Rappemonads are haptophyte phytoplankton

Masanobu Kawachi, Takuro Nakayama, Motoki Kayama, Mami Nomura, Hideaki Miyashita, Othman Bojo, Lesley Rhodes, Stuart Sym, Richard N Pienaar, Ian Probert, et al.

Report

Current Biology

Rappemonads are haptophyte phytoplankton

Highlights

- A novel alga closely related to the enigmatic rappemonads was isolated
- It is cosmopolitan and seemingly an important contributor to global production
- Organellar genome analyses placed it as an independent branch in the Haptophyta
- A new haptophyte class, Rappephyceae, was erected based on its unique morphology

Authors

Masanobu Kawachi, Takuro Nakayama, Motoki Kayama, ..., Ian Probert, Isao Inouye, Ryoma Kamikawa

Correspondence

kamikawa.ryoma.7v@kyoto-u.ac.jp

In brief

Kawachi et al. find the phylogenetic home of the enigmatic environmental DNA clade, the rappemonads. Using phylogenetic analyses, an environmental DNA survey, and detailed morphological observation, a newly established algal strain was revealed to be widespread and the closest relative of rappemonads, in a new haptophyte class, the Rappephyceae.

Kawachi et al., 2021, Current Biology *31*, 2395–2403
June 7, 2021
https://doi.org/10.1016/j.cub.2021.03.012

CellPress

Current Biology

CellPress
OPEN ACCESS

Report

Rappemonads are haptophyte phytoplankton

Masanobu Kawachi,[1,11] Takuro Nakayama,[2,11] Motoki Kayama,[3] Mami Nomura,[4] Hideaki Miyashita,[3] Othman Bojo,[5] Lesley Rhodes,[6] Stuart Sym,[7] Richard N. Pienaar,[7] Ian Probert,[8] Isao Inouye,[9] and Ryoma Kamikawa[10,12,*]

[1]National Institute for Environmental Studies, Ibaraki 305-8506, Japan
[2]Graduate School of Life Sciences, Tohoku University, Sendai, Miyagi 980-8578, Japan
[3]Graduate School of Human and Environmental Studies, Kyoto University, Kyoto 606-8501, Japan
[4]Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8501, Japan
[5]Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia
[6]Cawthron Institute, 98 Halifax St East, PB 2, Nelson 7010, New Zealand
[7]School of Animal, Plant and Environmental Science, University of the Witwatersrand, 1 Jan Smuts Avenue, Braamfontein 2001, Johannesburg, South Africa
[8]Sorbonne University/CNRS, FR2424 Station Biologique de Roscoff, 29680 Roscoff, France
[9]Faculty of Life and Environmental Sciences, University of Tsukuba, Ibaraki 305-8501, Japan
[10]Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan
[11]These authors contributed equally
[12]Lead contact
*Correspondence: kamikawa.ryoma.7v@kyoto-u.ac.jp
https://doi.org/10.1016/j.cub.2021.03.012

SUMMARY

Rapidly accumulating genetic data from environmental sequencing approaches have revealed an extraordinary level of unsuspected diversity within marine phytoplankton,[1–11] which is responsible for around 50% of global net primary production.[12,13] However, the phenotypic identity of many of the organisms distinguished by environmental DNA sequences remains unclear. The rappemonads represent a plastid-bearing protistan lineage that to date has only been identified by environmental plastid 16S rRNA sequences.[14–17] The phenotypic identity of this group, which does not confidently cluster in any known algal clades in 16S rRNA phylogenetic reconstructions,[15] has remained unknown since the first report of environmental sequences over two decades ago. We show that rappemonads are closely related to a haptophyte microalga, *Pavlomulina ranunculiformis* gen. nov. et sp. nov., and belong to a new haptophyte class, the Rappephyceae. Organellar phylogenomic analyses provide strong evidence for the inclusion of this lineage within the Haptophyta as a sister group to the Prymnesiophyceae. Members of this new class have a cosmopolitan distribution in coastal and oceanic regions. The relative read abundance of Rappephyceae in a large environmental barcoding dataset was comparable to, or greater than, those of major haptophyte species, such as the bloom-forming *Gephyrocapsa huxleyi* and *Prymnesium parvum*, and this result indicates that they likely have a significant impact as primary producers. Detailed characterization of *Pavlomulina* allowed for reconstruction of the ancient evolutionary history of the Haptophyta, a group that is one of the most important components of extant marine phytoplankton communities.

RESULTS AND DISCUSSION

A new algal strain is sister to the rappemonads

The clonal algal strain NIES-3900 was established in 1991 from a sample obtained at a seaport in Japan. The cells of this strain with yellow-brown plastids (Figure 1A) exhibited a morphology resembling no other formally described alga, although an informal previous record does exist (Figure S1).[18] We established additional strains, CAWP21 from New Zealand and RCC3430 from Japan, with identical morphological features and almost identical 18S rRNA gene sequences. These three strains were recovered as a monophyletic clade in a maximum-likelihood (ML) tree of an 18S rRNA gene, inferred under the GTR+Γ+I model (Figure S2A). The haptophyte strain CG5 (EE-2014; HG970975.1) isolated from South Africa[19] and two environmental DNA sequences KJ762982.1 and JX841518.1 from the eastern North Pacific (33°13′30″N, 118°12′40″W)[20,21] are also grouped with the clade (Figure S2A), suggesting a broad distribution of the species. Hereafter, we use NIES-3900 as the representative strain of the lineage.

Our ML analysis of plastid 16S rRNA gene sequences with the GTR+Γ+I model and Bayesian analysis with the CAT-GTR+Γ model revealed that NIES-3900 is monophyletic with environmental DNA sequences of rappemonads (Figures 1B, S2B, and S2C). This monophyletic relationship is supported by a 97% ML bootstrap value (MLBV) and 0.99 PhyloBayes posterior probability (PPP). The NIES-3900/rappemonad clade is sister to the Haptophyta clade comprising the Prymnesiophyceae and the Pavlovophyceae, although this relationship is not supported by MLBVs and PPPs (Figures 1B, S2B, and S2C). Deep-branching patterns within the haptophytes were also unresolved in previous phylogenetic analyses.[14,15,22] The NIES-3900 sequence is not

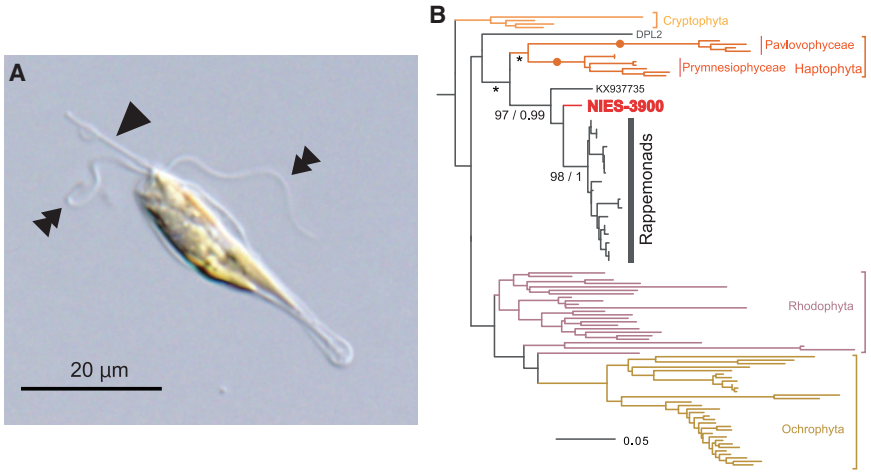

Figure 1. A novel alga closely related to the rappemonads

(A) A cell of the novel algal strain NIES-3900 bearing a haptonema (arrowhead) between two flagella (double arrowheads).

(B) 16S rRNA gene phylogeny of plastids in red algae (Rhodophyta) and red-alga-derived plastids. Maximum-likelihood bootstrap values (MLBVs; left) and PhyloBayes posterior probabilities (PPPs; right) are described only for the monophyly of Pavlovophyceae, Prymnesiophyceae, rappemonads, the NIES-3900/rappemonad clade, and for their relationships. The tree was rooted with Glaucophyta and Chloroplastida (see Figure S1). Asterisks show branches with MLBV <70% and PPP <0.9. Closed circles show branches with 100% MLBV and 1.0 PPP. Deep plastid lineage 2 (DPL2) is an environmental DNA sequence with unknown phylogenetic and taxonomic affinity reported in Choi et al.[22]

See also Figures S1 and S2.

completely identical to rappemonad sequences but rather has a close sister relationship. Another environmental sequence (GenBank accession number KX937735.1) that has been reported in Choi et al.[22] from the Atlantic Ocean is included in this clade but is distinct from NIES-3900 and the rappemonads, suggesting that this plastid-bearing lineage is more diverse than previously thought,[15] comprising three or more sub-lineages.

NIES-3900, a globally distributed cosmopolitan alga

Using the plastid 16S rRNA gene, Kim et al.[15] found rappemonads distributed in the North Pacific and North Atlantic, as well as UK coastal and freshwater samples. The lack of molecular data for this eukaryotic group other than plastid 16S rRNA gene sequences has precluded evaluation of the global distribution and relative abundance of rappemonads in the rapidly expanding databases of environmental 18S rRNA gene sequences.

