies might be functional partners of the NifU-FdhD fusion protein (Pelo_10620). Our experiments with heterologous localization in *S. cerevisiae* indeed revealed mitochondrial localization of the NifU Pelo_19958, but surprisingly also of the NifU-FdhD fusion protein (Pelo_10620), and at the same time, both hypothesized NifS sequences were localized in the cytosol (Fig. 3). Although the situation with the NIF system in *P. schiedti* cells remains unclear, the mitochondrial localization of two NifU homologues in yeast supports the presence of the FeS cluster assembly in the *P. schiedti* MRO with the identity of its NifS partner unresolved.

Sulfate activation pathway

The sulfate activation pathway produces PAPS, which is necessary for sulfolipid synthesis [23]. It is present in *E. histolytica* [23, 57] and *M. balamuthi* [25] MROs, and we also identified all of its components in *P. schiedti* (Fig. 2; Additional file 1: Table S2). This pathway requires two transporters. A sodium/sulfate symporter (NaS) is necessary for substrate delivery; however, its homologues in *P. schiedti* (Additional file 1: Table S2) are unrelated to *E. histolytica* mitochondrial NaS [23] (Additional file 2: Fig. S6a) making their role unclear. The PAPS exporter belongs to the mitochondrial carrier family (MCF) and, indeed, one of the *P. schiedti* MCF proteins branched sister to a clade of PAPS transporters of *E. histolytica* and *M. balamuthi* [29, 58] (Additional file 2: Fig. S6b). As this transporter exchanges PAPS with ATP, it plays a role in supplementing the ATP pool in the MRO, yet it cannot provide a net ATP

gain, because two ATP molecules are required for the production of one PAPS.

Anaerobic peroxisomes

We also investigated the presence of anaerobic peroxisomes, which were recently characterized in *M. balamuthi* and *E. histolytica* [17, 18]. *P. schiedti* encodes genes for 13 proteins required for peroxisome biogenesis (peroxins, Pexs) strongly supporting the presence of peroxisomes. The identified Pexs include Pex5 and Pex7, which are required for the recognition of peroxisomal targeting signals 1 and 2 (PTS1 and PTS2), respectively; Pex13 and 14, which mediate protein import; Pex1, 2, 6, 10, and 12, which are receptor-recycling Pexs; Pex3, 16, and 19, which are involved in protein import to the peroxisomal membrane; and Pex11, which participates in peroxisome fission (Fig. 5; Additional file 1: Table S3). Interestingly, the sets of Pexs in *P. schiedti* and *M. balamuthi* are identical, with the exception of two paralogs of Pex11 being found in *M. balamuthi* [17]. Thus, it seems that this set of Pexs is common to free-living Archamoebae. In contrast, in the parasitic *Entamoeba*, either the set of Pexs was reduced to only seven components (*E. histolytica*), or peroxisomes were completely lost (*E. invadens*) [18].

Prediction of putative peroxisomal matrix proteins based on the presence of PTS1/PTS2, revealed 67 candidates (Fig. 5; Additional file 1: Table S3). Of these, only four candidates had previously been found in the anaerobic peroxisomes of *M. balamuthi*. These include peroxisomal processing peptidase (PPP), *myo*-inositol 2-dehydrogenase (*myo*-IDH), nudix hydrolase, and D-lactate dehydrogenase (D-LDH), all of which had clear support for localization in *P. schiedti* peroxisomes (Additional file 1: Table S3). Two of them (*myo*-IDH and D-LDH) were selected for heterologous expression with an N-terminal mCherry tag in yeast, expressing the integrated GFP-tagged peroxisomal marker protein acyl-CoA oxidase (Pox1). Fluorescence microscopy revealed that both *myo*-IDH and D-LDH colocalized with Pox1 in round organelles in support of our predictions (Fig. 6).

Unlike in *M. balamuthi*, *P. schiedti* peroxisomes possibly contain pyridoxamine 5'-phosphate oxidase (PNPO; Additional file 1: Table S3), which utilizes molecular oxygen as an electron acceptor to catalyze the last step of the pyridoxal 5'-phosphate (PLP) biosynthesis with concomitant formation of ammonia and H₂O₂. The presence of PNPO raises the question of how H₂O₂ is detoxified as typical antioxidant enzymes, such as catalase and peroxidase, are not present. However, H₂O₂ could also be decomposed nonenzymatically by antioxidants, such as 2-oxoglutarate, in which the ketone group of the α -carbon atom reacts with H₂O₂ to form succinate, CO₂, and water [59]. Indeed, *P. schiedti* contains a putative

glutamate dehydrogenase that may produce 2-oxoglutarate and that possesses an –SKL triplet, a typical PTS1 (Additional file 1: Table S3) although its peroxisomal targeting was not supported by the PTS predictor. The other proteins with predicted peroxisomal localization include several enzymes for amino acid synthesis and degradation, carbohydrate metabolism, and hydrolases, albeit without clear biochemical context. More experimental studies are required to verify our predicted localizations and to delineate the function of peroxisomes in *P. schiedti*.

