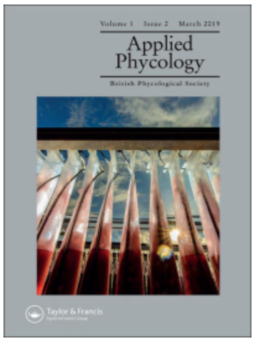


Applied Phycology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tapy20

Genomic and meta-genomic insights into the functions, diversity and global distribution of haptophyte algae

Mathias Penot, Joel B. Dacks, Betsy Read & Richard G. Dorrell

To cite this article: Mathias Penot, Joel B. Dacks, Betsy Read & Richard G. Dorrell (2022): Genomic and meta-genomic insights into the functions, diversity and global distribution of haptophyte algae, Applied Phycology, DOI: 10.1080/26388081.2022.2103732

To link to this article: https://doi.org/10.1080/26388081.2022.2103732

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

View supplementary material

Published online: 28 Jul 2022.

Submit your article to this journal

Article views: 303

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=tapy20

APPLIED PHYCOLOGY
https://doi.org/10.1080/26388081.2022.2103732

British Phycological Society
Understanding and using algae

Taylor & Francis
Taylor & Francis Group

REVIEW

OPEN ACCESS Check for updates

Genomic and meta-genomic insights into the functions, diversity and global distribution of haptophyte algae

Mathias Penot[a,b], Joel B. Dacks[c], Betsy Read[d] and Richard G. Dorrell[a,b]

[a]Institut de Biologie de l'ENS (IBENS), Département de Biologie, École Normale Supérieure, CNRS, INSERM, Université PSL, Paris, France; [b]CNRS Research Federation for the study of Global Ocean Systems Ecology and Evolution, Paris, France; [c]Division of Infectious Disease, Department of Medicine, University of Alberta and Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada; [d]Department of Biological Sciences, California State University San Marcos, San Marcos, California, USA

ABSTRACT
Haptophytes are an environmentally important phylum of eukaryotic phytoplankton, forming the second most abundant algal group after diatoms in recent estimates of ocean biodiversity. Haptophytes are phylogenetically and functionally diverse, including globally distributed and bloom-forming calcifying species such as *Emiliania* and *Coccolithus*, and non-calcifying orders that may form important components of phytoplankton communities in polar (*Phaeocystis, Chrysochromulina*) through to sub-tropical latitudes (*Pavlova*). In this review, we synthesize available phylogenetic, genomic and environmental information concerning the diversity of haptophyte life, considering the origins and placement on the eukaryotic tree; the diversity of the five major orders (Pavlovophyceae, Phaeocystales, Prymnesiales, the CSZ clade, and Isochrysidales); and the contrasting biogeographical distributions of haptophyte groups across different *Tara* Oceans sampling stations and size fractions. We additionally consider outstanding questions within the fields of haptophyte diversity and biology, particularly in the context of newly discovered and largely uncultured major groups (DPL lineages and Rappemonads), and current gaps in our knowledge of genomic content and niche adaptation across the haptophyte tree.

ARTICLE HISTORY
Received 19 December 2021
Accepted 21 June 2022

KEYWORDS
18S rDNA; coccolithophorid; dispersal; meta-genomics; niche adaptation; prymnesiophyte; systematics

Global importance of haptophyte algae in the world ocean

The entire ocean ecosystem and its associated micro- and macro-biota are supported by a diverse range of photosynthetic cyanobacteria and eukaryotic algae, which are distributed across the tree of life (Guidi et al., 2016; Ustick et al., 2021). Despite often being very small, these organisms significantly impact on planetary ecology, assimilating as much carbon dioxide from the atmosphere as plants (Guidi et al., 2016; Malhi & Grace, 2000) and enabling the trophic transfer and dynamic cycling of nutrients across the entire marine food chain (Chaffron et al., 2021). Recently, the environmental distributions of some of the key marine algal groups have been revealed at unprecedented levels of resolution through environmental sequencing (meta-genomic) initiatives such as Bio-GO-SHIP (Ustick et al., 2021) and the *Tara* Oceans Expedition at large (Ibarbalz et al., 2019; Planes et al., 2019; Sommeria-Klein et al., 2021).

Haptophytes are a phylum of eukaryotic algae of major environmental importance. Some haptophytes (the calcareous members, see below) are also referred to as coccolithophorids (Walker, Dorrell, Schlacht, & Dacks, 2011). Most of these organisms are photosynthetic, are unicellular and have a planktonic lifestyle. To date, 1,176 haptophyte species have been described within AlgaeBase (accessed 26/10/21) (Guiry et al., 2014), although these (as per other algal groups, c.f. (Williams, 2021)) may include synonymous and invalid names. Even greater haptophyte diversity remains to be formally classified, particularly in marine environments. For example, the initial *Tara* Oceans and *Tara* Polar Circle Expeditions, which between 2006 and 2013 mapped eukaryotic algal and protist diversity across over 200 sites within the global ocean (Bork et al., 2015), recorded 2550 distinct haptophyte 18S V9 OTUs (Operational Taxonomic Units, corresponding to species orsubspecies) as registered within the Ocean Barcode Atlas (OBA) (Vernette et al., 2021). Alongside this predominantly marine unexplored diversity, a small number of freshwater haptophyte species have been described both via classical isolation and taxonomic approaches (Deodato, Barlow, Hovde, & Cattolico, 2019;

CONTACT Mathias Penot penot@bio.ens.psl.eu; Richard G. Dorrell dorrell@bio.ens.psl.eu

Supplemental data for this article can be accessed online at https://doi.org/10.1080/26388081.2022.2103732

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Plancq, Couto, Ijaz, Leavitt, & Toney, 2019) and metagenomic investigation of river, lake and soil microbial communities (Singer et al., 2020).

The *Tara* Oceans meta-barcode total abundances indicate that haptophytes are the second most abundant group of eukaryotic phytoplankton in the modern oceans, after diatoms (and, debatably, dinoflagellates; Fig 1a) (Liu et al., 2009; Pierella Karlusich et al., 2022; Vernette et al., 2021). Recent and more detailed studies of phytoplankton biogeography across the *Tara* Oceans dataset have revealed characteristic spatial distributions for individual haptophyte species, defined by environmental latitude (Sommeria-Klein et al., 2021) and, to some extent, oceanic ecoregions (Dorrell et al., 2021a). Other classical and molecular ecology studies have provided deeper levels of insight into haptophyte behaviour in the world ocean, demonstrating, for example, their important roles as mixotrophs as well as obligate phototrophs in the marine food chain (Anderson, Charvet, & Hansen, 2018; Unrein, Gasol, Not, Forn, & Massana,

