Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans

Hui Liu¹, Ian Probert⁴, Julia Uitz⁴, Hervé Claustre⁴, Stéphane Aris-Brosou⁵, Miguel Frada², Fabrice Not³, and Colombian de Vargas¹,², ¹

¹Centre National de la Recherche Scientifique, Unité Mixte de Recherche 7144 and Université Pierre et Marie Curie Paris 06, Equipe Evolutions du Plancton et Paléo-Océans, Station Biologique de Roscoff, 29682, France; ²Institute of Marine and Coastal Sciences, Rutgers University, New Brunswick, NJ 08901; ³Marine Physical Laboratory, Scripps Institution of Oceanography, University of California at San Diego, La Jolla, CA 92093-0238; ⁴Centre National de la Recherche Scientifique, Unité Mixte de Recherche 7093 and Université Pierre et Marie Curie Paris 06, Laboratoire d’Océanographie de Villefranche/Mer, 06234, France; and ⁵Department of Biology and Department of Mathematics and Statistics, University of Ottawa, Ottawa, ON, Canada K1N 6N5

Communicated by W. A. Berggren, Woods Hole Oceanographic Institution, Woods Hole, MA, June 2, 2009 (received for review December 18, 2008)

The current paradigm holds that cyanobacteria, which evolved oxygenic photosynthesis more than 2 billion years ago, are still the major light harvesters driving primary productivity in open oceans. Here we show that tiny unicellular eukaryotes belonging to the photosynthetic lineage of the Haptophyta are dramatically diverse and ecologically dominant in the planktonic photic realm. The use of Haptophyta-specific primers and PCR conditions adapted for GC-rich genomes circumvented biases inherent in classical genetic approaches to exploring environmental eukaryotic biodiversity and led to the discovery of hundreds of unique haptophyte taxa in 5 clone libraries from subpolar and subtropical oceanic waters. Phylogenetic analyses suggest that this diversity emerged in Palaeozoic oceans, thrived and diversified in the permanently oxygenated Mesozoic Panthalassa, and currently comprises thousands of ribotypic species, belonging primarily to low-abundance and ancient lineages of the “rare biosphere.” This extreme biodiversity coincides with the pervasive presence in the photic zone of the world ocean of 19-hexanoyloxyfucoxanthin (19-Hex), an accessory photosynthetic pigment found exclusively in chloroplasts of haptophyte origin. Our new estimates of depth-integrated relative abundance of 19-Hex indicate that haptophytes dominate the chlorophyll a-normalized phytoplankton standing stock in modern oceans. Their ecological and evolutionary success, arguably based on mixotrophy, may have significantly impacted the oceanic carbon pump. These results add to the growing evidence that the evolution of complex microbial eukaryotic cells is a critical force in the functioning of the biosphere.

Haptophyta | photosynthesis | protistan biodiversity | eukaryotic biodiversity

Oxygeonic photosynthesis, the most complex and energetically powerful molecular process in biology, originated in cyanobacteria more than 2 billion years ago in Archean oceans (1). Marine photosynthesis still contributes ~50% of total primary production on Earth (2). This revolutionary process was integrated, at least once, into an ancestral phagotrophic eukaryotic lineage through the evolution of chloroplasts, which themselves were redistributed to a large variety of aquatic eukaryote lineages via permanent secondary and tertiary endosymbioses (3). Despite this evolutionary trend from photosynthetic prokaryotes to eukaryotes, particularly visible in today’s coastal oceans where microalgae such as diatoms and dinoflagellates are omnipresent, cyanobacteria have been repeatedly claimed as the champions of photosynthesis in open ocean waters (4). This hypothesis followed the introduction of flow cytometry and molecular genetic approaches to biological oceanography in the 1980s, which revealed astonishing concentrations of minute cyanobacterial cells of the genera Prochlorococcus and Synechococcus in marine waters (5). The physiology, ecology, and functional and environmental genomics of these prokaryotes are subjects of ongoing intensive study (6).

