

Phylogenomic position of eupelagonemids, abundant, and diverse deep-ocean heterotrophs

Gordon Lax^{1,*}, Noriko Okamoto^{1,2}, Patrick J. Keeling¹

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

²Hakai Institute, Heriot Bay, BC, V0P 1H0, Canada

*Corresponding author: Gordon Lax, Department of Botany, University of British Columbia, 2212 Main Mall, Vancouver, BC, V6T 1Z4, Canada.

Email: gordon.lax@gmail.com

Abstract

Eupelagonemids, formerly known as Deep Sea Pelagic Diplonemids I (DSPD I), are among the most abundant and diverse heterotrophic protists in the deep ocean, but little else is known about their ecology, evolution, or biology in general. Originally recognized solely as a large clade of environmental ribosomal subunit RNA gene sequences (SSU rRNA), branching with a smaller sister group DSPD II, they were postulated to be diplomonads, a poorly studied branch of Euglenozoa. Although new diplomonads have been cultivated and studied in depth in recent years, the lack of cultured eupelagonemids has limited data to a handful of light micrographs, partial SSU rRNA gene sequences, a small number of genes from single amplified genomes, and only a single formally described species, *Eupelagonema oceanica*. To determine exactly where this clade goes in the tree of eukaryotes and begin to address the overall absence of biological information about this apparently ecologically important group, we conducted single-cell transcriptomics from two eupelagonemid cells. A SSU rRNA gene phylogeny shows that these two cells represent distinct subclades within eupelagonemids, each different from *E. oceanica*. Phylogenomic analysis based on a 125-gene matrix contrasts with the findings based on ecological survey data and shows eupelagonemids branch sister to the diplomonad subgroup Hemistasiidae.

Keywords: phylogenetics, protist, flagellate, multigene, pelagic

Eupelagonemids are among the most abundant and diverse eukaryotic heterotrophs in the ocean, particularly in the deep ocean, but completely lack the decades or centuries of study that provide other such lineages with biological context illuminating their role in the deep sea ecosystem. They were first recognized from environmental SSU rRNA gene sequences from deep pelagic water [1], and later determined to form two subgroups related to diplomonads (Deep Sea Pelagic Diplonemids, or DSPD I and DSPD II; [2]), arguably the least-studied subgroup of Euglenozoa. Recent research on other diplomonad groups has shown their distribution and ecology is more complex than previously assumed, including a freshwater radiation [3–5]. Likewise the importance of eupelagonemids was not widely recognized until the analysis of TARA-Oceans data showed that DSPD I in particular was extremely abundant and diverse in the deep pelagic ocean, rendering these organisms likely one of the most widespread on the planet [6, 7]. A previous study [8] reported the first light microscopical morphology of DSPD I cells, as well as limited molecular data based on single amplified genomes (SAGs). Subsequently, DSPD I was formally described as Eupelagonemidae, with one described species, *Eupelagonema oceanica* [9].

Despite their prevalence, our knowledge of eupelagonemids remains restricted to these scant data. No strains are in culture, and the vast majority of sequence data is limited to SSU rRNA gene surveys, which do not resolve the placement

of eupelagonemids (nor do available SAGs, which had few identifiable genes; [8]). Single-cell transcriptomes have proven a productive strategy to fill such gaps for other uncultivated protists [10–13]. Marine water was collected using a Niskin bottle at 300m depth off the central coast of British Columbia (Station KC10: Lat. 51.6505, Lon. –127.9516) on 3rd July 2022. Several single cells matching the overall description of eupelagonemids were collected and imaged (Supplementary Methods), of which two cells, Eupelagonemid 7 and Eupelagonemid 8, ultimately yielded transcriptomic libraries. Eupelagonemid 7 was oblong in shape, roughly 30µm long and 10µm wide, with an apical papilla protruding ~15 µm from the apical end of the cell (Fig. 1 and Supplementary Files). Eupelagonemid 8 was spindle-shaped, ~25µm long, and 15µm wide, with a round anterior and acute posterior end. When dying, Eupelagonemid 8 released cytoplasm from the round anterior end, suggesting that the cell might have a rigid cell wall with an apical opening. Eupelagonemid 8 has a yellow-brown spherical structure in the posterior region, reminiscent of a digestive vacuole.