To gain further insight into its ecological impact, a hyper-variable region of the nuclear 18S rRNA gene of NIES-3900 (V9 region, 134 bases long) was screened against the 18S V9 rRNA gene metabarcode data obtained during the *Tara* Oceans project.[1] This analysis detected two V9 sequences, one identical to the NIES-3900 sequence and the other different by one base (99.2% identity), from sampling sites in several oceans (inset, Figure 2A). On the basis of previous studies using the V9 barcode,[23] the two sequences with over 99% identity are likely to be derived from the same species, but in the absence of confirmation of this we treated them separately in this study. These sequences were detected in samples from 38 of the 46 *Tara* Oceans sampling stations (Figure 2A), and were relatively more abundant in the surface water layer (SRF) than in the deep chlorophyll maximum layer (DCM; Figure 2B). Although slightly less abundant than the V9 barcode of the well-known bloom-forming coccolithophorid haptophytes *Gephyrocapsa huxleyi*/*Gephyrocapsa oceanica*/*Gephyrocapsa parvula* (which shared an identical 18S V9 sequence and which accounted for an average of 0.0039% of the 18S V9 metabarcodes of the 46 *Tara* Oceans sampling stations), on average 0.0015% and 0.0020% of the metabarcodes corresponded to NIES-3900 and its near relative, respectively, in *Tara* Oceans samples (Figure 2C). In contrast, fewer V9 barcodes were detected from other bloom-forming haptophyte species such as *Prymnesium parvum* (0.0015%). The relative read abundance of NIES-3900 and its near relative in *Tara* Oceans samples was therefore comparable to, or greater than, that of major bloom-forming haptophyte species such as *Gephyrocapsa* spp. and *Prymnesium* spp. This indicates that this species likely has a significant impact on the microbial ecology of the marine environment as a primary producer, but this effect remains to be quantified.

Rappemonads are members of a deep-branching haptophyte lineage

Phylogenetic positions of the plastid and organismal lineages of NIES-3900 were estimated using organellar phylogenomic analyses, for which we retrieved the complete plastid genome and the nearly complete mitochondrial genome from the total DNA assembly data of NIES-3900 (see STAR Methods). The 83-plastid-protein dataset comprises 65 taxa and 16,234 sites and covers the diversity of red algae and red-alga-derived complex plastids excluding long-branching Alveolata sequences (Table S1). ML analysis of this dataset under the LG+Γ+F+C60-PMSF model and PhyloBayes analysis under the CAT-GTR+Γ model both reconstructed NIES-3900 as nested within the Haptophyta, and revealed that NIES-3900 is most closely related to the Prymnesiophyceae, with 92% MLBV and 0.99 PPP (Figures 3A, S2D, and S2E).

A similar tree topology to that derived from the plastid dataset was obtained in our analysis of the 14-mitochondrial-protein dataset containing 49 taxa and 3,366 sites, with slightly lower support values (83% MLBV and 0.9 PPP; Figures 3B, S2F, and S2G; Table S2). The reduced support values for this phylogeny might be partly due to the fact that fewer mitochondria-encoded proteins, which are used in phylogenetic analyses, are conserved in eukaryotes.[24,25] Even after removal of fast-evolving sites from the plastid and mitochondrial datasets, alternative relationships among Haptophyta lineages were not strongly, or even moderately, supported (Figures 3C and 3D). Congruence of these organellar phylogenomic analyses strongly suggests that NIES-3900 and the closely related rappemonads are members

CellPress
OPEN ACCESS

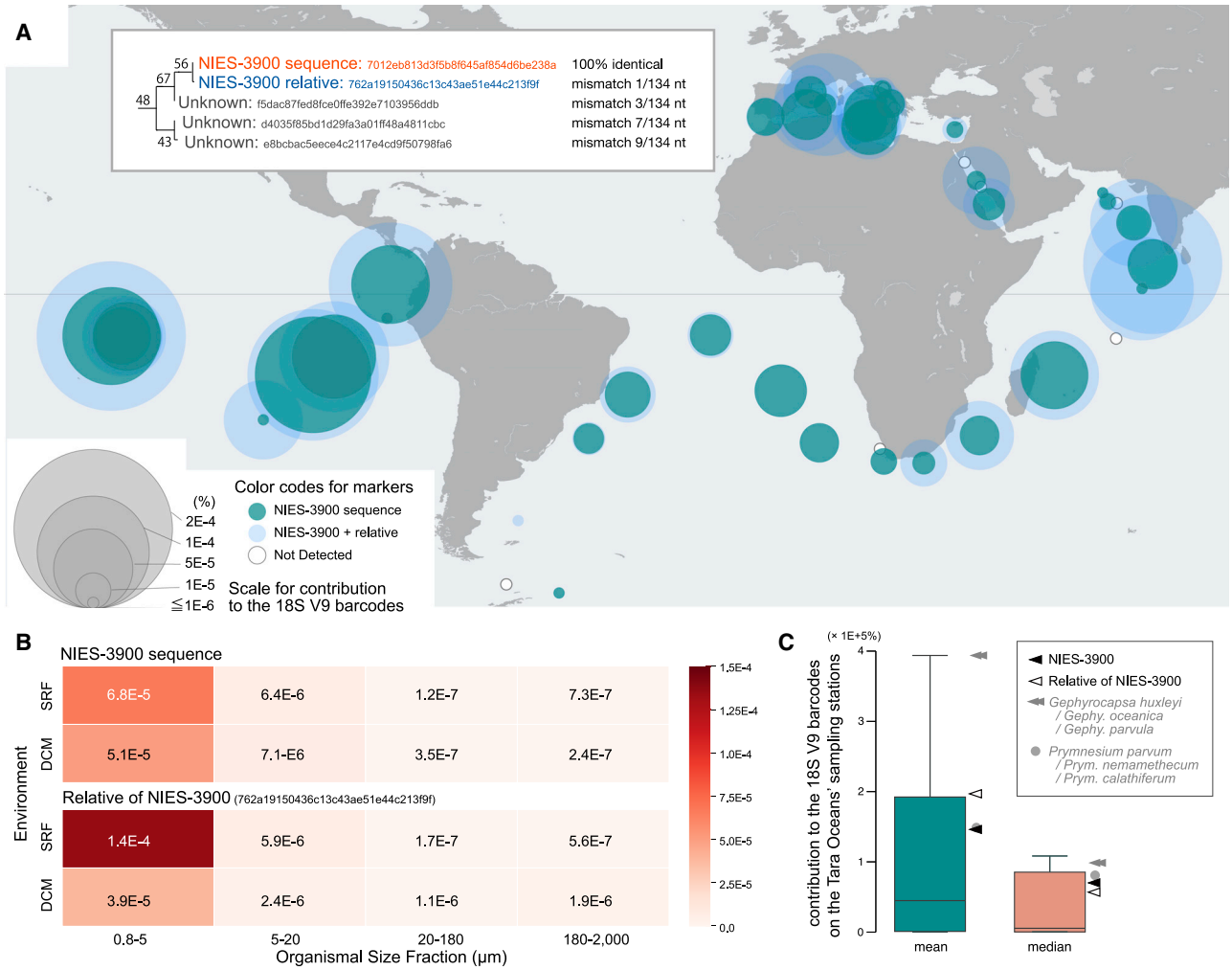

Figure 2. Distribution and abundances of the NIES-3900 lineage in natural environments, estimated based on the *Tara* Oceans project metabarcoding data

(A) Relative abundances of the 18S rRNA V9 sequences for NIES-3900 and its near relative (see inset) to total V9 sequences for this clade in each of 46 sampling stations. Green and pale blue circles show the relative abundance of the NIES-3900 sequence and of the sum of abundances for the two sequences (NIES-3900 and the near relative), respectively. Values for relative abundance (%) are described by *E* notation. Inset: a subtree of the ML tree, which shows the phylogenetic relationship between the NIES-3900 sequences and other barcode sequences, with similarity. The NIES-3900 sequence and the closest sequence (762a19150436c13c43ae51e44c213f9f) were found to differ by only one base and are regarded as near relatives of each other. The numbers above the branches show nonparametric bootstrap values.

(B) Relative abundances to total V9 sequences among the four size fractions and two environmental categories. Values for relative abundance (%) are described by *E* notation. SRF, surface water layer; DCM, deep chlorophyll maximum layer.

(C) Boxplot for mean and median abundances at 46 *Tara* Oceans project sampling stations for each of the 26 haptophyte species, including the two NIES-3900-related lineages. Values of representative species are indicated by arrowheads or circles. Outliers were omitted.

of the Haptophyta, and that their plastids were vertically inherited from the last common ancestor of this lineage. In light of the organismal phylogeny, it is highly likely that the NIES-3900 lineage is sister to the Prymnesiophyceae, with the Pavlovophyceae having diverged prior to the divergence of the Prymnesiophyceae and the NIES-3900/rappemonad lineage.

It is known that Cryptophyta and Haptophyta exclusively share a laterally transferred *rpl36* gene, called *rpl36*-c, distinct from the canonical plastid-type *rpl36*-p.[26,27] Supporting the phylogenetic position of NIES-3900 in Haptophyta, *rpl36* in NIES-3900 is of *rpl36*-c (Figure S3A). A comparison of the organellar genomes in Haptophyta indicates gene repertoires of NIES-3900 distinct from the other lineages of Haptophyta (Figures S3B–S3E), supporting its independent position. Nevertheless, the plastid genome of NIES-3900 shares more genes with the Prymnesiophyceae than the Pavlovophyceae (Figure S3C), which reinforces the sister relationship between these two lineages in phylogenetic analyses of plastid and mitochondrial proteins (Figures 3A and 3B).