Electron microscopy

Finally, we were interested in whether the two organelles described by the genomic data can be visualized by transmission electron microscopy. Careful inspection of micrographs revealed two populations of small vesicles, one bounded by a double membrane and the other by a single membrane (Fig. 7). We ascribe them respectively to putative MROs and peroxisomes characterized in this work in silico but leave confirmation of this for further studies.

Conclusions

Our bioinformatic survey of the putative proteome of the *Pelomyxa schiedti* MRO revealed several interesting insights and opened many questions for further investigation of this amoeba, which represent another, understudied lineage of Archamoebae. Most importantly, *P. schiedti* clearly does harbor an MRO with a very streamlined or lineage-specific set of protein translocases, and peroxisomes with a set of 13 soluble and membrane-associated peroxins. The MRO metabolism predicted in silico is similar to that of *M. balamuthi*, and it probably resembles the situation in the common free-living ancestor of this group of amoebae. Still, the situation with FeS cluster assembly in this amoeba seems rather complex and interesting as it contains the most diverse set of NIF pathway proteins of all previously investigated Archamoebae. These proteins very likely provide parallel FeS synthesis in MRO and cytosol, but in addition to this and uniquely in eukaryotes, some of them may be involved in the activation of formate dehydrogenase as seen in some prokaryotes. *P. schiedti* anaerobic peroxisomes, similarly to those of *M. balamuthi* and *E. histolytica*, lack enzymes for the β -oxidation of fatty acids and catalase, but the predicted set of enzymes differs considerably from other Archamoebae. Only four enzymes are shared between at least two species, and two of these (*myo*-IDH and D-LDH) were shown to localize in yeast peroxisomes, making them credible peroxisomal enzymes not only in *P. schiedti* but also in other Archamoebae. Although the role of these peroxisomes needs to be clarified, they seem

to be metabolically diversified among Archamoebae, and very distinct from their aerobic counterparts.

Material and methods

Cell culture

Polyxenic (and polyeukaryotic) cultures of *Pelomyxa schiedti* strain SKADARSKE were maintained in Sonneborn's *Paramecium* medium [60] as described previously [27].

Genome and transcriptome sequencing and assembly

Genome sequencing was performed from whole genome amplified DNA (WGA). Individual cells were picked by micromanipulation and washed twice in Trager U media [61]. Genomic DNA was amplified using the Illustra Single Cell GenomiPhi DNA Amplification Kit (GE Healthcare Life Sciences) according to the manufacturer's protocol and purified using ethanol precipitation. To confirm the presence of eukaryotic DNA, a fragment of the actin gene was amplified by PCR using specific primers (Additional file 1: Table S4). Sequencing libraries from seven positive samples were prepared using the Illumina TruSeq DNA PCR-Free kit (Illumina). Samples Pelo2 and Pelo5 were sequenced using the Illumina MiSeq (2 × 300 bp; Genomic Core Facility, Faculty of Science) and Nanopore (Oxford Nanopore Technologies) systems; samples P1–P5 were sequenced using the Illumina HiSeq X (Macrogen, Inc.). The Nanopore library was prepared using the Oxford Nanopore Technologies ligation sequencing kit (SQK-LSK108) from 4 μ g of T7 endonuclease I-treated (New England Biolabs) WGA. Sequencing was performed using a R9.4.1 Spot-On Flow cell (FLO-MIN106) for 48 h.

For transcriptome sequencing, single cells of *P. schiedti* were washed twice in Trager U medium, and transcriptomes were generated using the Smart Seq2 [62] protocol with 19 cycles of amplification. Five libraries were prepared using the Nextera XT DNA Library Preparation Kit (Illumina) and sequenced on the Illumina MiSeq platform (PE 2 × 300 bp; Genomic Core Facility, Faculty of Science).

Raw Illumina DNA- and RNA-Seq reads were quality- and adapter-trimmed using BBDuk v36.92 (part of BBTools suite: <https://jgi.doe.gov/data-and-tools/bbtools/>). Firstly, individual single-cell genome assemblies for Pelo2, Pelo5, and P1–P5 were generated with SPAdes v3.11.1 [63] using single-cell (--sc) mode and a *k*-mer size of 127. As all 18S rDNAs extracted from individual assemblies were identical, all reads (i.e., Illumina HiSeqs and MiSeq, and Nanopore) were assembled together by SPAdes v3.11.1 using --sc and *k*-mers of 21, 33, 55, 77, 99, and 121. The resulting assembly was binned and decontaminated using tetraESOM [64] and a BLASTing

strategy as described previously [65]. The final assembly was scaffolded using the P_RNA_scaffolder [66]. Prediction was done using Augustus v3.3.1 [67], and further improved by PASA and EVM [68] using our transcriptomic data. RNA-Seq reads were assembled using Trinity v2.6.5 [69] with default parameters, and contaminants were removed by BLASTing against the decontaminated genome assembly. RNA-Seq reads were mapped to the transcriptome using Bowtie2 v2.3.0 [70] and to the genome using HISAT2 v2.0.5 [71]. Genome and transcriptome completeness were assessed using BUSCO v3 with the eukaryota_odb9 dataset [72].