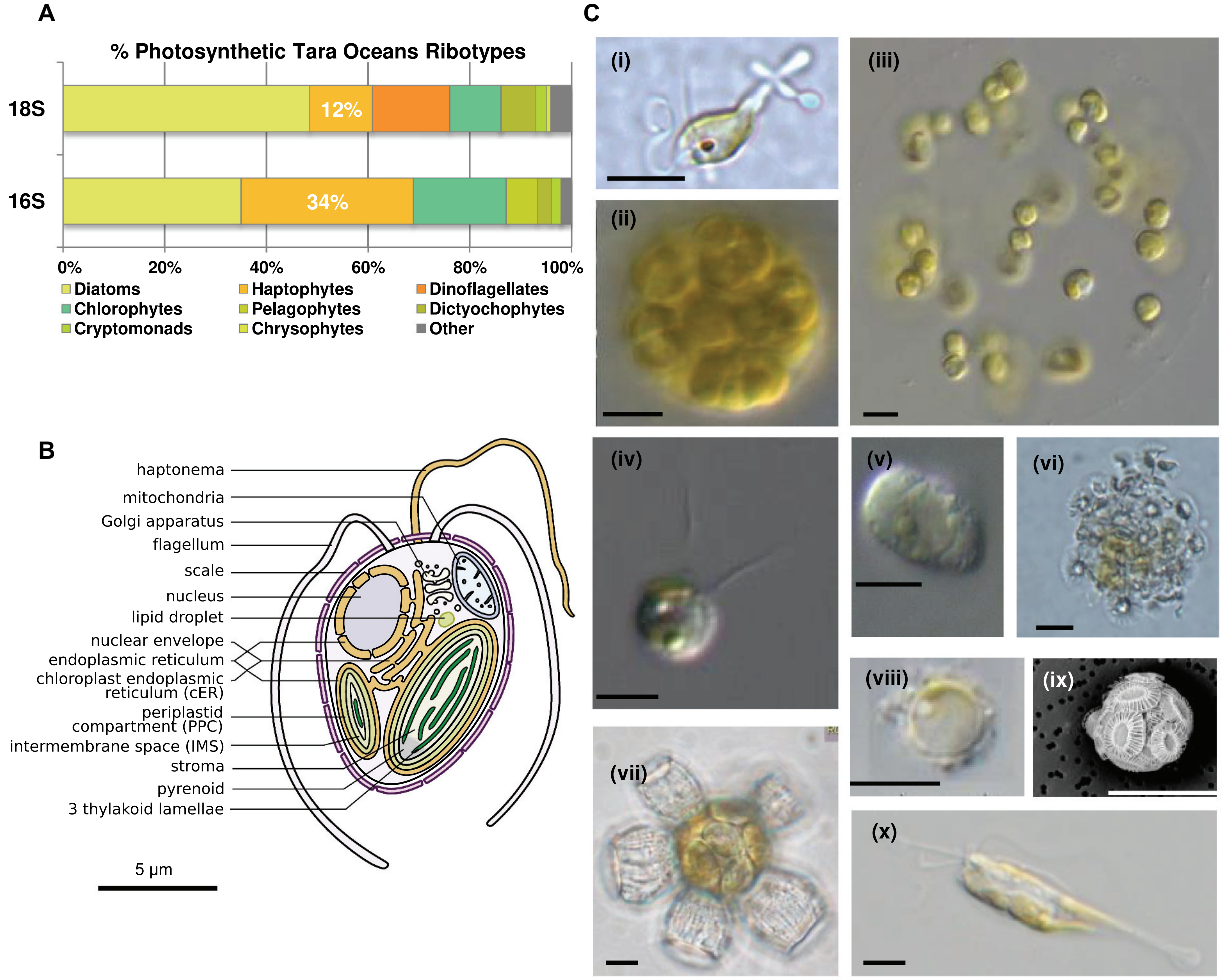


Figure 1. Importance and diversity of haptophytes in the world ocean. A: Total abundances of 18S V9 (following K Ibarbalz et al., (2019)) and plastid 16S V4-V5 (annotated by BLAST best hit analysis against a reference alignment previously published in (Karlusich et al., 2022)) ribotypes from *Tara* Oceans data, across all stations and size fractions that correspond to eukaryotic algae, with the relative contributions of haptophytes highlighted. Dinoflagellates are not shown in the 16S data due to the extreme divergence of their plastid 16S genes (Dorrell, Nisbet, Barbrook, Rowden, & Howe, 2019). B: A schematic haptophyte cell structure. C: Exemplarhaptophyte diversity (i) *Pavlova gyrans* CCAP940 (Pavlovophyceae); (ii) *Exanthemachrysis sp* RCC1532 (Pavlovophyceae); (iii) *Phaeocystis globosa* RCC851 (Phaeocystales); (iv) *Chrysochromulina camella* RCC1187 (Prymnesiales); (v) *Prymnesium simplex* RCC1387 (Prymnesiales); (vi) *Calcidiscus leptoporus* RCC1129 (CSZ clade); (vii) *Scyphosphaera apsteinii* RCC1480 (CSZ clade); (viii, ix) bright-field and SEM images of *Emiliania huxleyi* RCC914 (Isochrysidales); (x) *Pavlomulina ranunculoformis* NIES-3900 (Rappemonads). Image credits: El Mahdi Bendif (i); Priscillia Gourvil (ii, iv, v); Florence Le Gall (iii, viii); Margaux Carmichael (vi); Ian Probert (vii, ix); and Ryoma Kamikawa and Mami Nomura (x). Scale bar = 5 µm.

2014) and in their projected vulnerability and resilience to oceanic heating and acidification (Chaffron et al., 2021).

In this review, we first outline the evolutionary diversity of haptophytes within the eukaryotic tree of life and describe the key morphological features that unify the haptophyte lineage as a whole or differentiate individual haptophyte orders from one another. Next, we consider the biogeographical distributions of the major orders from meta-barcoding data within the *Tara* Oceans Barcode Atlas, highlighting differences in relative abundance, distribution and co-occurrence with different abiotic factors (Vernette et al., 2021). Finally, we establish the state of play in our current understanding of haptophyte genome and transcriptome diversity and the transformative potential of meta-genome and meta-transcriptome data from environmental sequencing initiatives for understanding haptophyte biological processes across the modern ocean (Carradec et al., 2018; Delmont et al., 2022).

Characteristic and diverse haptophyte morphology and ecology

Haptophytes display great morphological diversity, but with key features allowing for ready identification from light microscopy (Fig 1c). The characteristic haptophyte is a small eukaryotic unicellular organism (2–20 µm length) and typically covered with organic or mineralized scales (Eikrem et al., 2017). The haptonema, which gives the haptophytes their name ("háptô" = touch and "nema" = thread) and represents a synapomorphy for this phylum, plays a key role in phagocytosis as it is used to catch, aggregate and deliver prey to the anterior cell surface where they are consumed (Kawachi, Inouye, Maeda, & Chihara, 1991). It also allows cell attachment to substrates (e.g., sediments for benthic and terrestrial species) and might have a role in obstacle perception through coiling movements (Inouye & Kawachi, 1994). The length of the haptonema varies considerably, from 100 µm (*Chrysotila strobilus*) to 1 µm (*Isochrysis spp.)* (Eikrem et al., 2017), while some haptophytes (e.g., *Ochrosphaera*) only possess a vestigial haptonema (Fresnel & Probert, 2005). The haptonema structure is unusual compared to other eukaryotic flagellar organelles, with six or seven microtubules surrounded by a cylindrical endoplasmic reticulum, and very different from that of the true flagella (9 doublets + 2 singlets microtubules surrounded by the plasma membrane) (Eikrem et al., 2017). The swimming behaviour of motile haptophytes is afforded by two same-lengthened flagella (isokonts) that pull the cell forward by beating at the same time (homodynamic flagella) or in opposition (heterodynamic flagella) (Eikrem et al., 2017; Manton, 1968). Non-motile haptophytes are also prevalent, and a select few form colonies or short filaments (*Phaeocystis*) (Eikrem et al., 2017).

Haptophyte cells typically contain mitochondria with tubular lamellae and two yellow-brown chloroplasts derived from the direct or indirect endosymbiosis of red algae (*see below*, Fig 1b,c). Similar to the situation observed in other major groups of eukaryotic algae with chloroplasts of secondary red origin (e.g., cryptomonads and ochrophytes including diatoms), the haptophyte chloroplast contains three thylakoid membrane layers surrounded by four membranes, the outermost of which is contiguous to the endoplasmic reticulum (Andersen, 2004) (Fig 1b). Haptophyte chloroplasts are differentiated structurally from those of cryptomonads by the absence of a nucleomorph (vestigial red algal nucleus) (Cavalier-Smith, 2002) and from ochrophytes by the absence of a "girdle lamella" structure (thylakoid trilayer around the periphery of the chloroplast stroma that encloses other thylakoid membranes) (Andersen, Saunders, Paskind, & Sexton, 1993). Haptophyte chloroplasts contain multiple photosynthetic pigments including chlorophylls *a*, *c*1-3, fucoxanthin and 19'-hexanoyloxyfucoxanthin (Andersen, 2004), the latter of which is specific to haptophytes and their endosymbiotic derivatives (e.g., dinoflagellates within the Kareniaceae) (Takishita, Nakano, & Uchida, 1999) and accordingly used as a proxy to estimate their environmental abundances (Fuller et al., 2006). The main storage product of photosynthesis is chrysolaminarin, a small β-1-3-glucan (polymers of 20–50 glycosides) located in cytoplasmic vacuoles (Granum, Roberts, Raven, & Leegood, 2009). Lipid bodies containing fatty acids and sterols may be particularly important in some oleaginous species (e.g., *Chrysochromulina*) (Hovde et al., 2015).