Our above analyses indicate that the NIES-3900/rappemonad clade is a haptophyte lineage but that it is phylogenetically distinct from the two existing haptophyte classes, implying that

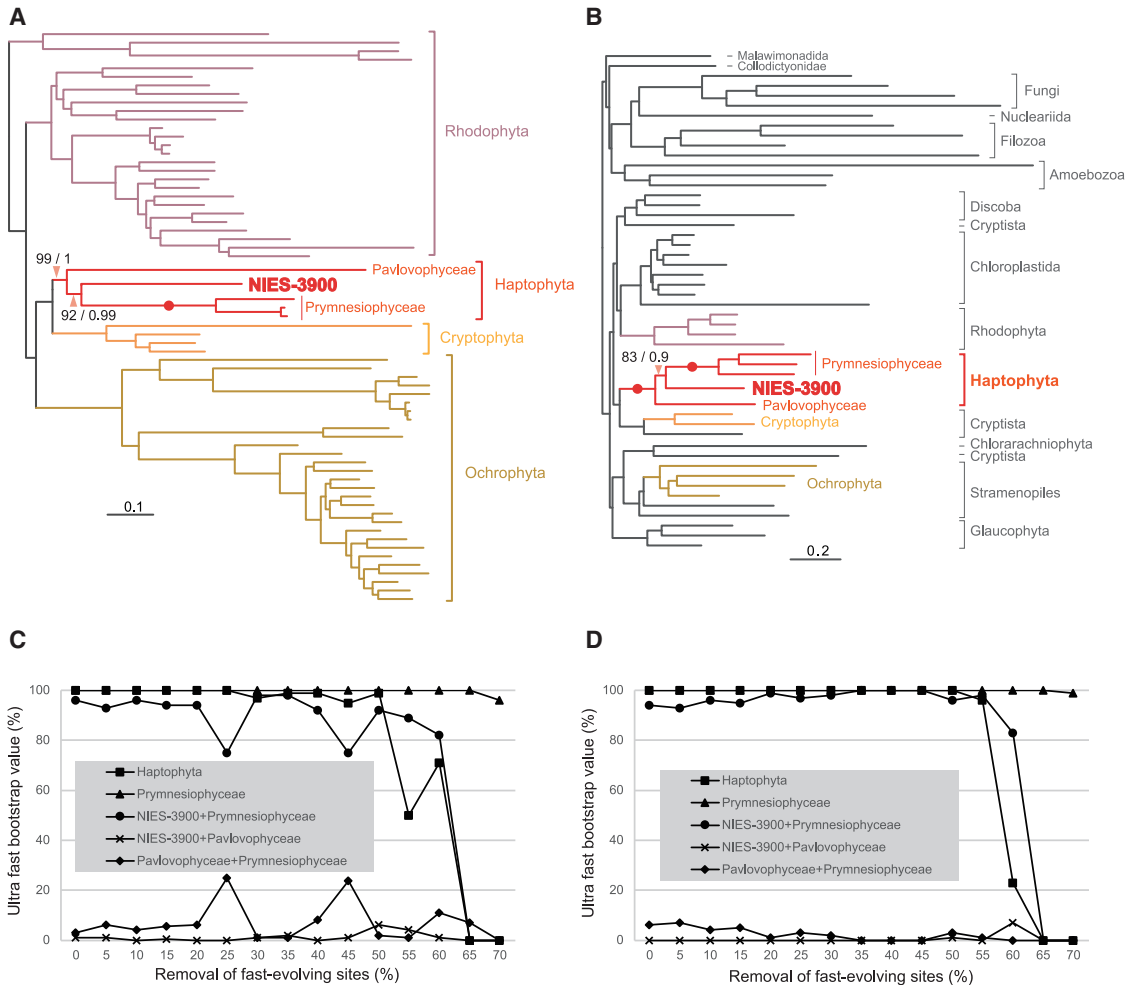

Figure 3. Organellar phylogenomic analyses

(A) ML tree inferred from the dataset of plastids in red algae (Rhodophyta) and from red-alga-derived plastids. The dataset comprises 65 taxa and 16,234 sites. MLBVs (left) and PPPs (right) are described only for those branches leading to the monophyly of Haptophyta, Prymnesiophyceae, and Prymnesiophyceae and NIES-3900. Closed circles show branches with 100% MLBV and 1.0 PPP.

(B) ML tree inferred from the mitochondrial dataset. The dataset is composed of 49 taxa and 3,366 sites. MLBVs (left) and PPPs (right) are described only for those branches leading to the monophyly of Haptophyta, Prymnesiophyceae, and Prymnesiophyceae and NIES-3900.

(C) Fluctuations in ultrafast bootstrap values for the monophyly of Haptophyta, the monophyly of Prymnesiophyceae, the monophyly of NIES-3900 and Prymnesiophyceae, and other alternative relationships as a function of the proportion of fast-evolving sites removed from the plastid dataset.

(D) Fluctuations of ultrafast bootstrap values for the monophyly of Haptophyta, Prymnesiophyceae, NIES-3900 and Prymnesiophyceae, and for other relationships, as a function of the proportion of fast-evolving sites removed from the mitochondrial dataset.

See also Figures S2 and S3 and Tables S1, S2, and S3.

these organisms are representatives of a novel class distinct from the Prymnesiophyceae and Pavlovophyceae. An additional analysis including sequences from plastid-bearing dinoflagellates with haptophyte-derived chloroplasts did not change this conclusion (Figure S2H; Table S3) because the haptophyte-derived plastid-bearing dinoflagellates were exclusively monophyletic with the Prymnesiophyceae, as shown in a previous study,[28,29] and the overall tree topology in the Haptophyta was not substantially different. This might suggest that a prymnesiophyte cell that emerged after the divergence from the NIES-3900/rappemonad lineage was engulfed by a dinoflagellate in a tertiary endosymbiotic event, giving rise to the haptophyte-derived plastid-bearing dinoflagellates, the Kareniaceae.[28,29] However, given the long branches of the Kareniaceae species in the tree (Figure S2H), the monophyletic clade of the Kareniaceae might be led by a long branch attraction artifact. It remains unclear whether their plastids have been derived from a single endosymbiotic event as discussed previously.[30]

Rappemonads, a new class of Haptophyta

Fluorescence *in situ* hybridization (FISH) analyses using oligonucleotide probes based on rappemonad 16S rRNA sequences revealed that cells were approximately 7 μm in length (i.e., in the nanoplankton size range) and contained two, three, or four plastids, with four being the most common number.[15] No further detailed morphological data have been obtained for this group.