Single-cell transcriptome libraries were generated using the Smart-Seq3 protocol [14], sequenced on a NextSeq 500 (Illumina) with 2 × 150 bp paired-end reads, and the data assembled and processed (see Supplementary Methods). Out of a public dataset of 240 phylogenetic marker genes, 54 and 36 orthologues were identified from Eupelagonemid 7 and 8, respectively. This is lower than often seen using these methods [10, 12], and another five

Received: 1 December 2023. Revised: 2 February 2024. Accepted: 6 March 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of the International Society for Microbial Ecology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

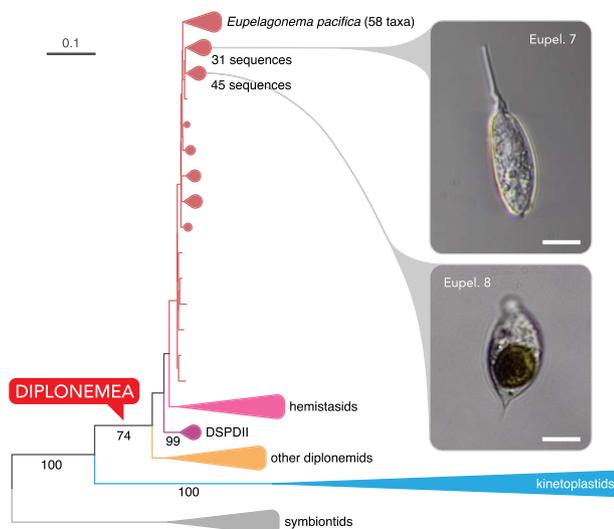


Figure 1. SSU rRNA gene phylogeny of eupelagonemids; ML phylogeny with a comprehensive sampling of diplomonads and two additional euglenozoan outgroups, estimated under GTR + gamma with 1000 non-parametric bootstrap replicates, and major clades are collapsed, for detailed phylogeny, see [Supplementary Fig. 1](#), see online supplementary material for a colour version of this figure, and circles on internodes denote full support (100%); micrographs show cells Eupelagonemid 7 and Eupelagonemid 8, both scale bars are 10 μm .

eupelagonemid cells we isolated yielded even fewer genes (≤ 10 ; data not shown). Taken with results from SAGs [8], this may indicate some intrinsic challenges to eupelagonemid genomics, but the current data nevertheless represent roughly a 10-fold increase in phylogenetic marker gene discovery for this group.

To see how Eupelagonemid 7 and 8 represent the diversity of the group, their evolutionary position was first tested within the breadth of full-length SSU rRNA gene sequences from other eupelagonemids. Sequences extracted from transcriptomes (Supplementary Methods) were added to an alignment of 237 eupelagonemids, and other euglenozoans and an unrooted maximum-likelihood (ML) phylogeny representing all major subdivisions of the group was inferred (Fig. 1). Eupelagonemids 7 and 8, and *E. oceanica* all fell within distinct subgroups in different parts of the tree. Eupelagonemids, Hemistasiidae, and DSPD II are all monophyletic (Fig. 1), but none of these groups, nor the branching order between them, is statistically supported. A similar topology was reported previously [15, 16], whereas other SSU rRNA gene analyses show hemistasiids as sister to eupelagonemids and DSPD II [17]. However, in none of these analyses, is the branching order resolved with statistical support, leaving the position of eupelagonemids both contradictory and unsupported.