Beyond these relatively conserved features, much diversity exists in haptophyte form and behaviour. The most famous group, coccolithophorids, are characterized by a cell surface covered by calcified scales named coccoliths that accumulated in the Cretaceous Seas of Europe and now form the chalk cliffs along the English Channel (Bown, 1998; O'Dea et al., 2014) (Fig 1c, panel ix). Other haptophyte groups may have non-calcareous scales, either totally organic or composed of silicon (Yoshida, Noël, Nakayama, Naganuma, & Inouye, 2006), and these scales may furthermore have the same shape or not, be present in one or several layers and cover the whole cell or be unevenly distributed dependent on the species (Eikrem et al., 2017). Scale characteristics represent helpful taxonomic criteria for haptophytes. Other haptophyte species are known for the production of high-value products (alkenones and omega-3-fatty acids) of interest to biofuel research (Araie et al., 2018; Shi, 2019), while others may produce toxins that cause significant damages to fisheries and natural ecosystems (Blossom et al., 2014; Valenti et al., 2010).

In the wild, haptophytes are classically thought of as K-strategists, characterized by high affinities for nutrients but lower growth rates, adapted to grow in highly

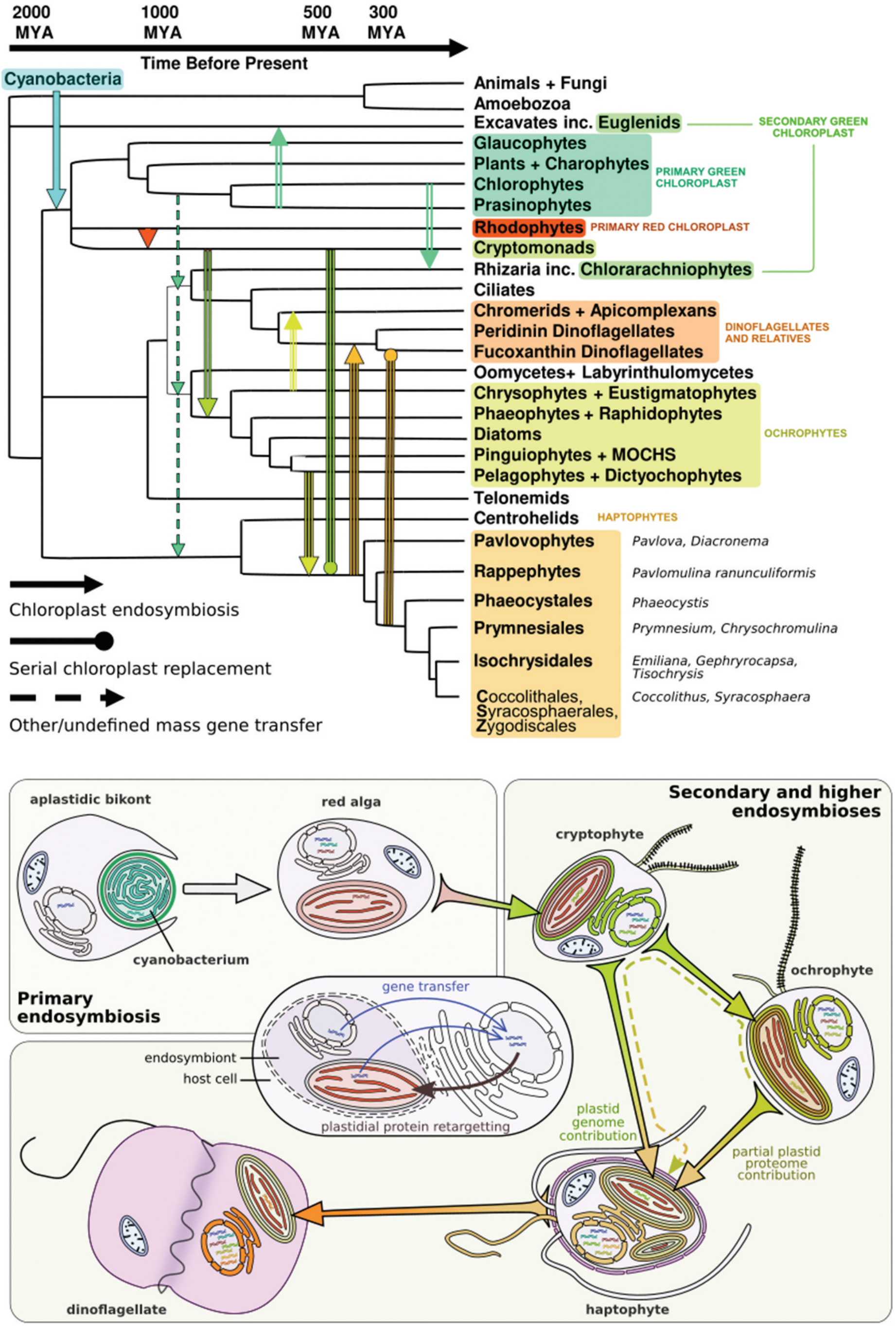


Figure 2. Convoluted evolutionary history of haptophyte nuclear and chloroplast genomes. Schematic tree of eukaryotic life, showing the evolutionary position of haptophytes (as sister to centrohelids and distantly related to other eukaryotic algae) and the timeline of their divergence from other groups (Burki et al., 2016; Strassert et al., 2021); the probable evolutionary history of their chloroplasts, which is ultimately of red eukaryotic origin but has likely been acquired from a cryptomonad-like organism (Rice & Palmer, 2006) and is supported by nucleus-encoded and chloroplast-targeted proteins acquired from pelagophyte or dictyochophyte algae within the stramenopiles via a possible cryptic endosymbiotic event (Dorrell et al., 2017, 2021b); the evolutionary relationships between six major haptophyte orders (Dorrell et al., 2021a; Kawachi et al., 2021); and the inferred origin points of endosymbiotic transfers from haptophytes into dinoflagellates (e.g., within the Kareniaceae) (Kawachi et al., 2021; Takishita et al., 1999). Arrows representing chloroplast transfers involved in haptophyte history are highlighted. The cartoon below summarise these events.

stratified waters where nutrient availability is low (Alexander et al., 2015; Endo, Ogata, & Suzuki, 2018). Nonetheless, some haptophyte species are known to be bloom-forming (e.g., *Emiliania, Gephyrocapsa* and *Phaeocystis* in temperate, tropical and polar regions, respectively) when light and nutrient conditions become favourable (Søgaard et al., 2021; Vincent, Sheyn, Porat, Schatz, & Vardi, 2021). During these blooming episodes, haptophytes may be able to outcompete diatoms, which are considered principally as r-strategists, for available nutrients (Alexander et al., 2015). These bloom dynamics are typically curtailed by viral infection and lysis (Frada, Probert, Allen, Wilson, & de Vargas, 2008; Vincent et al., 2021). Other haptophyte groups, including species adapted to oligotrophic conditions, may utilize mixotrophy to support their growth, with bacterivory observed in some laboratory species (*Chysocampanula spinifera, Chrysochromulina leadbeteri and Haptolina hirta*) (Al-Nahdi & Sayegh, 2021; Johnsen et al., 1999; Kawachi et al., 1991) and inferred in many others from environmental data (Unrein et al., 2014; Zubkov & Tarran, 2008).