Several lines of evidence in fact argue for eukaryotic supremacy over marine oxygeonic photosynthesis. Flow cytometric cell counts (7) show that picophototrophic protists (0.2–3 μm cell size) are indeed 1–2 orders of magnitude less abundant than cyanobacteria. However, biophysical and group-specific 14C-uptake measurements suggest that tiny eukaryotes can, through equivalent or higher growth rates of relatively larger cells, dominate carbon biomass and net production in both coastal (8) and oceanic (7) settings. High performance liquid chromatography (HPLC) analyses of group-specific accessory pigments have further stressed the ecologic prevalence of phototrophic protist taxa. In particular, 19-hexanoyloxysfucoxanthin (19-Hex) was originally estimated to account for 20–50% of total chlorophyll a (Chl a) biomass in tropical Atlantic and Pacific sites (9) and has since been consistently reported in open ocean photic zone waters, e.g., (10, 11), suggesting a ubiquitous occurrence of haptophytes in upper layers of the water column. Surveys of genetic diversity based on environmental ribosomal DNA libraries over the last decade have unveiled an unexpected diversity of tiny eukaryotes in all oceans (12). Paradoxically, most picoeukaryotic sequence diversity from photic layers represented novel heterotrophic (13) and parasitic (14) protists within phyla traditionally thought to be dominated by photoautotrophs. This bias indicated that marine protist diversity might be significantly skewed toward heterotrophic taxa (15), as appears to be the case for prokaryotes. However, the paucity of haptophyte nuclear rDNA sequences in these surveys contrasts strikingly with the abundance of 19-Hex in marine waters.

Here we use a combination of previously undescribed genetic, pigment, and microscopy data to unveil a dramatic and ancient diversity of unique photosynthetic picoplanktonic protists within the Haptophyta. This diversity could account for the mysteriously high concentration of 19-Hex in the photic layer of the world oceans, our calculations indicating that haptophytes contribute ~2-fold more than either cyanobacteria or diatoms to global oceanic Chl a standing stock. The phylogenetic position of these tiny haptophytes implies that they are photophagotrophic, coinciding with the recent discovery of dominant bacterivory by small eukaryotic phytoplankton in the oceans (16). Mixotrophy may provide a competitive advantage over both purely phototrophic microalgae (including cyanobacteria) and aplastidal protists, and the extreme genetic diversity of tiny haptophytes matches the cellular and behavioral complexity inherent in this mixed mode of nutrition.

Author contributions: C.d.V. designed research; H.L., I.P., J.U., H.C., M.F., and F.N. performed research; H.L., J.U., H.C., S.A.-B., F.N., and C.d.V. analyzed data; and I.P. and C.d.V. wrote the paper.

The authors declare no conflict of interest.

¹To whom correspondence should be addressed. E-mail: vargas@sdb-roscoff.fr.

This article contains supporting information online at www.pnas.org/cgi/content/full/0905841106/DCSupplemental.
Results and Discussion

A Massive Unique Diversity of Oceanic Picohaptophytes. We first show that previous nuclear rDNA PCR-based studies of eukaryotic communities were subject to severe selective amplification biases. Several groups of protists known to have long and/or GC-rich rDNA are virtually missing from environmental clone libraries produced by classical PCR amplification protocols using “general eukaryote” SSU rDNA primers (12, 13). This is the case for the haptophytes, the rDNA of which has a mean GC content of 57%. We therefore used haptophyte-specific primers and a PCR protocol designed for GC-rich genomes to amplify LSU rDNA D1–D2 fragments from bulk DNA extracted from the 0.2- to 3-μm fraction of seawater collected at 4 offshore stations in the Arctic and Indian oceans (Fig. S1, Table S1). Standard eukaryotic rDNA analyses of these samples yielded ~0.4–0.7% haptophyte sequences (10, 11). In contrast, our data reveal hundreds of previously undescribed rDNA sequences from tiny haptophytes. Rarefaction curves for individual clone libraries (Fig. 1) indicate that current sequencing effort is far from exhaustive, notably in subtropical waters where genetic diversity is particularly dramatic. Estimates of the number of unique ribotypes using the Chao1 estimator were 1098–1147 and 325–509, respectively, for the Indian and Arctic ocean samples (with rather large confidence intervals, see Table S2). The frequency distribution of unique ribotypes (Fig. 1) indicates higher species richness in subtropical waters, with a substantial number of orphan and deep-branching genotypes (see below) in both warm and cold waters. This parallels recent observations for marine prokaryotes of a “seed bank” of ancient and rare taxa, termed the “rare biosphere” (17).