To resolve this problem, we used the single-cell transcriptome assemblies to conduct multigene phylogenomic analyses. Amino acid sequences for all putative protein-coding genes were inferred from both assemblies, and these were added, along with an additional 12 other new euglenozoans (Supplementary Table 1) to PhyloFisher [18] to infer multigene trees. Ten published eupelagonemid SAGs [8] were initially included as well, but their coverage was too low to be informative. The resulting protein matrix with 9207 and 5856 amino acid sites (54 and 36 genes) for Eupelagonemid 7 and Eupelagonemid 8, respectively, was used to generate a 125-gene, 32780-site LG + C60 + F + G ML phylogeny of 33 taxa of euglenozoans and outgroups (Fig. 2). The two Eupelagonemids form a clade with full support, and with full

support branch sister to the hemistasiids. Although this conflicts with the topology of our SSU rRNA gene phylogeny (Fig. 1), it is congruent with some previous reports [18, 19], and none of the published SSU rRNA gene trees are supported. To robustly test our placement of eupelagonemids in the multigene framework, an additional phylogeny under LG + C60 + F + G using only the 19 genes that are available in both Eupelagonemid 7 and 8 (4913 sites total) resulted in the same topology but lower support for the monophyly of hemistasiids (93%; [Supplementary Fig. 2](#), see online supplementary material for a colour version of this figure.). An Approximately Unbiased (AU) test comparing this topology with an alternative found in some bootstrap trees (eupelagonemids branch within hemistasiids), rejected the latter in both datasets (see Supplementary Methods).

Hemistasiids are a much less diverse group with only three genera are known to date [17, 19] and environmental sequencing suggesting their diversity is a small fraction of that of eupelagonemids [7]. Hemistasiids are also marine, but are only known from surface waters, including coastal habitats, whereas eupelagonemids and DSPD II are predominantly found in deep marine water, and generally open ocean [7, 17]. The diversity of eupelagonemids might therefore represent a radiation from colonizing the deep sea.

Since their discovery, and even their recognition as a major component of the deep sea ecosystem [6, 7], the eupelagonemids have been stubbornly recalcitrant to study. The single-cell transcriptomes described here are the first representative genomic resources for any members of the group and have already redefined their phylogenetic position; placing them sister to another subgroup of diplomonads primarily associated with surface ocean waters. Even though this is an important step, additional data on their biological and ecological functions will require cultures of representative eupelagonemids, along with more depth and breadth of genetic data, and parallel advances for the even more elusive DSPD II group.

Acknowledgements

We thank the Hakai Institute and Hakai field technicians for support, especially Chris Mackenzie and Isabelle Desmarais for providing marine water samples. We also thank Dr Akinori Yabuki for sharing additional SSU rRNA gene sequences.

Supplementary material

[Supplementary material](#) is available at *The ISME Journal* online.

Conflicts of interest

The authors declare no competing interests.

Funding

This work was supported by grants to P.J.K. from the Hakai Institute, the Natural Sciences and Engineering Research Council of Canada (RGPIN-2014-03994) and the Gordon and Betty Moore Foundation (<https://doi.org/10.37807/GBMF9201>).

Data availability

Raw transcriptome reads are available under NCBI BioProject accession PRJNA1041876, SSU rRNA sequences under accessions OR831206 and OR831207. Assemblies, predicted proteomes, SSU

