



Evolutionary History of Mitochondrial Genomes in Discoba, Including the Extreme Halophile *Pleurostomum flabellatum* (Heterolobosea)

Khaoula Ettahi¹, Duckhyun Lhee¹, Ji Yeon Sung², Alastair G.B. Simpson^{3,4}, Jong Soo Park ^{2,5,*}, and Hwan Su Yoon ^{1,*}

¹Department of Biological Sciences, Sungkyunkwan University, Suwon, South Korea

²Department of Oceanography, Kyungpook Institute of Oceanography, School of Earth System Sciences, Kyungpook National University, Daegu, South Korea

³Department of Biology, Dalhousie University, Halifax, Nova Scotia, Canada

⁴Centre for Comparative Genomics and Evolutionary Bioinformatics, Dalhousie University, Halifax, Nova Scotia, Canada

⁵Research Institute for Dok-do and Ulleung-do Island, Kyungpook National University, Daegu, South Korea

*Corresponding authors: E-mails: jongsoopark@knu.ac.kr; hsyoon2011@skku.edu.

Accepted: 10 November 2020

Abstract

Data from Discoba (Heterolobosea, Euglenozoa, Tsukubamonadida, and Jakobida) are essential to understand the evolution of mitochondrial genomes (mitogenomes), because this clade includes the most primitive-looking mitogenomes known, as well some extremely divergent genome information systems. Heterolobosea encompasses more than 150 described species, many of them from extreme habitats, but only six heterolobosean mitogenomes have been fully sequenced to date. Here we complete the mitogenome of the heterolobosean *Pleurostomum flabellatum*, which is extremely halophilic and reportedly also lacks classical mitochondrial cristae, hinting at reduction or loss of respiratory function. The mitogenome of *P. flabellatum* maps as a 57,829-bp-long circular molecule, including 40 coding sequences (19 tRNA, two rRNA, and 19 *orfs*). The gene content and gene arrangement are similar to *Naegleria gruberi* and *Naegleria fowleri*, the closest relatives with sequenced mitogenomes. The *P. flabellatum* mitogenome contains genes that encode components of the electron transport chain similar to those of *Naegleria* mitogenomes. Homology searches against a draft nuclear genome showed that *P. flabellatum* has two homologs of the highly conserved Mic60 subunit of the MICOS complex, and likely lost Mic19 and Mic10. However, electron microscopy showed no cristae structures. We infer that *P. flabellatum*, which originates from high salinity (313‰) water where the dissolved oxygen concentration is low, possesses a mitochondrion capable of aerobic respiration, but with reduced development of cristae structure reflecting limited use of this aerobic capacity (e.g., microaerophily).

Key words: *Pleurostomum flabellatum*, Heterolobosea, *Naegleria* with cristae, MICOS complex, microaerophily.

Significance

Pleurostomum flabellatum (Discoba: Heterolobosea) is a microbial eukaryote that thrives in near-saturated brine at 40 °C, where dissolved oxygen concentration is several times lower than in typical seawater or freshwater—this, together with a lack of obvious mitochondrial cristae hinted that *P. flabellatum* could be an anaerobe. We sequenced the 57,829-bp mitochondrial genome of *P. flabellatum*, revealing a similar complement of electron transport chain protein-coding genes to the related aerobe *Naegleria*, and identified two homologs of the mitochondrial cristae organization protein Mic60 in preliminary nuclear genome data. In addition to demonstrating *P. flabellatum*'s capacity for aerobic respiration (possibly as a microaerophile), we add a mitochondrial genomic perspective to the picture of the evolution of Discoba's remarkable mitogenome diversity.

© The Author(s) 2020. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Introduction

Endosymbiosis is an important evolutionary process that has greatly impacted the evolution of life and continues to shape eukaryotic cells and their genomes (Lang et al. 1997). Mitochondria are eukaryotic organelles that originated from an endosymbiotic α -proteobacterium, perhaps ~ 1.8 billion years ago (Bhattacharya et al. 2004; Gray 2012). Because then, mitochondrial genomes (mitogenomes) evolved specific DNA structure and genetic functions (Gray et al. 1989, 1999; Lang et al. 1997; Gray 2012) through massive gene loss, by a combination of endosymbiotic gene transfer (EGT) and gene elimination (Adams and Palmer 2003). Endosymbiotic gene transfer occurs when a gene migrates from the organellar genome to the host nucleus (and is subsequently lost from the organellar genome). In contrast, gene elimination occurs when a gene is no longer required by the host–endosymbiont holobiont. This history of evolutionary change began before the last common (extant) eukaryotic ancestor, but has continued after the diversification of the major eukaryotic lineages, leading to substantial variation in gene content among the mitogenomes of extant eukaryotes (Hancock et al. 2010). During mitochondrial evolution, many homologous genes have frequently been lost in separate branches of the eukaryotic tree, making the pattern and the prevalence of mitogenome gene loss more complicated to understand (Gray 1998; Lang et al. 1999; Adams and Palmer 2003). Reconstructing an evolutionary history of mitogenomes in a given clade requires mitogenome sequences from taxa that represent the variety of the investigated clade, as well as an accurate phylogeny of the group (Chihade et al. 2000).

The taxon Discoba consists of four main lineages—Jakobida, Euglenozoa, Heterolobosea, and Tsukubamonadida—that have various unique mitogenome features (Gray et al. 2004; Simpson et al. 2006; Hampl et al. 2009). Jakobida has the most bacteria-like and gene-rich mitogenomes known. For instance, the mitogenome of *Andalucia godoyi* comprises 100 functionally assignable genes, with 66 and 34 genes encoding proteins and structural RNAs, respectively (Burger et al. 2013; Gray et al. 2020). Several unique and/or complex features characterize the mitogenomes of Euglenozoa (Kinetoplastea, Diplonemea, and Euglenida) (Roy et al. 2007; Dobáková et al. 2015). For example, kinetoplastid mitogenomes are present as the “kinetoplast” (thus, “kDNA”), which typically includes two different DNA molecule classes called maxicircles and minicircles (Lukes et al. 2002).

Heterolobosea (Discoba), meanwhile, is a major heterotrophic protist group containing ~ 150 described species. In addition to mesophiles that live in conventional freshwater, seawater, or soil systems, this group is notable for including many species that are halophiles, acidophiles, thermophiles, or anaerobes (Pánek et al. 2014; Park and Simpson 2015). It seems that anaerobic lineages arose at least three times in Heterolobosea. Psalteriomonadidae, *Creneis*, and

Dactylomonas represent distinct lineages that possess mitochondrion-related organelles that lack cristae and are known or suspected to function anaerobically, for example, as hydrogenosomes (Park et al. 2007; de Graaf et al. 2009; Barberà et al. 2010; Hanousková et al. 2019). It is suggested that the lack of cristae in anaerobic protists such as psalteriomonads is related to a loss of the mitochondrial contact site and cristae organizing system (MICOS) (see fig. 5 in Muñoz-Gómez, Slamovits, Dacks, Baier, et al. 2015), localized in the cristae junctions. The MICOS complex of yeast consists of six proteins: Mic10, Mic12, Mic19, Mic26, Mic28, and Mic60. The most ancient component of the MICOS subunits may be Mic60, which seems to have originated from α -proteobacterium, whereas Mic10 is the most widespread component among eukaryotes (Muñoz-Gómez, Slamovits, Dacks, and Wideman 2015; Muñoz-Gómez, Slamovits, Dacks, Baier, et al. 2015; Gray et al. 2020). Some Discoba genomes were previously reported to encode Mic10, Mic19, and Mic60 homologs (fig. 5; Muñoz-Gómez, Slamovits, Dacks, Baier, et al. 2015; Gray et al. 2020). The well-studied heterolobosean *Naegleria gruberi* is a free-living amoeba, closely related to the human pathogen *Naegleria fowleri* that is the causative agent of the deadly human disease primary amoebic meningoencephalitis. Muñoz-Gómez, Slamovits, Dacks, Baier, et al. (2015) examined the phylogenetic distribution of MICOS subunits and identified two MICOS subunits (Mic60 and Mic19) in *N. gruberi*. Furthermore, *N. gruberi* was suggested to contain genes related to anaerobic energy metabolism (e.g., [FeFe]-hydrogenase, Fritz-Laylin et al. 2010), although a more recent study suggests that hydrogen production occurs exclusively in the cytoplasm (Tsaousis et al. 2014). Thus, there may be a broad and complex diversity of functions in aerobic, anaerobic, or microaerobic Heterolobosea.

Pleurostomum flabellatum is an obligate extremely halophilic heterolobosean. It can grow optimally in more than 300‰ salinity water at 40 °C, where theoretical oxygen saturation (1.26 mg l^{-1}) is 3.5 times lower than in freshwater (4.47 mg l^{-1}) at 40 °C (Battino et al. 1983). This organism is reported to lack mitochondrial cristae (Park et al. 2007) and is sometimes listed as an additional (fourth) lineage of anaerobic heteroloboseids (Pánek et al. 2014; Hanousková et al. 2019). However, nothing is known about the function of mitochondrion-related organelles in *P. flabellatum* (Park et al. 2007). Codon usage by extremophiles, namely halophiles and thermophiles, has been reported to be significantly different from non-extremophile organisms, at least in prokaryotes (Khan and Patra 2018). Because *P. flabellatum* is both a candidate anaerobe and confirmed extreme halophile, it is a particularly interesting species to examine to better understand genome evolution in anaerobic and/or halophilic eukaryotes.