Taxonomy and Evolutionary History of the Previously Undescribed Diversity. The 674 novel environmental LSU rDNA sequences were aligned with 64 orthologous gene sequences from clonal culture strains representing a cross-section of known haptophyte biodiversity. Phylogenetic analyses indicate that all environmental sequences belong to the Haptophyta (Fig. 2), a eukaryotic phytoplankton division classically considered as nanoplankton (3–20 μm) and including the calcifying coccolithophores (18). However, not a single environmental sequence was strictly identical to any of the taxonomically defined sequences. The vast majority of environmental sequences form new clusters branching deep in the haptophyte phylogeny, most being related to Chrysochromulina species from clade B2 within the order Prymnesiales (19). The described representatives of this clade are non-lithogenic and known almost exclusively from coastal and shelf environments (20); our data show they are in fact derived...
from open ocean picoplanktonic taxa (Fig. 2). Note that the 2 other major prymnesiophyte lineages, the Phaeocystales and the Calcihaptophycidae, also appear to emerge from clusters of picohaptophyte sequences.

Calibration of our tree with the stratigraphic record of coccolithophore taxa (Fig. 2B) suggests that tiny haptophyte biodiversity emerged more than 250 Ma, in Paleozoic oceans (Fig. S2), before the evolution of intracellular biomineralization in the Calcihaptophycidae, which according to both fossil and molecular clock data occurred ≈220 Ma (18). The phylogenetic depth of most picohaptophyte clades argues for a Mesozoic diversification of the group, which may have thrived in the newly permanently oxygenated and largely oligotrophic Panthalassic Ocean, conditions which served as a selection matrix for a wide range of chlorophyll a + c containing protists (21). Many genotypes or genotype clusters were found exclusively in either subarctic or subtropical oceans, supporting significant lineage partitioning between cold mixed and warmer stratified waters (Fig. 3). This phylogeographic distribution of ribotypes suggests that tropical waters were the original center of diversification, with biodiversity spreading secondarily into higher latitudes, a scenario that fits the putative early radiation of the group in the warm Panthalassa.

Ecologic Relevance of the Picohaptophytes. Our genetic survey positions the haptophytes as the most diverse group of picophytoplankton in modern open oceans. Recent exploration of chloroplastic SSU rDNA in pelagic (22) and coastal (23) environments supports this conclusion. Haptophytes dominate the emerging chloroplastic view* of marine tiny eukaryotic phytoplankton in terms of both diversity and abundance. In a year-round data set from the Gulf of Naples (23), >45% total and >70% unique eukaryote chloroplastic rDNA sequences were of haptophyte origin, 55% of them belonging to the Prymnesiales clade-B2 (Fig. S3). This extreme diversity coincides with a numerical significance. Group-specific fluorescent in situ hybridization data from various oceanic settings indicate that haptophytes represent up to 35% of total picoeukaryotic cell numbers (24). Dot blot hybridizations using group-specific chloroplastic rDNA probes indicated a mean dominance of ≈45% of haptophytes among other eukaryotic divisions during a 2-year survey of ultraphytoplankton (<5 μm cell size) in Mediterranean waters (23). To assess whether these localized observations are representative of a global trend, we evaluated the contribution of haptophytes to oceanic phototrophic biomass using an empirical model based on >2,400 worldwide vertical profiles of HPLC pigment data integrated through monthly ocean-color composites of surface Chl a concentrations measured by the SeaWiFS satellite sensor in the year 2000 (Fig. 4). This analysis revealed that 19-Hex was the dominant accessory pigment in the oceans over this period, representing about twice the standing stocks of either fucoxanthin (diatoms) or zeaxanthin (prokaryotes) when normalized to Chl a. Haptophytes appear thus to represent the background oceanic light harvesters, contributing from 30 to 50% of total photosynthetic standing stock across the world ocean.