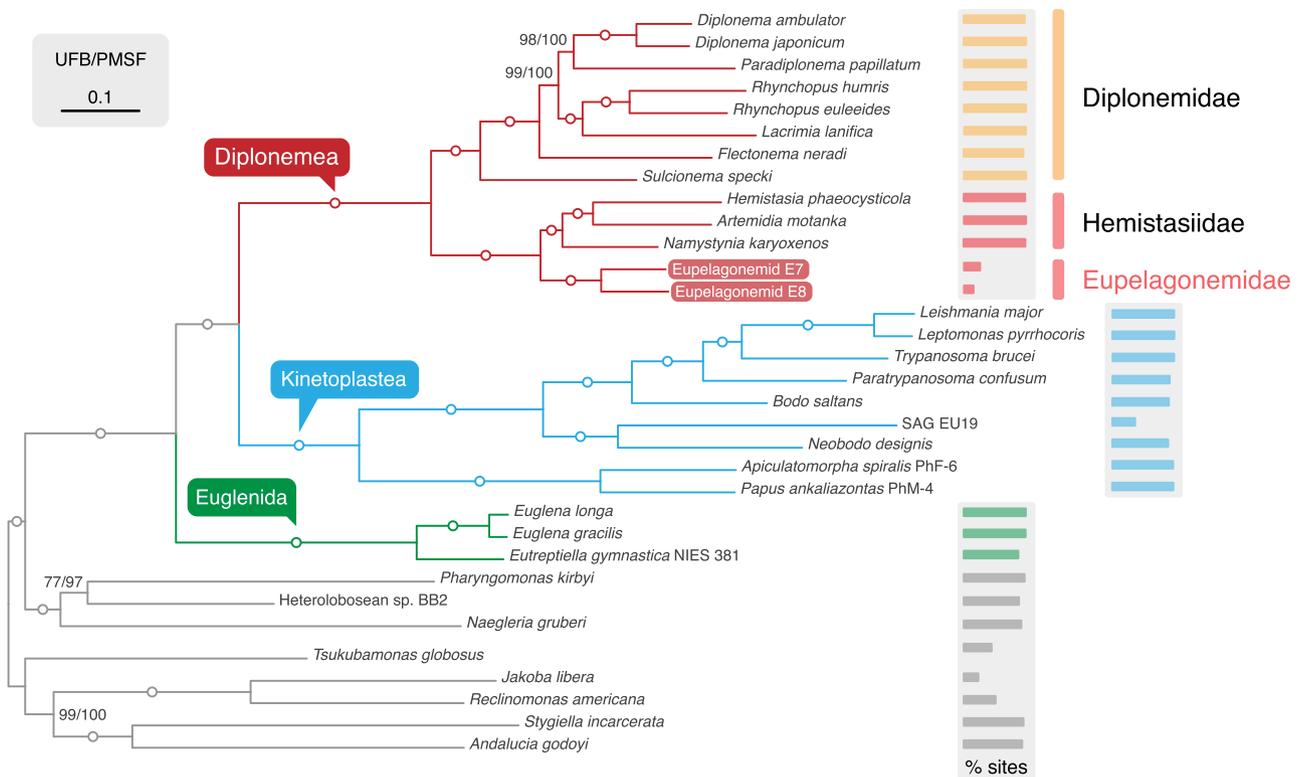


Figure 2. 125-gene phylogeny of euglenozoans; ML-phylogeny of 32 780 sites from 33 taxa, including all major subgroups of euglenozoa, except symbiontids (for which there are very limited data available), with heteroloboseans and jakobids as outgroups, and tree estimated under the LG + C60 + F + G model with 1000 ultrafast bootstraps, and 200 non-parametric bootstraps with LG + C60 + F + G + PMSF (PMSF), and circles on internodes denote full support in both analyses; the percentage of sites recovered from each taxon is represented by bars on the right.

rRNA gene alignment and tree data, and all multigene alignments and trees (single gene and concatenated), and PhyloFisher database folder are deposited under DataDryad accession <https://doi.org/10.5061/dryad.hqbzkh1pj>.