In this study, we assembled the mitogenome of *P. flabellatum* and analyzed its gene composition and gene arrangements. We performed comparative and phylogenetic

analyses of the *P. flabellatum* mitogenome together with those of other discobids, including six heteroloboseids. Also, we studied codon usage bias to detect possible variation between heterolobosean mitogenomes and performed a GC skewness analysis to compare inferred replication patterns of mitochondria in Discoba. *P. flabellatum* is sister to *Naegleria* spp. in our phylogenetic analysis of mitogenome data and is likely a microaerophile rather than an anaerobe based on the mitogenome gene repertoire. The apparent absence of cristae-like structures may be due to the lack of genes encoding several MICOS components.

Materials and Methods

Isolation, Cultivation, and Ultrastructure

Pleurostomum flabellatum strain EHF1 (10–14 μm -long) was initially isolated from 313‰ salinity water collected from the Soosung solar saltern located at Seosin on the west coast of the Republic of Korea (36°09'36"N, 126°40'44"E; Park et al. 2007). The culture of *P. flabellatum* was maintained in 200‰ salinity media, incubated at 37 °C and subcultured every 4 weeks for 12 years. In brief, 0.1 ml of fluid from the culture was used to inoculate 5 ml of 200‰ salinity liquid media, made by dilution of Medium V (300‰; 272 g NaCl, 7.6 g KCl, 17.8 g MgCl₂, 1.8 g MgSO₄•7H₂O, 1.3 g CaCl₂ l⁻¹ water; see Park 2012) with sterile distilled water, and supplemented with Marine Broth 2216 (final concentration of 0.5%; Difco) plus autoclaved barley grains to grow indigeneous prokaryotes in the culture as prey. Before DNA extraction, approximately 2.75 l of well-grown culture (15,000 *P. flabellatum* cells per ml) in liquid media (200‰ salinity) was prepared. For ultrathin sectioning, the culture of *P. flabellatum* was grown in 250‰ salinity media. All procedures for transmission electron microscopy (TEM) were as described in Park et al. (2007).

DNA Extraction and Sequencing

After the culture was pre-filtered using a 38- μm mesh sieve, cells were harvested using a 10- μm pore size membrane filter (Advantec, Tokyo, Japan). A DNeasy Plant Mini Kit (Qiagen, Santa Clarita, CA) was used for DNA extraction. A sequencing library was constructed using the Truseq Nano DNA Prep Kit (Insert 550-bp) based on the manufacturer's manual. Samples were sequenced as 150-bp paired-end reads on the Illumina Novaseq6000 platform at DNA-Link Inc. (Seoul, Korea).

Mitogenome Assembly and Annotation

The mitogenome of *P. flabellatum* was assembled using the *de novo* assembler NOVOPlasty 2.7.2 (Dierckxsens et al. 2017) from the Illumina whole-genome sequencing data. Here, the mitochondrial gene sequences from the closely related species *N. gruberi* (GenBank accession number = AF288092) was

used as the seed sequence for NOVOPlasty. The processed reads were mapped to the assembly to identify potential assembly errors, followed by coverage evaluation. Open reading frames (*orfs*) were identified by Geneious 8.1.2 (<https://www.geneious.com>, last accessed November 1, 2018), with the universal genetic code, as this code is used in closely related species within Heterolobosea. Annotation of putative protein-coding sequences (CDSs) was inferred first with BLASTX (e-value $\leq 1.0\text{e}^{-05}$) by searching against the non-redundant (nr) protein database for sequence similarities. Less well-preserved genes (i.e., that were not annotated with BLASTX) were subjected to HMMER searches (<http://hmmer.janelia.org>, last accessed November 1, 2018) derived from models of all available mitochondrion-encoded proteins. Multiple protein-alignment examinations were further used to approve or reject *orfs* with significant or close-to-significant sequence similarity.

Large subunit (*rnl*) and small subunit (*rns*) ribosomal RNA (rRNA) genes were identified by RNAmmer 1.2 (www.cbs.dtu.dk/services/RNAmmer, last accessed November 1, 2018) and transfer RNA (tRNA) genes were detected using ARAGORN (<http://mbio-serv2.mbioekol.lu.se/ARAGORN>, last accessed November 1, 2018). Structural comparison of mitogenomes was performed using progressive Mauve (Darling et al. 2010). The circular mitogenome map was visualized using DNAPlotter 17.0.1 (Carver et al. 2009). Genome sequences were deposited in the NCBI GenBank database under the following accession number MT843578.

Phylogenetic Analysis

For the phylogenetic analysis of Discoba, 20 complete mitogenomes were obtained from GenBank as follows: seven heteroloboseids, including *P. flabellatum* (this study), *Naegleria gruberi* (Fritz-Laylin et al. 2010, 2011; AF288092), *Naegleria fowleri* (KX580902; Herman et al. 2013), *Acrasia kona* (KJ679272; Fu et al. 2014), *Stachyamoeba lipophora* (KP165388, Valach et al. 2014), *Pharyngomonas kirbyi* (NC_034798; Park and Simpson 2011; Yang et al. 2017), and the undescribed amoeba "Heteroloboseid sp. BB2" (KY379823; Yang et al. 2017); *Tsukubamonas globosa* (Tsukubamonadida) (AB854048; Kamikawa et al. 2014); five euglenozoans: *Diplonema papillatum* (EU123537; Marande et al. 2005), *Diplonema ambulator* (MF436761, Valach et al. 2017), *Rhynchopus euleeides* (JF698667), *Trypanosoma rangeli* (KJ803830; Stoco et al. 2014), and *Angomonas deanei* (KJ778684); and seven jakobids: *Jakoba bahamiensis* (KC353354; Burger et al. 2013), *Jakoba libera* (KC353355; Burger et al. 2013), *Reclinomonas americana* (KC353356; Lang et al. 1997), *Histiona aroides* (KC353353; Burger et al. 2013), *Seculamonas eudoriensis* (KC353355; Burger et al. 2013), *Andalucia godoyi* (KC353352; Burger et al. 2013), and *Ophirina amphinema* (LC369600; Yabuki et al. 2018).

Conserved protein sequences shared by the discobid groups were retrieved (14 proteins total: *nad1-5*, *nad7-9*, *nad4L*, *cox1-3*, *atp6*, *rps12*). The amino acid sequences were aligned and concatenated using MAFFT 7.1.10 (Kato and Standley 2013). Aligned mitogenome genes were concatenated for multigene phylogenetic analysis. The phylogenetic trees were inferred from the resulting concatenated amino acid alignment. A maximum likelihood (ML) tree was constructed using IQ-TREE 1.6.7 with LG+F+I+G4 as the best-fit model of sequence evolution (as chosen by ModelFinder) and with robustness evaluated with a 1000-replicate UFBoot bootstrap approximation (Minh et al. 2013; Flouri et al. 2015; Nguyen et al. 2015; Kalyaanamoorthy et al. 2017). A Bayesian analysis was carried out using PhyloBayes v8.28 (Lartillot et al. 2009) under the CAT+GTR model. Two independent chains were run with discrete gamma distribution with four categories, and a burn-in of 1000.

Mitogenome Replication and Codon Bias

The Relative Synonymous Codon Usage (RSCU) values from mitogenome CDSs were calculated to characterize synonymous codon usage, and correspondence analysis (COA) was performed, both using codonW (<http://codonw.sourceforge.net/culong.html>, last accessed November 15, 2018).

Cumulative GC-skew values were examined to find the origin point (*oriC*) of replication with window/step settings of 1000/300 (Eppinger et al. 2004). This method measures the asymmetric strand distribution of G and C using the formula $GC\text{-skew} = [G - C] / [G + C]$ and accumulates the obtained values (Perna and Kocher 1995). The GC skewness has been used in several studies to assess the strand bias of nucleotide composition of different eukaryotic mitogenomes (e.g., Burger et al. 2013; Jackson and Reyes-Prieto 2014). This strand bias varies gradually along a genome, and the region with the lowest GC skew value indicates where the origin of replication might be located. GC-skew values for selected discobids with circular mitogenomes were obtained using *fasta_gc_skew.py* program (https://github.com/shenwei356/bio_scripts, last accessed November 15, 2018).