**Mixotrophy, the Key to the Success of Tiny Haptophytes in Open Oceans?** The phylogenetic position of the majority of the picohaptophytes in the Prymnesiales strongly suggests that they are mixotrophic, i.e., able to supplement their phototrophic regime with uptake and assimilation of organic nutrients. Laboratory experiments have shown that members of the Prymnesiales are

---

*Note that chloroplast genomes are typically not GC biased, meaning they are amenable to standard PCR protocols and may provide a more accurate view of the real photoplanktonic diversity.
typically capable of ingesting organic particles and prey (e.g., refs. 25–27). Their third flagellum-like appendix, the haptonema, which is particularly long relative to cell size in all described members of Prymnesiales clade B2, can be used to catch prey and transfer them to the cell membrane for active phagocytosis (28). Even in calciphagophytes, highly modified coccoliths may be involved in harvesting preys (29).

Field studies on bacterivory by plastid-containing protists have demonstrated the dominance of tiny haptophyte-like cells in oceanic mixotrophy, e.g., see refs. 16 and 30. Recent quantitative evidence (16) revealed that eukaryotic algae with cell size \( \leq 5 \mu m \), expected to be mostly haptophytes, carry out most of the bacterivory in the ephotic layer of both the temperate and tropical Atlantic Ocean. Furthermore, significant 19-Hex concentrations were recorded in 200- to 300-m-deep layers of the clearest waters on Earth in the South Pacific gyre (31), depths where irradiance at noon is not even sufficient for photosynthesis to cover basic cellular metabolic requirements. The complex combination of phagocytotic and photosynthetic modes of nutrition can be postulated to have allowed haptophytes to attain relatively large size and morphological complexity while maintaining prokaryote-like growth rates, and thus to have radiated into a wide diversity of ecotypes. The nutritional flexibility offered by mixotrophy is likely to have equipped the tiny haptophytes with a significant competitive advantage over both purely phototrophic and aplastidial cells under different light (depth) and nutrient regimes†.

Concluding Remarks. Besides their unanticipated diversity and abundance, the unveiled haptophytes display morphological features that suggest they play critical roles in organic carbon fluxes on a global scale. Size analyses of cells identified by haptophyte-specific fluorescent probes revealed a mode of \( \approx 4 \mu m \), with largest sizes of \( 8–9 \mu m \) (Fig. S4 and Table S4). In terms of volume, haptophytes are thus typically 300–3,000 times larger than \textit{Prochlorococcus}, the most abundant marine cyanobacteria. The few available electron microscopy images of these open ocean tiny haptophytes indicate that they do produce organic plate scales (Fig. S5), a plesiomorphic character common to the overwhelming majority of prymnesiophytes. Interestingly, abundant and diverse \textit{Chrysochromulina} spp. scales were recently observed in Atlantic surface sediments collected at 4,850 m (32). The taxonomic origin and pristine preservation of these scales, previously overlooked in deep-sea sediments because of their minute size (\( \leq 1 \mu m \)), suggest that they were rapidly transported to the sea floor. Eukaryotic scales made of proteins embedded into cellulose and other polysaccharides potentially provide abundant resistant and sticky matter to enhance aggregation and flux of marine snow particles to the deep ocean, contributing to the largely underestimated role of coagulation of small phytoplankters in the biological pump (33). Thus, the tiny haptophytes may have been essential mediators of carbon fluxes from the atmosphere to the deep oceans and the lithosphere throughout much of the Phanerozoic Eon.

