Phylogenomics of novel ploeotid taxa contribute to the backbone of the euglenid tree

Gordon Lax [ORCID] | Anna Cho | Patrick J. Keeling

Department of Botany, University of British Columbia, Vancouver, V6T 1Z4, British Columbia, Canada

Correspondence
Gordon Lax, Department of Botany, University of British Columbia, 2212 Main Mall, Vancouver, V6T 1Z4, British Columbia, Canada.
Email: gordon.lax@gmail.com

Abstract

Euglenids are a diverse group of flagellates that inhabit most environments and exhibit many different nutritional modes. The most prominent euglenids are phototrophs, but phagotrophs constitute the majority of phylogenetic diversity of euglenids. They are pivotal to our understanding of euglenid evolution, yet we are only starting to understand relationships amongst phagotrophs, with the backbone of the tree being most elusive. Ploeotids make up most of this backbone diversity—yet despite their morphological similarities, SSU rDNA analyses and multigene analyses show that they are non-monophyletic. As more ploeotid diversity is sampled, known taxa have coalesced into some subgroups (e.g. Alistosa), but the relationships amongst these are not always supported and some taxa remain unsampled for multigene phylogenetics. Here, we used light microscopy and single-cell transcriptomics to characterize five ploeotid euglenids and place them into a multigene phylogenetic framework. Our analyses place *Decastava* in Alistosa; while *Hemiolia* branches with *Liburna*, establishing the novel clade Karavia. We describe *Hemiolia limna*, a freshwater-dwelling species in an otherwise marine clade. Intriguingly, two undescribed ploeotids are found to occupy pivotal positions in the tree: *Chelandium granulatum* nov. gen. nov. sp. branches as sister to *Olkasia*, and *Gaulosia striata* nov. gen. nov. sp. remains an orphan taxon.

KEYWORDS
biodiversity, flagellates, freshwater, marine, phagotrophic, phylogeny, protist

INTRODUCTION

EUGLENIDS are a large and diverse group of Euglenozoa with phototrophic, osmotrophic, and phagotrophic species (Kostygov et al., 2021; Leander et al., 2017). Most known euglenids have one or two flagella and usually either glide on substrate or swim (Lax & Simpson, 2020; Leander et al., 2017). One key characteristic is the euglenid pellicle—proteinaceous parallel strips underlying the cell membrane that are backed by a system of microtubules (Cavalier-Smith, 2017; Leander, 2004). The number of pellicle strips can vary drastically in euglenids, some phototrophic euglenids have upwards of 100 pellicle strips, whereas some phagotrophic petalomonads have only four pellicle strips (Leander et al., 2017). Species with over 20 or so pellicle strips can exhibit 'euglenid metaboly', rendering the cells highly flexible (Leander et al., 2007), in contrast to taxa with fewer pellicle strips, which are all rigid.

Both the phototrophic Euglenophyceae and osmotrophic Aphagea arose from phagotrophic euglenids, yet until recently relatively few molecular sequences of phagotrophs were available (Lee & Simpson, 2014a, 2014b). Even the selection of SSU rDNA sequences was not representative of phagotrophic euglenid diversity, with many genera completely unsampled, and most only having 1–2 sequences available (Leander et al., 2017). Recent work has greatly expanded the taxon-sampling of

SSU rDNA from phagotrophs, leading to many insights into the relationships between these species, but the inherent limitations of single-gene trees, together with the fact that euglenoid SSU rDNA is divergent in nature, has complicated these analyses (Busse et al., 2003; Cavalier-Smith et al., 2016; Lax et al., 2019; Lax & Simpson, 2020; Paerschke et al., 2017). To fully clarify these relationships, large multigene analyses are required, but such data has only recently been acquired (Lax et al., 2021).

Currently known phagotrophic euglenids can be divided into three main groups: generally flexible Spirocuta with 20+ pellicle strips, rigid petalomonads with ≤10 pellicle strips, and rigid ploeotids with 10–12 pellicle strips (Leander et al., 2007, 2017). Spirocuta also contains phototrophic and primary osmotrophic euglenids and their closest living phagotrophic relatives, and petalomonads have been identified as the deepest branch of euglenids to date (Lax et al., 2021). Meanwhile ploeotids are a phylogenetically diverse group of phagotrophs that share certain morphological characters. They are all rigid cells with 10–12 pellicle strips that can feed on either eukaryotes and/or bacteria (Lax et al., 2019) and glide on their posterior flagellum (Lax & Simpson, 2020; Leander et al., 2017). Despite these similarities, questions have been raised as to whether these taxa are indeed monophyletic. Recent SSU rDNA analyses with a focused taxon sampling on ploeotids have shown that their evolutionary relationships appear to be more complicated (Lax et al., 2019; Lax & Simpson, 2020; Lee & Simpson, 2014a; Paerschke et al., 2017). Moreover, a recent multigene-analysis showed similar results: rather than falling into a single monophyletic clade, ploeotids were divided into up to four discrete groups, indeed mostly each known genus forming its own group (Lax et al., 2021). The sole exception to date is the well-supported clade Alistosa, which contains the genera *Keelungia*, *Lentomonas*, *Ploeotia*, and *Serpenomonas*. The other ploeotid clades for which multigene-data are currently available are *Entosiphon*, *Liburna*, and *Olkasia*.

Several ploeotid taxa which—from SSU rDNA sequences—appear to fall in key phylogenetic positions (Lax et al., 2019) are still missing from multigene analyses. A single SSU sequence of soil-dwelling *Decastava* exists, but its relationship to other ploeotids is ambiguous (Cavalier-Smith et al., 2016) and no multigene data exists for this taxon. *Hemiolia* is morphologically very similar to *Liburna*: Both are only known to be marine, and both exhibit a characteristic fast gliding and sudden-stop motion (Ekebom et al., 1995; Larsen, 1987). SSU rDNA analyses do not decisively show their evolutionary placement amongst euglenids, and multigene analyses so far only included *Liburna* (Lax et al., 2021). Moreover, two seemingly important ploeotid strains are only known from single-cell sampled SSU rDNA sequences: 'SMS7' and 'CARR5' (Lax et al., 2019). Again, their placement in SSU-analyses is highly unstable and no additional molecular data exists.

To further our understanding of phagotrophic euglenid phylogenetics, we generated five single-cell transcriptomes of ploeotid taxa (one *Decastava*, two *Hemiolia*, two previously 'Unidentified ploeotids' closely related to single-cells SMS7 and CARR5 from a previous study). We use the transcriptomic data to update the available multigene phylogenetic data for euglenids. Based on this phylogeny, we evaluate the relationships with other euglenid groups, propose a new clade Karavia that encompasses *Liburna* and *Hemiolia*, and propose the novel taxa *Hemiolia limna* nov. sp. (the first freshwater member of this clade), *Gaulosia striata* nov. gen. nov. sp. (related to CARR5), and *Chelandium granulatum* nov. gen. nov. sp. (related to SMS7).

MATERIALS AND METHODS

Sample collection and preparation

Marine intertidal samples were collected from various sites in British Columbia, Canada: Centennial Beach Park at Boundary Bay (Vancouver), and Powell River. A single soil sample was collected from Galliano Island in British Columbia, Canada, and a single freshwater sample was collected from Quadra Island, British Columbia, Canada. See Table S1 for additional information.

Marine samples were treated as described previously (Larsen & Patterson, 1990). Briefly, samples were placed in trays 1–2 cm high, covered with a Kimwipe tissue, and #1 coverslips added onto them. After 24–72 h in a natural day-night cycle, the coverslips were examined with the downside facing upwards on a Leica DLIM inverted microscope and imaged with a Sony a7RIII camera. For the Galliano sample (SPO2), minute amounts of soil were placed in a petri dish or on a coverslip, and a couple of drops of 0.2 μm-filtered tapwater was added. These samples were examined and imaged as above.

Single cell isolation

Single cells of phagotrophic euglenids were identified and imaged at 630×, and subsequently picked using manual single-cell isolation. Cells were washed 2–5× in 0.2 μm-filtered seawater from the same source (marine samples), 0.2 μm-filtered tapwater (soil samples), or PCR-grade water (freshwater samples), before being placed in 0.2 mL PCR-tubes containing 2 μL of Smart-seq2 lysis buffer (Picelli et al., 2014).

To lyse the cells, tubes were subjected to 3–5 freeze–thaw cycles (alternating between −80°C and room temperature). The Smart-seq2 protocol was followed to generate cDNA using poly-A selection, with 24 PCR cycles for final amplification (Picelli et al., 2014).

The resulting cDNA was then quantified using a Qubit HS DNA assay (Invitrogen), and libraries were

prepared using Illumina Nextera XT or Illumina DNA Library Preparation kits (Table S2). Sequencing for all samples was done on an Illumina NextSeq instrument with 2×150-bp paired end reads, whereas sample BB14 was additionally resequenced on an Illumina MiSeq with 2×250-bp reads.