A complex and ancient origin of the haptophyte genome

Within the eukaryotes, haptophytes are distantly related to plants, animals and other major algal groups such as red and green algae, cryptomonads, dinoflagellates and ochrophytes (Fig 3). The exact phylogenetic position of haptophytes remains the subject of debate, although multigene phylogenies typically resolve haptophytes as either sister group to the SAR clade of stramenopiles (including ochrophytes), alveolates (including dinoflagellates) and rhizaria (Burki et al., 2016; Strassert, Irisarri, Williams, & Burki, 2021) or as part of a CCTH group (previously termed "Hacrobia"), including cryptomonads, which show greater evolutionary proximity to green algae, red algae and plants (Okamoto, Chantangsi, Horák, Leander, & Keeling, 2009) (Fig 2). The single closest relatives to the haptophytes, the centrohelids, are a group of obligately heterotrophic phagotrophs (Burki et al., 2016; Cavalier-Smith & von der Heyden, 2007), with no documented evolutionary history of a chloroplast (Fig 2).

Haptophyte chloroplasts are surrounded by four membranes and contain the accessory pigment chlorophyll *c*, both of which are characteristic of chloroplasts of secondary or higher red (eukaryotic) algal endosymbiotic origin (Andersen, 2004). These features indeed unify cryptomonad, ochrophyte and dinoflagellate chloroplasts (although dinoflagellate chloroplasts are only surrounded by three membranes), which have historically been proposed to have a common endosymbiotic origin within the red algae (Cavalier-Smith, 1999). However, more recent phylogenomic studies have indicated independent endosymbiotic acquisitions of the cryptomonad, haptophyte, ochrophyte and dinoflagellate chloroplasts, with the haptophyte chloroplast (and its associated genome) most likely having been acquired from a cryptomonad alga that itself possessed a secondary (or higher) red algal chloroplast, indicating a tertiary origin (Rice & Palmer, 2006; Stiller et al., 2014) (Fig 2).

This already complicated evolutionary picture is further muddied by equivalent studies of chloroplast-associated genes encoded in the haptophyte nuclear genome, which suggest that the haptophyte host may have also performed a tertiary (or higher) endosymbiotic acquisition of an ochrophyte, most likely within

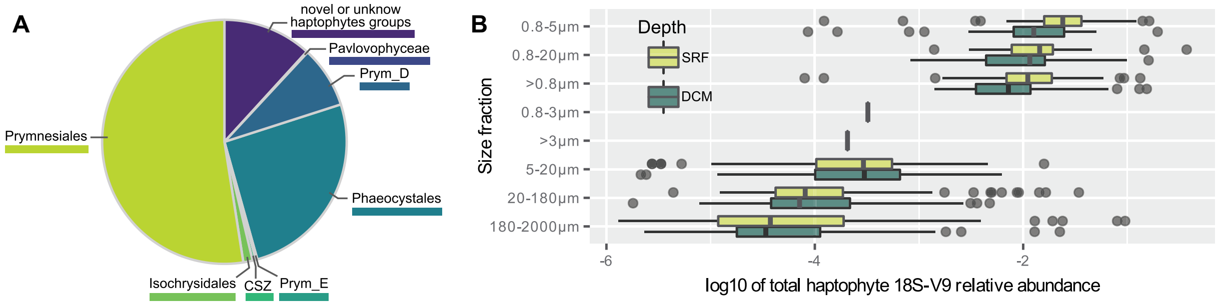


Figure 3. Total occurrence of haptophytes in *Tara* Oceans data. **A**: Pie chart of the total contribution of five major cultured haptophyte orders (Pavlovophyceae, Phaeocystales, Prymnesiales, the CSZ clade and Isochrysidales) alongside three uncultured groups (corresponding to Prymnesiophyte clades D and E and other unknown groups) (X. L. Shi et al., 2009) to total 18S V9 sequence abundances across all stations and size fractions within the *Tara* Oceans Barcode Atlas (Ibarbalz et al., 2019; Vernette et al., 2021). **B**: Boxplot distributions of $\log_{10}$ haptophyte relative abundances across all stations for different combinations of depth and size fractions, demonstrating negligible differences between surface and DCM stations but much greater relative abundances in small (0.8–20 μm) than larger (5–2000 μm) size fractions.

the pelagophytes or dictyochophytes (Dorrell et al., 2017, 2021b; Stiller et al., 2014) (Fig 2). This complicated series of evolutionary events renders the haptophyte cell an evolutionary mishmash, containing (most probably) genes from the direct ancestors of the eukaryotic nuclear genome, alongside genes of proteobacterial (mitochondrial) and cyanobacterial (chloroplast) origin, genes from the red algal ancestor of the haptophyte chloroplast and genes from whatever other eukaryotic algal groups (cryptomonads, pelagophytes) this chloroplast has passed through on its way to being acquired by the haptophyte cell (Fig 2). Finally, haptophytes have passed their chloroplasts on at least once by further (quaternary) endosymbioses to dinoflagellates within the Kareniaceae through at least one acquisition (Hehenberger, Gast, & Keeling, 2019; Takahashi, Benico, Lum, & Iwataki, 2019; Takishita et al., 1999) and potentially other dinoflagellate groups (e.g., peridinin-containing species) as well (Dorrell et al., 2021b; Yoon, Hackett, & Bhattacharya, 2002).

The exact timing of all of these events, while not entirely certain, is most definitely ancient. Despite the appearance of haptophytes in the fossil records only 200–300 million years before the present (Medlin, Saez, & Young, 2008), molecular estimates place their divergence from centrohelids (and, indeed, acquisition or acquisitions of chloroplast) well into the pre-Cambrian (Liu, Aris-Brosou, Probert, & de Vargas, 2010; Strassert et al., 2021) and potentially stretching back to a billion years (Yoon, Hackett, Ciniglia, Pinto, & Bhattacharya, 2004) (Fig 2). The ancient separation (and distinct phylogenetic position) of haptophytes from other algal groups on the tree of life creates the rich potential for the evolution of unique genome contents and eco-physiological life strategies. Moreover, while the internal radiation dates of the haptophyte groups remain poorly understood, their ancient origin suggests an internal genetic diversity as great as the evolutionary distances between angiosperms and mosses or mammals and fish (Liu et al., 2010; Medlin et al., 2008) (Fig 2). Understanding the genetic and functional components of haptophyte diversity will be instrumental in explaining their varied distributions across the contemporary ocean.

Individual biology and eco-physiology of the major haptophyte orders

Phylogenetic data positions several major groups or «orders» within the haptophytes, each with very different biological strategies and biogeographical trends (Figs 1 c; 2). Here, we describe key features associated with five haptophyte orders that are well established in laboratory culture: the Pavlovales within the Pavlovophyceae and the Phaeocystales, Prymnesiales, the «CSZ» clade and Isochrysidales, which together form the Prymnesiophyceae.

The Pavlovophyceae is the most distantly related haptophyte group, with the divergence between these and the Prymnesiophyceae occurring potentially over five hundred million years ago (Parfrey, Lahr, Knoll, & Katz, 2011; Strassert et al., 2021) (Fig 1c, panels I, ii; Fig 2). The only described order within the Pavlovophyceae, the Pavlovales, bears one chloroplast and two flagella of unequal lengths. An eyespot is often present in the chloroplast and plays a role in photoreception (Foster & Smyth, 1980; J. C. Green, 1980) (Fig 1c, panel i). They are covered in small scales ("knob scales") that are diversely arranged on the cell surface depending on the species. These organisms are referenced as common in nearshore areas, estuaries and coastal environments (Simon, Lopez-Garcia, Moreira, & Jardillier, 2013). They are found in the phytoplankton fraction or in benthic communities, and some species have also been described in lakes and ponds, with one species (*Diacronema* sp. CCMP2436) even known from ice-influenced Arctic ecosystems (Dorrell et al., 2021a). However, they are described as scarce in the open ocean and there are no cultured strains originating from oceanic gyres among the 49 available in the Roscoff culture collection (accessed 20/10/2021) (Vaulot, Le Gall, Marie, Guillou, & Partensky, 2004).