Sequence searches and localization predictions

Proteins predicted to localize in the *M. balamuthi* hydrogenosome, *E. histolytica* mitosome, and *A. castellanii* mitochondria served as queries for BLAST v2.6.0+ [73] searches in *P. schiedti* assemblies. Sensitive searches for components of the TOM/TIM machinery were done using HMMER v3.3 [74]. Protein domains were predicted by InterProScan [75] as implemented in Geneious Prime v2020.2.3 [76].

Potentially mitochondrion-targeted proteins were identified using TargetP v2 [77], PSORT II [78], MultiLoc2 [79], and NomPred [80] tools. Since *P. schiedti* does not harbor plastid, the plant setting from TargetP and MultiLoc2 was omitted. NomPred was used in the MRO and in the *Dictyostelium* settings. A protein was considered as mitochondrial if predicted by at least one setting of MultiLoc2 or NomPred.

Peroxisins were identified by BLAST searches using *M. balamuthi* queries. Peroxisomal matrix proteins were predicted by searching for peroxisomal targeting signals (PTS). The tripeptides SRI and [SAP][KR][LM] (excluding AKM, PKM, and PRM) were used to search for the C-terminal PTS1. A proline at position -3 and a methionine at position -1 were included based on experimental verification in *M. balamuthi* [17]. Two nanopptides R[LI](x5)HL were used for N-terminal PTS2 searches [81]. All putative transmembrane proteins determined by TMHMM Server v2.0 [82], were filtered out. PTS1 candidates were submitted to the PTS1 Predictor [83] using the GENERAL function evaluating twelve C-terminal residues.

Phylogenetic analyses

The 18S rRNA gene dataset was aligned by MAFFT v7 [84] server with the G-INS-i algorithm using default settings. The alignment was manually edited in BioEdit v7.0.4.1 [85] resolving 1437 positions. A phylogenetic tree was constructed using maximum-likelihood by RAxML v8.0.0 [86] under the GTRGAMMAI model, 100 starting trees, and 1000 standard bootstrap pseudoreplicates.

For selected proteins, datasets were aligned by MAFFT v7.313 [84], trimmed by trimAl v1.4 [87], and maximum-likelihood trees were inferred by IQ-TREE v1.6.8 [88] using the posterior mean site frequency method [89] and the LG+C20+F+G model, with the guide tree inferred under the LG+F+G model. Branch supports were obtained by the ultrafast bootstrap approximation [90] with 1000 replicates. For NIF components, standard bootstrapping with 100 replicates was used as well.

Localization assays

Most of the localized genes were amplified from cDNA using specific primers (Additional file 1: Table S4) and PrimeSTAR[®] Max DNA Polymerase (Takara Bio, Inc.) premix. Only the P5CR (Pelo_6477) gene sequence was synthesized in vitro (GenScript). For mitochondrial localization, the genes were cloned into the pUG35 vector containing C-terminal green fluorescence protein (GFP) and transformed into *S. cerevisiae* strain BY4742 using the lithium acetate method [91]. Transformants were grown on a selective medium without uracil (SD-URA) at 30 °C. For localization, the transformed cells were incubated with MitoTracker Red CMXRos (1:10,000; Thermo Fisher Scientific) for 10 min, followed by two washes with PBS, and mounted in 1% low-melting agarose. Cells were imaged using a Leica SP8 confocal microscope. Deconvolution was performed using Huygens Professional v17.10 and ImageJ v1.50b. For heterologous expression of peroxisomal candidates in yeast, the genes were cloned into the pTVU100 vector for protein expression, with an N-terminal mCherry tag, in *S. cerevisiae* strain BY4742:POX1-EGFP, expressing the GFP-tagged peroxisomal marker protein acyl-CoA oxidase (Pox1) [17, 92]. Selected transformants were inoculated in oleate medium to stimulate peroxisome formation, followed by microscopy [17].

Transmission electron microscopy

A grown culture of *P. schiedti* was pelleted by centrifugation and fixed for 1 h on ice with 2.5% glutaraldehyde in 0.1 M cacodylate buffer (pH 7.2). After washing in 0.1 M cacodylate buffer, the cells were postfixed for 1 h on ice with 1% OsO₄. After washing with distilled water, the fixed cells were dehydrated in a graded series of ethanol, transferred to acetone, and embedded in EPON resin. Alternatively, the protocol was modified several times using different buffers (PHEM and PBS buffer) and a fixation time of 30 min. Ultrathin sections were prepared on an ultramicrotome (Reichert-Jung Ultracut E) with a diamond knife. Sections were stained with uranyl acetate and lead citrate and examined using a JEOL 1011 transmission electron microscope.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12915-022-01247-w>.