CellPress
OPEN ACCESS

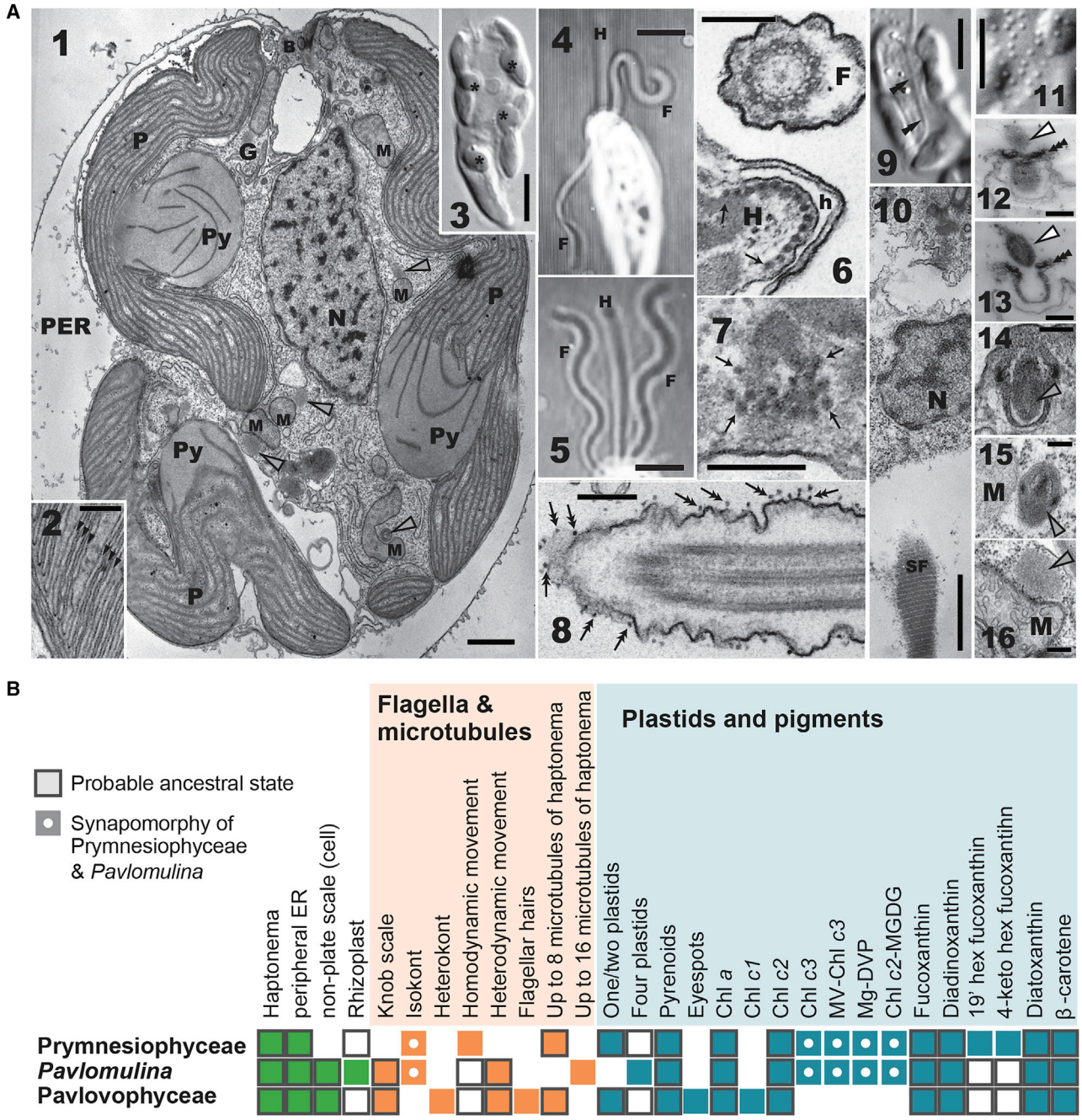

Figure 4. Morphological features

(A) Light and transmission electron microscopy of NIES-3900. (1) Longitudinal section of a whole cell; plastid (P) with a projecting pyrenoid (Py), nucleus (N), mitochondrion (M), basal body (B), Golgi body (G), and peripheral endoplasmic reticulum (PER). There are no girdle lamellae. Note electron-dense structures (gray arrowheads) near the mitochondrion. (2) A plastid with three thylakoid lamellae (arrowheads). (3) A fixed cell showing four chloroplasts each with a projecting pyrenoid (*), and a nucleus with a nucleolus between the two anterior chloroplasts. (4) A swimming cell showing the haptonema (H) and the heterodynamic motions of two flagella (F). (5) A swimming cell showing the haptonema (H) and the synchronized backward flagellar beating (F). (6) A section showing the flagellum (F) and the haptonematal base (H) with the haptonematal endoplasmic reticulum (h). Note that the 14 haptonematal microtubules are filled with electron-dense material (arrows). (7) A section showing components of the haptonematal base with 16 microtubules (arrows). (8) A section showing the knob scales on the flagellum, indicated by arrows with double heads. (9) The thick fibrous structure (double arrowheads) connecting the flagellar bases with the posterior end of the cell. (10) A section showing the striated fibrous structure (SF). (11) Small granules on the surface of the cell, representing the pot-shaped scale (PSS). (12 and 13) Serial sections of a PSS, showing two components, a vase-like structure with a collar (triple arrowheads) and a spherical lid-like structure (open arrowheads). Note the cell membrane is locally raised to wrap around the vase-like structure. (14) Electron-dense structure (gray arrowhead) resembling a PSS just before release from a cell invagination near the cell membrane. (15 and 16) Electron-dense structures (gray arrowhead) likely in the process of development of PSSs near the mitochondrion (M). Scale bars, 1 μm (1 and 10), 0.2 μm (2, 6–8, and 12–16), and 5 μm (3–5, 9, and 11).

(legend continued on next page)

Here we describe the morphological features of NIES-3900 as a representative of the new class including the rappemonads (Figure 4). Consistent with the organellar phylogenomic analyses, light microscopy observation identified NIES-3900 as a haptophyte because cells possess a conspicuous haptonema, a typical feature of the Haptophyta, reaching ca. 20 μm in length and extending in the direction of swimming (Figure 1A). The structure of the chloroplasts is typical of haptophytes, with three thylakoid lamellae and no girdle lamellae (Figure 4A, 1 and 2). A Golgi body with typical dilated cisternae is located between the nucleus and the basal bodies (Figure 4A, 1). NIES-3900 contains the typical chlorophyll *a*, chlorophyll *c2*, and fucoxanthin-based pigment content (Figure S4; Table S4) of haptophytes,[31] although there is no trace of 19′-hexanoyloxyfucoxanthin, which is often used as a marker pigment for haptophytes, but which is in fact absent in many species within this division.[31] Some of the other morphological features of NIES-3900 are unique, and have not previously been described in the Haptophyta.[32,33] Cells typically contain four plastids, each with a projected pyrenoid (Figure 4A, 3), whereas other haptophytes typically contain one or two plastids. Rappemonad cells notably also typically possess four plastids,[15] indicating that the number of plastids is a characteristic shared with NIES-3900, and likely a synapomorphy in the NIES-3900/rappemonad lineage. The two flagella are of equal length (isokont), ca. 35 μm, and, in forward swimming, exhibit heterodynamic motion, one extending to the posterior with a recurved flagellar beating motion and the other to the anterior exhibiting an S-shaped beating motion (Figures 1A and 4A, 4). In this respect, NIES-3900 is intermediate between the heterodynamic motion of the anisokont flagella of the Pavlovophyceae and the homodynamic motion of the isokont flagella of the Prymnesiophyceae. In backward swimming, the flagella reorientate directly anterior to the cell and exhibit homodynamic flagellar beating (Figure 4A, 5). The non-coiling haptonema of NIES-3900 comprises 14–16 microtubules filled with electron-dense material and the haptonematal endoplasmic reticulum (Figure 4A, 6 and 7). The number of microtubules is almost twice as many as observed in any other haptophyte. The flagella of NIES-3900 are covered with knob-like scales (Figure 4A, 8), a distinctive feature of the pavlovophyceans.[32] NIES-3900 also possesses a rhizoplast-like, striated fibrous structure (Figure 4A, 9 and 10) that has not been reported in other haptophytes. The cell surface is covered with small granules (Figure 4A, 11) that have extrusome-like features (Figure 4A, 12–16) but which are referred to here as "pot-shaped scales" (PSSs). The PSSs seem to be developed from an intracellular electron-dense structure originating around the mitochondrion and are released to the cell surface (Figure 4A, 14–16). These differ in structure and synthesis location from the knob scales present in pavlovophytes and the plate scales in prymnesiophytes.

On the basis of overall comparison with other haptophytes (Figure 4B), we propose herein the new class Rappephyceae comprising the new species described here, *Pavlomulina ranunculiformis* gen. et sp. nov., and the closely related rappemonads (see Taxonomic diagnoses in STAR Methods).

Evolutionary history of the haptophytes

The detailed characterization of *P. ranunculiformis* NIES-3900 and comparative analyses (Figure 4) provide insights into the ancient shift to new ecological niches in the haptophytes, a group that makes an important contribution to extant marine phytoplankton communities.[34] Given the tree topology for the three haptophyte classes (Figure 3), the Pavlovophyceae was the first to diverge from the other lineages. Pavlovophytes are known almost exclusively from coastal, brackish, and in some cases freshwater locations,[32] whereas many prymnesiophytes, notably the coccolithophores, are predominantly found in oceanic regions. *Pavlomulina* sequences have been found in several coastal locations around the world (Figure 2A) and environmental sequences of rappemonads have been reported from coastal and freshwater sites.[15] Our analyses indicate that the distribution of *Pavlomulina* is not restricted to these regions, being present and apparently abundant across oceans (Figure 2A). This tends to reinforce the view that ancestral haptophytes were predominantly coastal organisms, with colonization of oceanic ecosystems in lineages that diverged from the Pavlovophyceae. In the Haptophyta, some pigments have been previously observed exclusively in prymnesiophycean species, such as chlorophyll *c3*, chlorophyll *c2*-monogalactosyl diacylglyceride ester, monovinyl-chlorophyll *c3*, and Mg-2,4-divinyl phaeoporphyrin *a5* monomethyl ester (Figures 4B and S4), and they are also found in *P. ranunculiformis*, suggesting that these pigments existed prior to the divergence of the Prymnesiophyceae and the Rappephyceae. In *G. huxleyi* the quantity of these pigments is modified according to light intensity and quality,[35] suggesting that they play a role in adaptation to different ecological niches through photo-acclimation. The evolutionary shift of pigment composition in the ancestor of the Prymnesiophyceae and Rappephyceae was accompanied by morphological innovations, such as the evolution of isokont flagella and a prominent haptonema, which presumably also facilitated adaptation to more varied niches, including open-ocean conditions. However, to date there has been no report of blooms formed by the Rappephyceae although it is well known that several prymnesiophytes, including coccolithophorids such as *G. huxleyi*, form massive blooms in temperate and subpolar regions.[36] Coccolithophorids are covered by extracellular calcified coccoliths that may serve a protective function, and the intracellular biosynthesis thereof mitigates against a build-up of cytotoxic levels of Ca^{2+} ions.[37] In addition to the calcified coccoliths of coccolithophorids, many other metabolic, genetic, and physiological features underpin the success of the Prymnesiophyceae as oceanic primary producers.[36] The future characterization of such facets in the Rappephyceae and comparative analysis with the bloom-forming members of the Prymnesiophyceae would provide further insight into how this group of primary producers has radiated

(B) Features in three haptophyte lineages. Boxes with green, orange, and blue colors show characters observed in each of the Prymnesiophyceae, Pavlovophyceae, and *Pavlomulina*. Chl, chlorophyll; Chl *c2*-MGDG, chl *c2*-monogalactosyl diacylglyceride ester; ER, endoplasmic reticulum; hex, hexanoyl; MV-Chl *c3*, monovinyl-chl *c3*; Mg-DVP, Mg-2,4-divinyl phaeoporphyrin *a5* monomethyl ester.
See also Figure S4 and Table S4.

and, in some cases, come to dominate extant oceanic phytoplankton communities.