Prymnesiophyceae are divided in several orders (Fig 2), the main ones being the Phaeocystales, the Prymnesiales, the CSZ (Coccolithales-Syracosphaerales-Zygodiscales) and the Isochrysidales. Early in their evolutionary history, a close relative of the common ancestor of Prymnesiophyceae was acquired via a tertiary (or higher) endosymbiotic event by a specific dinoflagellate group, the Kareniaceae, although free-living extant relatives of this lineage have yet to be identified (Kawachi et al., 2021; Takishita et al., 1999) (Fig 2).

The first evolutionary split within cultured Prymnesiophyceae occurred between the Phaeocystales and all other groups (Fig 2). Phaeocystales are planktonic, photosynthetic, mixotrophic and non-calcified organisms encountered in marine environments worldwide, from tropical to polar latitudes (Assmy et al., 2017; Schoemann, Becquevort, Stefels, Rousseau, & Lancelot, 2005; Vogt et al., 2012). They include motile and non-motile free-living species, as well as several species forming gelatinous ("palmelloid"), round, motile or non-motile colonies during specific life stages (Fig 1c; panel iii).

Some species produce massive blooms (*P.* antarctica and *P. pouchetti*) (Assmy et al., 2017; DiTullio et al., 2000), which may be harmful for marine animals (*P. globosa*) (Wang, Song, Wang, & Chen, 2021). Some Phaeocystales are also known to be endosymbionts of zooplanctonic Acantharians, within the Rhizaria, and non-photosynthetic dinoflagellates related to but distinct from the Kareniaceae (Fig 2) (Decelle et al., 2019; Hehenberger et al., 2019).

Prymnesiales is an order comprising the *Chrysochomulinaceae* and the *Prymnesiaceae* families (Fig 1c, panels iv, v). While not unified by 18S rDNA phylogenies, they are typically resolved as monophyletic by concatenated multigene trees (Dorrell et al., 2021a; Strassert et al., 2021). They are found in all marine and some freshwater environments (Deodato et al., 2019), with select species capable of forming blooms. Extraordinary metabolic diversity is found within the Prymnesiales, including many species with siliceous biomineralization (*Prymnesium polylepis*), and/or toxin-producing (*Prymnesium parvum*) capabilities (Blossom et al., 2014; Yoshida et al., 2006). Several studies describe the relatively high abundance and the diversity of Prymnesiales in the oligotrophic oceans compared to other phytoplanktonic eukaryotes (Fuller et al., 2006). This success may in part be due to their tremendous capacity for mixotrophy, with active phagocytic cells able to ingest particles using characteristically long haptonemas (Anderson et al., 2018; Liu et al., 2009) (Fig 1c, panel iv).

The CSZ clade is a group containing members of the *Coccolithales, Syracosphaerales* and *Zygodiscales*, which may either be monophyletic or paraphyletic excluding the Isochrysidales (below) (De Vargas, Aubry, Probert, & Young, 2007; Dorrell et al., 2021a) (Fig 1c, panels vi, vii; Fig 2). This group is composed of calcifying species, including aesthetically renowned members (e.g., *Coccolithus pelagicus*), and are mainly free planktonic cells in marine ecosystems (Edvardsen, Egge, & Vaulot, 2016). Because of the density of their coccoliths, these algae sink at a higher rate than non-mineralized species and thus play an important role in the biological carbon pump and nutrient export from the photic zone (Guidi et al., 2016). Some members of this group (e.g., *Braarudosphaera*) possess diminished cyanobacterial endosymbionts (within the UCYN-A group), which lack the capacity to perform photosynthesis independently. The cyanobacteria provide *Braarudosphaera* with fixed organic nitrogen in exchange for organic carbon (Mills et al., 2020; Thompson et al., 2012).

Finally, the Isochrysidales is an order that includes both calcified and non-calcified species. The most studied species within this group is the coccolith-bearing *Emiliana huxleyi*, which is known to form massive blooms during summer and autumn, sometimes visible from space (Read et al., 2013; Zondervan, 2007) (Fig 1c, panels viii, ix). The *Emiliania* genus complex contains both haplo-diploid members (*Emiliania*) and permanently diploid members (*Gephyrocapsa*), with contrasting latitudinal preferences. *Emiliania* prefers more temperate waters, whereas *Gephyrocapsa* resides primarily in tropical environments (Bendif et al., 2019; Liu et al., 2018). These distributions are related to the prevalence of the *Emiliania* virus that only infects diploid *Emiliania* cells but shows heat-dependent inactivation. This constrains the obligately diploid subspecies to subtropical waters (Bendif et al., 2019). The Isochrysidales order also includes terrestrial and euryhaline species (*Ruttnera*) (Green & Parke, 1974), along with other species (*Tisochrysis, Chrysotila*). Members of this group of haptophytes produce the very long-chain (C_{35} -C_{39}) fatty acids known as alkenones, the saturation index of which has long since been used as a proxy to reconstruct paleotemperatures. Their unusual fatty acids have also been targeted for industrial purposes including biofuels, cosmetic products and novel materials (Araie et al., 2018; Richter et al., 2019; Shi, 2019).

Uncultured and emergent haptophyte orders

Historical understanding of haptophyte biology and diversity has been limited by their cultivability. Many environmentally important species remain difficult to establish in culture either due to undiscovered auxotrophies (Absolon, Smith, & Helliwell, 2019; Nef et al., 2019) or nutrient requirements that may be different (e.g., more oligotrophic) than standard laboratory media (Andersen, 2005). Environmental sequencing is therefore an important tool for understanding haptophyte biodiversity in the real world, giving us insight into the environmental abundances of major orders. It allows the molecular classification of species complexes that cannot be differentiated via morphological synapomorphies (Ibarbalz et al., 2019) and the identification of new, uncultured haptophyte groups, which may ultimately then become established as laboratory model organisms (Kawachi et al., 2021).

One such example is the Rappemonads, a lineage first detected in environmental sequence data surveys from the Pacific Ocean (Rappé, Suzuki, Vergin, & Giovannoni, 1998; Shi, Marie, Jardillier, Scanlan, & Vaulot, 2009) and subsequently detected also in the North Atlantic (Kim et al., 2011) and globally within the temperate and sub-tropical ocean, as part of the *Tara* Oceans Expedition (Kawachi et al., 2021). Phylogenetic analysis of the rappemonad rDNA operon suggested that this lineage forms either a sister group to all Prymnesiophytes to the exclusion of Pavlovophyceae or to all haptophytes entirely (Kim et al., 2011; Shi et al., 2009), and fluorescence microscopy of environmental rappemonad isolates sorted by FISH revealed the presence of two to four chloroplasts per cell, indicating either photosynthetic capacity or the retention of stable photosynthetic endosymbionts (Kim et al., 2011). Most recently, a previously isolated haptophyte species, *Pavlomulina ranunculiformis*, has been re-identified as the first laboratory rappemonad culture by Kawachi et al. (2021). The *Pavlomulina* cell contains ultrastructural synapomorphies that link it to Prymnesiophyceae and not Pavlovophyceae (e.g., flagella of equal length, presence of chlorophyll *c*3 and not *c*1 and absence of an eyespot) but lacks 19' hexanoyl-fucoxanthin as found in core Prymnesiophycean groups (Kawachi et al., 2021) and indeed contains structural features (four plastids per cell) unique among haptophytes (Fig 1c, panel x). Phylogenetic analysis of the *Pavlomulina* organelle (plastid and mitochondria) genomes confirms a probable sister group position to all Prymnesiophyceae (Kawachi et al., 2021), to the exclusion of the ancestor of the Kareniacean dinoflagellate plastid (Fig 2), although alternative phylogenetic positions (e.g., a sister-group position to all Pavlovophyceae) have also been proposed (Song, Chen, Liu, & Chen, 2021).