Transcriptome assembly

Raw reads were quality-checked with FastQC 0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and error-corrected with rcorrector 1.0.4 (Song & Florea, 2015). Trimmomatic 0.39 (Bolger et al., 2014) was used to adapter- and quality trim resulting reads, using Nextera adapter sequences (5′-CTGTCTCTTAT ACACATCTCCGAGCCCACGAGAC-3′, 5′-CTGTC TCTTATACACATCTGACGCTGCCGACGA-3′) and Smart-seq2 adapter sequences (TSO: 5′-AAGCAGTG GTATCAACGCAGAGTACATGGG-3′, oligo-dT: 5′-AAGCAGTGGTATCAACGCAGAGTACTTTT TTTTTTTTTTTTTTTTTTTTTTTTTTT-3′, ISPCR: 5′-AAGCAGTGGTATCAACGCAGAGT-3′) as input with options ILLUMINACLIP: 2:30:10 LEADING:5 SLIDINGWINDOW:5:16 MINLEN:60. Trimmed paired and unpaired sequences were assembled with rnaSPAdes 3.14.1 (Bushmanova et al., 2018) using default settings.

General assembly metrics of each transcriptome were collected with QUAST 5.2.0 (Mikheenko et al., 2018), and completeness was assessed with BUSCO 5.4.3 (Manni et al., 2021) using the eukaryota_odb10 database. Outputs from FastQC, trimmomatic, QUAST, and BUSCO were summarized using MultiQC 1.13 (Ewels et al., 2016).

SSU rRNA phylogenetics

barrnap 0.9 (https://github.com/tseemann/barrnap/) was used to extract any eukaryotic SSU rRNA sequences from the assemblies. The extracted sequences were then blasted against NCBI's nr/nt database to confirm any euglenid sequences. These sequences were added to a modified, previously published dataset (Lax & Simpson, 2020), aligned using MAFFT E-INS-I 7.475 (Katoh & Standley, 2013), manually corrected for errors, and masked with trimAl 1.2rev59 (Capella-Gutiérrez et al., 2009) (-gt 0.88 -st 0.001). The resulting masked dataset was composed of 218 taxa with 1036 positions. Sequences used in the following analyses are listed in Table S3.

Maximum likelihood (ML) phylogenies were estimated using RAxML-NG 1.1.0 (Kozlov et al., 2019) under the GTR+GAMMA model and 1000 non-parametric bootstrap replicates. We also conducted a Bayesian analysis in MrBayes 3.2.7a (Ronquist et al., 2012) under the GTR+GAMMA model running four parallel chains

with default heating parameters and 50,000,000 generations each (run on the CIPRES webserver: https://www.phylo.org/index.php). Trees were sampled every 10,000 generations, and 25% were discarded as burn-in. Chain convergence was assessed by PSRF values (Potential Scale Reduction Factor) approaching 1.0. Both analyses are summarized in Figure S1.

Phylogenomic analyses

We base our dataset on a recently published study that investigated euglenid phylogenetics with a 20-gene dataset (Lax et al., 2021). We also added two recently published members of the Prokinetoplastina (Tikhonenkov et al., 2021).

Coding regions of all transcriptomes were determined with Transdecoder 5.5.0 (Grabherr et al., 2011). Relevant marker genes were extracted with blastp from these translated peptide assemblies as queries against the 20 single-gene seeds. Hits were subsequently checked with blastp against Swiss-Prot. The surviving top five hits for each transcriptome were then appended to the existing 20 single-gene datasets, aligned with MAFFT L-INS-I, and trimmed with BMGE 1.12 (-m BLOSUM30 -h 0.5 -g 0.2; Criscuolo & Gribaldo, 2010). A phylogeny was estimated for each single-gene alignment using the LG4X model in IQ-TREE 1.6.12 (Nguyen et al., 2015) with 1000 ultrafast bootstraps (UFB). Each single-gene tree was then carefully checked, and contaminant, paralogous, or otherwise long-branching sequences removed from the untrimmed alignments. After re-aligning, re-trimming, and re-estimating the trees, this process was repeated one more time.

We noticed some deep paralogies amongst euglenids in PABPC4 trees, so we subsequently excluded this protein from our analyses. The cleaned, aligned, and trimmed single-genes were then concatenated, producing a dataset with 19 genes, 62 taxa, and 6194 positions. For this '19-base' dataset, a phylogeny was reconstructed under the LG+C60+F+G model in IQ-TREE 2.2.0 (Minh et al., 2020), with 1000 UFB (Minh et al., 2013). All taxa included in this and the following analyses are listed in Table S4. We also generated 500 non-parametric bootstraps for the same dataset under the posterior mean site frequency model (PMSF) in IQ-TREE, using the previous LG+C60+F+G output as a guide tree (Wang et al., 2018). The '19-base' dataset was run in PhyloBayes-MPI v1.8 (Lartillot et al., 2009) under the CAT+GTR model with four independent cold chains, each running for 30,000 cycles with a burn-in of 25%, sampled every 10 trees (all four chains converged). Both analyses are depicted in Figure 2.

To investigate the impact of fast-evolving sites on our phylogeny, a '19-FSR' dataset was created, where fast-evolving sites were removed in 4%-increments as described previously (Lax et al., 2021). For each increment,

a LG+C20+F+G phylogenetic tree with 1000 UFBoot replicates was constructed. Figure S2 plots the bootstrap values for select relevant nodes for each of the incremental fast-site removal steps.

A rogue taxon analysis with RogueNaRok using the RNR algorithm (Aberer et al., 2013) showed diplonemid 'SAG D1' and euglenid *Dinema litorale* UB26 as unstable, or 'rogue' taxa. We excluded these taxa in a '19-noRogue' dataset, which was then run under LG+C60+F+G in IQ-TREE with 1000 UFB (Figure S3).

RESULTS

Description of cells
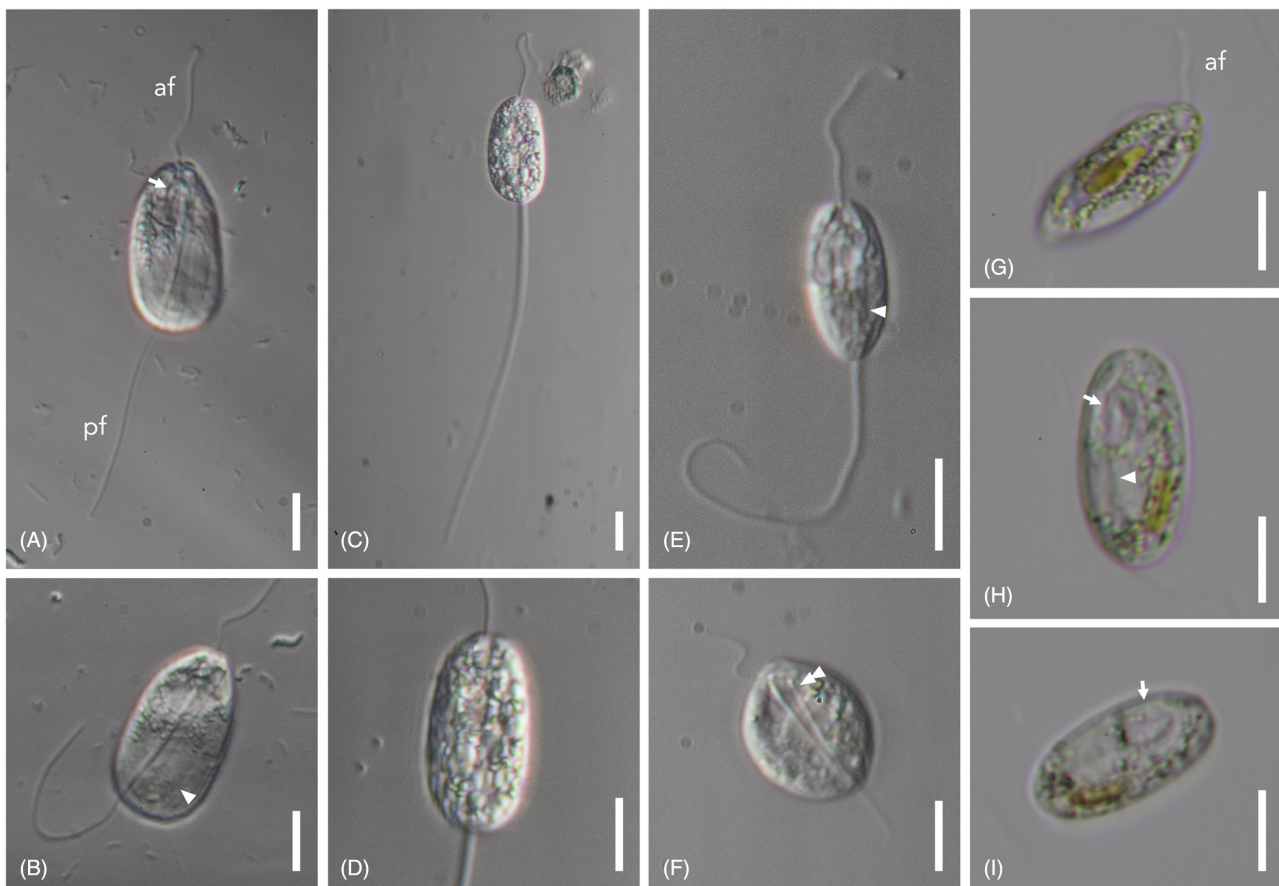
Gaulosia striata (*BB9*)