In this study, in addition to describing *P. ranunculiformis* gen. nov. et sp. nov., we demonstrate that the previously unidentified rappemonad environmental DNA clade belongs to a novel class of Haptophyta, the Rappephyceae, which is cosmopolitan and seemingly an important contributor to coastal and oceanic primary production. Because *P. ranunculiformis* is apparently distinct from the rappemonads in plastid 16S rRNA gene sequences (Figure 1B), it should be stressed that not all the phenotypic characteristics of *P. ranunculiformis* (Figures 1A and 4) are necessarily shared by rappemonads. In addition, there are still numerous unidentified environmental 18S and plastid 16S rRNA gene sequences thought to be derived from unknown, deep-branching haptophyte species[19,22,33,38,39] (e.g., Figure S2A), as in the majority of other phytoplankton lineages. Further efforts to relate nucleotide sequence information to biology, via culture isolation (as in this study) or cultivation-independent methods, are essential for improving the capacity of meta-genomic techniques to provide critical insights into the diversity, ecology, biogeochemical impact, and evolution of extant oceanic primary producers.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - *Pavlomulina ranunculiformis*
  - Taxonomic diagnoses
- METHOD DETAILS
  - DNA-sequencing, assembly, and organellar genome annotation
  - Organellar phylogenomics
  - Phylogenetic analyses of single genes
  - Distribution of the NIES-3900 lineage based on Tara Oceans metabarcoding data
  - HPLC analysis
  - Microscopy
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2021.03.012.

ACKNOWLEDGMENTS

We thank Azusa Itoh (Kyoto University) for his technical support in this project. The original phylogenomic datasets for plastids and mitochondria were kindly provided by Dr. Sergio A. Muñoz-Gómez and Dr. Chris Jackson, respectively. This work was supported in part by Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (B) (awarded to R.K. [19H03274] and T.N. [20H03305]), by the National BioResource Project (NBRP) from the Japan Agency for Medical Research and Development (AMED) (to Masanobu Kawachi), and by the French National Research Agency (ANR) PhenoMap project (to I.P.). We thank the *Tara* Oceans consortium and sponsors who supported the *Tara* Oceans Expedition for making the data accessible. We also thank the MBI and MCC for maintaining the NIES-3900 strain, and Dr. Jahn Throndsen for allowing us to include his drawing.

AUTHOR CONTRIBUTIONS

R.K. and Masanobu Kawachi conceived this research. R.K. and T.N. analyzed the molecular data. Masanobu Kawachi, M.N., O.B., L.R., S.S., R.N.P., I.P., I.I., and R.K. obtained the morphological data of the cells. Motoki Kayama, H.M., and R.K. analyzed the pigment compositions. R.K., T.N., Masanobu Kawachi, S.S., and I.P. wrote the manuscript. All of the authors commented on the first draft and approved the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 11, 2021
Revised: February 12, 2021
Accepted: March 3, 2021
Published: March 26, 2021

REFERENCES

1. de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., et al.; *Tara* Oceans Coordinators (2015). Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. Science *348*, 1261605.
2. Massana, R., del Campo, J., Sieracki, M.E., Audic, S., and Logares, R. (2014). Exploring the uncultured microeukaryote majority in the oceans: reevaluation of ribogroups within stramenopiles. ISME J. *8*, 854–866.
3. Edvardsen, B., Egge, E.S., and Vaulot, D. (2016). Diversity and distribution of haptophytes revealed by environmental sequencing and metabarcoding—a review. Perspect. Phycol. *3*, 77–91.
4. Flegontova, O., Flegontov, P., Malviya, S., Audic, S., Wincker, P., de Vargas, C., Bowler, C., Lukeš, J., and Horák, A. (2016). Extreme diversity of diplonemid eukaryotes in the ocean. Curr. Biol. *26*, 3060–3065.
5. Lopes Dos Santos, A., Pollina, T., Gourvil, P., Corre, E., Marie, D., Garrido, J.L., Rodríguez, F., Noël, M.H., Vaulot, D., and Eikrem, W. (2017). Chloropicophyceae, a new class of picophytoplanktonic prasinophytes. Sci. Rep. *7*, 14019.
6. Leblanc, K., Quéguiner, B., Diaz, F., Cornet, V., Michel-Rodriguez, M., Durrieu de Madron, X., Bowler, C., Malviya, S., Thyssen, M., Grégori, G., et al. (2018). Nanoplanktonic diatoms are globally overlooked but play a role in spring blooms and carbon export. Nat. Commun. *9*, 953.
7. Del Campo, J., Heger, T.J., Rodríguez-Martínez, R., Worden, A.Z., Richards, T.A., Massana, R., and Keeling, P.J. (2019). Assessing the diversity and distribution of apicomplexans in host and free-living environments using high-throughput amplicon data and a phylogenetically informed reference framework. Front. Microbiol. *10*, 2373.
8. Ratnasingham, S., and Hebert, P.D. (2007). bold: The Barcode of Life Data System (http://www.barcodinglife.org). Mol. Ecol. Notes *7*, 355–364.
9. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. *41*, D590–D596.
10. Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., Decelle, J., et al. (2013). The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. Nucleic Acids Res. *41*, D597–D604.
11. Decelle, J., Romac, S., Stern, R.F., Bendif, M., Zingone, A., Audic, S., Guiry, M.D., Guillou, L., Tessier, D., Le Gall, F., et al. (2015). PhytoREF: a

reference database of the plastidial 16S rRNA gene of photosynthetic eukaryotes with curated taxonomy. Mol. Ecol. Resour. *15*, 1435–1445.