Beyond rappemonads, there are multiple other groups of haptophytes described from environmental data that have yet to be established in laboratory culture. Choi et al. (2017) reported the occurrence of two novel haptophyte groups, termed Deep-Branching Plastid Lineages (DPL) 1 and 2, from analysis of chloroplast environmental 16S rDNA. Both DPL groups are of low abundance in *Tara* Oceans data but show global distributions across temperate and sub-tropical stations as per rappemonads. DPL1 appears to resolve as an immediate sister group to the Prymnesiophyceae to the exclusion of rappemonads and the ancestors of the Kareniaceaen plastid, whereas DPL2 may form a sister group to all cultured haptophyte orders (Choi et al., 2017). The veracity of these positions rests on the isolation and cultivation of DPL1 and DPL2 group species, and particularly the resolution of their phylogenetic relationships to Prymnesiophyceae, haptophytes at large and potentially even centrohelids via multigene phylogeny (Burki et al., 2016; Strassert et al., 2021). Additional haptophyte groups, detectable within 18S rDNA barcodes from the *Tara* Oceans Expedition, include Clade U2, projected to form a sister group to all Prymnesiophyceae, but distinct from rappemonads and potentially corresponding to a DPL group (X. L. Shi et al., 2009) and Prymnesiophyte Clades D and E, which may correspond to new orders that are sisters to all other Prymnesiophyceae or to all Prymnesiophyceae excluding Phaeocystales, following previous phylogenetic analyses of the complete 18S rDNA sequence (Wu, Huang, Liao, & Sun, 2014; Wu, Wang, Liao, & Huang, 2015).

Nonetheless, further new haptophyte species, genera or orders may still await our discovery in environmental sequence data, dependent on seasonality of sampling and masking members of the rare marine biosphere by sequences from more abundant organisms.

Global distribution of the major haptophyte orders in *Tara* Oceans data

While the biogeographical distribution of individual haptophyte species (Bendif et al., 2016; von Dassow et al., 2015) and community composition across spatial and temporal transects have been studied and reviewed elsewhere (Alexander et al., 2015; Egge et al., 2015), relatively little is known to date about the distributions of the specific orders on a global scale. Here, we attempt to resolve this knowledge gap by considering the patterns of relative abundance for eight distinct haptophyte categories (Pavlovophyceae, Phaeocystales, Prymnesiales, CSZ clade, Isochrysidales, clades D and E prymnesiophyceae and haptophytes *incertae sedis*), following the classification scheme erected using the PR^2 (protist ribosomal rDNA) Reference Database in 18S V9 meta-barcode data within the *Tara* Oceans Barcode Atlas (Guillou et al., 2013; Vernette et al., 2021) (Figs 3–5). We acknowledge that these insights are preliminary and should ideally be supported with additional barcode data (*e.g.*,, 18S v4 or 16S V4-V5), alongside meta-transcriptome-based methods for evaluating haptophyte relative abundances, particularly given the limited efficiency of common V9 primers in amplifying 18S rDNA from specific haptophyte orders (Liu et al., 2009; Karlusich et al., 2022).

Abundance, size and depth distributions: Across all *Tara* Oceans stations, Prymnesiophyceae represented approximately 98% of haptophyte total relative abundance, with a slight majority (52%) from Prymnesiales. Phaeocystales also contributed a substantial proportion (26%), with <2% total relative abundance attributable each to Pavlovophyceae, the CSZ clade and Isochrysidales (Fig 3a). Considering their total relative abundances across *Tara* Oceans as previously calculated (Kawachi et al., 2021) (3.2×10^{-4} per million total ribotypes), Rappemonads represent only 0.004% of total haptophyte abundance. Relatively little differences in relative abundances exist between surface and deep chlorophyll maximum (DCM) stations, reflecting