A single cell was isolated from oxygenated marine intertidal sediment, oblong-pyriform in shape with a rounded posterior end (Figure 1A). The cell measured 29.3×14.5 μm in size (Table 1) and was gliding on its thickened posterior flagellum (2.2× cell length) with its

anterior flagellum being 'cast out' in front of the cell (like a fishing line; 0.8× cell length; Movie S1). Five straight, broad pellicle strips were found on the dorsal and five on the ventral side, on a slight angle compared to the cell outline (Figure 1B). The cell changed direction after slowly moving back on its posterior flagellum. The posterior half of the cell was smooth and presumably housed the nucleus. The small, slim flagellar pocket was hard to see but was situated along the midline in the anterior quarter of the cell (Figure 1A). We did not observe any feeding apparatus. The BUSCO completeness score was 43.2% complete single and duplicated, 14.9% fragmented, and 41.9% missing. The SSU rDNA sequence of this cell is deposited under GenBank accession number OQ331014.

Chelandium granulatum (*BB14*)

From oxygenated marine intertidal sediment, we isolated a single cell, oblong in shape and heavily granulated (Figure 1C). It measured 27.1×14.6 μm in size and was gliding on its thickened posterior flagellum (4.4×

FIGURE 1 Micrographs of isolated cells. (A, B) *Gaulosia striata* nov. gen. nov. sp. BB9, arrow points towards flagellar pocket, arrowhead shows pellicle strips. (C, D) *Chelandium granulatum* nov. gen. nov. sp. BB14. (E) *Hemiolia trepidum* PRC4, arrowhead points onto pellicle strip. (F) *Decastava* sp. SPO2, with a double-arrowhead denoting the feeding apparatus. (G–I) *Hemiolia limna* nov. sp. ML5-ML7, arrows in (H) and (I) pointing towards feeding apparatus, arrowhead in H shows pellicle strips. 'af' marks the anterior flagellum, 'pf' the posterior flagellum. All scale bars are 10 μm.

TABLE 1 Cell measurements of single cells isolated in this study.

Cell	Length (μm)	Width (μm)	Af length	Af ratio	Pf length	Pf ratio
BB9	29.3	14.5	24 μm	0.8×	64 μm	2.2×
BB14	27.1	14.6	29.5 μm	1×	120 μm	4.4×
ML5	24.4	9.8	?[a]	–	?[a]	–
ML7	24.4	12	?[a]	–	?[a]	–
PRC4	17.3	8.8	23.5 μm	1.3×	62 μm	3.5×
SPO2	21.5	17.3	18 μm	0.8×	30 μm	1.4×

[a]Flagellum truncated during isolation.

cell length; Table 1), while the anterior flagellum was slowly repeatedly being 'cast out' similar to *Gaulosia* (1× cell length; Movie S2). Neither pellicle strips nor feeding apparatus could be seen, likely due to the presence of refractile granules in the interior of the cell (Figure 1D). For this cell, 3.5% complete single-copy and duplicate, and 2.4% fragmented BUSCOs were recovered, with 94.1% missing. The corresponding SSU rDNA sequence for this cell can be found under GenBank accession number OQ331015.

Hemiolia trepidum (*PRC4*)

One cell was observed and isolated from oxygenated marine intertidal sediment. The cell was oblong, with a slightly tapered posterior, slightly dorso-ventrally flattened, and measured 17.3 × 8.8 μm (Figure 1E; Table 1). The thickened posterior flagellum (3.5× cell length) forming a hook-shape, was used for a fast gliding motion, which was stopped intermittently, at which time both flagella largely stopped moving. The anterior flagellum (1.3× cell length) was usually held in front of the cell, towards the right, and was trembling along its distal half (Movie S5). The cell occasionally stopped and jerked back on its posterior flagellum after which it resumed gliding in a different direction. 3–4 faint pellicle strips could be seen on its dorsal side (Figure 1E). No feeding apparatus was visible at 630× magnification. We recovered 6.7% complete single-copy and duplicate, and 3.9% fragmented BUSCOs, with 89.4% of BUSCOs missing. The corresponding SSU rDNA sequence for this cell is deposited under GenBank accession number OQ331010.

Decastava sp. (*SPO2*)

We observed and isolated one cell from soil and forest litter. The cell was irregularly lentil-shaped, dorso ventrally flattened and measured 21.5 × 17.3 μm (Figure 1F; Table 1). It was gliding on its posterior flagellum of ~1.4× cell length, while the anterior flagellum (~0.8× cell length) was flailing in front of the cell (Movie S6). A prominent hook-shaped feeding apparatus extends almost completely down the whole length of the cell, along its midline (Figure 1F). No pellicle strips were observed with light microscopy at 630× magnification. For this assembly, 31.4% complete single-copy and duplicate BUSCOs were recovered, 10.2% fragmented, and the remaining 58.4% of BUSCOs missing. Its corresponding SSU rDNA sequence is deposited under GenBank accession number OQ331009.

Hemiolia limna (*ML5* and *ML7*)

We isolated two cells from the same freshwater sediment. Cells ranged from 24.4/24.4 × 9.8/12 μm in size were oblong with a pointed posterior (ML5 only) and slightly ventrally flattened (Figure 1G–I; Table 1). Cell ML5 had its anterior flagellum truncated and posterior flagellum sheared prior to measurement (Figure 1G), whereas cell ML7 lost its anterior flagellum completely and likely had its posterior flagellum truncated. We observed a flagellar pocket in both cells (1/3× cell length; Figure 1H,I), and 3–4 faint dorsal pellicle strips could be seen on both cells (ML7: Figure 1H; ML5: Movie S3). Both ML5 and ML7 had some internal yellow-green material, presumably ingested algal material (Figure 1G–I; Movies S3 and S4). This co-assembly had 14.9% single-copy and duplicate, and 20% fragmented BUSCOs recovered, with 65.1% missing. The SSU rDNA sequence for ML5 is deposited under GenBank accession number OQ331011, ML7 under OQ331012, and the sequence from co-assembly ML5-ML7 under OQ331013 (all three SSU rDNAs are identical).