Homology Search

Proteins homologous to MICOS components were searched for using BLAST against the list of MICOS protein sequences from Muñoz-Gómez, Slamovits, Dacks, Baier, et al. (2015), that is, *Mic10*, *Mic12*, *Mic19*, *Mic25*, *Mic26*, *Mic27*, and *Mic60*. The sequences that were hit with an *e*-value lower than 0.10 were considered as significant. The two *Mic60* sequences were compared against protein-coding regions within the *P. flabellatum* transcript assembly set ($N_{50} = 3,485$), using BLAST searches with an *e*-value cut-off $\leq 1.0e^{-05}$. Then, we searched for protein signatures in the InterPro (Mitchell et al. 2019) and Pfam (El-Gebali et al. 2019) databases using InterProScan (<https://www.ebi.ac.uk/>

interpro, last accessed May 20, 2020) and HMMER (<http://hmmer.janelia.org>, last accessed May 20, 2020), respectively. The two *Mic60*-like sequences were deposited in the NCBI GenBank (accession numbers: MW019459 and MW019460).

Results and Discussion

Pleurostomum flabellatum Contains Classical Mitochondrial Genome

General Features

We assembled the complete mitogenome of *P. flabellatum* into a single circular map of 57.8-kbp (fig. 1). This size is within the previously reported range for heterolobosean mitogenomes (49.5–75.7 kbp; [supplementary table S1, Supplementary Material](#) online), except for the early-diverged Heteroloboseid sp. BB2, which has a larger mitogenome (119.3-kbp), mostly due to a 49-kbp inverted repeat (IR) region (Yang et al. 2017). The proportion of coding regions in the *P. flabellatum* mitogenome (86.1%) is similar to that of other heteroloboseans (81–93.2%), and jakobids (78.5–94.1%) (Burger et al. 2013; Yabuki et al. 2018), but higher than in *Tsukubamonas* (65.7%) ([supplementary table S1, Supplementary Material](#) online). The *P. flabellatum* mitogenome is AT-rich (71.1%), which is also consistent with other heteroloboseids (AT content = 69.3–87.5%) ([supplementary table S1, Supplementary Material](#) online).

The mitogenome of *P. flabellatum* contains 40 CDSs, two rRNAs, 20 tRNAs. In addition to CDSs with known functions, the *P. flabellatum* mitogenome contains 19 predicted *orfs*, including *orf145*, which also occurs in *Naegleria* mitogenomes ([supplementary fig. S1, Supplementary Material](#) online). The *orf145* from *P. flabellatum* shares 59.3% amino acid sequence similarity with *N. gruberi* and 58.8% with *N. fowleri*. However, the function of *orf145* remains unknown. The gene set of the *P. flabellatum* mitogenome is similar to those of other heteroloboseans (tRNAs: 11–24, CDSs: 26–42) ([supplementary tables S2 and S3, Supplementary Material](#) online). All the genes are located on the same strand except *ccmF* and *trnY* (fig. 1).

Codon Usage

Codon usage in halophilic (*P. flabellatum* and *P. kirbyi*) and thermophilic (Heteroloboseid sp. BB2) heterolobosean mitogenomes was compared with that of the mesophiles *N. gruberi*, *N. fowleri*, *A. kona*, and *S. lipophora*. A [supplementary table S4, Supplementary Material](#) online, summarizes the use of amino acids and relative synonymous codon usage (RSCU) values in the CDS sequences. The RSCU values of all heteroloboseans revealed that AT-rich codons are favored over synonymous codons with lower A/T content ([supplementary fig. S2, Supplementary Material](#) online).

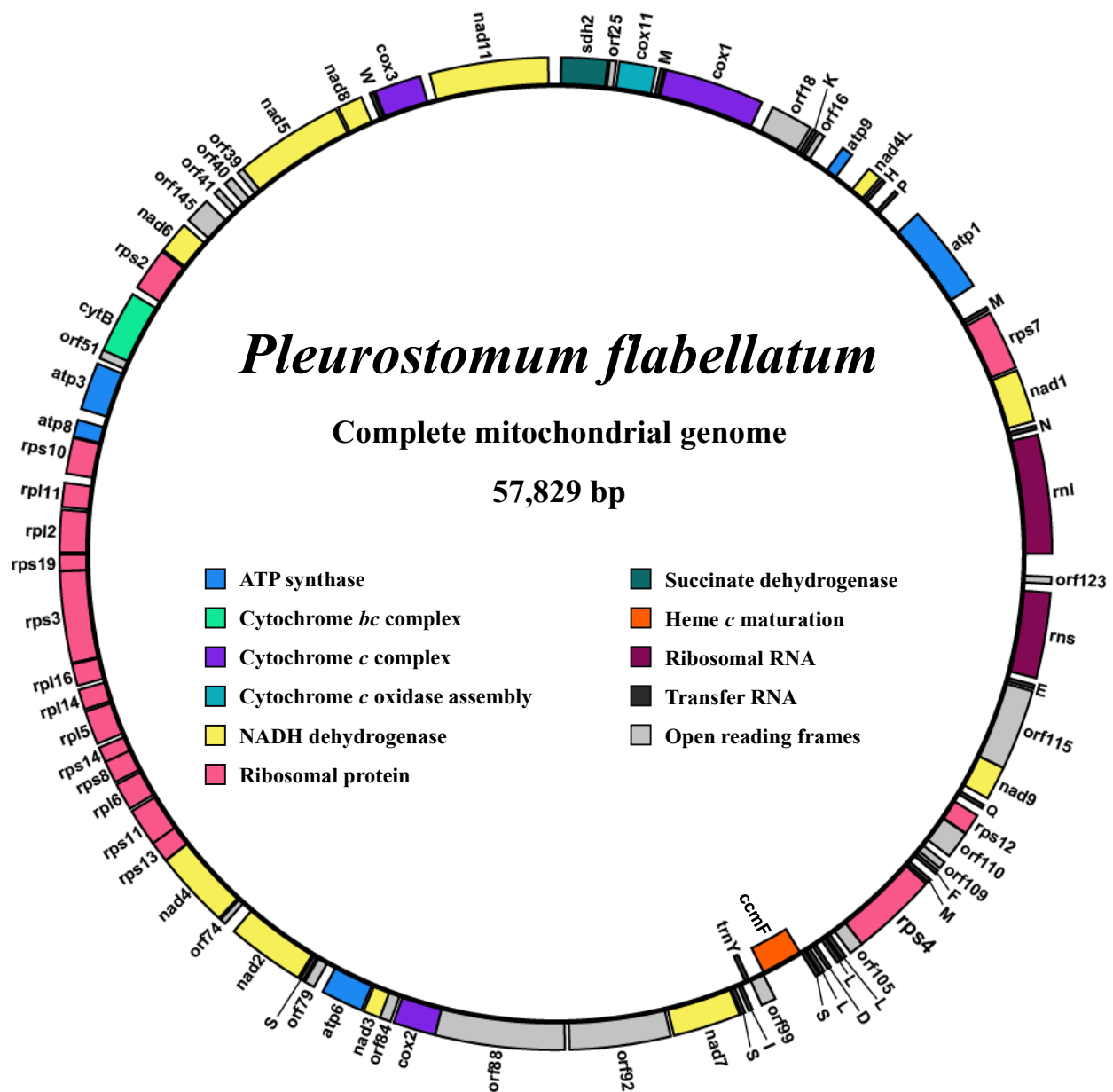


Fig. 1.—Gene map of the *Pleurostomum flabellatum* mitogenome. Functional gene groups are color coded. Genes drawn inside of the circle indicate the transcriptionally clockwise strand, whereas genes on outside indicate the counterclockwise strand.

The examination of the correspondence (COA) was carried out based on RSCU values of heterolobosean mitogenomes. Correlation analysis was conducted on the difference in codon patterns represented by the first and second major COA axis (supplementary fig. S3, Supplementary Material online). Heterolobosean mitogenomes showed low variability in codon usage, with no specific patterns between extremophiles and mesophiles. Overall, we found no significant difference between the mitogenome codon usage of heterolobosean extremophiles and non-extremophiles.

Phylogeny and Evolutionary Inferences Based on Mitochondrial Proteins

The Discoba organismal phylogeny was inferred from the concatenated amino acid data set of 14 conserved mitochondrial proteins (fig. 2A). Our phylogeny is consistent with the immediate sister relationship between the Heterolobosea and Euglenozoa clades to the exclusion of Jakobida and Tsukubamonadida (Hampl et al. 2009; Kamikawa et al. 2014; Yang et al. 2017), although the heterolobosean clade itself was poorly supported (BS = 62%; PP = 0.9).