References

- López-García P, Rodríguez-Valera F, Pedrós-Alió C et al. Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* 2001;**409**:603–7. <https://doi.org/10.1038/35054537>
- Lara E, Moreira D, Vereshchaka A et al. Pan-oceanic distribution of new highly diverse clades of deep-sea diplomonads. *Environ Microbiol* 2009;**11**:47–55. <https://doi.org/10.1111/j.1462-2920.2008.01737.x>
- Mukherjee I, Salcher MM, Andrei A et al. A freshwater radiation of diplomonads. *Environ Microbiol* 2020;**22**:4658–68. <https://doi.org/10.1111/1462-2920.15209>
- Kostygov AY, Karnkowska A, Votýpka J et al. Euglenozoa: taxonomy, diversity and ecology, symbioses and viruses. *Open Biol* 2021;**11**:200407. <https://doi.org/10.1098/rsob.200407>
- Tashyreva D, Simpson AGB, Prokopchuk G et al. Diplonemids – a review on “new” flagellates on the oceanic block. *Protist* 2022;**173**:125868. <https://doi.org/10.1016/j.protis.2022.125868>
- de Vargas C, Audic S, Henry N et al. Eukaryotic plankton diversity in the sunlit ocean. *Science* 2015;**348**:1261605. <https://doi.org/10.1126/science.1261605>
- Flegontova O, Flegontov P, Malviya S et al. Extreme diversity of diplomonad eukaryotes in the ocean. *Curr Biol* 2016;**26**:3060–5. <https://doi.org/10.1016/j.cub.2016.09.031>
- Gawryluk RMR, del Campo J, Okamoto N et al. Morphological identification and single-cell genomics of marine diplomonads. *Curr Biol* 2016;**26**:3053–9. <https://doi.org/10.1016/j.cub.2016.09.013>
- Okamoto N, Gawryluk RMR, del Campo J et al. A revised taxonomy of diplomonads including the Eupelagonemidae n. fam. and a type species, *Eupelagonema oceanica* n. gen. & sp. *J Euk Micro* 2019;**66**:519–24. <https://doi.org/10.1111/jeu.12679>
- Lax G, Eglit Y, Eme L et al. Hemimastigophora is a novel supra-kingdom-level lineage of eukaryotes. *Nature* 2018;**564**:410–4. <https://doi.org/10.1038/s41586-018-0708-8>
- Mathur V, Wakeman KC, Keeling PJ. Parallel functional reduction in the mitochondria of apicomplexan parasites. *Curr Biol* 2021;**31**:2920–2928.e4. <https://doi.org/10.1016/j.cub.2021.04.028>
- Cooney EC, Holt CC, Jacko-Reynolds VKL et al. Photosystems in the eye-like organelles of heterotrophic warnowiid dinoflagellates. *Curr Biol* 2023;**33**:4252–4260.e3. <https://doi.org/10.1016/j.cub.2023.08.052>
- Holt CC, Hehenberger E, Tikhonenkov DV et al. Multiple parallel origins of parasitic marine Alveolates. *Nat Commun* 2023;**14**:7049. <https://doi.org/10.1038/s41467-023-42807-0>
- Hagemann-Jensen M, Ziegenhain C, Chen P et al. Single-cell RNA counting at allele and isoform resolution using smart-seq3. *Nat Biotechnol* 2020;**38**:708–14. <https://doi.org/10.1038/s41587-020-0497-0>
- Yabuki A, Kawato M, Nagano Y et al. New barcoding primers for the efficient monitoring of diplomonad diversity. *J Protistol* 2020;**52**:e003.
- Yabuki A, Kawato M, Nagano Y et al. Structural comparison of diplomonad communities around the Izu peninsula,

- Japan. *Microbes Environ* 2021;**36**. <https://doi.org/10.1264/jsme2.ME21012>
17. Prokopchuk G, Tashyreva D, Yabuki A et al. Morphological, ultra-structural, motility and evolutionary characterization of two new Hemistasiidae species. *Protist* 2019;**170**:259–82. <https://doi.org/10.1016/j.protis.2019.04.001>
 18. Tice AK, Žihala D, Pánek T et al. PhyloFisher: a phylogenomic package for resolving eukaryotic relationships. *PLoS Biol* 2021; **19**:e3001365. <https://doi.org/10.1371/journal.pbio.3001365>
 19. Yabuki A, Tame A. Phylogeny and reclassification of *Hemistasia phaeocysticola* (Scherffel) Elbrächter & Schnepf, 1996. *J Eukaryot Microbiol* 2014;**62**:426–9. <https://doi.org/10.1111/jeu.12191>