Phylogenetics

We recovered euglenids as a monophyletic clade in all of our multigene trees, with high support (85%–99% BS, 1 pp; Figure 2; Figure S3; also see graph in Figure S2). Major groups of euglenids that were recovered are Petalomonads as the deepest branch (100% BS, 1 pp), Spirocuta as a large clade encompassing Euglenophyceae, Aphagea, and many phagotrophs (100% BS, 1 pp), and Alistosa as a clade containing various ploeotid taxa (100%
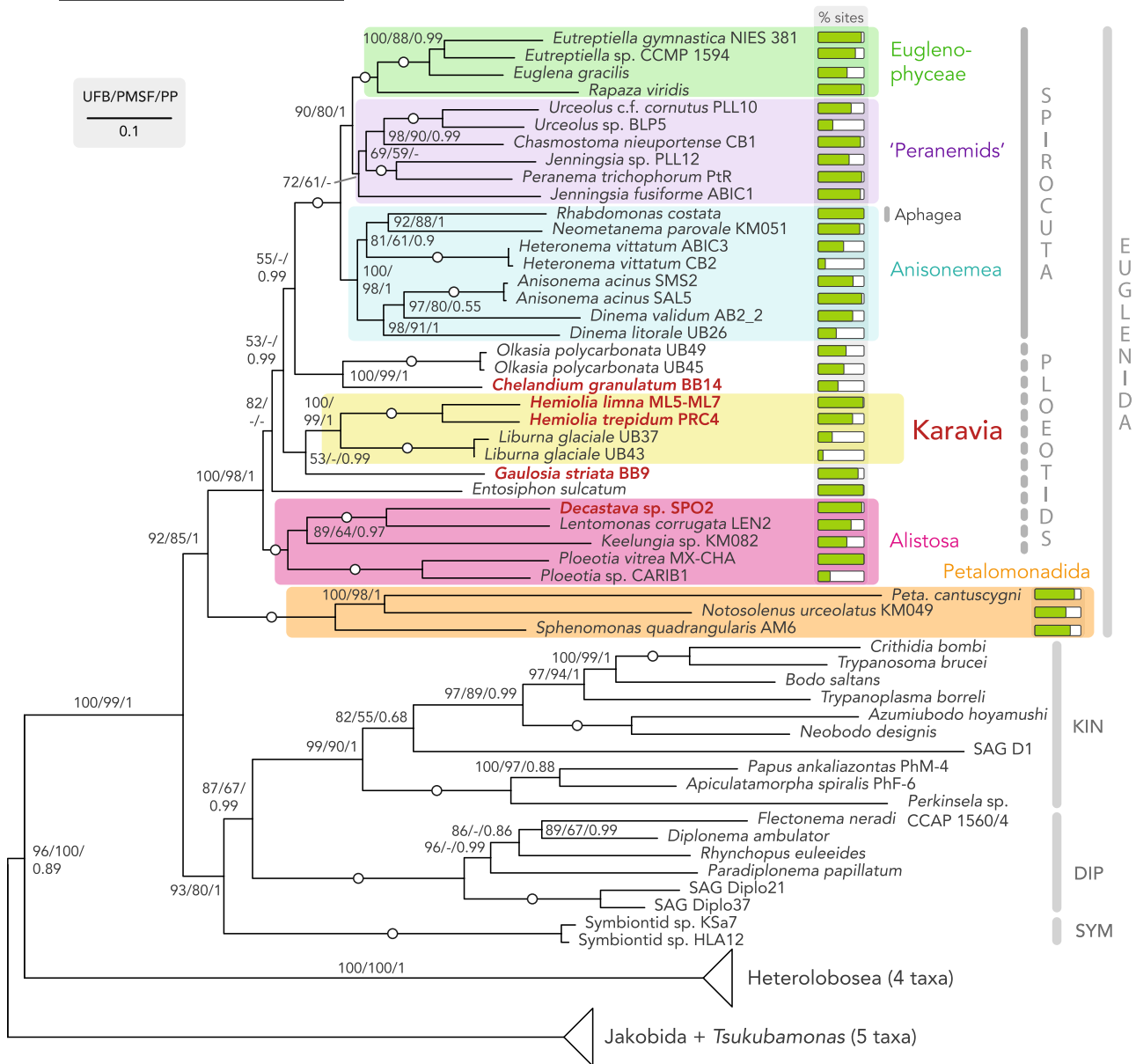
FIGURE 2 19-gene phylogenetic tree of Discoba, estimated with Maximum likelihood in IQ-TREE2 under the LG+C60+F+G model and 1000 UFB. New taxa are bolded and red. UFB values, non-parametric bootstrap values derived from 200 BS under the PMSF model, and posterior probabilities derived from a PhyloBayes analysis under the CAT+GTR model are mapped onto each node (PP). Nodes that received full support under all methods are denoted with a circle. A dash (−) denotes a missing split in that particular analysis. For euglenid taxa, we also report the percentage of amino acid sites recovered from the 19-gene dataset (out of 6194 aa total). Ploeotids are recovered as non-monophyletic. Kinetoplastea are abbreviated as 'KIN', Diplonemea as 'DIP', and Symbiontida as 'SYM'.

BS, 1 pp). Some of our analyses also recovered Olkaspira, which includes spirocute taxa and *Olkasia*, albeit with very low support (54%–55% BS, 0.99 pp, not recovered in main PMSF analysis). The ploeotid taxa *Hemiolia* and *Liburna* were always recovered as a highly to fully-supported clade, for which we propose the name Karavia (99%–100% BS, 1 pp), with *Gaulosia* sometimes falling sister to this clade, but always with low support (52%–53% BS, 0.99 pp, not recovered in main PMSF analysis).

The SSU rDNA analysis was, not surprisingly, less well resolved. We recovered euglenids with low support (Figure S1), with symbiontids branching inside euglenids.

An additional analysis excluding all symbiontid sequences did not change the major topology of the tree (data not shown). Spirocuta is recovered with low to moderate support, while maximally supported Petalomonadida branches amongst ploeotids. Ploeotids fall into several clades, as in previous SSU rDNA analyses (Lax et al., 2019; Lax & Simpson, 2020; Schoenle et al., 2019). Clade Alistosa breaks apart into 3 main branches (*Lentomonas*+*Decastava*; *Keelungia*; Ploeotiidae), rendering the group polyphyletic. Karavia is non-monophyletic, with long-branching *Hemiolia* forming the sister clade to *Lentomonas*+*Decastava* clade, whereas *Liburna* branches

as part of a poorly supported clade of Petalomonadida, *Gaulosia*, and Ploeotiidae. *Chelandium* falls sister to Spirocuta with low support, whereas *Olkasia* is recovered as the deepest branch of euglenids with low support.

Taxonomic descriptions

Karavia Lax, Cho and Keeling

Taxonomic summary. Eukaryota; Discoba; Euglenozoa; Euglenida.

Diagnosis. The smallest crown-clade comprising *Liburna glaciale* Lax et al., 2019 and *Hemiolia trepidum* Lax et al., 2019, but not *Euglena gracilis* Klebs (Olkaspira; Spirocuta; Euglenophyceae), *Entosiphon sulcatum* Dujardin, *Olkasia polycarbonata* Lax et al., 2019 (Olkaspira), and *Ploeotia vitrea* Dujardin 1841 (Alistosa). This is a minimum crown-clade definition with external specifiers.

Etymology. From 'caravela' (Portuguese, singular), a class of small sailing ships, owing to the fact that both current genera of this clade, *Liburna* and *Hemiolia*, are named after historical ships.

Reference phylogeny. Figure 2 in this paper.

Comments. This is a zoological name above the level of family and falls outside the zoological (and botanical) codes of nomenclature.

The following nomenclatural work has been registered in Zoobank, under LSID urn:lsid:zoobank.org:pub:E952938D-563B-4FDA-8C0C-7E83D72ACB1B.

Gaulosia gen. nov. Lax, Cho and Keeling (ICZN).

Taxonomic summary. Eukaryota; Discoba; Euglenozoa; Euglenida.

Diagnosis. Free-living, rigid, biflagellated heterotrophic euglenid, oblong profile with a markedly rounded, blunt posterior. Cell slightly flattened ventrally. Anterior flagellum ~0.8× cell length, thickened posterior flagellum ~2.2× cell length. Slow gliding speed, when under stress slowly retracts on posterior flagellum. 10 faint broad pellicle strips can be seen with light microscopy.

Type species. Gaulosia striata Lax, Cho and Keeling (see below).

Etymology. From 'gaulos' (Greek, singular), a Phoenician trading ship. Refers to the rounded, non-flattened appearance and visible pellicle strips resembling wooden ship laps, and relative slow gliding speed.

Transfer of existing taxa. Transfer 'Unidentified Ploeotid CARR5' (Lax et al., 2019) to genus *Gaulosia* as *Gaulosia* sp. based on morphological similarities and phylogenetic affinity.

ZooBank Accession. LSID urn:lsid:zoobank.org:act:D04E8FEE-5DE3-42AB-859A-FED4F603ADFF.

Gaulosia striata sp. nov. Lax, Cho and Keeling (ICZN).

Diagnosis. Oblong cell profile with rounded blunt posterior, 29.3 μm long and 14.5 μm wide, ventrally flattened.

Type material. Cell imaged in Figure 1A,B.

Type habitat. Marine oxic intertidal sediment.

Type locality. Centennial Beach, Boundary Bay, Vancouver, British Columbia, Canada (49°0′36.1188″N, 123°2′11.1984″W).

Etymology. From Latin 'striata' for 'striped', owing to the visible pellicle strips in light microscopy.

Gene sequence. The SSU-rDNA sequence extracted from the single-cell transcriptome assembly of cell BB9 is deposited under GenBank Accession number OQ331014.

ZooBank Accession. LSID urn:lsid:zoobank.org:act:E264F20C-55D6-43ED-A880-CD47DC10708F.

Chelandium gen. nov. Lax, Cho and Keeling (ICZN).

Taxonomic summary. Eukaryota; Discoba; Euglenozoa; Euglenida.

Diagnosis. Free-living, rigid, biflagellated hetertrophic euglenid with an oblong outline, slightly flattened ventrally, cell is full of refractile granules. Anterior flagellum ~1× cell length, considerably thickened posterior flagellum ~4.4× cell length. Cell glides relatively fast on the posterior flagellum and in a straight line.

Type species. Chelandium granulatum Lax, Cho and Keeling (see below).

Etymology. From 'chelandium' (Latin, singular), the latinized name of 'chelandion' (greek), a war and transport galley of the Byzantine navy.