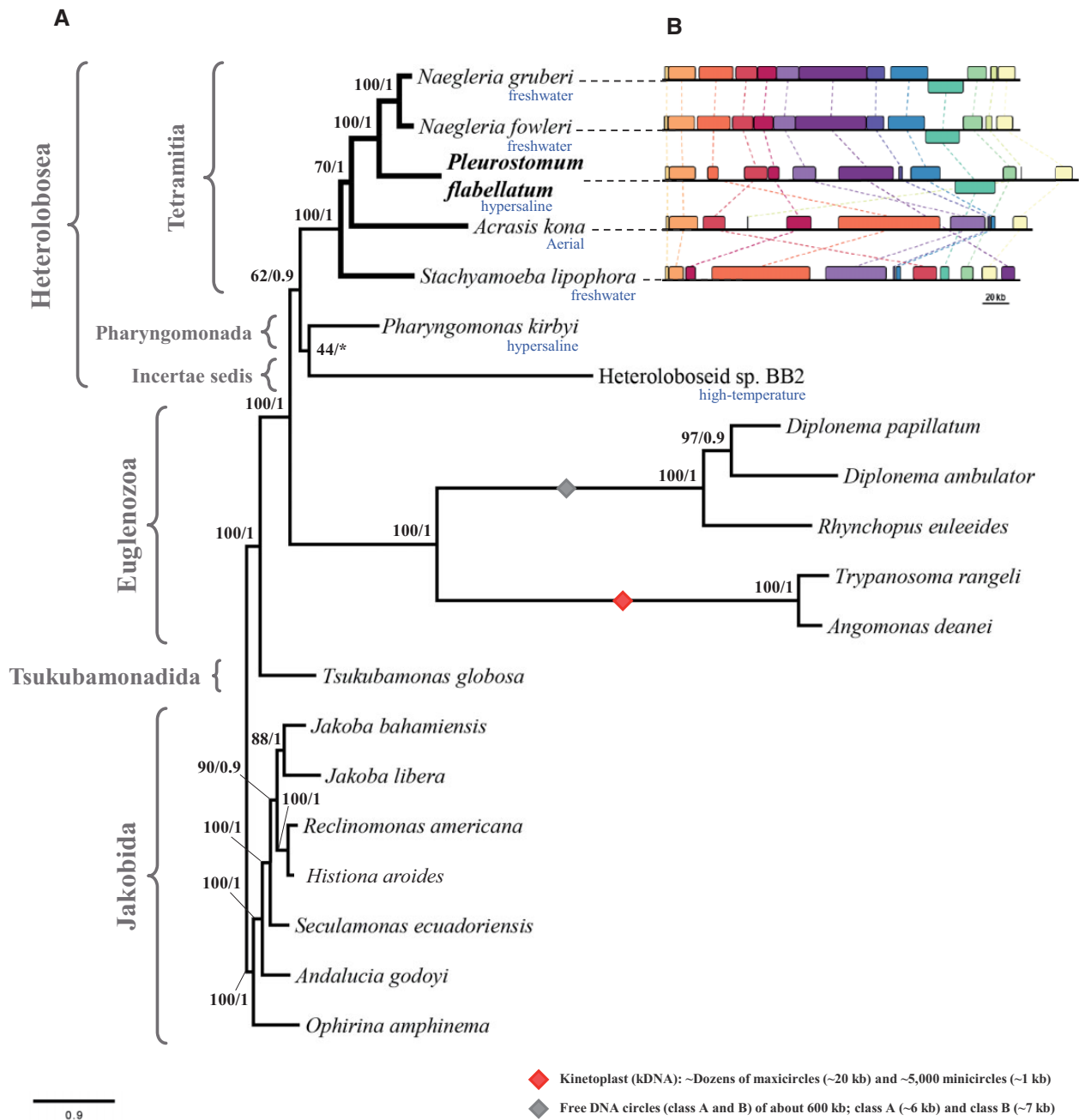


FIG. 2.—Phylogenetic tree of the Discoba based on mitogenome data and synteny comparisons among the Heterolobosea. (A) Phylogenetic relationships between heteroloboseids and other discobids deduced from the analysis of 14 concatenated mitochondrial genes from 20 discobids. The tree was inferred based on amino acid sequences using maximum-likelihood (ML) under the LG+F+I+G4 model and Bayesian phylogenetic analysis under the CAT+GTR model. Numbers at nodes indicate UFBoot ML bootstrap approximation percentages and Bayesian posterior probability, slash separated and “*” means not recovered. The tree is rooted with the Jakobida as sister to other Discoba following Yang et al. (2017) and Yabuki et al. (2018). Habitats of heteroloboseid taxa are indicated below each species name. Structures of euglenozoan mitogenomes are listed following Faktorová et al. (2016). (B) Synteny comparison among five Tetramitia species (*Naegleria gruberi*, *N. fowleri*, *Pleurostomum flabellatum*, *Acrasis kona*, and *Stachyamoeba lipophora*) using the Mauve program.

Pleurostomum and *Naegleria* are clustered with strong statistical support (BS = 100%; PP = 1) and fall along with *Acrasis* and *Stachyamoeba* into a robust Tetramitia clade (BS = 100%; PP = 1). This *Pleurostomum*+*Naegleria* cluster

is consistent with previous inferences from 18S rRNA genes (Park et al. 2007, 2012; Harding et al. 2013; Tyml et al. 2017; Jhin and Park 2019). Heteroloboseid BB2 and *Pharyngomonas* are monophyletic in the ML tree with low support

(BS = 44%), but paraphyletic in the Bayesian inference (supplementary fig. S4, Supplementary Material online). The phylogenetic positions of BB2 and *Pharyngomonas* have not been resolved by a previous phylogenetic analysis based on 18S rRNA genes (Harding et al. 2013), whereas a phylogenomic analysis of 252 nucleus-encoded genes also showed BB2 and *Pharyngomonas* are forming a clade (with full bootstrap support; Yang et al. 2017). Thus, as with previous studies, our phylogeny robustly shows *P. flabellatum* and *P. kirbyi* as phylogenetically distant, despite both surviving in hypersaline ecosystems (Pánek et al. 2014; Kirby et al. 2015; Jhin and Park 2019). This result implies that halophilic characters evolved independently in these two heteroloboseids.

Mitogenome Rearrangement

The alignment of heterolobosean mitogenomes was inspected to investigate the arrangement of the *P. flabellatum* mitogenome (fig. 2B). Synteny appeared to be highly conserved between *P. flabellatum* and *Naegleria* spp., where a total of 13 syntenic blocks are well preserved in the same order. *P. flabellatum* has more regions containing *orfs* of unknown function (figs. 1 and 2B). In contrast, several rearrangements are evident among *P. flabellatum*, *A. kona*, and *S. lipophora*, which suggest that the shuffling has occurred after the split of these lineages. In terms of mitogenome size, *A. kona* is close to *P. flabellatum* and *Naegleria* spp. but has a very different organization and gene content (Fu et al. 2014) (fig. 2B). Conserved synteny blocks in *A. kona* and *S. lipophora* counts for 10 and 12 blocks, respectively. In addition, syntenies of *P. flabellatum* with *P. kirbyi* showed four low homology blocks where the first two *P. flabellatum* syntenic blocks are inverted in *P. kirbyi* (supplementary fig. S5, Supplementary Material online). The structures of the heterolobosea mitogenomes are thus quite diverse.

Characterization of GC Skewness in Discoba

There is a negative AT skew over the *P. flabellatum* mitogenome (-0.094) indicating a slight bias towards the use of T (38.9%) over A (32.2%), as well as a positive GC-skew ($+0.214$) showing a preference toward the use of G (17.6%) over C (11.4%) on the leading strand. This nucleotide composition difference between the two strands is typical for mitogenomes and likely reflects asymmetric arrangements of mutational variations between strands (Lobry 1995; Sueoka 1995). Nevertheless, the codon bias in protein-coding regions can also cause a strand-specific GC-skew to a minor extent (Tillier and Collins 2000).

GC skewness of one strand of a bacterial chromosome provides a good indicator for identifying the origin and terminus of the replication (Grigoriev 1998; Salzberg et al. 1998). This method is already used to infer the origin of replication in eukaryotic mitogenomes, such as jakobids (Burger et al. 2013). In bacteria with circular chromosomes, replication

starts from specific replication origins by a unidirectional theta-type intermediate and ends at the replication terminus, nearly opposite of the origin (Lobry 1996; Hines and Ray 2011). Likewise, the kinetoplast DNA molecules—both minicircles and maxicircles—replicate from specific replication origins via unidirectional theta-type intermediates (Hines and Ray 2011).

Here, we analyzed the cumulative GC-skew plots of the circular-mapping mitogenomes of discobids (fig. 3). Heterolobosean GC-skew trajectories and the cumulative GC-skew curves map simply to organismal phylogeny (fig. 3). Cumulative GC-skew of the *P. flabellatum* mitogenome shows a homogenous and positive skew curve along the sequence, resembling the situation in *A. kona*, *Naegleria* spp., and *S. lipophora*. A similar pattern is seen in some jakobids; *J. bahamiensis* and *O. amphinema*, as well as *Tsukubamonas* (*T. globosa*) (fig. 3B; Burger et al. 2013). This linear skewness was also observed in the mitogenomes of the glaucophyte *Cyanoptyche gloeocystis* (Jackson and Reyes-Prieto 2014), and in the yeast *Candida glabrata*, where this latter mitogenome was experimentally shown to replicate by a rolling circle (Maleszka et al. 1991; Burger et al. 2013). This skewness is also usually found in circular plasmids, known to replicate by unidirectional replication of the rolling circle (RCR) (Arakawa et al. 2009).