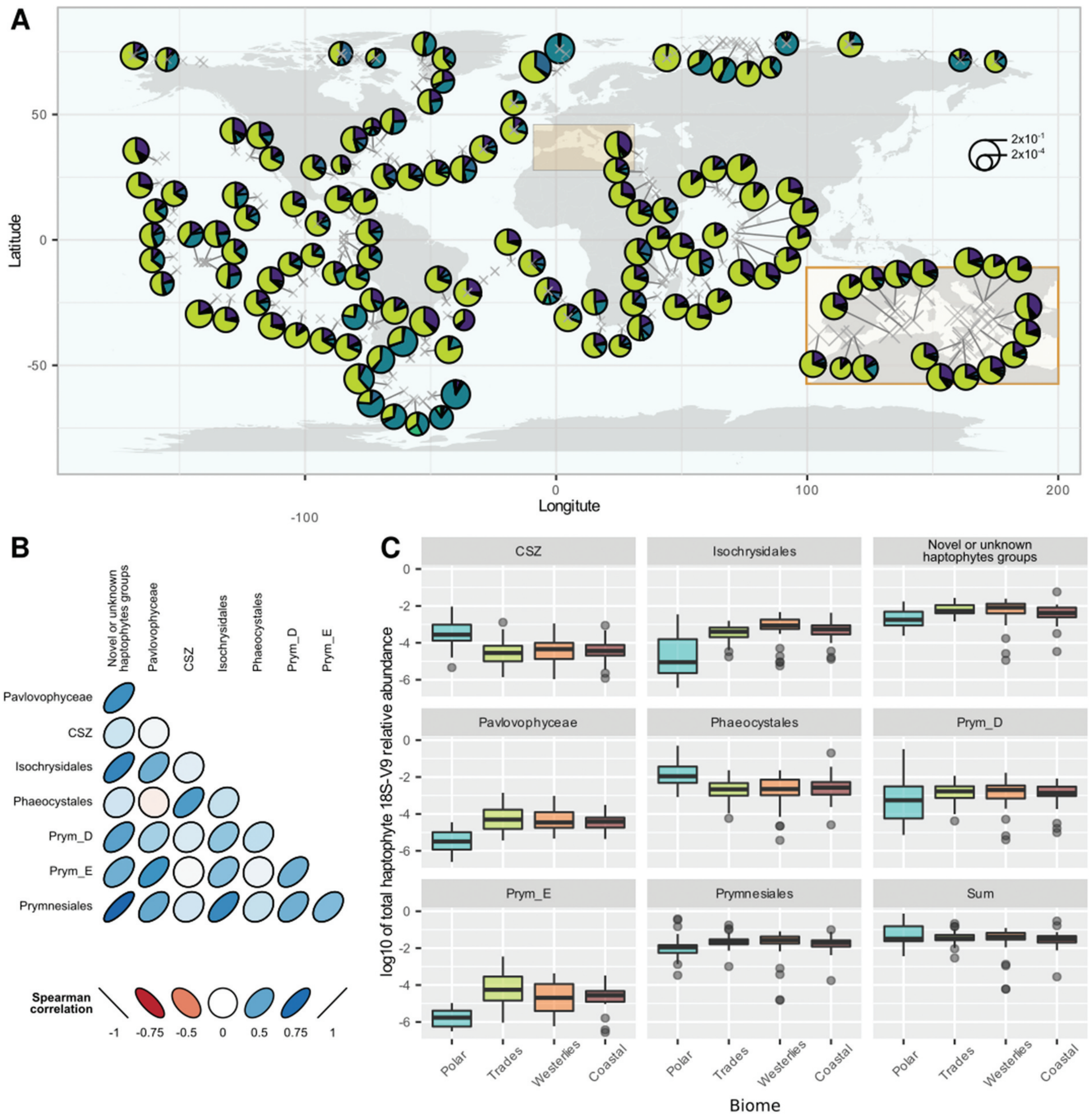


Figure 4. Biogeographical distributions of haptophyte orders across the world ocean. **A**: Map of *Tara* Oceans stations, showing the proportional contribution of different haptophyte orders at the surface and in nano- and pico- (>0.8, 0.8–3, 0.8–5, 0.8–20 μm) size fraction combinations to each station (Vernette et al., 2021). **B**: Heatmap, showing the strength of co-association of different haptophyte orders to one another, as per previous studies (Lima-Mendez et al., 2015; Vincent & Bowler, 2020), across all station, depth and size fraction combinations, showing separate partitioning of CSZ clade and Phaeocystalean haptophytes to all other haptophyte orders. **C**: Boxplots of relative haptophyte abundances in nano- and pico-size fractions of each haptophyte order across *Tara* Oceans stations divided into biomes per Longhurst (2006), demonstrating in particular contrasting enrichments (CSZ clade and Phaeocystales) and scarcity (Pavlovophyceae and Isochrysidales) in Polar Biome occurrence of different haptophyte orders.

haptophyte physiological flexibility between nutrient-limited (oligotrophic) surface habitats and light-limited (mesotrophic) DCM habitats. However, substantial differences are apparent in haptophyte relative abundance across size fractions (Fig 3b). Typically, haptophyte relative abundance is greatest in nano- and picoplanktonic (0.8–20, 0.8–5 and 5–20 μm) size fractions, with approximately 100 times greater relative abundance than observed in the larger (5–20, 20–180 and 180–2000 μm) size fractions (Fig 3b). The disparity in haptophyte relative abundance across size fractions may hint to the greater relative diversity found in small

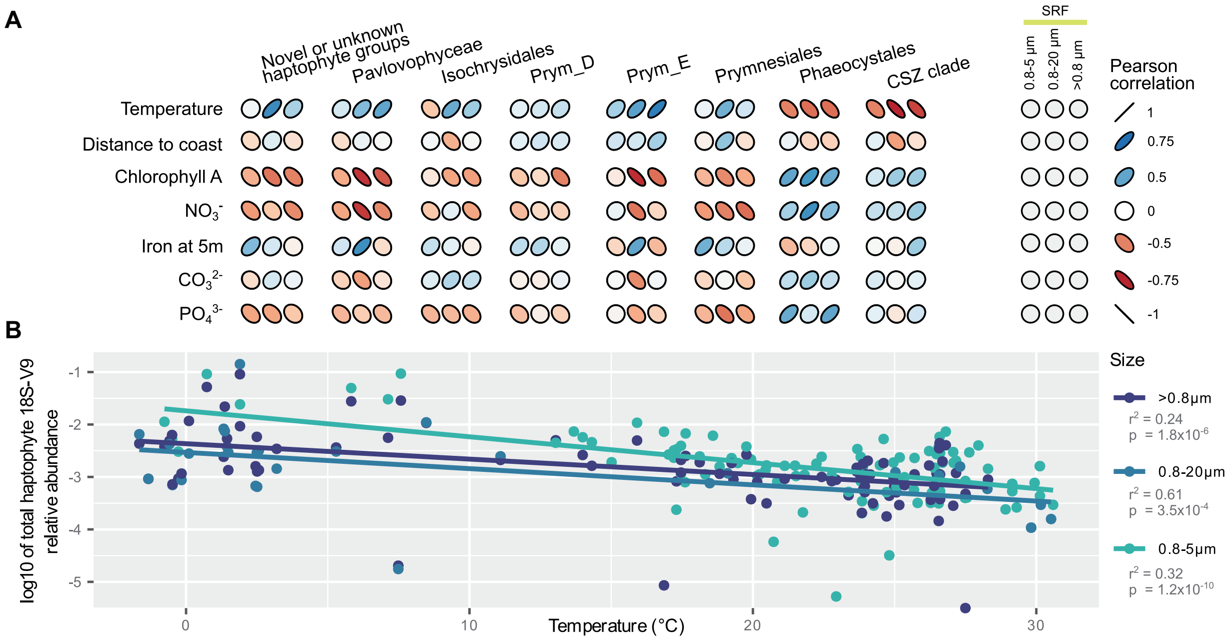


Figure 5. Environmental factors influencing haptophyte biogeographical distributions. **A**: Correlation heatmap, as per Fig 4b, of the correlation coefficients of the relative abundances in the 0.8–5, 0.8–20 and >0.8 µm size fractions of different haptophyte orders and seven exemplar environmental variables (temperature, coastal distance, estimated iron and measured nitrate, phosphate, carbonate and chlorophyll concentrations) in surface samples, as described in PANGAEA Pesant et al., 2015, across all *Tara* Oceans stations. Full outputs for all size fractions and environmental variables are provided in Supplementary table S1. **B**: Exemplar scatterplots of relative abundance of Phaeocystales in surface samples for three size fractions (0.8–5, 0.8–20 and >0.8 µm), demonstrating greater relative abundances in stations with low environmental temperatures.

and non-calcifying species than larger, colonial and calcareous haptophytes and/or a greater occurrence of free-living over symbiotic species in marine communities (Liu et al., 2009).

Spatial distributions: Distinctive patterns in haptophyte distribution were observed across the *Tara* Oceans dataset (Fig 4a). Most stations were dominated by Prymnesiales, particularly members of *Chrysochromulina sp.* (70% of Prymnesiales relative abundance), with the exception of stations in the Southern Ocean, which principally contained ribotypes from the Phaeocystales (Fig 4a). Phaeocystales additionally showed strong presences in the Arctic and Southern Atlantic Oceans, although the vast majority of these ribotypes did not correspond to classified species (*Phaeocystis antarctica, cordata, globosa;* Fig 4a). Other haptophyte lineages (CSZ, Isochrysidales, Pavlovophyceae) showed globally low abundances, albeit with some regional asymmetries (e.g., greater relative occurrence of Isochrysidales in stations along the Gulf Stream and the North Atlantic Drift; Fig 4a). Reflecting these asymmetries, Phaeocystales and CSZ clade haptophytes showed positive co-associations (i.e., co-preferences for polar oceanic regions) (Lima-Mendez et al., 2015) across *Tara* Oceans data, to the exclusion of Pavlovophyceae, Prymnesiales, Isochrysidales and uncultured lineages, which all showed varying degrees of positive co-occurrence to one another (Fig 4b).

We note the occurrence of haptophytes in most *Tara* Oceans stations, underscoring their global importance to the marine ecosystem, although with striking asymmetries between stations (Fig 4c). Using partitions (Polar, Coastal, Trades and Westerlies) for each station defined in previous studies (Longhurst, 2006; Sommeria-Klein et al., 2021), we noted more than two to five times greater relative abundances of haptophytes in Polar stations than in other regions (Fig 4c). No obvious seasonal effect on abundances were observed across the entire dataset, although we note that different ocean regions were sampled in different seasons. For example, Arctic sampling was conducted during spring, summer and autumn, and Southern Ocean sampling occurred during spring and summer, when blooms are more likely to happen. Different haptophyte orders showed preferential accumulations in different oceanic biomes, likely reflective of both environmental preferences and sampling time relative to different blooming events. Pavlovophyceae appeared to be 10 to 25 times less abundant on average in Polar biomes relative to temperate and tropical (Coastal, Westerlies and Trades biomes) (Fig 4c). Isochrysidales also tends to be less abundant in Polar biomes, whereas Prymnesiales showed the same relative abundances in all biomes and regions (Fig. 4c). In contrast, Phaeocystales and CSZ clade haptophytes were 10–15 times more abundant in Polar than other biomes, perhaps due to blooming events (Fig 4c).