Transfer of existing taxa. Transfer 'Unidentified Ploeotid SMS7' (Lax et al., 2019) to genus *Chelandium* as *Chelandium* sp. based on morphological similarities and phylogenetic affinity.

ZooBank Accession. LSID urn:lsid:zoobank.org:act:337DC96F-D9A3-4A91-8773-257CB357476C.

Chelandium granulatum sp. nov. Lax, Cho and Keeling (ICZN).

Diagnosis. Cell is oblong, 27.1 μm long and 14.6 μm wide and ventrally flattened.

Type material. Cell imaged in Figure 1C,D.

Type habitat. Marine oxic intertidal sediment.

Type locality. Centennial Beach, Boundary Bay, Vancouver, British Columbia, Canada (49°0′36.1188″N, 123°2′11.1984″W).

Etymology. From Latin 'granulatum' for 'granulate', referring to the granulated appearance due to refractile granules.

Gene sequence. The SSU-rDNA sequence extracted from the single-cell transcriptome assembly of cell BB14 is deposited under GenBank Accession number OQ331015.

ZooBank Accession. LSID urn:lsid:zoobank.org:act:1BF5DCF2-5F50-4438-9091-A9D53FDB140C.

Hemiolia limna sp. nov. Lax, Cho and Keeling (ICZN).

Taxonomic summary. Eukaryota; Discoba; Euglenozoa; Euglenida; Karavia.

Diagnosis. Cell ML7 is oblong, ventrally flattened, 24.4 μm by 12 μm. Phylogenetically closely related, but distinct from *Hemiolia trepidum*.

Type material. Cell ML7 imaged in Figure 1G–I.

Type habitat. Oxic freshwater sediment and gravel.

Type locality. Mud Lake, Quadra Island, British Columbia, Canada (50°7′58.368″N, 125°17′23.964″W).

Etymology. From Greek 'limne' for 'lake', since it was found in a freshwater environment.

Gene sequence. The SSU-rDNA sequence extracted from the single-cell transcriptome assembly of cell ML7 is deposited under GenBank Accession number OQ331012.

ZooBank Accession. LSID urn:lsid:zoobank.org:act:86478CE6-C4EF-434C-9C30-8340273B22F3.

DISCUSSION

General phylogenetic relationships

Most phylogenetic relationships recovered in our analyses confirm those of the most recent multigene phylogeny of euglenids (Lax et al., 2021). We recover euglenids as a monophyletic group with high support in all of our analyses, which is in agreement with several recent multigene analyses that included euglenids (Lax et al., 2021; Wideman et al., 2019; Záhonová et al., 2021). Spirocuta, a large clade that includes many phagotrophic euglenids, but also the diverse clade of Euglenophyceae (phototrophic euglenids) and Aphagea (primary osmotrophic euglenids), is fully supported in our analyses.

Petalomonadida, likely one of the most species-rich groups of euglenids (Huber-Pestalozzi, 1955; Leander et al., 2017), remain the deepest branching clade of euglenids to date (Figure 2). Petalomonads have fewer than 10 pellicle strips, one or two emergent flagella, and always have a rigid pellicle (Leander et al., 2007, 2017). While

the exact relationships within this group are still somewhat unclear and disputed (Cavalier-Smith, 2016; Lax & Simpson, 2020; Lee & Simpson, 2014b), they are always recovered as a strongly supported group in all recent SSU rDNA phylogenies and in a multigene phylogeny (Cavalier-Smith, 2016; Cavalier-Smith et al., 2016; Lax et al., 2021; Lax & Simpson, 2020; Lee & Simpson, 2014b; Paerschke et al., 2017; Schoenle et al., 2019).

The middle section of the euglenid tree backbone—nestled between Spirocuta and Petalomonadida—is occupied by a variety of ploeotid taxa and groups (Figure 2). While the phylogenetic placement of some ploeotid taxa like *Gaulosia* remain unclear, the majority of sampled ploeotids to date have now coalesced into two main groups in multigene studies: Alistosa and Karavia.

Decastava is an alistosan ploeotid

Our *Decastava* sp. SPO2 transcriptome is the first multigene-grade data for this genus, enabling its placement in a phylogenomic context. *Decastava* branches sister to *Lentomonas* with full support, as part of Alistosa in our multigene analyses (Figure 2). When included in previous SSU rDNA analyses *Decastava* branched sister to *Lentomonas*, but with low to moderate support only (Lax et al., 2019; Lax & Simpson, 2020) or to *Keelungia* when *Lentomonas* was not included (Schoenle et al., 2019).

Alistosa was formally proposed as molecularly defined clade in 2021 and was inferred to include the genera *Ploeotia*, *Serpenomonas*, *Lentomonas*, *Keelungia*, and *Decastava* (Lax et al., 2021). Our analysis here confirms this inferred composition. While these genera all share certain morphological characters—such as 10 pellicle strips (often prominent), an anterior flagellum presumably used for prey capture (Lax et al., 2019; Leander et al., 2017), and a posterior flagellum used for gliding—it is unclear what, if any morphological characters might be unique to them, since those listed above apply equally well to other ploeotid taxa (e.g. *Olkasia*). Additional ultrastructural studies might be needed to further our understanding of any potential shared characters in this group like the structure of the feeding apparatus or pellicle strips. Past studies have shown that *Ploeotia*, *Serpenomonas*, and *Lentomonas* all appear to have a similar 'type II' feeding apparatus (Farmer & Triemer, 1994; Linton & Triemer, 1999). This might be observable with light microscopy—all of the taxa examined have a prominent hook-shaped feeding apparatus (Figure 1F). They all have bifurcations on their abutting pellicle joints, although the non-alistosan *Entosiphon* also has the same character (Cavalier-Smith et al., 2016; Chan et al., 2013; Farmer & Triemer, 1988, 1994; Triemer, 1986).

Our new *Decastava* cell has an SSU sequence that is 97.1% similar to the two existing *Decastava* SSU sequences from culture CCAP 1265/2. While this new cell

might represent a new species of *Decastava*, we think it is impossible to tell from sequence similarity alone, and our light microscopy imagery is not sufficient to see any potential morphological differences between our cell and the culture. Additionally, while all existing *Decastava* have been found in soil samples only, the sampling location for our *Decastava* was different than the other sequences (Canada vs. Scotland/Europe).

Hemiolia and *Liburna* form novel group Karavia

While a previous multigene analysis has established *Liburna* as a ploeotid taxon fitting somewhere in the backbone of the euglenid tree, its sister taxon *Hemiolia* has been absent from these analyses (Lax et al., 2021). SSU rRNA analyses show both taxa sometimes forming a clade (Lax et al., 2019; Lax & Simpson, 2020), which is in agreement with their shared morphology. Both *Hemiolia* and *Liburna* are fast gliders that use their posterior flagellum, and exhibit a characteristic sudden stop-and-resume motion (Lax et al., 2019). Now that single cell transcriptomic data for *Hemiolia* is available (cell PRC4), our multigene phylogeny perhaps unsurprisingly shows them both forming a strongly to fully supported clade in all our analyses (Figure 2; Figures S2 and S3).

As the *Hemiolia* described here appears to be unlike other described species from this genus, we propose a novel species *H. limna* to reflect the position of cells ML5 and ML7 in our analyses (cells ML5 and ML7 are co-assembled in our dataset). The new species branches with *H. trepidum* sequences in our SSU phylogenies (Figure S1), but *H. limna* is clearly distinct from any sequence attributed to *H. trepidum* and merits a new species of *Hemiolia* based on this, and the fact that *H. limna* is currently the only freshwater member of this group (see below).

Indeed, the tree now also shows the SSU rDNA sequences from *H. trepidum* are not monophyletic: *H. limna* ML5 and ML7 branch specifically with a maximally supported clade of several *H. trepidum* (Figure S1), to the exclusion of '*H. trepidum*' STS7. We propose that the currently identified *H. trepidum* represent multiple distinct species and further revisions will be required. Since *Hemiolia* and *Liburna* were both established mainly on their distinct phylogenetic position (Lax et al., 2019), morphological markers other than a distinct mode of motility are crucial in furthering our understanding of these genera—e.g. the fine structure of the pellicle strips, or feeding apparatus structure (Triemer & Farmer, 1991). Unfortunately, these electron microscopy-based studies are predicated on culturing the organism in question.

To reflect the highly supported branching of *Liburna* and *Hemiolia* in our multigene trees, and their morphological similarities, we propose a new name for this clade, Karavia.