However, *R. americana* and *H. aroides* display prominent bimodal GC-skew curves—more distinctly in *R. americana*—which is characteristic of the classical bidirectional theta mode (fig. 3; Burger et al. 2013). The GC-skew graphs of the remaining discobids (*P. kirbyi*, BB2, *A. godoyi*, '*S. ecuadoriensis*') mitogenomes are insufficient to draw meaningful inferences about the replication mechanism (fig. 3; Burger et al. 2013). Interestingly, similar GC trajectories have been observed in bacterial genomes (Saha et al. 2014) and have been shown to be induced by genomic rearrangement types mixing leading and lagging sequences that influence trends in local base usage.

Distribution of Gene Losses in Discoba

Heterolobosea *sensu lato* is divided into two subgroups: Tetramitia (represented here by four genera) and *Pharyngomonas* + BB2 (fig. 2), albeit the latter grouping was paraphyletic in our Bayesian analysis (Cavalier-Smith and Nikolaev 2008; Harding et al. 2013; Yang et al. 2017; Petru et al. 2018). The cluster *Pharyngomonas*+BB2 includes two extra CDS genes (*rpl10*, *rpl32*) and three tRNAs (*trnA*, *trnC*, and *trnG*). Four additional CDSs are distributed unequally over the mitogenomes of the Tetramitia genera (fig. 4A, supplementary tables S2 and S3, Supplementary Material online); three cytochrome c maturase subunits; the ABC-transporter *ccmA* and heme transfer proteins *ccmC* and *ccmF*, along with ribosomal protein S4 (*rps4*). The twin-arginine translocase *tatC* was recently identified

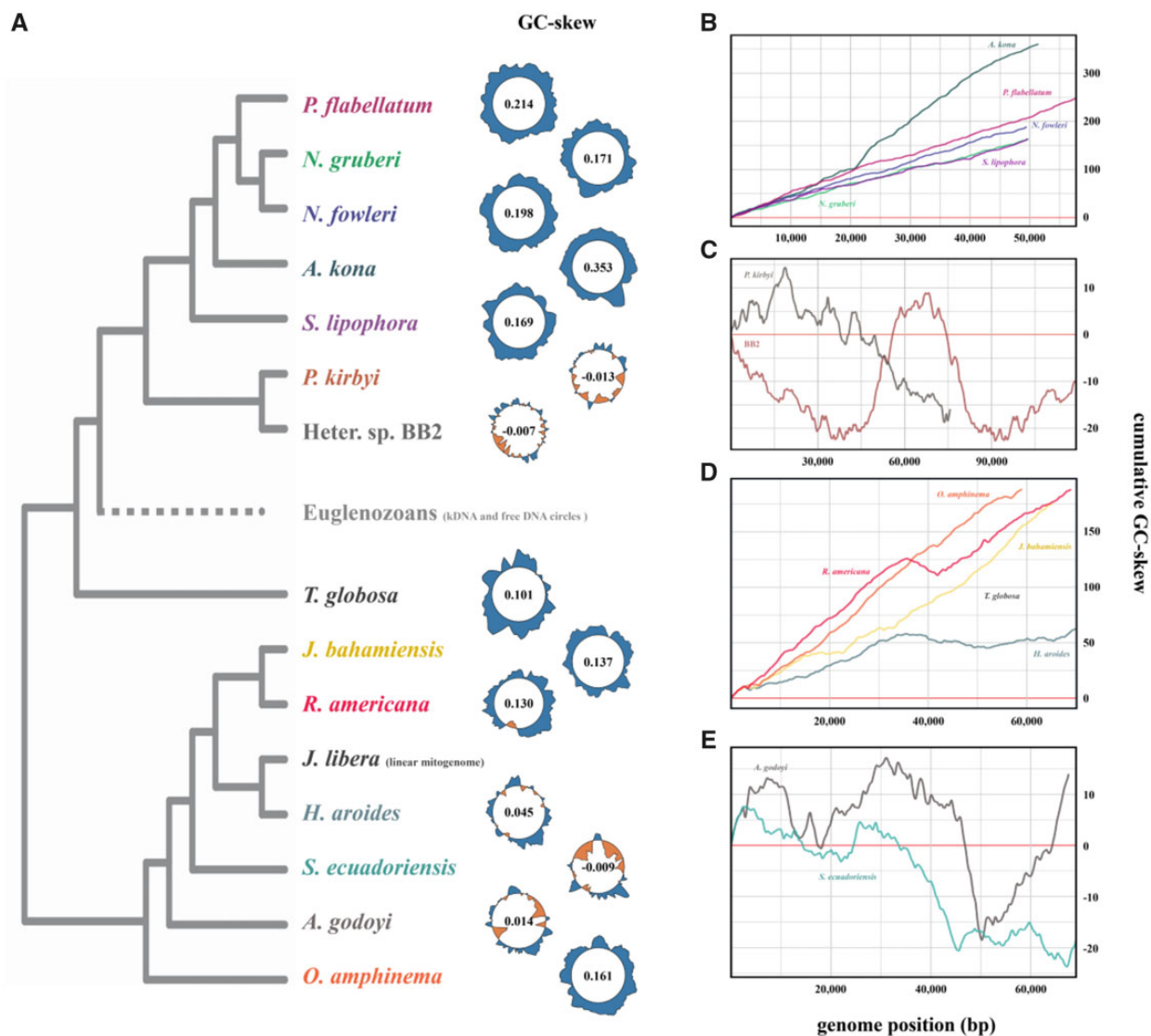


Fig. 3.—GC-skew of the Discoba mitogenomes and cumulative GC-skew trajectories, cumulative GC-skew schemes. (A) The GC-skew is calculated as $(G - C)/(G + C)$, with a window size and a step size of 1000 and 300 nucleotides, respectively. For GC-skew trajectories, blue-filled regions outside the circle line denote positive GC-skew values whereas orange-filled regions inside the circle line denote negative GC-skew values. GC-skew is mapped onto the phylogenetic relationships among discobids (as shown in fig. 2). The complete GC-skew value of each mitogenome is written inside the GC-skew plots. (B–E) The x-axis represents position in the genome, and the y-axis represents cumulative GC-skew. (B) Heterolobosean mitogenomes showing a homogenous curve with positive skew along the sequence. (C) Heterolobosean mitogenomes showing multimodal curve with positive and negative skews along the sequence. (D) Jakobids mitogenomes showing a homogenous/bimodal curve with positive skew along the sequence. (E) Jakobids mitogenomes showing multimodal curve with both positive and negative skew along the sequence.

as present among both Tetramitida and Pharyngomonada (Petrů et al. 2018, see fig. 4), having been overlooked in the original description of the *Pharyngomonas kirbyi* mitogenome (Yang et al. 2017). We reconstructed single-gene phylogenies to examine the evolutionary history of these five genes among the eukaryotes, none of which showed strong evidence against a common origin of mitochondrial forms across eukaryotes (supplementary figs. S6–S10, Supplementary Material online). The presence of *ccmA*, *C*, *F*, *rps4*, and *tatC* genes in some but not all mitogenomes

in various distantly related eukaryote lineages suggests that these genes were lost at several independent points during eukaryotic/mitogenome evolution.

Electron Transport Chain Components in *Pleurostomum flabellatum*

The typical mitochondrion of aerobic eukaryotes contains the electron transport chain (ETC; Complexes I–IV) and an ATP-synthase (Complex V) within the inner mitochondrial