Additional file 1: Table S1. Statistics of *Pelomyxa schiedti* assemblies were compared with those of *Mastigamoeba balamuthi*. **Table S2.** Proteins targeted to MRO of *Pelomyxa schiedti*. Localization of proteins was predicted by several tools, as listed in columns E - J. Mitochondrial predictions are highlighted by white text on blue background. Column K shows final inferred prediction of localization. Abbreviations: cyt, cytosolic; ER, endoplasmic reticulum; extracell, extracellular; mit, mitochondrial; nuc, nuclear; Other, other localization; perox, peroxisomal; PM, plasma membrane; sec, secretory system; SP, signal peptide; -, protein not localized in MRO; +, protein localized in MRO; +?, protein localized in MRO with low confidence. **Table S3.** Proteins required for peroxisome biogenesis and targeted to peroxisome. Proteins identified in *Mastigamoeba balamuthi* are highlighted by white text on blue background. Proteins were considered peroxisome-targeted, if they contained PTS1 (SRI or [SAP][KR][LM]) or PTS2 motif (R[LIV](x5)HL), and/or were predicted by PTS1 predictor [83]. **Table S4.** Primers used in this study.

Additional file 2: Figure S1. Phylogenetic analysis of amoebozoan 18S rDNA. Maximum-likelihood tree places *Pelomyxa schiedti* in monophyletic Pelomyxidae group inside monophyletic Archamoebae. Standard bootstrap support values shown when $\geq 50\%$. **Figure S2.** BUSCO analysis of *Pelomyxa schiedti* transcriptome and predicted proteins. Completeness of *P. schiedti* datasets were assessed using odb9_eukaryota dataset and compared with completeness of predicted proteins from *Mastigamoeba balamuthi*. **Figure S3.** Phylogenetic analysis of HSP70 proteins. Maximum-likelihood phylogenetic tree documents that one *Pelomyxa schiedti* HSP70 sequence is related to mitochondrial orthologues from other eukaryotes. Ultrafast bootstrap support values shown when $\geq 75\%$. ER, endoplasmic reticulum. **Figure S4.** Phylogenetic analysis of PFO enzymes. Maximum-likelihood phylogenetic tree identifies PFO version putatively operating in *Pelomyxa schiedti* MRO. Hydrogenosomal PFO copies of *Mastigamoeba balamuthi* marked with stars. Number in parentheses shows number of species in the collapsed clade. Ultrafast bootstrap support values shown when $\geq 75\%$. **Figure S5.** Sequences of *Pelomyxa schiedti* components of Nif system. a) *P. schiedti* protein Pefo_10620, composed of NifU N-terminal domain fused to FdhD (formate dehydrogenase accessory sulfurtransferase) C-terminal domain, as determined by InterProScan. b-c) Sequence alignment of NifU (b) and NifS (c) proteins from *P. schiedti* and *Mastigamoeba balamuthi* in comparison with bacterial homologues from *Thermotoga maritima*. Amino acid residues necessary for function of NifU and NifS are labeled according to graphical legend. **Figure S6.** Phylogenetic analysis of transporters involved in sulfate activation pathway. a) Phylogenetic analysis did not resolve which of the sodium/sulfate symporters of *Pelomyxa schiedti* is related to *Entamoeba histolytica* mitochondrial transporter. b) Maximum-likelihood phylogenetic tree confirms one *P. schiedti* transporter as PAPS (3'-phosphoadenosine 5'-phosphosulfate) transporter, while two others belong to broader family of mitochondrial carrier transporters. Experimentally proven mitochondrial transporters of *E. histolytica* marked with stars. Ultrafast bootstrap support values shown when $\geq 75\%$.

Acknowledgements

We thank Dr. Zoltán Füssy (Charles University, BIOCEV, Vestec, Czech Republic) and Dr. Courtney W. Stairs (Lund University, Lund, Sweden) for helpful discussions and Dr. Aaron A. Heiss (Charles University, BIOCEV, Vestec, Czech Republic) for proofreading the manuscript.

Authors' contributions

SCT cultured the cells, prepared the sequencing libraries, and performed the Nanopore sequencing. KZ and SCT assembled the genome and transcriptome, and predicted the proteins. IČ provided the culture of *P. schiedti* and conducted the 18S rDNA phylogeny. KZ performed most of the bioinformatic and phylogenetic analyses. KZ, IŠ-S, JT, and VH analyzed the metabolism of the mitochondrion-related organelle. TL and JT analyzed the peroxisomal metabolism. SCT and TL performed the heterologous localizations. PH performed the

microscopic observations. JT and VH supervised the project. KZ, JT, and VH wrote the manuscript. All authors read and approved the final manuscript.