Freshwater- and marine-dwelling Karavia

Hemiolia limna falls into the Karavia clade and is the first known member of this group that has been found in freshwater. Initial descriptions of the only other currently known members of this group, *Hemiolia trepidum* and *Liburna glaciale*, are from marine intertidal sediments (Larsen, 1987; Larsen & Patterson, 1990). The delineation between freshwater and marine boundaries and their transition rates amongst euglenids has received little attention to date. Transition rates between marine and non-marine environments vary between major eukaryotic groups appears to be quite variable: For example, cercozoans and fungi seem to have a higher number of transitions from marine to non-marine, whereas the opposite is true for dinoflagellates (Jamy et al., 2022). The same study unfortunately excluded any sequence data of Discoba due to poor phylogenetic resolution and ambiguous rate estimates (Jamy et al., 2022), so the following discussion is based on unquantified observations.

Curiously the phototrophic euglenid genera *Eutreptiella* and *Eutreptia* (order Eutreptiales) are predominantly marine, whereas the remainder (order Euglenales) have long been assumed to be predominantly freshwater (Kostygov et al., 2021; Marin et al., 2003). Using an environmental sequencing and metagenomics approach, a recent study has found indications of several low-abundance marine or brackish clades within Euglenales (Lukešová et al., 2020). This suggests that in Euglenophyceae, the transition from marine to freshwater seems to have occurred only rarely, the main event being from Eutreptiales to Euglenales. This is in contrast to phagotrophic euglenids where this event likely happened several times, if we assume the predecessor of euglenids to be marine. Most phagotrophic genera are in fact found in both freshwater and marine environments, like *Anisonema*, *Dinema*, *Jenningsia*, *Urceolus*, *Heteronema*, and *Peranema* (all part of Spirocuta). *Entosiphon* is currently the only phagotrophic clade that has been observed solely in freshwater or brackish environments (Kolisko et al., 2020; Triemer & Fritz, 1987). The distribution of freshwater taxa seems to be focussed more on spirocute taxa, with only *Entosiphon* and *Hemiolia limna* amongst ploeotids, and a large variety of petalomonads having been found in freshwater (Lee, 2022; Triemer & Fritz, 1987). Since ploeotids appear to harbor much of the phylogenetic diversity of euglenids, it is possible that additional freshwater-only clades might be found. Though it is noteworthy that the largest effort to examine freshwater euglenid diversity lists only five *Entosiphon* species and no other ploeotids as freshwater-dwelling (Huber-Pestalozzi, 1955). While we cannot know for sure without additional studies, it seems likely Petalomonadida is the richest in freshwater taxa amongst euglenids, but despite this,

no clear delineations of freshwater-only or marine-only clades have emerged in this group either (Huber-Pestalozzi, 1955; Leander et al., 2017), likely due to insufficient taxon sampling.

Gaulosia and *Chelandium*

A previous study has generated SSU rRNA data for 'Unidentified Ploeotid' cells SMS7 and CARR5, but did not describe any new formal taxa, since their placement in SSU phylogenies is unclear and no real affiliation could be determined (Lax et al., 2019). Cells *Gaulosia* BB9 and *Chelandium* BB14 are morphologically very similar and branch sister to SMS7 and CARR5, respectively, in our SSU rDNA phylogeny (Figure S1). Furthermore, our single cell transcriptomes enabled multigene phylogenetic analyses of these taxa.

Chelandium BB14 branches solidly as sister to the ploeotid *Olkasia*, and as part of Olkaspira (Figure 2). *Chelandium* and *Olkasia* seem to share few morphological characteristics, other than both appearing to glide on their posterior flagellum and waving the anterior flagellum around (Figure 1; Movie S2), both of which are traits of ploeotids in general (Lax et al., 2019). We could not observe any pellicle strips on our *Chelandium* isolate, potentially due to the limitations of the microscope used to isolate single cells (e.g. neither DIC optics nor 1000× magnification). *Olkasia* has large, well-defined pellicle strips (Lax et al., 2019). The refractive granules inside *Chelandium* are not found in *Olkasia* (Figure 1), and its posterior flagellum is much longer than that of *Olkasia*. Despite their strong phylogenetic grouping, we at this point refrain from formally describing a novel clade that includes *Olkasia* and *Chelandium*, chiefly because of their apparent morphological differences. We think *Chelandium* needs some further investigation with higher resolution light microscopy and possibly electron microscopy to warrant the establishment of a new group.

The previously reported 'Unidentified Ploeotid SMS7' cell is branching with full support with *Chelandium* BB14 in our SSU tree (Figure S1), and generally shares the same morphology. While SMS7 is half the size (16 × 8 μm) compared to BB14 (27 × 14.5 μm), both show the same gliding pattern on a very long thickened posterior flagellum and are heavily granulated. Owing to the size difference, difference in sampling locality (SMS7 was isolated from marine subtidal sediment in Curaçao), and their branch lengths, we consider SMS7 as *Chelandium*, but likely a different species. We thus propose to transfer 'Unidentified Ploeotid SMS7' to '*Chelandium* sp. SMS7'.

The phylogenetic position of *Gaulosia* on the other hand unfortunately remains unclear—its position as sister to Karavia is tentative at best in any of our analyses and does not always get recovered (Figure 2; Figure S2). Neither are there many morphological similarities between BB9 and Karavia. Both glide on their posterior

flagellum, but the typical fast gliding and stop-and-go pattern of both *Hemiolia* and *Liburna* are missing in *Gaulosia* (Larsen, 1987; Lax et al., 2019), it instead has a slower gliding pace. The protein sampling for our *Gaulosia* BB9 isolate is high (18 out of 19 proteins, with 87% of sites), so either increased taxon sampling for this part of the tree or additional and/or different phylogenetic marker genes are likely needed to fully resolve the position of this novel taxon.

The SSU rDNA sequence of cell 'Unidentified Ploeotid CARR5' could not be placed confidently amongst ploeotids in previous SSU analyses and was treated as an orphan ploeotid taxon without any known close relatives (Lax et al., 2019). Our *Gaulosia* BB9 SSU sequence forms a maximally supported clade with CARR5 and a short environmental sequence from Antarctic lake sediment (Figure S1). The branch lengths separating *Gaulosia* BB9 and CARR5 are not particularly large. In terms of morphology, both cells are more or less the same size (BB9 is 29 × 14.5 μm vs. CARR5 is 32 × 20 μm; Table 1), but more importantly they show the same number of easily visible pellicle strips and share the same pyriform shape. The posterior flagellum in CARR5 was reported to be 4.5× cell length, while BB9 only has a 2.2× cell length long flagellum (Lax et al., 2019). Like with *Chelandium*, based on the morphological similarities and phylogenetic affinities, we propose to transfer 'Unidentified Ploeotid CARR5' to *Gaulosia* and treat it as an undetermined species '*Gaulosia* sp. CARR5'.

BLASTing our *Gaulosia* BB9 sequence against GenBank's nr/nt database, we recovered a short environmental sequence AB510392. This sequence and our BB9 form clade in our SSU rDNA analysis with *Gaulosia* CARR5 falling sister to them both, together forming a maximally supported clade (Figure S1). This environmental sequence was generated from sediment from hypersaline lake Suribati in Antarctica and was classified as an embryophyte by SILVA. This demonstrates the need for a more comprehensive sampling of phagotrophic euglenids, considering there might be additional 'hidden' phagotrophic euglenid sequences in public databases that currently are not being classified correctly because the reference databases used underrepresent euglenids and other taxa with highly divergent SSU sequences (Forster et al., 2016).

Ploeotids harbor much of the phylogenetic diversity of phagotrophic euglenids

While the sampling efforts in a recent paper (Lax et al., 2021) and in this study have significantly furthered our understanding of the phylogenetic history of euglenids, some of euglenid diversity likely remains to be sampled, as the discovery of *Chelandium* and *Gaulosia* emphasizes. The backbone of the tree consists of Petalomonadida (currently the deepest branching

euglenids) and several ploeotid subgroups. We now are only starting to discover all the different taxa in this part of the tree, but are still some time away from understanding how they are related to each other.

Considering the apparent species richness of ploeotids (Lax et al., 2019), we can assume at least many of them should fall into some of the currently established clades, Alistosa (contains *Ploeotia vitrea*), Karavia (*Hemiolia* and *Liburna*), and the *Olkasia+Chelandium* clade. Members of Alistosa likely share certain morphological characteristics, like the hook-shaped feeding apparatus (Lax et al., 2019, 2021; Linton & Triemer, 1999). As such, we expect taxa with a similar hook-shaped feeding apparatus and 10 pellicle strips—like *Ploeotia adherens*, *P. discoides*, and *P. plana*—to be part of Alistosa, if molecular data will ever be collected for these species (Larsen & Patterson, 1990; Lee, 2012; Schroeckh et al., 2003). However, given the phylogenetic diversity of ploeotids, it is not unlikely that some undescribed taxa might not fall into established clades. A prominent example is 'Unidentified Ploeotid WF2_3' (Lax et al., 2019), for which SSU data exists but its placement in the phylogenies is unresolved, other than it likely being a ploeotid. Multigene phylogenetics would likely help determine its placement and help to further our understanding of the backbone of the euglenid tree. It is also not unlikely that new taxa will be found to fall into unexpected parts of the tree, like we now know *Chelandium* does. Its phylogenetic position to its closest relative *Olkasia* might seem surprising when examining cells with light microscopy only, but based on the tree, we now suspect scanning and transmission electron microscopy studies of both organisms might reveal shared morphological characters (e.g. structure of pellicle joints, of feeding apparatus structure). This underscores the need for deeper investigations of the ultrastructure of the euglenid groups that are currently being established.