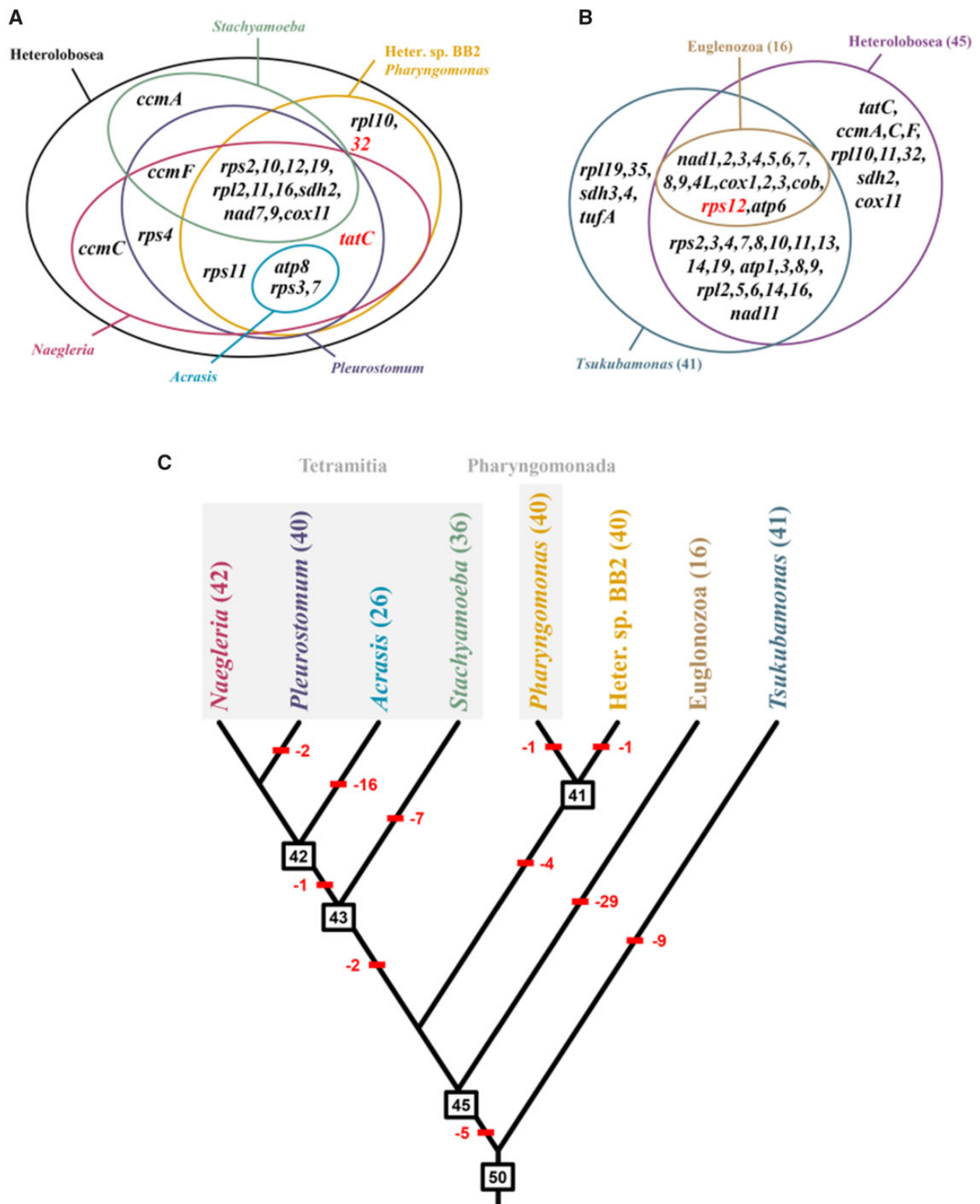


FIG. 4.—A comparison of mitogenome gene contents of three lineages of Discoba (Heterolobosea, Euglenozoa, and *Tsukubamonas*). (A) Venn diagram representation of the variable portions of the protein-coding gene repertoires of available heterolobosean mitogenomes, that is, excluding the protein-coding genes present in all Heteroloboseid mitogenomes (*nad1-6*, *nad*, 8, 11, *nad4L*, *cytB*, *cox1,2,3*, *atp1,3,6,8,9*, *rps13,14*, *rpl5,6,14*). For each taxon, the current sets of functionally assignable CDSs in each species mitogenomes are shown in parentheses. The gene *rpl32* (in red) was lost in *Pharyngomonas kirbyi*. The gene *tatC* (in red) was identified in *Pharyngomonas kirbyi* by Petú et al. (2018), however, there is no evidence of its presence in the heteroloboseid sp. BB2. (B) Venn diagram representing the minimal inferred gene content of the mitogenome of the last common heterolobosean ancestor, along with representative euglenozoans and *Tsukubamonas*. The gene in red (*rps12*) is present in *Trypanosoma* but absent from *Diplonema*. (C) Inferred gene losses are shown on tree branches in red. Numbers in square boxes indicate the minimum inferred number of genes in ancestral species. Parentheses after genus names show the number of genes retained by extant mitogenomes. Gene loss events were predicted on the phylogeny of discobids, as extracted from figure 2.

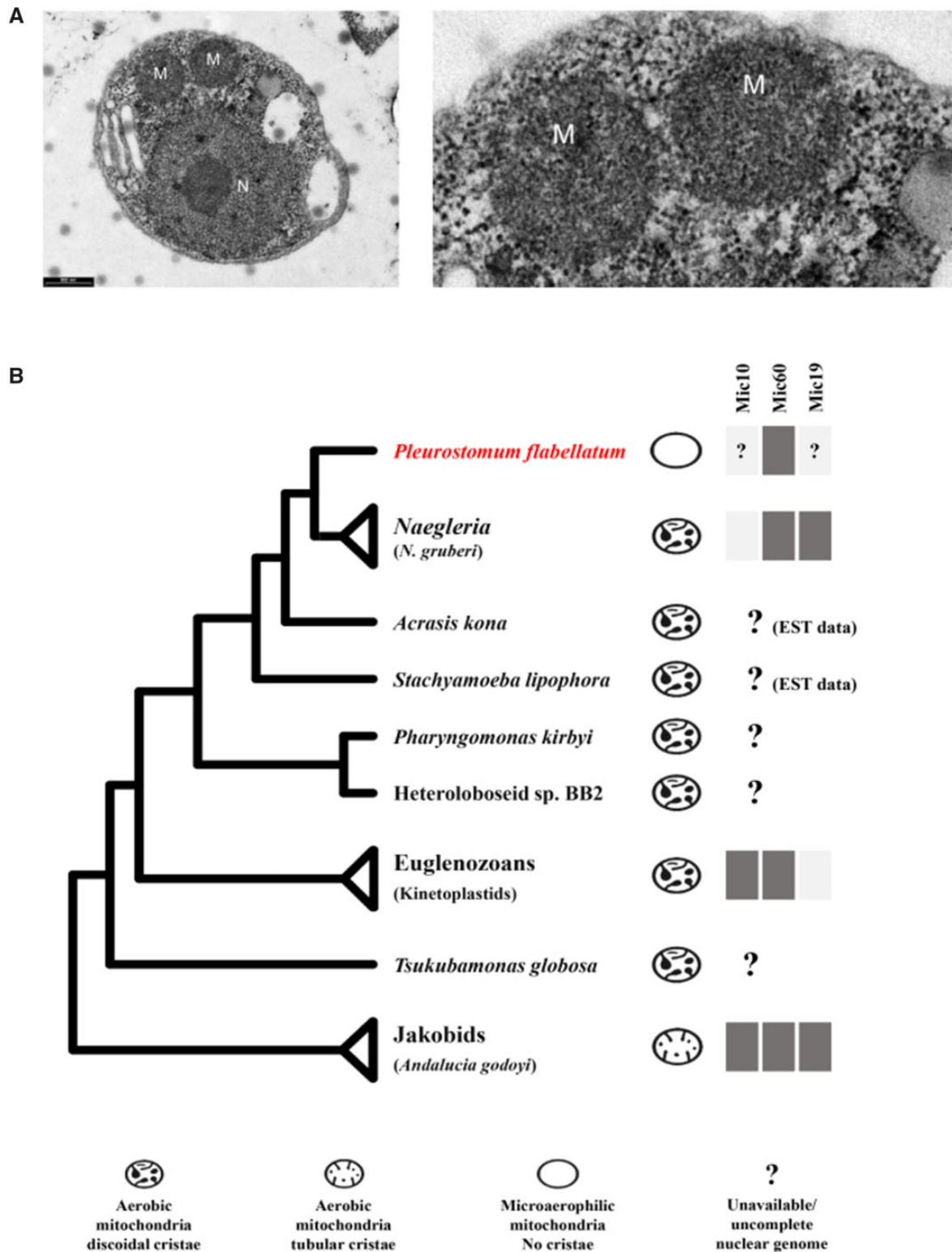


FIG. 5.—Mitochondria cristae morphotype and MICOS proteins distribution in the Discoba. (A) TEM micrographs of *P. flabellatum* (N, nucleus; M, mitochondrion; scale bar: 500-nm). (B) Distribution of MICOS subunits across the Discoba with mitochondria cristae morphotype. The presence of certain MICOS components (Mic10, Mic60, Mic19) are shown by gray rectangles. Other lineages examined in this analysis and characterized by “EST data” had only EST data sets available. The question mark indicates unavailable/uncomplete nuclear genome data.

membrane (Gray 2012). Some of the proteins that make up this complex system are invariably encoded on the mitogenome in such aerobes, even when the mitogenome is unusually small and the ETC is reduced (Flegontov et al. 2015). The *P. flabellatum* and *Naegleria* mitogenomes contain the same genes that encode electron transport chain components (supplementary table S2, Supplementary Material online), indicating that *Pleurostomum* mitochondria must be either aerobic or microaerophilic, contrary to some suggestions in the literature (Park et al. 2007; Pánek et al. 2014; Hanousková et al. 2019).

Mic60 Suggests the Potential for Mitochondrial Cristae Formation

We identified two genes coding putative homologs of the highly conserved Mic60 protein (GenBank accession numbers: MW019459 and MW019460) from the draft nuclear genome as well as transcriptome data of *P. flabellatum* (unpublished). Mic60 protein sequences of *P. flabellatum* contain the mitofilin domain architecture found in *N. gruberi* Mic60 (Naegru_Mic60_64657 XP_002683319.1, see supplementary fig. S11, Supplementary Material online). The transcript sequences include an extension of approximately 800-bp and poly-A tail in 5'- and 3'-UTR regions, respectively, suggesting active function of Mic60 genes. This implies that *P. flabellatum*, which lacks typical cristae, nonetheless shares at least the Mic60 subunit with cristae-bearing heteroloboseans.