Methods

Sampling, DNA Extraction, and Construction of LSU rDNA Clone Libraries. At each sampling station (Fig. S1 and Table S1) 5–15 L of seawater was immediately prefiltered through a 200-μm-cylinder mesh and collected in an acid-washed carboy. The water was then filtered, using peristaltic pumping, through a 3-μm-pore-size Nucleopore polycarbonate filter (Millipore), before recovery of picoplanktonic cells in 0.2-μm-pore-size Sterivex filter units (Millipore). Filters were preserved in lyss buffer (40 mM EDTA, 50 mM Tris-HCl, 0.75 M sucrose) and stored at \(-80^\circ C\) until genomic DNA extraction was performed as in ref. 34. Approximately 1,000 bp nuclear LSU rDNA fragments including the D1–D2 domains were PCR amplified using the forward haptophyte-specific primer Hapto-4 (5’-atcgcgaatgagcagcagc-3’), and the reverse general eukaryote primer Euk-34r (5’-gcagtgacatgtgctgacct-3’). PCR reactions (98 °C for 30 s, 50 °C for 30 s, and 72 °C for 60 s, with initial denaturation and final extension steps) were performed over a maximum of 30 cycles to limit formation of chimeric sequences (35) using the Phusion high-fidelity PCR DNA polymerase (New England BioLabs), which is specifically suited for amplification of GC-rich DNA. PCR products were purified using the MinElute gel extraction kit (Qiagen) and 3’-A-overhangs were bound to DNA fragments by adding 0.2 mM dATP, 1 unit of Taq DNA polymerase, and 1× Taq DNA polymerase buffer to the purified PCR product, and incubating for 20 min at 72 °C. Classical TA-cloning into OneShot DH5α competent bacteria using the TOPO TA kit (Invitrogen) was then performed according to the manufacturer’s instructions. Clone libraries were checked by PCR using the M13 forward and reverse primers and sequencing of \(-25–35\) random clones in both directions. The entire process of library construction was repeated until \( >85\% \) of white clones yielded high-quality sequences. Libraries were then sent to High-Throughput Sequencing Solutions (www.htseq.org) for random automatic picking of 200 clones, plasmid mini-preps, and automatic sequencing of both strands of 150–200 LSU rDNA fragments per library.

Molecular Biodiversity, Phylogenetic, Molecular Clock, and Biogeographic Analyses. The choice of the LSU rDNA D1–D2 fragment over the classically used SSU rDNA to assess haptophyte diversity was motivated by the inability of the latter marker to distinguish closely related species (Table S5). D1–D2 LSU rDNA fragments show virtually no intraspecific variations while discriminating morphospecies that split in the Pleistocene. Unambiguous LSU rDNA sequences were first screened for chimeras using CheckChimera (46), then double checked by thorough visual inspection of all sequences producing abnormally long branches in neighbor-joining trees (37) as described in Hugenholtz and Huber (38). This conservative approach led to the removal of \(-13\% \) of putative chimeric sequences from subsequent analyses. The remaining 674 environmental sequences were added to 64 nuclear LSU rDNA sequences obtained from taxonomically identified clonal haptophyte cultures from the Roscoff Culture Collection (http://www.sb-roscoff.fr/Phyto/RCC), and 3 sequences from the LSU rDNA sequencing database. The resulting alignment was manually inspected in Genetic Data Environment 2.2 (40). The Akaike Information Criterion (41) was used to select the most appropriate model of nucleotide substitution: The general time-reversible model plus \( \Gamma \) and invariable sites. For each of the libraries, PAUP* 4.0b10 (42) was used to build pairwise maximum likelihood distance matrices under the model selected above for estimation of rarefaction curves and rDNA richness based on the average neighbor algorithm implemented in DOTUR (43). Phylogenies including both environmental and culture sequences were reconstructed using MrBayes v3.1.2 (44), with 2 independent samplers, 10⁷ steps, tempering with 1 cold and 3 heated chains, and burn-in of 10⁶ steps. A Bayesian analysis implemented in BEAST v1.4.6 (45) was performed to construct the phylogeny while estimating divergence times (Fig. S2). This relaxed clock analysis included the 184 sequences from the total alignment that were \( >95\% \) divergent. Absolute time calibration was based on the earliest geological record for the evolution of calcification \([i.e., \sim 220\ Mya (46)]\) and 4 minimum divergence dates derived from stratigraphic data: in the order Coccolithus, the \(-65-million-year-old\) first appearance of the genus \textit{Cocco-}