One hypothesis as to why the backbone of the euglenid tree remains unresolved is that some key taxa/groups are missing in our current sampling, e.g. 'Unidentified Ploeotid WF2_3' or potentially other undescribed species that remain to be sampled. The hope is that continuous sampling of euglenids in general, and ploeotids in particular, will eventually resolve the tree. Fortunately, single-cell transcriptomics is well-suited for generating multigene-grade data of phagotrophic euglenids. Even our worst transcriptome reported here (*Chelandium granulatum* BB14 with 94.1% BUSCOs missing, compared to the *Euglena gracilis* GEFR01 transcriptome assembly with only 12.6% BUSCOs missing) still yielded 43% of sites in our multigene dataset (Table S4). By contrast, *Gaulosia striata* BB9 had 87% of multigene sites and 41.9% BUSCOs missing, yet could not be confidently placed in our trees (e.g. Figure 2), again hinting at the need to sample additional taxa. Another potential way forward is increasing the number of phylogenetic marker genes. Our study uses 19 markers, which resolves most

of the tree. However, most recent analyses of deeper branching relationships use upwards of 150, which could negate some systematic errors in the data if used with appropriately complex models (Gawryluk et al., 2019; Rodríguez-Ezpeleta et al., 2007; Sierra et al., 2022; Tice et al., 2021).

DATA AVAILABILITY STATEMENT

Raw transcriptome reads are deposited under GenBank BioProject PRJNA929644, SSU rDNA sequences under GenBank NCBI accessions OQ331009–OQ331015. Assembled transcriptomes and their predicted protein sequences, alignment files, tree files, additional micrographs, and multiQC summary can be found under DataDryad accession doi: 10.5061/dryad.w3r2280vv.

ORCID

Gordon Lax https://orcid.org/0000-0003-3986-3068

REFERENCES

Aberer, A.J., Krompass, D. & Stamatakis, A. (2013) Pruning rogue taxa improves phylogenetic accuracy: an efficient algorithm and Webservice. *Systematic Biology*, 62, 162–166.

Bolger, A.M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114–2120.

Bushmanova, E., Antipov, D., Lapidus, A. & Prjibelski, A.D. (2018) rnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. *bioRxiv*, 32, 1009–1015.

Busse, I., Patterson, D.J. & Preisfeld, A. (2003) Phylogeny of Phagotrophic Euglenids (Euglenozoa): a molecular approach based on culture material and environmental samples. *Journal of Phycology*, 39, 828–836.

Capella-Gutiérrez, S., Silla-Martínez, J.M. & Gabaldón, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25, 1972–1973.

Cavalier-Smith, T. (2016) Higher classification and phylogeny of Euglenozoa. *European Journal of Protistology*, 56, 250–276. Available from: https://doi.org/10.1016/j.ejop.2016.09.003

Cavalier-Smith, T. (2017) Euglenoid pellicle morphogenesis and evolution in light of comparative ultrastructure and trypanosomatid biology: semi-conservative microtubule/strip duplication, strip shaping and transformation. *European Journal of Protistology*, 61, 137–179. Available from: https://doi.org/10.1016/j.ejop.2017.09.002

Cavalier-Smith, T., Chao, E.E. & Vickerman, K. (2016) New phagotrophic euglenoid species (new genus Decastava; Scytomonas saepesedens; Entosiphon oblongum), Hsp90 introns, and

putative euglenoid Hsp90 pre-mRNA insertional editing. *European Journal of Protistology*, 56, 147–170. Available from: https://doi.org/10.1016/j.ejop.2016.08.002

Chan, Y.-F., Moestrup, Ø. & Chang, J. (2013) On Keelungia pulex nov. gen. Et nov. sp., a heterotrophic euglenoid flagellate that lacks pellicular plates (Euglenophyceae, Euglenida). *European Journal of Protistology*, 49, 15–31. Available from: https://doi.org/10.1016/j.ejop.2012.04.003

Criscuolo, A. & Gribaldo, S. (2010) BMGE (block mapping and gathering with entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evolutionary Biology*, 10, 210. Available from: https://doi.org/10.1186/1471-2148-10-210

Ekebom, J., Patterson, D.J. & Vørs, N. (1995) Heterotrophic flagellates from coral reef sediments (great barrier reef, Australia). *Archiv für Protistenkunde*, 146, 251–272. Available from: https://doi.org/10.1016/S0003-9365(96)80013-3

Ewels, P., Magnusson, M., Lundin, S. & Käller, M. (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. *Genome Biology and Evolution*, 32, 3047–3048.

Farmer, M.A. & Triemer, R.E. (1988) A Redescription of the genus Ploeotia Duj. (Euglenophyceae). *Taxon*, 37, 319–325. Available from: https://doi.org/10.2307/1222141

Farmer, M.A. & Triemer, R.E. (1994) An ultrastructural study of Lentomonas applanatum (Preisig) N. G. (Euglenida). *The Journal of Eukaryotic Microbiology*, 41, 112–119.

Forster, D., Dunthorn, M., Mahé, F., Dolan, J.R., Audic, S., Bass, D. et al. (2016) Benthic protists: the under-charted majority. *FEMS Microbiology Ecology*, 92, fiw120.

Gawryluk, R.M.R., Tikhonenkov, D.V., Hehenberger, E., Husnik, F., Mylnikov, A.P. & Keeling, P.J. (2019) Non-photosynthetic predators are sister to red algae. *Nature*, 572, 240–243. Available from: https://doi.org/10.1038/s41586-019-1398-6

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I. et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29, 644–652. Available from: https://doi.org/10.1038/nbt.1883

Huber-Pestalozzi, G. (1955) In: Thienemann, A. (Ed.) *Das Phytoplankton des Süßwassers: Systematik und Biologie*. Stuttgart: Schweizerbart'sche Verlagsbuchhandlung.

Jamy, M., Biwer, C., Vaulot, D., Obiol, A., Jing, H., Peura, S. et al. (2022) Global patterns and rates of habitat transitions across the eukaryotic tree of life. *Nature Ecology and Evolution*, 6, 1458–1470. Available from: https://doi.org/10.1038/s41559-022-01838-4

Katoh, K. & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772–780. Available from: https://doi.org/10.1093/molbev/mst010

Kolisko, M., Flegontova, O., Karnkowska, A., Lax, G., Maritz, J.M., Pánek, T. et al. (2020) EukRef-excavates: seven curated SSU ribosomal RNA gene databases. *Genome Biology and Evolution*, 2020, 1–11. Available from: https://doi.org/10.1093/database/baaa080/5996027

Kostygov, A.Y., Karnkowska, A., Votýpka, J., Tashyreva, D., Maciszewski, K., Yurchenko, V. et al. (2021) Euglenozoa: taxonomy, diversity and ecology, symbioses and viruses. *Open Biology*, 11, 200407. Available from: https://doi.org/10.1098/rsob.200407

Kozlov, A.M., Darriba, D., Flouri, T., Morel, B. & Stamatakis, A. (2019) RAxML-NG: a fast, scalable, and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, 35, 4453–4455. Available from: https://doi.org/10.1093/bioinformatics/btz305

Larsen, J. (1987) Algal studies of the Danish Wadden Sea. IV. A taxonomic study of the interstitial euglenoid flagellates. *Nordic Journal of Botany*, 7, 589–607. Available from: https://doi.org/10.1111/j.1756-1051.1987.tb02028.x

Larsen, J. & Patterson, D.J. (1990) Some flagellates (Protista) from tropical marine sediments. *Journal of Natural History*, 24, 801–937.

Lartillot, N., Lepage, T. & Blanquart, S. (2009) PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics*, 25, 2286–2288. Available from: https://doi.org/10.1093/bioinformatics/btp368

Lax, G., Kolisko, M., Eglit, Y., Lee, W.J., Yubuki, N., Karnkowska, A. et al. (2021) Multigene phylogenetics of euglenids based on single-cell transcriptomics of diverse phagotrophs. *Molecular Phylogenetics and Evolution*, 159, 107088. Available from: https://doi.org/10.1016/j.ympev.2021.107088

Lax, G., Lee, W.J., Eglit, Y. & Simpson, A. (2019) Ploeotids represent much of the phylogenetic diversity of euglenids. *Protist*, 170, 233–257. Available from: https://doi.org/10.1016/j.protis.2019.03.001

Lax, G. & Simpson, A.G.B. (2020) The molecular diversity of Phagotrophic Euglenids examined using single-cell methods. *Protist*, 171, 125757. Available from: https://doi.org/10.1016/j.protis.2020.125757

Leander, B.S. (2004) Did trypanosomatid parasites have photosynthetic ancestors? *Trends in Microbiology*, 12, 251–258.