Funding

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 771592 to VH), from the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic (CR) in the Centre for research of pathogenicity and virulence of parasites (project No. CZ.02.1.01/0.0/0.0/16_019/0000759 to VH and JT), from the Czech Science Foundation (grant No. 19-19297S to IČ), and from the University Research Centre of Charles University (UNCCE, SCI/12 to JT). We acknowledge the Imaging Methods Core Facility at BIOCEV, supported by the MEYS CR (Large RI Project LM2018129 Czech-Biolmaging) and ERDF (project No. CZ.02.1.01/0.0/0.0/16_013/0001775 and CZ.02.1.01/0.0/0.0/18_046/0016045) for their support with obtaining imaging data. Microscopic observations in a transmission electron microscope were performed in the Imaging Methods Core Facility of Biology Section of Faculty of Science – Viničná. Computational resources were supplied by the project “e-Infrastruktura CZ” (e-INFRA LM2018140) project, supported by the MYES CR.

Availability of data and materials

All data needed to evaluate the conclusions in the paper are present in the paper and/or the supplementary information files. The raw sequencing data and final assemblies are available at NCBI (<https://www.ncbi.nlm.nih.gov/>) as BioProject PRJNA672820 [93].

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Parasitology, Faculty of Science, Charles University, BIOCEV, Vestec, Czech Republic. ²Institute of Parasitology, Biology Centre, Czech Academy of Sciences, České Budějovice (Budweis), Czech Republic. ³Department of Biochemistry, Faculty of Natural Sciences, Comenius University, Bratislava, Slovakia. ⁴Department of Zoology, Faculty of Science, Charles University, Prague, Czech Republic.

Received: 16 July 2021 Accepted: 3 February 2022

Published online: 01 March 2022

References

- Müller M, Mentel M, van Hellemond JJ, Henze K, Woehle C, Gould SB, et al. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol Mol Biol Rev.* 2012;76:444–95.
- Roger AJ, Muñoz-Gómez SA, Kamikawa R. The origin and diversification of mitochondria. *Curr Biol.* 2017;27:R1177–92.
- Gawryluk RMR, Stairs CW. Diversity of electron transport chains in anaerobic protists. *Biochim Biophys Acta Bioenerg.* 2021;1862:148334.
- Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Di B, Juzokaite L, Vancaester E, et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature.* 2017;541:353–8.
- Sagan L. On the origin of mitosing cells. *J Theor Biol.* 1967;14:255–74.
- Martijn J, Vosseberg J, Guy L, Offre P, Ettema TJG. Deep mitochondrial origin outside the sampled alphaproteobacteria. *Nature.* 2018;557:101–5.
- Gawryluk RMR, Kamikawa R, Stairs CW, Silberman JD, Brown MW, Roger AJ. The earliest stages of mitochondrial adaptation to low oxygen revealed in a novel rhizarian. *Curr Biol.* 2016;26:2729–38.
- Leger MM, Kolisko M, Kamikawa R, Stairs CW, Kume K, Čepička I, et al. Organelles that illuminate the origins of *Trichomonas* hydrogenosomes and *Giardia* mitochondria. *Nat Ecol Evol.* 2017;1:0092.