The evolutionary history of mitochondria in eukaryotes generally showed a correlation between the occurrence of cristate mitochondria and the presence of genes encoding MICOS subunits, although there are some exceptions likely due to the lack of high-quality genome data (Muñoz-Gómez, Slamovits, Dacks, Baier, et al. 2015; Huynen et al. 2016). Mic60 has been inherited from α -proteobacterium through endosymbiotic gene transfer and is regarded as a core subunit for cristae formation (Muñoz-Gómez, Slamovits, Dacks, Baier, et al. 2015; Huynen et al. 2016). Cristae-bearing discobids generally encode at least two MICOS subunits, minimally Mic60 and either Mic10 or Mic19 (see fig. 5B). For instance, the heteroloboseans *N. gruberi* and *Percolomonas* species retain Mic60 and Mic19 proteins (Muñoz-Gómez, Slamovits, Dacks, Baier, et al. 2015), but all trypanosomatids (euglenozoans) contain a putative Mic60 and two Mic10 paralogs (Eichenberger et al. 2019; Hashimi 2019; Pánek et al. 2020). Furthermore, the jakobid *Andalucia godoyi* has Mic60, Mic10, and Mic19 proteins (Gray et al. 2020). *P. flabellatum* lacks typical cristae (fig. 5A; Park et al. 2007) and may retain only Mic60, hinting that Mic60 alone being insufficient to support cristae formation in Discoba. This idea, however, needs to be tested with more data, because, outside Discoba, the cristae-bearing ciliates *Tetrahymena thermophila* and *Paramecium tetraurelia*

and amoebozoan *Acanthamoeba castellanii* all lack Mic60 but possess Mic10 (Huynen et al. 2016).

Conclusions

Comparison of the complete *P. flabellatum* mitogenomic sequence with other discobid taxa shows that the proportion of coding regions and the AT content are comparable to other heteroloboseans. There is coherence between mitochondrial gene loss and phylogenetic position in Heterolobosea. For instance, the *P. flabellatum* and its close relative *Naegleria* mitogenomes share considerable similarity in terms of mitogenome content, architecture, and synteny. Heterolobosean mitochondrion-encoded genes show minimal variation in codon usage, including the two halophilic heteroloboseids (i.e., *Pharyngomonas* and *Pleurostomum*). *P. flabellatum* cumulative GC-skew reveals a positive skew plot along with other heteroloboseans except for early-diverging *P. kirbyi* and Heteroloboseid sp. BB2.

The mitogenome *P. flabellatum* includes the same genes that encode components of the electron transport chain as the *Naegleria* mitogenome, indicating that *P. flabellatum* mitochondria are either aerobic or microaerophilic. Although no cristae structure has been shown by electron microscopy, the draft nuclear genome sequence of *P. flabellatum* revealed at least two putative, but actively transcribed Mic60/Mic60 homologs. These findings suggest that the Mic60 subunit alone is not enough to form the cristae junction. We conclude that *P. flabellatum* has an unusual mitochondrial structure due to its tolerance to low oxygen conditions, which prevail under extreme hypersalinity. Further analysis of the transcriptome and nuclear genome of *P. flabellatum* will help elucidate the functions of this organelle within the cell, and these examinations may provide useful insights into generating energy under suboxic conditions.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

This study was supported by the Collaborative Genome Program of the Korea Institute of Marine Science and Technology Promotion (KIMST) funded by the Ministry of Oceans and Fisheries (MOF) (20180430), the National Research Foundation of Korea (NRF-2017R1A2B3001923), the Next-generation BioGreen21 Program (PJ01389003) from the RDA (Rural Development Administration), Korea to Hwan Su Yoon, and the National Research Foundation of Korea (NRF-2016R1A6A1A05011910 and 2019R1A2C2002379) to Jong Soo Park.

Data Availability

The mitogenome sequence has been deposited at GenBank under the accession number MT843578, as well as two two Mic60Mic60-like sequences (accession numbers: MW019459 and MW019460).

Literature Cited

- Adams KL, Palmer JD. 2003. Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Mol Phylogenet Evol.* 29(3):380–395.
- Arakawa K, Suzuki H, Tomita M. 2009. Quantitative analysis of replication related mutation and selection pressures in bacterial chromosomes and plasmids using generalised GC skew index. *BMC Genomics.* 10(1):640.
- Barberà MJ, et al. 2010. *Sawyeria marylandensis* (Heterolobosea) has a hydrogenosome with novel metabolic properties. *Eukaryot Cell.* 9(12):1913–1924.
- Battino R, Rettich TR, Tominaga T. 1983. The solubility of oxygen and ozone in liquids. *J Phys Chem Ref Data.* 12(2):163–178.
- Bhattacharya D, Yoon HS, Hackett JD. 2004. Photosynthetic eukaryotes unite: endosymbiosis connects the dots. *Bioessays.* 26(1):50–60.
- Burger G, Gray MW, Forget L, Lang BF. 2013. Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. *Genome Biol Evol.* 5(2):418–438.
- Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. 2009. DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics.* 25(1):119–120.
- Cavalier-Smith T, Nikolaev S. 2008. The zooflagellates *Stephanopogon* and *Percolomonas* are a clade (Class Percolatea: phylum Percolozoa). *J Eukaryot Microbiol.* 55(6):501–509.
- Chihade JW, Brown JR, Schimmel PR, Ribas De Pouplana L. 2000. Origin of mitochondria in relation to evolutionary history of eukaryotic alanyl-tRNA synthetase. *Proc Natl Acad Sci U S A.* 97(22):12153–12157.
- Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One.* 5(6):e11147.
- de Graaf RM, et al. 2009. The hydrogenosomes of *Psalteriomonas lanterna*. *BMC Evol Biol.* 9(1):287.
- Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45(4):e18.
- Dobáková E, Flegontov P, Skalický T, Lukeš J. 2015. Unexpectedly streamlined mitochondrial genome of the euglenozoan *Euglena gracilis*. *Genome Biol Evol.* 7(12):3358–3367.
- Eichenberger C, et al. 2019. The highly diverged trypanosomal MICOS complex is organized in a nonessential integral membrane and an essential peripheral module. *Mol Microbiol.* 112(6):1731–1743.
- El-Gebali S, et al. 2019. The Pfam protein families database in 2019. *Nucleic Acids Res.* 47(D1):D427–32.
- Eppinger M, Baar C, Raddatz G, Huson DH, Schuster SC. 2004. Comparative analysis of four Campylobacteriales. *Nat Rev Microbiol.* 2(11):872–885.
- Faktorová D, Dobáková E, Peña-Díaz P, Lukeš J. 2016. From simple to supercomplex: mitochondrial genomes of euglenozoan protists. *F1000Res.* 5:392.
- Flegontov P, et al. 2015. Divergent mitochondrial respiratory chains in phototrophic relatives of apicomplexan parasites. *Mol Biol Evol.* 32(5):1115–1131.
- Flouri T, et al. 2015. The phylogenetic likelihood library. *Syst Biol.* 64(2):356–362.
- Fritz-Laylin LK, et al. 2010. The genome of *Naegleria gruberi* illuminates early eukaryotic versatility. *Cell.* 140(5):631–642.
- Fritz-Laylin LK, Ginger ML, Walsh C, Dawson SC, Fulton C. 2011. The *Naegleria* genome: a free-living microbial eukaryote lends unique insights into core eukaryotic cell biology. *Res Microbiol.* 162(6):607–618.
- Fu CJ, Sheikh S, Miao W, Andersson SG, Baldauf SL. 2014. Missing genes, multiple ORFs & C-to-U type RNA editing in *Acrasis kona* (Heterolobosea, Excavata) mitochondrial DNA. *Genome Biol Evol.* 6(9):2240–2257.
- Gray MW. 1998. Genome structure and gene content in protist mitochondrial DNAs. *Nucleic Acids Res.* 26(4):865–878.
- Gray MW. 2012. Mitochondrial evolution. *Cold Spring Harb Perspect Biol.* 4(9):a011403.
- Gray MW, Burger G, Lang BF. 1999. Mitochondrial evolution. *Science.* 283(5407):1476–1481.
- Gray MW, Cedergren R, Abel Y, Sankoff D. 1989. On the evolutionary origin of the plant mitochondrion and its genome. *Proc Natl Acad Sci U S A.* 86(7):2267–2271.
- Gray MW, Lang BF, Burger G. 2004. Mitochondria of protists. *Annu Rev Genet.* 38(1):477–524.
- Gray MW, et al. 2020. The draft nuclear genome sequence and predicted mitochondrial proteome of *Andalucia godoyi*, a protist with the most gene-rich and bacteria-like mitochondrial genome. *BMC Biol.* 18(1):35.
- Grigoriev A. 1998. Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res.* 26(10):2286–2290.
- Hampel V, et al. 2009. Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic “supergroups”. *Proc Natl Acad Sci U S A.* 106(10):3859–3864.
- Hancock L, Goff L, Lane C. 2010. Red algae lose key mitochondrial genes in response to becoming parasitic. *Genome Biol Evol.* 2:897–910.
- Hanousková P, Taborsky P, Cepicka I. 2019. *Dactylomonas* gen. nov., a novel lineage of heterolobosean flagellates with unique ultrastructure, closely related to the amoeba *Selenaiion kontiopes* Park, De Jonckheere & Simpson, 2012. *J Eukaryot Microbiol.* 66:369–370.
- Harding T, et al. 2013. Amoeba stage in the deepest branching heteroloboseans, including *Pharyngomonas*: evolutionary and systematic implications. *Protist.* 164(2):272–286.
- Hashimi H. 2019. A parasite’s take on the evolutionary cell biology of MICOS. *PLoS Pathog.* 15(12):e1008166.
- Herman EK, et al. 2013. The mitochondrial genome and a 60-kb nuclear DNA segment from *Naegleria fowleri*, the causative agent of primary amoebic meningoencephalitis. *J Eukaryot Microbiol.* 60(2):179–191.
- Hines JC, Ray DS. 2011. A second mitochondrial DNA primase is essential for cell growth and kinetoplast minicircle DNA replication in *Trypanosoma brucei*. *Eukaryot Cell.* 10(3):445–454.
- Huynen MA, Mühlmeister M, Gotthardt K, Guerrero-Castillo S, Brandt U. 2016. Evolution and structural organization of the mitochondrial contact site (MICOS) complex and the mitochondrial intermembrane space bridging (MIB) complex. *Biochim Biophys Acta.* 1863(1):91–101.
- Jackson CJ, Reyes-Prieto A. 2014. The mitochondrial genomes of the glaucophytes *Gloeochaete wittrockiana* and *Cyanoptyche gloeocystis*: multilocus phylogenetics suggests a monophyletic Archaeplastida. *Genome Biol Evol.* 6(10):2774–2785.
- Jhin SH, Park JS. 2019. A new halophilic heterolobosean flagellate, *Aurem hypersalina* gen. n. et sp. n., closely related to the *Pleurostomum-Tulamoeba* clade: implications for adaptive radiation of halophilic eukaryotes. *J Eukaryot Microbiol.* 66(2):221–231.
- Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 14(6):587–589.
- Kamikawa R, et al. 2014. Gene content evolution in Discobid mitochondria deduced from the phylogenetic position and complete