|Complex mixotrophic regimes may also explain why none of the open ocean haptophytes are currently available in culture collections. |
derived pigment profiles, the ones from stratified waters were discriminated from those from well-mixed waters, based on the ratio of the euphotic layer depth $Z_{eu}$ (the depth at which photosynthetically available radiation is reduced to 1% of its surface value) to the mixed layer depth $Z_m$. $Z_{eu}$ was computed from the vertical profiles of $[\text{Chl}a]$ using bio-optical models (48, 49), while $Z_m$ was extracted from the Levitus global monthly mean climatology. For both stratified and mixed waters, the vertical profiles of $19$-Hex, Fuco, and Zea were sorted into “trophic categories” defined by successively increasing values of $[\text{Chl}a]_{surf}$ values. Average profiles were first computed independently for each trophic category and each pigment. Because the average pigment profiles display a deterministic behavior in terms of vertical shape and magnitude along the trophic gradient, they could be modeled and parameterized as a function of $[\text{Chl}a]_{surf}$. The predictive skill of the parameters was successfully tested using an independent data set (47). Our empirical model was then applied to monthly composite pigment profiles, Zea, values for the year 2000, on a pixel-by-pixel basis. $Z_{eu}$ was first computed from $[\text{Chl}a]_{surf}$ by using the log-log linear relationship linking $[\text{Chl}a]_{surf}$ to the euphotic layer-integrated Chla content (Eq. 8 in ref. 47) and the relationship linking this last parameter to $Z_{eu}$ (49). The euphotic depth was then compared to the mixed layer depth to determine whether the water column was stratified (i.e., $Z_{eu} = Z_m$) or mixed (i.e., $Z_{eu} < Z_m$). For stratified waters, $[\text{Chl}a]_{surf}$ was used to produce dimensionless profiles (with respect to depth and biomass) of $19$-Hex, Fuco, and Zea, which were then restored to physical units by multiplying depths by $Z_{eu}$ and concentrations by the average Chla concentration within the euphotic layer. For mixed water conditions, the surface concentration of each pigment was inferred from $[\text{Chl}a]_{surf}$ and extrapolated within the euphotic layer to generate uniform vertical profiles. This procedure yielded monthly depth-resolved fields of $19$-Hex, Fuco, Zea, and Chla for the world ocean, which were then integrated over the euphotic zone. For each pixel, the resulting monthly $19$-Hex, Fuco, and Zea integrated contents were converted into Chla equivalents using the appropriate pigment to Chla ratios determined by multiple regression analysis performed on the global pigment database (47). The obtained Chla biomasses attributed to each group were averaged over the year to estimate annual mean values. These values were normalized to the annual mean euphotic layer-integrated Chla content to determine the relative contribution (%) of each phytoplankton group to the total phytoplankton chlorophyll-based biomass (Fig. 4). Finally, for each of the 3 phytoplankton groups, an annual mean Chla standing stock was calculated as the sum of the annually averaged value of each multiplied by the pixel multiplied by the pixel surface area. Coastal areas (bathymetry < 200 m), large lakes, and inland seas were not considered in this analysis.**

**ACKNOWLEDGMENTS.** We thank Swati Narayan-Yadav for practical help to obtain genetic data; Ramon Massana for sharing pico-DNA extractions; Jeremy Young for providing SEM images; and Marie-Pierre Aubry, Paul Falkowski, and David Gruber for critical discussions. This work was supported by National Science Foundation Grants DEB-0415351 and IRES-0652093, an Action Thématique et Incitative sur Programme fellowship from the Centre National de la Recherche Scientifique (to C.D.V.), and the Canadian Natural Sciences and Engineering Research Council (S.A.B.). It is part of the pluridisciplinary project Biodiversity of Open Ocean Microcaliﬁers (BOOM), funded by the Institut Français de la Biodiversité via the Agence Nationale de la Recherche, Grant ANR-05-BDIV-004.