Leander, B.S., Esson, H.J. & Breglia, S.A. (2007) Macroevolution of complex cytoskeletal systems in euglenids. *BioEssays*, 29, 987–1000. Available from: https://doi.org/10.1002/bies.20645

Leander, B.S., Lax, G., Karnkowska, A. & Simpson, A.G. (2017) Euglenida. In: Archibald, J.M., Simpson, A.G. & Slamovits, C.H. (Eds.) *Handbook of Protists*, Vol. 2, pp. 1047–1088. Cham: Springer Nature. Available from: https://doi.org/10.1007/978-3-319-28149-0_13

Lee, W.J. (2012) Free-living benthic heterotrophic euglenids from Botany Bay, Australia. *Marine Biology Research*, 8, 3–27. Available from: https://doi.org/10.1080/17451000.2011.596545

Lee, W.J. (2022) Three newly recorded heterotrophic euglenids (Protist), Entosiphon oblongum, Euglena longa and Keelungia pulex from South Korea. *Journal of Species Research*, 11, 89–93. Available from: https://doi.org/10.12651/JSR.2022.11.2.089

Lee, W.J. & Simpson, A.G.B. (2014a) Ultrastructure and molecular phylogenetic position of Neometanema parovale sp. nov. (Neometanema gen. Nov.), a marine Phagotrophic Euglenid with skidding motility. *Protist*, 165, 452–472. Available from: https://doi.org/10.1016/j.protis.2014.05.001

Lee, W.J. & Simpson, A.G.B. (2014b) Morphological and molecular characterisation of Notosolenus urceolatus Larsen and Patterson 1990, a member of an understudied deep-branching Euglenid group (Petalomonads). *The Journal of Eukaryotic Microbiology*, 61, 463–479. Available from: https://doi.org/10.1111/jeu.12126

Linton, E.W. & Triemer, R.E. (1999) Reconstruction of the feeding apparatus in Ploeotia costata (Euglenophyta) and its relationship to other euglenoid feeding apparatuses. *Journal of Phycology*, 35, 313–324. Available from: https://doi.org/10.1046/j.1529-8817.1999.3520313.x

Lukešová, S., Karlicki, M., Hadariová, L.T., Szabová, J., Karnkowska, A. & Hampl, V. (2020) Analyses of environmental sequences and two regions of chloroplast genomes revealed the presence of new clades of photosynthetic euglenids in marine environments. *Environmental Microbiology Reports*, 12, 78–91. Available from: https://doi.org/10.1111/1758-2229.12817

Manni, M., Berkeley, M.R., Seppey, M., Simão, F.A. & Zdobnov, E.M. (2021) BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology and Evolution*, 38, 4647–4654. Available from: https://doi.org/10.1093/molbev/msab199

Marin, B., Palm, A., Klingberg, M. & Melkonian, M. (2003) Phylogeny and taxonomic revision of plastid-containing Euglenophytes based on SSU rDNA sequence comparisons and Synapomorphic signatures in the SSU rRNA secondary structure. *Protist*, 154, 99–145.

Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D. & Gurevich, A. (2018) Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics*, 34, i142–i150. Available from: https://doi.org/10.1093/bioinformatics/bty266

Minh, B.Q., Nguyen, M.A.T. & Haeseler, A. (2013) Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution*, 30, 1188–1195. Available from: https://doi.org/10.1093/molbev/mst024

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Haeseler, A. et al. (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37, 1530–1534.

Nguyen, L.-T., Schmidt, H.A., Haeseler, A.v. & Minh, B.Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32, 268–274. Available from: https://doi.org/10.1093/molbev/msu300

Paerschke, S., Vollmer, A.H. & Preisfeld, A. (2017) Ultrastructural and immunocytochemical investigation of paramylon combined with new 18S rDNA-based secondary structure analysis clarifies phylogenetic affiliation of Entosiphon sulcatum (Euglenida: Euglenozoa). *Organisms, Diversity and Evolution*, 59, 1–12. Available from: https://doi.org/10.1007/s13127-017-0330-x

Picelli, S., Faridani, O.R., Björklund, Å.K., Winberg, G., Sagasser, S. & Sandberg, R. (2014) Full-length RNA-seq from single cells using smart-seq2. *Nature Protocols*, 9, 171–181. Available from: https://doi.org/10.1038/nprot.2014.006

Rodríguez-Ezpeleta, N., Brinkmann, H., Roure, B., Lartillot, N., Lang, B.F. & Philippe, H. (2007) Detecting and overcoming systematic errors in genome-scale phylogenies. *Systematic Biology*, 56, 389–399. Available from: https://doi.org/10.1080/10635150701397643

Ronquist, F., Teslenko, M., Mark, P.v.d., Ayres, D.L., Darling, A., Höhna, S. et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61, 539–542.

Schoenle, A., Živaljić, S., Prausse, D., Voss, J., Jakobsen, K. & Arndt, H. (2019) New phagotrophic euglenids from deep sea and surface waters of the Atlantic Ocean (Keelungia nitschei, Petalomonas acorensis, Ploeotia costaversata). *European Journal of Protistology*, 69, 102–116. Available from: https://doi.org/10.1016/j.ejop.2019.02.007

Schroeckh, S., Lee, W.J. & Patterson, D.J. (2003) Free-living heterotrophic euglenids from freshwater sites in mainland Australia. *Hydrobiologia*, 493, 131–166.

Sierra, R., Mauffrey, F., Cruz, J., Holzmann, M., Gooday, A.J., Maurer-Alcalá, X. et al. (2022) Taxon-rich transcriptomics supports higher-level phylogeny and major evolutionary trends in foraminifera. *Molecular Phylogenetics and Evolution*, 174, 107546. Available from: https://doi.org/10.1016/j.ympev.2022.107546

Song, L. & Florea, L. (2015) Rcorrector: efficient and accurate error correction for Illumina RNA-seq reads. *GigaScience*, 4, 1–8. Available from: https://doi.org/10.1186/s13742-015-0089-y

Tice, A.K., Žihala, D., Pánek, T., Jones, R.E., Salomaki, E.D., Nenarokov, S. et al. (2021) PhyloFisher: a phylogenomic package for resolving eukaryotic relationships. *PLoS Biology*, 19, e3001365. Available from: https://doi.org/10.1371/journal.pbio.3001365

Tikhonenkov, D.V., Gawryluk, R.M.R., Mylnikov, A.P. & Keeling, P.J. (2021) First finding of free-living representatives of Prokinetoplastina and their nuclear and mitochondrial genomes. *Scientific Reports*, 11, 1–21. Available from: https://doi.org/10.1038/s41598-021-82369-z

Triemer, R.E. (1986) Light and electron microscopic description of a colorless Euglenoid, Serpenomonas costata n. g., n. sp. *The Journal of Protozoology*, 33, 412–415. Available from: https://doi.org/10.1111/j.1550-7408.1986.tb05632.x/abstract

Triemer R. E. & Farmer M. A. 1991. The ultrastructural organization of the heterotrophic euglenids and its evolutionary implications. In: Larsen J., Patterson D. J., Association S. & Durham G. B. (Eds.), *The biology of free-living heterotrophic flagellates*. Oxford: Clarendon Press. pp. 185–204.

Triemer, R.E. & Fritz, L. (1987) Structure and operation of the feeding apparatus in a colorless Euglenoid, Entosiphon sulcatum. *The Journal of Protozoology*, 34, 39–47. Available from: https://doi.org/10.1111/j.1550-7408.1987.tb03129.x

Wang, H.-C., Minh, B.Q., Susko, E. & Roger, A.J. (2018) Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate Phylogenomic estimation. *Systematic Biology*, 67, 216–235.

Wideman, J.G., Lax, G., Leonard, G., Milner, D.S., Rodríguez-Martínez, R., Simpson, A.G.B. et al. (2019) A single-cell genome reveals diplonemid-like ancestry of kinetoplastid mitochondrial gene structure. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, 374, 20190100. Available from: https://doi.org/10.1098/rstb.2019.0100

Záhonová, K., Lax, G., Sinha, S.D., Leonard, G., Richards, T.A., Lukeš, J. et al. (2021) Single-cell genomics unveils a canonical origin of the diverse mitochondrial genomes of euglenozoans. *BMC Biology*, 19, 1–14. Available from: https://doi.org/10.1186/s12915-021-01035-y

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.