9. Karnkowska A, Vacek V, Zubáčová Z, Treitli SC, Petrželková R, Eme L, et al. A eukaryote without a mitochondrial organelle. *Curr Biol*. 2016;26:1274–84.
10. Panigrahi AK, Ogata Y, Zíková A, Anupama A, Dalley RA, Acestor N, et al. A comprehensive analysis of *Trypanosoma brucei* mitochondrial proteome. *Proteomics*. 2009;9:434–50.
11. Lee CP, Taylor NL, Harvey MA. Recent advances in the composition and heterogeneity of the *Arabidopsis* mitochondrial proteome. *Front Plant Sci*. 2013;4:4.
12. Gawryluk RMR, Chisholm KA, Pinto DM, Gray MW. Compositional complexity of the mitochondrial proteome of a unicellular eukaryote (*Acanthamoeba castellanii*, supergroup Amoebozoa) rivals that of animals, fungi, and plants. *J Proteome*. 2014;109:400–16.
13. Žárský V, Tachezy J. Evolutionary loss of peroxisomes - not limited to parasites. *Biol Direct*. 2015;10:74.
14. Schlüter A, Fourcade S, Ripp R, Mandel JL, Poch O, Pujol A. The evolutionary origin of peroxisomes: an ER-peroxisome connection. *Mol Biol Evol*. 2006;23:838–45.
15. Gabaldón T. Peroxisome diversity and evolution. *Philos Trans R Soc B Biol Sci*. 2010;365:765–73.
16. Gabaldón T, Ginger ML, Michels PAM. Peroxisomes in parasitic protists. *Mol Biochem Parasitol*. 2016;209:35–45.
17. Le T, Žárský V, Nývltová E, Rada P, Harant K, Vancová M, et al. Anaerobic peroxisomes in *Mastigamoeba balamuthi*. *Proc Natl Acad Sci U S A*. 2020;117:2065–75.
18. Verner Z, Žárský V, Le T, Narayanasamy RK, Rada P, Rozbeský D, et al. Anaerobic peroxisomes in *Entamoeba histolytica* metabolize myo-inositol. *PLoS Pathog*. 2021;17:e1010041.
19. Pánek T, Zadrožilková E, Walker G, Brown MW, Gentekaki E, Hroudová M, et al. First multigene analysis of Archamoebae (Amoebozoa: Conosa) robustly reveals its phylogeny and shows that Entamoebidae represents a deep lineage of the group. *Mol Phylogenet Evol*. 2016;98:41–51.
20. Kang S, Tice AK, Spiegel FW, Silberman JD, Pánek T, Čepička I, et al. Between a pod and a hard test: the deep evolution of Amoeboae. *Mol Biol Evol*. 2017;34:2258–70.
21. Pearce XG, Annesley SJ, Fisher PR. The *Dictyostelium* model for mitochondrial biology and disease. *Int J Dev Biol*. 2019;63:497–508.
22. Leger MM, Gawryluk RMR, Gray MW, Roger AJ. Evidence for a hydrogenosomal-type anaerobic ATP generation pathway in *Acanthamoeba castellanii*. *PLoS One*. 2013;8:e69532.
23. Mi-ichi F, Yousuf MA, Nakada-Tsukui K, Nozaki T. Mitosomes in *Entamoeba histolytica* contain a sulfate activation pathway. *Proc Natl Acad Sci*. 2009;106:21731–6.
24. Gill EE, Diaz-Triviño S, Barberá MJ, Silberman JD, Stechmann A, Gaston D, et al. Novel mitochondrion-related organelles in the anaerobic amoeba *Mastigamoeba balamuthi*. *Mol Microbiol*. 2007;66:1306–20.
25. Nývltová E, Stairs CW, Hrdý I, Ridl J, Mach J, Pačes J, et al. Lateral gene transfer and gene duplication played a key role in the evolution of *Mastigamoeba balamuthi* hydrogenosomes. *Mol Biol Evol*. 2015;32:1039–55.
26. Nývltová E, Šuták R, Harant K, Šedinová M, Hrdý I, Paces J, et al. 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*. 2013;110:7371–6.
27. Zadrožilková E, Walker G, Čepička I. Morphological and molecular evidence support a close relationship between the free-living Archamoebae *Mastigella* and *Pelomyxa*. *Protist*. 2015;166:14–41.
28. Seravin L, Goodkov A. Cytoplasmic microbody-like granules of the amoeba *Pelomyxa palustris*. *Tsitologiya*. 1987;29:600–3.
29. Žárský V, Klimeš V, Pačes J, Vlček Č, Hradilová M, Beneš V, et al. The *Mastigamoeba balamuthi* genome and the nature of the free-living ancestor of *Entamoeba*. *Mol Biol Evol*. 2021;38:2240–59.
30. Wang D, Hancock J. IntronDB: a database for eukaryotic intron features. *Bioinformatics*. 2019;35:4400–1.
31. Bolender N, Sickmann A, Wagner R, Meisinger C, Pfanner N. Multiple pathways for sorting mitochondrial precursor proteins. *EMBO Rep*. 2008;9:42–9.
32. Makiuchi T, Mi-ichi F, Nakada-Tsukui K, Nozaki T. Novel TPR-containing subunit of TOM complex functions as cytosolic receptor for *Entamoeba* mitochondrial transport. *Sci Rep*. 2013;3:1129.
33. Dolezal P, Dagley MJ, Kono M, Wolynec P, Likič VA, Foo JH, et al. The essentials of protein import in the degenerate mitochondrion of *Entamoeba histolytica*. *PLoS Pathog*. 2010;6:e1000812.
34. Schneider A. Evolution of mitochondrial protein import - lessons from trypanosomes. *Biol Chem*. 2020;401:663–76.