- mitochondrial genome of *Tsukubamonas globosa*. *Genome Biol Evol.* 6(2):306–315.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Khan MF, Patra S. 2018. Deciphering the rationale behind specific codon usage pattern in extremophiles. *Sci Rep.* 8(1):15548.
- Kirby WA, et al. 2015. Characterization of *Tulamoeba bucina* n. sp., an extremely halotolerant amoeboflagellate heterolobosean belonging to the *Tulamoeba-Pleurostomum* clade (Tulamoebidae n. fam.). *J Eukaryot Microbiol.* 62(2):227–238.
- Lang BF, et al. 1997. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* 387(6632):493–497.
- Lang BF, Gray MW, Burger G. 1999. Mitochondrial genome evolution and the origin of eukaryotes. *Annu Rev Genet.* 33(1):351–397.
- Lartillot N, Lepage T, Blanquart S. 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25(17):2286–2288.
- Lobry JR. 1995. Properties of a general model of DNA evolution under no-strand-bias conditions. *J Mol Evol.* 40(3):326–330.
- Lobry JR. 1996. Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol Biol Evol.* 13(5):660–665.
- Lukes J, et al. 2002. Kinetoplast DNA network: evolution of an improbable structure. *Eukaryot Cell.* 1(4):495–502.
- Maleszka R, Skelly PJ, Clark-Walker GD. 1991. Rolling circle replication of DNA in yeast mitochondria. *Embo J.* 10(12):3923–3929.
- Marande W, Lukes J, Burger G. 2005. Unique mitochondrial genome structure in diplomonads, the sister group of kinetoplastids. *Eukaryot Cell.* 4(6):1137–1146.
- Minh BQ, Nguyen MA, von Haeseler A. 2013. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol.* 30(5):1188–1195.
- Mitchell AL, et al. 2019. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* 47(D1):D351–D360.
- Muñoz-Gómez SA, Slamovits CH, Dacks JB, Wideman JG. 2015. The evolution of MICOS: ancestral and derived functions and interactions. *Commun Integr Biol.* 8(6):e1094593.
- Muñoz-Gómez SA, Slamovits CH, Dacks JB, Baier KA, et al. 2015. Ancient homology of the mitochondrial contact site and cristae organizing system points to an endosymbiotic origin of mitochondrial cristae. *Curr Biol.* 25(11):1489–1495.
- 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(1):268–274.
- Pánek T, Eliáš M, Vancová M, Lukeš J, Hashimi H. 2020. Returning to the fold for lessons in mitochondrial crista diversity and evolution. *Curr Biol.* 30(10):R575–R588.
- Pánek T, Ptackova E, Cepicka I. 2014. Survey on diversity of marine/saline anaerobic Heterolobosea (Excavata: Discoba) with description of seven new species. *Int J Syst Evol Microbiol.* 64(Pt_7):2280–2304.
- Park JS. 2012. Effects of different ion compositions on growth of obligately halophilic protozoan *Halocafeteria seosinensis*. *Extremophiles* 16(1):161–164.
- Park JS, Simpson AG. 2011. Characterization of *Pharyngomonas kirbyi* (= "*Macropharyngomonas halophila*" nomen nudum), a very deep-branching, obligately halophilic heterolobosean flagellate. *Protist* 162(5):691–709.
- Park JS, Simpson AG. 2015. Diversity of heterotrophic protists from extremely hypersaline habitats. *Protist* 166(4):422–437.
- Park JS, De Jonckheere JF, Simpson AGB. 2012. Characterization of *Selenaion koniopes* n. gen., n. sp., an amoeba that represents a new major lineage within Heterolobosea, isolated from the Wieliczka salt mine. *J Eukaryot Microbiol.* 59(6):601–613.
- Park JS, Simpson AGB, Lee WJ, Cho BC. 2007. Ultrastructure and phylogenetic placement within Heterolobosea of the previously unclassified, extremely halophilic heterotrophic flagellate *Pleurostomum flabellatum* (Ruinen 1938). *Protist* 158(3):397–413.
- Perna NT, Kocher TD. 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J Mol Evol.* 41(3):353–358.
- Petú M, et al. 2018. Evolution of mitochondrial TAT translocases illustrates the loss of bacterial protein transport machines in mitochondria. *BMC Biol.* 16(1):141.
- Roy J, Faktorová D, Lukeš J, Burger G. 2007. Unusual mitochondrial genome structures throughout the Euglenozoa. *Protist* 158(3):385–396.
- Saha SK, Goswami A, Dutta C. 2014. Association of purine asymmetry, strand-biased gene distribution and PoC within Firmicutes and beyond: a new appraisal. *BMC Genomics.* 15:430.
- Salzberg SL, Salzberg AJ, Kerlavage AR, Tomb JF. 1998. Skewed oligomers and origins of replication. *Gene* 217(1–2):57–67.
- Simpson AGB, Inagaki Y, Roger AJ. 2006. Comprehensive multigene phylogenies of Excavate protists reveal the evolutionary positions of "primitive" eukaryotes. *Mol Biol Evol.* 23(3):615–625.
- Stoco PH, et al. 2014. Genome of the avirulent human-infective trypanosome – *Trypanosoma rangeli*. *PLoS Negl Trop Dis.* 8(9):e3176.
- Sueoka N. 1995. Intrastrand parity rules of DNA base composition and usage biases of synonymous codons. *J Mol Evol.* 40(3):318–325.
- Tillier ER, Collins RA. 2000. The contributions of replication orientation, gene direction, and signal sequences to base-composition asymmetries in bacterial genomes. *J Mol Evol.* 50(3):249–257.
- Tsaousis AD, Nývltová E, Šuták R, Hrdý I, Tachezy J. 2014. A nonmitochondrial hydrogen production in *Naegleria gruberi*. *Genome Biol Evol.* 6(4):792–799.
- Tymł T, Lares-Jimenez LF, Kostka M, Dykova I. 2017. *Neovahlkampfia nana* n. sp. reinforcing an underrepresented subclade of Tetramitida, Heterolobosea. *J Eukaryot Microbiol.* 64(1):78–87.
- Valach M, Burger G, Gray MW, Lang BF. 2014. Widespread occurrence of organelle genome-encoded 5S rRNAs including permuted molecules. *Nucleic Acids Res.* 42(22):13764–13777.
- Valach M, Moreira S, Hoffmann S, Stadler PF, Burger G. 2017. Keeping it complicated: mitochondrial genome plasticity across diplomonads. *Sci Rep.* 7(1):14166.
- Yabuki A, Gyaltsen Y, Heiss AA, Fujikura K, Kim E. 2018. *Ophirina amphinema* n. gen., n. sp., a new deeply branching Discobid with phylogenetic affinity to Jakobids. *Sci Rep.* 8(1):16219.
- Yang J, Harding T, Kamikawa R, Simpson AGB, Roger AJ. 2017. Mitochondrial genome evolution and a novel RNA editing system in deep-branching heteroloboseids. *Genome Biol Evol.* 9(5):1161–1174.

Associate editor: Geoff McFadden