12. Field, C.B., Behrenfeld, M.J., Randerson, J.T., and Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. Science *281*, 237–240.
13. Falkowski, P.G., Barber, R.T., and Smetacek, V. (1998). Biogeochemical controls and feedbacks on ocean primary production. Science *281*, 200–207.
14. Rappé, M.S., Suzuki, M.T., Vergin, K.L., and Giovannoni, S.J. (1998). Phylogenetic diversity of ultraplankton plastid small-subunit rRNA genes recovered in environmental nucleic acid samples from the Pacific and Atlantic coasts of the United States. Appl. Environ. Microbiol. *64*, 294–303.
15. Kim, E., Harrison, J.W., Sudek, S., Jones, M.D., Wilcox, H.M., Richards, T.A., Worden, A.Z., and Archibald, J.M. (2011). Newly identified and diverse plastid-bearing branch on the eukaryotic tree of life. Proc. Natl. Acad. Sci. USA *108*, 1496–1500.
16. Archibald, J.M. (2015). Genomic perspectives on the birth and spread of plastids. Proc. Natl. Acad. Sci. USA *112*, 10147–10153.
17. Sibbald, S.J., and Archibald, J.M. (2020). Genomic insights into plastid evolution. Genome Biol. Evol. *12*, 978–990.
18. Throndsen, J. (1983). Ultra- and nanoplankton flagellates from coastal waters of southern Honshu and Kyushu, Japan. In Working Party on Taxonomy in the Akashiwo Mondai Kenkyukai, M. Chihara, and I. Haruhik, eds. (Japan Fisheries Agency), pp. 1–62.
19. Egge, E.S., Eikrem, W., and Edvardsen, B. (2015). Deep-branching novel lineages and high diversity of haptophytes in the Skagerrak (Norway) uncovered by 454 pyrosequencing. J. Eukaryot. Microbiol. *62*, 121–140.
20. Kim, D.Y., Countway, P.D., Yamashita, W., and Caron, D.A. (2012). A combined sequence-based and fragment-based characterization of microbial eukaryote assemblages provides taxonomic context for the terminal restriction fragment length polymorphism (T-RFLP) method. J. Microbiol. Methods *91*, 527–536.
21. Lie, A.A., Liu, Z., Hu, S.K., Jones, A.C., Kim, D.Y., Countway, P.D., Amaral-Zettler, L.A., Cary, S.C., Sherr, E.B., Sherr, B.F., et al. (2014). Investigating microbial eukaryotic diversity from a global census: insights from a comparison of pyrotag and full-length sequences of 18S rRNA genes. Appl. Environ. Microbiol. *80*, 4363–4373.
22. Choi, C.J., Bachy, C., Jaeger, G.S., Poirier, C., Sudek, L., Sarma, V.V.S.S., Mahadevan, A., Giovannoni, S.J., and Worden, A.Z. (2017). Newly discovered deep-branching marine plastid lineages are numerically rare but globally distributed. Curr. Biol. *27*, R15–R16.
23. Tragin, M., Zingone, A., and Vaulot, D. (2018). Comparison of coastal phytoplankton composition estimated from the V4 and V9 regions of the 18S rRNA gene with a focus on photosynthetic groups and especially Chlorophyta. Environ. Microbiol. *20*, 506–520.
24. Roger, A.J., Muñoz-Gómez, S.A., and Kamikawa, R. (2017). The origin and diversification of mitochondria. Curr. Biol. *27*, R1177–R1192.
25. Gray, M.W., Burger, G., and Lang, B.F. (1999). Mitochondrial evolution. Science *283*, 1476–1481.
26. Rice, D.W., and Palmer, J.D. (2006). An exceptional horizontal gene transfer in plastids: gene replacement by a distant bacterial paralog and evidence that haptophyte and cryptophyte plastids are sisters. BMC Biol. *4*, 31.
27. Hovde, B.T., Starkenburg, S.R., Hunsperger, H.M., Mercer, L.D., Deodato, C.R., Jha, R.K., Chertkov, O., Monnat, R.J., Jr., and Cattolico, R.A. (2014). The mitochondrial and chloroplast genomes of the haptophyte *Chrysochromulina tobin* contain unique repeat structures and gene profiles. BMC Genomics *15*, 604.
28. Tengs, T., Dahlberg, O.J., Shalchian-Tabrizi, K., Klaveness, D., Rudi, K., Delwiche, C.F., and Jakobsen, K.S. (2000). Phylogenetic analyses indicate that the 19′hexanoyloxy-fucoxanthin-containing dinoflagellates have tertiary plastids of haptophyte origin. Mol. Biol. Evol. *17*, 718–729.
29. Gabrielsen, T.M., Minge, M.A., Espelund, M., Tooming-Klunderud, A., Patil, V., Nederbragt, A.J., Otis, C., Turmel, M., Shalchian-Tabrizi, K., Lemieux, C., and Jakobsen, K.S. (2011). Genome evolution of a tertiary dinoflagellate plastid. PLoS ONE *6*, e19132.
30. Hehenberger, E., Gast, R.J., and Keeling, P.J. (2019). A kleptoplastidic dinoflagellate and the tipping point between transient and fully integrated plastid endosymbiosis. Proc. Natl. Acad. Sci. USA *116*, 17934–17942.
31. Zapata, M., Jeffrey, S.W., Wright, S.W., Rodriguez, F., Garrido, J.L., and Clementson, J.A. (2004). Photosynthetic pigments in 37 species (65 strains) of Haptophyta: implications for oceanography and chemotaxonomy. Mar. Ecol. Prog. Ser. *270*, 83–102.
32. Bendif, M., Probert, I., Hervé, A., Billard, C., Goux, D., Lelong, C., Cadoret, J.P., and Véron, B. (2011). Integrative taxonomy of the Pavlovophyceae (Haptophyta): a reassessment. Protist *162*, 738–761.
33. Edvardsen, B., Eikrem, W., Green, J.C., Andersen, R.A., Moon-van der Staay, S., and Medlin, L.K. (2000). Phylogenetic reconstructions of the Haptophyta inferred from 18S ribosomal DNA sequences and available morphological data. Phycologia *39*, 19–35.
34. Eikrem, E., Medlin, L.K., Henderiks, J., Rokitta, S., Rost, B., Probert, I., Throndsen, J., and Edvardsen, B. (2017). Haptophyta. In Handbook of the Protists, Second Edition, *Volume 2*, J.M. Archibald, A.G.B. Simpson, and C.H. Slamovits, eds. (Springer International Publishing), pp. 893–954.
35. Garrido, J.L., Brunet, C., and Rodríguez, F. (2016). Pigment variations in *Emiliania huxleyi* (CCMP370) as a response to changes in light intensity or quality. Environ. Microbiol. *18*, 4412–4425.
36. Taylor, A.R., Brownlee, C., and Wheeler, G. (2017). Coccolithophore cell biology: chalking up progress. Annu. Rev. Mar. Sci. *9*, 283–310.
37. Müller, M.N., Barcelos e Ramos, J., Schulz, K.G., Riebesell, U., Kázmierczak, J., Gallo, F., Mackinder, L., Li, L., Nesterenko, P.N., Trull, T.W., and Hallegraeff, G.M. (2015). Phytoplankton calcification as an effective mechanism to alleviate cellular calcium poisoning. Biogeosciences *12*, 6493–6501.
38. Shi, X.L., Marie, D., Jardillier, L., Scanlan, D.J., and Vaulot, D. (2009). Groups without cultured representatives dominate eukaryotic picophytoplankton in the oligotrophic south east Pacific Ocean. PLoS ONE *4*, e7657.
39. Shalchian-Tabrizi, K., Reier-Røberg, K., Ree, D.K., Klaveness, D., and Bråte, J. (2011). Marine-freshwater colonizations of haptophytes inferred from phylogeny of environmental 18S rDNA sequences. J. Eukaryot. Microbiol. *58*, 315–318.
40. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. *17*, 10–12.
41. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.
42. Zerbino, D.R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. *18*, 821–829.
43. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780.
44. Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. *17*, 540–552.
45. Nguyen, L.T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. *32*, 268–274.
46. Lartillot, N., Lepage, T., and Blanquart, S. (2009). PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics *25*, 2286–2288.
47. Valach, M., Burger, G., Gray, M.W., and Lang, B.F. (2014). Widespread occurrence of organelle genome-encoded 5S rRNAs including permuted molecules. Nucleic Acids Res. *42*, 13764–13777.
48. Lowe, T.M., and Eddy, S.R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. *25*, 955–964.
49. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

50. Pesant, S., Not, F., Picheral, M., Kandels-Lewis, S., Le Bescot, N., Gorsky, G., Iudicone, D., Karsenti, E., Speich, S., Troublé, R., et al.; *Tara* Oceans Consortium Coordinators (2015). Open science resources for the discovery and analysis of *Tara* Oceans data. Sci. Data *2*, 150023.

51. Watanabe, M.M., Kasai, F., and Sudo, R. (1988). NIES-Collection. List of Strains, Second Edition, Microalgae and Protozoa (National Institute for Environmental Studies).

52. Bojo, O. (2002). Systematic studies of New Zealand nanoflagellates with a special reference to members of the Haptophyta. PhD thesis (University of Otago).

53. Rhodes, L., Edwards, A.R., Bojo, O., and Chang, F.H. (2011). Phylum Haptophyta. In New Zealand Bioinventory of Biodiversity, *Volume 3*, D.P. Gordon, ed. (Canterbury University Press), pp. 312–321.

54. Nishimura, Y., Kamikawa, R., Hashimoto, T., and Inagaki, Y. (2014). An intronic open reading frame was released from one of group II introns in the mitochondrial genome of the haptophyte *Chrysochromulina* sp. NIES-1333. Mob. Genet. Elements *4*, e29384.

55. Muñoz-Gómez, S.A., Mejía-Franco, F.G., Durnin, K., Colp, M., Grisdale, C.J., Archibald, J.M., and Slamovits, C.H. (2017). The new red algal subphylum Proteorhodophytina comprises the largest and most divergent plastid genomes known. Curr. Biol. *27*, 1677–1684.e4.

56. Jackson, C.J., and 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*, 2774–2785.

57. Klinger, C.M., Paoli, L., Newby, R.J., Wang, M.Y., Carroll, H.D., Leblond, J.D., Howe, C.J., Dacks, J.B., Bowler, C., Cahoon, A.B., et al. (2018). Plastid transcript editing across dinoflagellate lineages shows lineage-specific application but conserved trends. Genome Biol. Evol. *10*, 1019–1038.

58. Bendif, E.M., Nevado, B., Wong, E.L.Y., Hagino, K., Probert, I., Young, J.R., Rickaby, R.E.M., and Filatov, D.A. (2019). Repeated species radiations in the recent evolution of the key marine phytoplankton lineage *Gephyrocapsa*. Nat. Commun. *10*, 4234.

59. Zapata, M., Rodríguez, F., and Garrido, J.L. (2000). Separation of chlorophylls and carotenoids from marine phytoplankton: a new HPLC method using a reversed phase C_8 column and pyridine-containing mobile phases. Mar. Ecol. Prog. Ser. *195*, 29–45.

60. Spurr, A.R. (1969). A low-viscosity epoxy resin embedding medium for electron microscopy. J. Ultrastruct. Res. *26*, 31–43.