35. Makki A, Rada P, Žárský V, Kereiče S, Kováčik L, Novotný M, et al. Triplet-pore structure of a highly divergent TOM complex of hydrogenosomes in *Trichomonas vaginalis*. *PLoS Biol*. 2019;17:e3000098.
36. Garg S, Stöltgen J, Zimorski V, Rada P, Tachezy J, Martin WF, et al. Conservation of transit peptide-independent protein import into the mitochondrial and hydrogenosomal matrix. *Genome Biol Evol*. 2015;7:2716–26.
37. Hao H-X, Khalimonchuk O, Schradlers M, Dephoure N, Bayley J-P, Kunst H, et al. SDH5, a gene required for flavination of succinate dehydrogenase, is mutated in paraganglioma. *Science*. 2009;325:1139–42.
38. Stairs CW, Eme L, Muñoz-Gómez SA, Cohen A, Dellaire G, Shepherd JN, et al. Microbial eukaryotes have adapted to hypoxia by horizontal acquisitions of a gene involved in rhodoquinone biosynthesis. *Elife*. 2018;7:e34292.
39. Castro-Guerrero NA, Jasso-Chávez R, Moreno-Sánchez R. Physiological role of rhodoquinone in *Euglena gracilis* mitochondria. *Biochim Biophys Acta*. 2005;1710:113–21.
40. Coustou V, Besteiro S, Rivière L, Biran M, Biteau N, Franconi JM, et al. A mitochondrial NADH-dependent fumarate reductase involved in the production of succinate excreted by procyclic *Trypanosoma brucei*. *J Biol Chem*. 2005;280:16559–70.
41. Jitrapakdee S, St Maurice M, Rayment I, Cleland WW, Wallace JC, Attwood PV. Structure, mechanism and regulation of pyruvate carboxylase. *Biochem J*. 2008;413:369–87.
42. Leger MM, Eme L, Hug LA, Roger AJ. Novel hydrogenosomes in the microaerophilic jakobid *Stygiella incarcerata*. *Mol Biol Evol*. 2016;33:2318–36.
43. Kikuchi G. The glycine cleavage system: composition, reaction mechanism, and physiological significance. *Mol Cell Biochem*. 1973;1:169–87.
44. Dale RA. Catabolism of threonine in mammals by coupling of L-threonine 3-dehydrogenase with 2-amino-3-oxobutylate-CoA ligase. *Biochim Biophys Acta*. 1978;544:496–503.
45. Schertl P, Braun H-P. Respiratory electron transfer pathways in plant mitochondria. *Front Plant Sci*. 2014;5:163.
46. Thomson JM, Gaucher EA, Burgan MF, De Kee DW, Li T, Aris JP, et al. Resurrecting ancestral alcohol dehydrogenases from yeast. *Nat Genet*. 2005;37:630–5.
47. Hügl M, Wirsén CO, Fuchs G, Taylor CD, Sievert SM. Evidence for autotrophic CO₂ fixation via the reductive tricarboxylic acid cycle by members of the ε subdivision of proteobacteria. *J Bacteriol*. 2005;187:3020–7.
48. Jackson JB, Peake SJ, White SA. Structure and mechanism of proton-translocating transhydrogenase. *FEBS Lett*. 1999;464:1–8.
49. Yu Y, Samuelson J. Primary structure of an *Entamoeba histolytica* nicotinamide nucleotide transhydrogenase. *Mol Biochem Parasitol*. 1994;68:323–8.
50. Yousuf MA, Mi-ichi F, Nakada-Tsukui K, Nozaki T. Localization and targeting of an unusual pyridine nucleotide transhydrogenase in *Entamoeba histolytica*. *Eukaryot Cell*. 2010;9:926–33.
51. Kuchenreuther JM, Britt RD, Swartz JR. New insights into [FeFe] hydrogenase activation and maturase function. *PLoS One*. 2012;7:e45850.
52. Tachezy J, Doležal P. Iron-sulfur proteins and iron-sulfur cluster assembly in organisms with hydrogenosomes and mitosomes. In: Martin WF, Müller M, editors. *Origin of mitochondria and hydrogenosomes*. Berlin: Springer Berlin Heidelberg; 2007. p. 105–33.
53. Stairs CW, Eme L, Brown MW, Mutsaers C, Susko E, Dellaire G, et al. A SUF Fe-S cluster biogenesis system in the mitochondrion-related organelles of the anaerobic protist *Pygusua*. *Curr Biol*. 2014;24:1176–86.
54. Thomé R, Gust A, Toci R, Mendel R, Bittner F, Magalon A, et al. A sulfurtransferase is essential for activity of formate dehydrogenases in *Escherichia coli*. *J Biol Chem*. 2012;287:4671–8.
55. Mansy SS, Wu G, Surerus KK, Cowan JA. Iron-sulfur cluster biosynthesis. *Thermotoga maritima* IscU is a structured iron-sulfur cluster assembly protein. *J Biol Chem*. 2002;277:21397–404.
56. Kaiser JT, Clausen T, Bourenkov GP, Bartunik HD, Steinbacher S, Huber R. Crystal structure of a NifS-like protein from *Thermotoga maritima*: implications for iron sulphur cluster assembly. *J Mol Biol*. 2000;297:451–64.
57. Mi-ichi F, Makiuchi T, Furukawa A, Sato D, Nozaki T. Sulfate activation in mitosomes plays an important role in the proliferation of *Entamoeba histolytica*. *PLoS Negl Trop Dis*. 2011;5:e1263.

58. Mi-ichi F, Nozawa A, Yoshida H, Tozawa Y, Nozaki T. Evidence that the *Entamoeba histolytica* mitochondrial carrier family links mitochondrial and cytosolic pathways through exchange of 3'-phosphoadenosine 5'-phosphosulfate and ATP. *Eukaryot Cell*. 2015;14:1144–50.
59. Liu S, He L, Yao K. The antioxidative function of alpha-ketoglutarate and its applications. *Biomed Res Int*. 2018;2018:3408467.
60. Sonneborn TM. Methods in the general biology and genetics of *Paramecium aurelia*. *J Exp Zool*. 1950;113:87–147.
61. Trager W. The cultivation of a cellulose-digesting flagellate, *Trichomonas termopsisidis*, and of certain other termite protozoa. *Biol Bull*. 1934;66:182–90.
62. Picelli S, Faridani OR, Björklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc*. 2014;9:171–81.
63. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19:455–77.
64. Dick GJ, Andersson AF, Baker BJ, Simmons SL, Thomas BC, Yelton AP, et al. Community-wide analysis of microbial genome sequence signatures. *Genome Biol*. 2009;10:R85–5.
65. Treitli SC, Kolisko M, Husník F, Keeling PJ, Hampl V. Revealing the metabolic capacity of *Streblomastix strix* and its bacterial symbionts using single-cell metagenomics. *Proc Natl Acad Sci U S A*. 2019;116:19675–84.
66. Zhu B-H, Xiao J, Xue W, Xu G-C, Sun M-Y, Li J-T. P_RNA_scaffold: a fast and accurate genome scaffold using paired-end RNA-sequencing reads. *BMC Genomics*. 2018;19:175.
67. Stanke M, Schöffmann O, Morgenstern B, Waack S. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics*. 2006;7:62.
68. Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, et al. Automated eukaryotic gene structure annotation using EVIDENCEModeler and the Program to Assemble Spliced Alignments. *Genome Biol*. 2008;9:R7–7.
69. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29:644–52.
70. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357–9.
71. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. 2019;37:907–15.
72. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 2015;31:3210–2.
73. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215:403–10.
74. Eddy SR. A new generation of homology search tools based on probabilistic inference. *Genome Inf*. 2009;23:205–11.
75. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, et al. InterPro-Scan 5: genome-scale protein function classification. *Bioinformatics*. 2014;30:1236–40.
76. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28:1647–9.
77. Almagro Armenteros JJ, Salvatore M, Emanuelsson O, Winther O, von Heijne G, Elofsson A, et al. Detecting sequence signals in targeting peptides using deep learning. *Life Sci Alliance*. 2019;2:e201900429.
78. Horton P, Nakai K. Better prediction of protein cellular localization sites with the k nearest neighbors classifier. *Proc Int Conf Intell Syst Mol Biol*. 1997;5:147–52.
79. Blum T, Briesemeister S, Kohlbacher O. MultiLoc2: integrating phylogeny and gene ontology terms improves subcellular protein localization prediction. *BMC Bioinformatics*. 2009;10:274.
80. Kume K, Amagasa T, Hashimoto T, Kitagawa H. NommPred: prediction of mitochondrial and mitochondrion-related organelle proteins of nonmodel organisms. *Evol Bioinformatics Online*. 2018;14:1176934318819835.
81. Reumann S. Specification of the peroxisome targeting signals type 1 and type 2 of plant peroxisomes by bioinformatics analyses. *Plant Physiol*. 2004;135:783–800.
82. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*. 2001;305:567–80.
83. Neuberger G, Maurer-Stroh S, Eisenhaber B, Hartig A, Eisenhaber F. Prediction of peroxisomal targeting signal 1 containing proteins from amino acid sequence. *J Mol Biol*. 2003;328:581–92.
84. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30:772–80.
85. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser*. 1999;41:95–8.
86. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30:1312–3.
87. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. 2009;25:1972–3.
88. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 2015;32:268–74.
89. Wang H-C, Minh BQ, Susko E, Roger AJ. Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. *Syst Biol*. 2018;67:216–35.
90. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol*. 2017;35:518–22.
91. Gietz RD, Woods RA. Yeast transformation by the LiAc/SS carrier DNA/PEG method. In: Xiao W, editor. *Yeast Protocols*. Totowa, NJ: Humana Press; 2006. p. 107–20.
92. Malinská K, Malinský J, Opekarová M, Tanner W. Visualization of protein compartmentation within the plasma membrane of living yeast cells. *Mol Biol Cell*. 2003;14:4427–36.
93. Single-cell genome and transcriptome sequencing of *Pelomyxa schiedti*. NCBI. 2020. <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA672820>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