61. Reynolds, E.S. (1963). The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. J. Cell Biol. *17*, 208–212.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical commercial assays		
TruSeq Nano DNA Library Prep Kit	Illumina	Cat#20015965
Plant DNA Preparation - Solution Kit	Jena Bioscience	Cat#PP-207S
Deposited data		
Plastid genome	This paper	DDBJ LC564891
Mitochondrial genome	This paper	DDBJ LC564892- LC564893
18S rRNA genes	This paper	DDBJ LC599498, LC603170, and LC603171
Phylogenetic datasets	This paper	https://doi.org/10.5061/dryad.x0k6djhhw
Experimental models: Organisms/strains		
Pavlomulina ranunculiformis NIES-3900	This paper	https://mcc.nies.go.jp/index_en.html
Pavlomulina ranunculiformis CAWP21	This paper	N/A
Pavlomulina ranunculiformis RCC3430	This paper	N/A
Pavlomulina ranunculiformis CG5	[19]	N/A
Software and algorithms		
Cutadapt1.1	[40]	https://cutadapt.readthedocs.io/en/stable/
Trimmomatic0.32	[41]	http://www.usadellab.org/cms/index.php?page=trimmomatic
Velvet1.2.08	[42]	https://www.ebi.ac.uk/~Ezerbino/velvet/
MAFFT	[43]	https://mafft.cbrc.jp/alignment/software/
Gblocks	[44]	http://molevol.cmima.csic.es/castresana/Gblocks.html
IQ-TREE	[45]	http://www.iqtree.org/
PhyloBayes	[46]	https://github.com/bayesiancook/pbmpi
MFannot	[47]	https://github.com/BFL-lab/MFannot
tRNAScan-SE	[48]	https://github.com/UCSC-LoweLab/tRNAscan-SE
PHYLIP package	https://evolution.genetics.washington.edu/phylip.html	N/A
BLAST	[49]	https://blast.ncbi.nlm.nih.gov/Blast.cgi
Expasy	https://web.expasy.org/translate/	N/A
Other		
Tara Oceans V9 rDNA metabarcoding dataset	[1,50]	https://store.pangaea.de/Publications/DeVargas_et_al_2015/Database_W4_barcode_occurences.tsv/Database_W4_barcode_occurences.tsv.zip

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ryoma Kamikawa (kamikawa.ryoma.7v@kyoto-u.ac.jp)

Materials availability

The newly isolated, type strain of *P. ranunculiformis* was deposited and maintained in the National Institute for Environmental Studies. *P. ranunculiformis* CAWP21 and RCC3430 are also maintained in the Cawthron Institute and the Roscoff Culture Collection, respectively. These strains can be purchased from these institutes.

Data and code availability

The accession numbers for the sequence data reported in this paper are available from DNA Data Bank of Japan: accession numbers LC564891 – LC564893 for the plastid and mitochondrial genome sequences of *Pavlomulina ranunculiformis* strain NIES-3900, and LC599498, LC603170, and LC603171 for the 18S rRNA genes of *P. ranunculiformis* strains NIES-3900, CAWP21, and RCC3430, respectively. Phylogenetic datasets are available from Dryad (https://doi.org/10.5061/dryad.x0k6djhhw).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Pavlomulina ranunculiformis

Seawater samples were collected from the Saeki Port, Oita, Japan (32°58′39.4″N 131°54′13.1″E) in June 1991 and incubated in ESM medium[51] containing germanium dioxide to inhibit diatom growth. Single live cells were isolated from the sample using a glass-micropipette. The established culture strain was deposited in the Microbial Culture Collection (https://mcc.nies.go.jp) at the National Institute for Environmental Study (NIES), Japan (the collection strain number NIES-3900). We established culture strains of *P. ranunculiformis* by isolating during routine toxic-phytoplankton monitoring[52,53] at Tapeka Point, New Zealand (strain CAWP21; 35°14′42″S; 174°07′04″E) in 1997; in False Bay, December 1996 (34°10′17″S; 18°45′28″E), Gansbaai, April 1998 (34°34′45″S; 19°17′28″E), and Groot Brak River Beach near Mossel Bay, January 18 2010 (34°3′27″S; 22°14′27″E) on the South African coast (strain CG5; 18S rRNA sequence GenBank accession number HG970975.1).[19] Another culture corresponding to *P. ranunculiformis* (deposited in the Roscoff Culture Collection as RCC3430) was also established from offshore Japan (38°0'N, 142°0'E) in 2013. The strains were confirmed to be the same species as *P. ranunculiformis* by light microscopic observation, such as tadpole-like cells, each with a long non-coiling haptonema and two equal heterodynamic flagella, and 18S rRNA gene sequences (DDBJ: LC599498, LC603170, and LC603171 for NIES-3900, CAWP21, and RCC3430, respectively).

Taxonomic diagnoses

Haptophyta

Rappephyceae M. Kawachi, R. Kamikawa & T. Nakayama classis nov.

Diagnosis: Cells found widely in marine waters. One nucleus, two to four chloroplasts surrounded by four membranes. Peripheral endoplasmic reticulum beneath the cell membrane. 19′-hexanoyloxyfucoxanthin and 4-keto-hexanoyloxyfucoxanthin absent. Flagellate cells with two flagella and a haptonema. Flagella with knob scales.

Pavlomulinales M. Kawachi & I. Inouye ord. nov.

Diagnosis: With characters of class.

Pavlomulinaceae M. Kawachi & I. Inouye fam. nov.

Diagnosis: With characters of class.

Pavlomulina M. Kawachi & I. Inouye gen. nov.

Diagnosis: With characters of class. Flagellate cells with a long haptonema and two isokont flagella showing heterodynamic behavior, either S-shaped motion or recurved flagellar beat: 14–16 microtubules in the haptonema. Four chloroplasts.

Etymology: Named for the cell features it shares with both *Pavlova* (the S-shaped flagellar motion in one flagellum) and *Chrysochromulina* (a conspicuous haptonema and isokont flagella). The name is created by synthesizing parts of two names, ''Pavlo'' from *Pavlova* and ''mulina'' from *Chrysochromulina.*

Type species: *Pavlomulina ranunculiformis*

Pavlomulina ranunculiformis S. Sym, R. Pienaar & M. Kawachi sp. nov.

Diagnosis: With characters of class and genus. Solitary, slender tadpole-shaped cells with a tail, 40–55 μm long and 8–12 μm wide. Rounded cells, 20–25 μm long and 13–18 μm wide. Two equal heterodynamic flagella with knob scales, 30–40 μm long. The haptonema 18–22 μm long, non-coiling. Cells covered with layers of extrusome-like bodies termed pot-shaped scales (PSS), the mature encapsulated form in peripheral cytoplasm 0.3 μm wide and 0.5 μm long, and the released PSS 0.35 μm wide and 0.45 μm long subtended by a raised portion of the cell membrane. Four golden-brown chloroplasts with a projected pyrenoid. A striated, fibrous, rhizoplast-like structure connects the nucleus and posterior end of the cell.

Etymology: *ranunculiformis* (L) –ranunculus meaning little frog (= tadpole); formis meaning shaped

Holotype: Figure 1A. Cells embedded in resin specimens, deposited in the Natural Science Museum, Tsukuba, Japan, as the permanent slide (TNS-AC-58979) and the block (TNS-AC-58979tb).

Type locality: Saeki Port, Oita, Japan (32°58′39.4″N 131°54′13.1″E).

Type culture: The culture strain was established in 1991 and is deposited at the National Institute for Environmental Studies, Japan, as NIES-3900.

METHOD DETAILS

DNA-sequencing, assembly, and organellar genome annotation

Total DNA, extracted from cells using Plant DNA Preparation - Solution Kit (Jena Science), of NIES-3900 was subjected to library preparation using a TruSeq Nano DNA Library Prep Kit (350 bp insert; Illumina), and to sequencing using an Illumina Hiseq 2500 System (Hokkaido System Science Co.). This resulted in 46.9 million paired-end reads of 100 bp. After quality filtering procedures using

cutadapt1.1[40] and Trimmomatic 0.32[41] (with the options -phred33 LEADING:0 TRAILING:0 SLIDINGWINDOW:20:20 MINLEN:50), 45.6 million paired-end reads remained. A circularly mapped c. 110 kb complete plastid genome of NIES-3900 was generated by assembling the paired-end reads using Velvet 1.2.08[42] with the default settings, followed by homology-based search of plastid-encoded genes by BLASTN[49] and gap filling between contigs by PCR and Sanger sequencing. We also obtained two sequences of 11 and 10 kb, respectively, carrying genes encoded in many mitochondrial genomes, such as cytochrome *c* oxidase subunit 1 and cytochrome *b*. However, we failed to fill the gaps between the mitochondria-derived sequences, possibly because there are long repeated regions dispersed over the mitochondrial genomes of haptophyte species.[54]

Organellar phylogenomics

Genes in organellar DNA contigs were identified using MFannot (https://megasun.bch.umontreal.ca/RNAweasel/) and tRNAScan-SE.[43] From those plastid and mitochondrial sequences, we prepared two distinct phylogenetic datasets, i.e., a plastid and a mitochondrial dataset. The plastid-encoded protein sequences obtained by *in silico* translation (Expasy; https://web.expasy.org/translate/) were added to the single protein datasets used in Muñoz-Gómez et al.[55] Some redundant taxa in the red algae were removed prior to further analysis. Alignment and removal of ambiguously aligned positions were performed as described above, except that the Gblocks option was set to t = p. Eighty-three single-protein datasets were concatenated; the concatenated dataset comprised 65 taxa and 16,234 amino acid sites (Table S1). Preparation of the mitochondrial dataset was similarly performed, by adding mitochondrial protein sequences of NIES-3900 to the fourteen single protein datasets from Jackson and Reyes-Prieto.[56] The concatenated mitochondrial dataset comprised 49 taxa and 3,366 sites (Table S2). The plastid dataset was subjected to IQ-TREE analyses,[45] under the LG+Γ+F+C60 model. The generated tree was then used as the initial tree for IQ-TREE analyses,[45] under the LG+Γ+F+C60-PMSF model, with 100 bootstrap iterations. The same dataset was also analyzed using PhyloBayes[46] under the CAT-GTR+Γ model, with two independent MCMCs that were run for 5,500 trees with burn in of 1,400 trees. Two chains converged with maxdiff = 0.259. Subsequently, the consensus tree with branch lengths and BPPs was calculated from the rest of the sampled trees. The mitochondrial dataset was subjected to IQ-TREE analyses[45] under the LG+Γ+F+C60 model. The generated tree was used as the initial tree for IQ-TREE analyses,[45] under the LG+Γ+F+C60-PMSF model, with 100 bootstrap analyses. The same dataset was also analyzed using PhyloBayes[46] under the CAT-GTR+Γ, model with two independent MCMC chains that were run for 7,500 trees with burn in of 2,500 trees. Two chains converged with maxdiff = 0.249. Subsequently, the consensus tree with branch lengths and BPPs was calculated from the rest of the sampled trees. To evaluate the functions of fast-evolving sites in the plastid and mitochondrial protein datasets, we progressively removed the fastest evolving sites at steps of 5% sites removed at a time. Among-site evolutionary rates were inferred using IQ-TREE,[45] with the -wsr option, under the LG+Γ+F+C60 model. We reconstructed phylogenetic trees using IQ-TREE[45] under the LG+Γ+F+C60-PMSF model, with the -bb 1,000 option. To calculate ultrafast bootstrap values for alternative topologies in the 1,000 bootstrapped trees, we used CONSENSE from the PHYLIP package (Phylogeny Inference Package v. 3.6; University of Washington, Seattle, WA, 1999).

To generate the plastid phylogeny, we prepared another dataset. We added sequences of the haptophyte-derived plastid-bearing dinoflagellate species *Karenia mikimotoi* and *Karlodinium veneficum* to the plastid dataset. Because the dinoflagellate organellar transcripts undergo RNA editing, mRNA sequences reported in Klinger et al.,[57] rather than genome sequences, were used in this analysis. Alignment and removal of ambiguously aligned positions were performed as described above. We removed single protein datasets lacking the Kareniaceae sequences, and concatenated the resultant 56 single protein datasets, resulting in 67 taxa and 12,726 sites (Table S3). The dataset including the dinoflagellates was subjected to IQ-TREE analyses[45] as described for the plastid dataset.

Phylogenetic analyses of single genes

Plastid 16S rRNA genes of red algae, haptophytes, cryptophytes, ochrophytes, and environmental DNA, including rappemonads, were retrieved from GenBank. We added the plastid 16S rRNA gene of NIES-3900 to the data and aligned it in MAFFT[43] using the L-INS-i method. Removal of ambiguously aligned positions was performed using Gblocks[44] (options: -t = d and -b5 = h). The trimmed dataset comprising 97 taxa and 1295 nucleotide positions was subjected to maximum likelihood analyses using IQ-TREE[45] under the GTR+Γ+I model, with 100 bootstrap iterations. The same trimmed dataset was also analyzed using PhyloBayes[46] under the CAT-GTR+Γ model, with two independent Markov chain Monte Carlo chains (MCMC) that were run for 50,000 trees with burn in of 13,000 trees. Two chains converged with maxdiff = 0.102. Subsequently, the consensus tree with branch lengths and Bayesian posterior probabilities (BPPs) was calculated from the rest of the sampled trees.

Nuclear 18S rRNA genes of haptophytes and environmental DNA reported to be unidentified deep-branching haptophytes[3,19,22,33,38,39] as well as centrohelids (outgroup taxa) were retrieved from GenBank. We added the nuclear 18S rRNA gene of the *P. ranunculiformis* strains to the data, and performed alignment using MAFFT[43] with the L-INS-i method. Removal of ambiguously aligned positions was performed as described for the plastid 16S rRNA analyses. The trimmed dataset comprising 47 taxa and 948 nucleotide sites was subjected to maximum likelihood analyses using IQ-TREE,[45] under the GTR+Γ+I model with 100 bootstrap iterations.

Rpl36 of NIES-3900 was added to the dataset used previously,[27] followed by removal of redundant taxa and addition of the Rpl36 sequence of the Prymnesiophyceae *Isochrysis galbana* (NC_049168.1). Alignment was performed with MAFFT[43] as described above. Removal of unaligned positions were performed manually. The resultant dataset comprising 75 taxa and 37 amino acid sites was subjected to maximum likelihood analyses using IQ-TREE,[45] under the LG+Γ+F model with 100 bootstrap iterations.

Distribution of the NIES-3900 lineage based on Tara Oceans metabarcoding data

In order to estimate the abundance of NIES-3900 and its near-relatives in the natural environment, we utilized the *Tara* Oceans metabarcoding dataset.[1,50] The dataset, which consists of counts of unique 18S rRNA V9 sequences per *Tara* Oceans sample, was downloaded from PANGAEA, a data library for earth system science (https://store.pangaea.de/Publications/DeVargas_et_al_2015/Database_W4_barcode_occurences.tsv/Database_W4_barcode_occurences.tsv.zip). To obtain V9 sequences which are identical and/or similar to the 18S rRNA gene sequence of NIES-3900, a BLASTN similarity search using the NIES-3900 sequence was conducted against V9 sequences obtained from the *Tara* Oceans metabarcoding dataset and sequences showing similarity were extracted from the database. Five V9 sequences, including one identical to the NIES-3900 18S rRNA sequence, comprised a monophyletic clade based on a preliminary, maximum likelihood phylogenetic analysis. The remaining hits (restricted to cryptophytes and haptophytes) were numerous but of little use in grounding affiliations with any known taxa. To address this, each of them was used as a nucleotide query against deposited sequences in GenBank and only those sequences that had a complete match (obviously limited to the V9 sequence) were retained as better identifiable OTUs. Those without a match were discarded as uninformative, a step that also served to reduce the chances of incorporating data generated by sequencing error. All data recovered in this manner, including of course NIES-3900 and its near-relatives (81 V9 sequences), were subjected to a further round of phylogenetic analysis. They were aligned using MAFFT,[43] with the L-INS-i method, and subjected to phylogenetic analysis using IQ-TREE,[45] under the TIM3e+I+Γ model. Statistical support for the bipartitions in the maximum likelihood tree was assessed using a nonparametric bootstrap analysis (100 replicates).

To compare abundances among the *Tara* Oceans sampling stations, size fractions, and sample environments, we calculated read numbers for two 18S rRNA V9 sequences, one identical to and the other different by only one base from the NIES-3900 sequence (sequence IDs: 7012eb813d3f5b8f645af854d6be238a and 762a19150436c13c43ae51e4, respectively) in four size-fractionated water samples (0.8–5 μm, 5–20 μm, 20–180 μm, and 180–2,000 μm) for two types of environments (SRF: surface water layer, DCM: the deep chlorophyll maximum layer), for each of 46 stations. Stations with any missing values were not used in the analysis. Relative abundances of the NIES-3900 sequence and the related sequence (762a19150436c13c43ae51e4) to the total V9 barcode sequences in each sample category (e.g., station, size-fraction or environment type) were regarded as their contributions to the micro-eukaryotic community of each category. Comparison of abundances for the two NIES-3900 lineage sequences with those of other haptophytes were conducted using 22 V9 sequences that showed 100% identity to the 18S rRNA gene sequences of 33 haptophyte species in GenBank. *Emiliania huxleyi* and *Reticulofenestra parvula* were treated as *Gephyrocapsa huxleyi* and *Gephyrocapsa parvula*, respectively, according to Bendif et al.[58]

HPLC analysis

Pigments of NIES-3900, *G. huxleyi* NIES-2697, and *Pavlova gyrans* NIES-623 were extracted with 150 μL of 100% methanol by ultrasonication for 1 min in Branson 5510 (Yamato). The methanol extract was centrifuged at 12,000 rpm to remove debris. Each pigment was separated using a C8-HPLC column[59] equipped in the Separations Module Waters 2695 (Waters) and then detected by the Photodiode Array Detector Waters 2996 (Waters). Carotenoids were identified on the basis of their retention time and characteristic absorption spectra, as per Zapata et al.[59]

Microscopy

Light microscopy observations of living and fixed specimens were conducted using a Nikon Optiphot ZF-NT microscope with differential interference contrast. Living cell movements were recorded using a 3CCD camera (DXC750NS; Sony Co.) and control unit, with an Hi8 video recorder. Sectioned material for TEM observations was prepared by two fixation methods: 1) fixation for 1 h with 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer (pH 7.2) containing 0.25 M sucrose at room temperature (20°C); and 2) 5 h fixation with 0.1% OsO_4 and 2.5% glutaraldehyde fixation in 0.1 M sodium cacodylate buffer (pH 7.2) at 4°C. The latter was used for observations of the rhizoplast and the pot-shaped scales. Both were rinsed with 0.1 M sodium cacodylate buffer (pH 7.2), and post-fixed in 1% OsO_4 in the same buffer for 2 h at 4°C. After rinsing with the same buffer, the cells were dehydrated in a graded alcohol series and then embedded in Spurr's resin.[60] Thin sections were cut using a diamond knife and were double-stained with 2% aqueous uranyl acetate and lead citrate.[61] Specimens were observed using a JEOL 100 CXII transmission electron microscope.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical support for phylogenies was obtained using 100 non-parametric bootstraps (using the GTR+Γ+I model, the TIM3e+I+Γ model, the LG+Γ+F model, and the LG+Γ+F+C60-PMSF model), 1000 ultrafast bootstraps (using the LG+Γ+F+C60-PMSF model), and Baysian posterior probabilities (GTR-CAT+Γ model, two chains, chain bipartition discrepancies: max difference < 0.3) (Figures 1, 2, and 3).