Correlations to environmental variables: Finally, we calculated Spearman correlations between the relative abundances of each major haptophyte order across physical variables in *Tara* Ocean stations (Pesant et al., 2015). In Fig 5, we present correlation heatmaps for the >0.8, 0.8–5, and 0.8–20 µm size fractions and surface depths, for seven environmental characteristics: 1) temperature, 2) coastal distance to distinguish coastal versus offshore species, 3) NO_3^- and 4) PO_4^- concentrations, as a measure of oligotrophy/eutrophy), 5) CO_3^{2-} levels, as an indicator of coccolith formation in calcareous species (Zondervan, 2007) 6) chlorophyll *a*, as an index of primary production, and 7) estimated Fe concentration at 5 m depth, which is typically negatively correlated with diatom abundance, and hence might serve as a proxy of haptophyte/diatom co-occurrence in the open ocean (Caputi et al., 2019). Full tabulated correlations are found in Table S1.

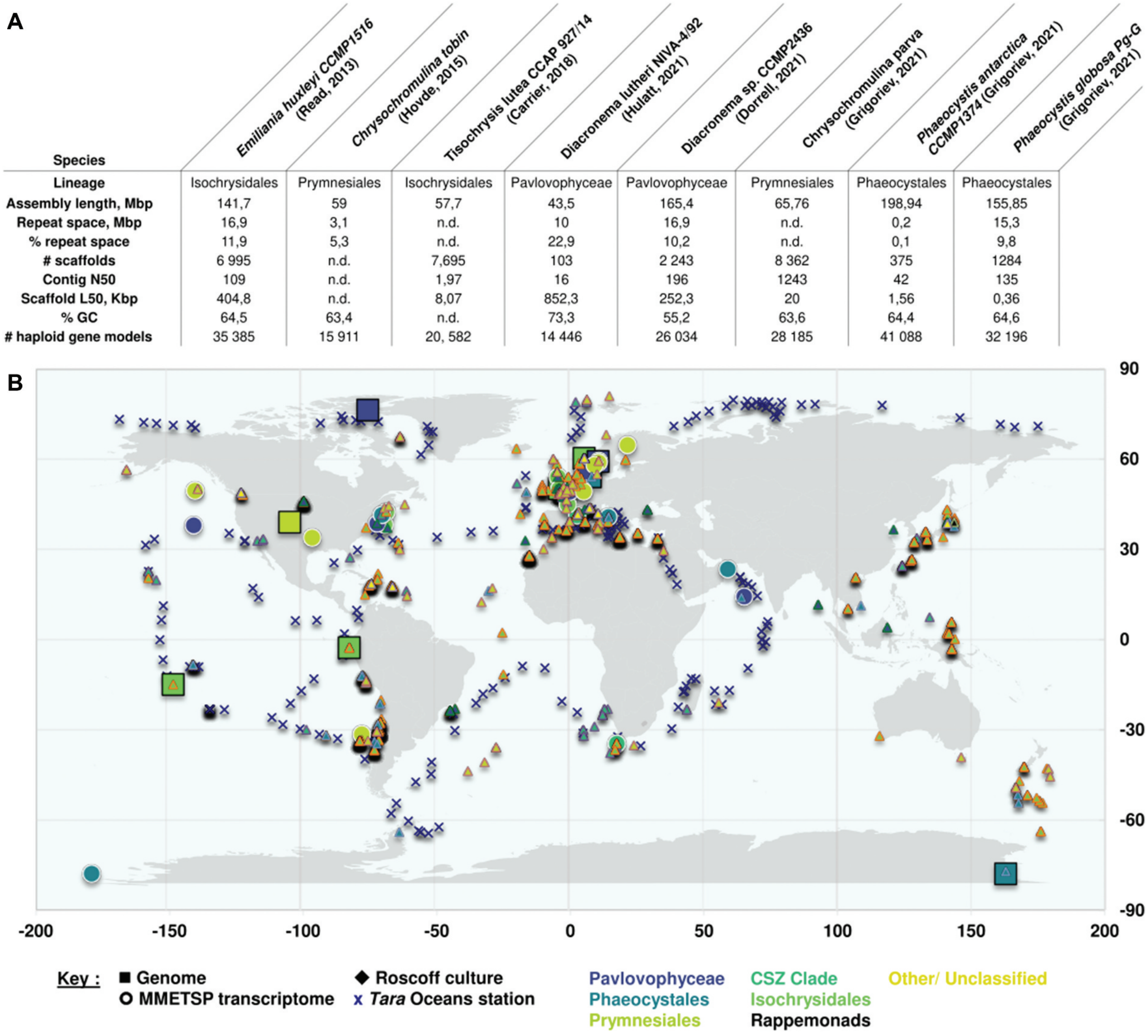
The relative abundances of Pavlovophyceae, Prymnesiales, Isochrysidales and basal haptophyte orders negatively correlated with nutrients (nitrate and phosphate) and chlorophyll content and positively correlated with temperature (Fig 5a). Pavlovophyceae showed the strongest negative correlations to nitrate and phosphate and a strong positive correlation (in 0.8–20 µm size fractions, covering the Arctic Ocean, Red Sea and Indian Ocean) to iron. While Pavlovophyceae showed positive correlations to temperature in surface fractions, mixed trends were observed within the DCM (Supplementary table S1), suggesting that this lineage is likely to be ecologically heterogeneous, including previously undescribed open-ocean and polar components (Bendif et al., 2011; Dorrell et al., 2021a). Contrasting with the historical understanding of Pavlovophyceae as a predominantly coastal order, we did not globally detect an enrichment in their abundance in coastal *Tara* Oceans stations, considering either biome distribution or coastal distance (Figs 4c, 5a) (De Vargas et al., 2007). Prymnesiales exhibit positive correlations with both temperature and (for 0.8–20 µm size fractions) coastal distance, suggesting their importance to subtropical and oligotrophic open-ocean waters (Fig 5a). In contrast to Pavlovophyceae and Prymnesiales, Isochrysidales showed an additional positive correlation (particularly in DCM fractions) to carbonate concentrations, reflecting the global importance of calcification for this lineage (Kegel, John, Valentin, & Frickenhaus, 2013; O'Dea et al., 2014; Read et al.,2013; Zondervan, 2007) (Fig 5a; Supplementary table S1).

In contrast, both the CSZ clade and Phaeocystales are more abundant in cold, nutrient and chlorophyll-rich waters, consistent with their positive co-association with one another and preferences with polar biomes (Figs 4b, c, a). These trends were particularly notable in Phaeocystales, which were most abundant (relative abundance $>10^{-3}$) in stations with environmental temperatures below 8°C (Fig 5b). Despite the reputation of CSZ clade haptophytes as calcifiers, they showed a strong negative correlation to CO_3^- concentrations in DCM size fractions. This raises questions concerning the relative environmental abundances of non-calcareous species (e.g., silicifiers such as *Prymnesium neolepis*) (Yoshida et al., 2006) or cells (e.g., haploid members of haplo-diploid species) within this clade (Frada et al., 2008) (Supplementary table S1).

Knowns and unknowns of haptophyte genomes, transcriptomes and environmental genomics

Genomic knowledge of haptophytes started with the sequencing of the «pan» genome (13 globally distributed strains, with four selected for deep sequencing) of the Isochrysidalean *Emiliania huxleyi* in 2013 (Read et al.,2013). The *Emiliania* genome, as a typical haptophyte genome assembly, comprises 141.7 Mbp sequences assembled over 6,995 scaffolds, with 35,385 annotated gene models (Read et al.,2013) (Fig 6a). Comparative genomic analysis of the four *Emiliania* strains revealed considerable intra-specific variation, with as much as one-tenth of the protein coding genome unique to individual strains and highlighting the importance of micro-evolution for understanding the genetic diversification of haptophytes and marine algae in general (Blanc-Mathieu et al., 2017; Rastogi et al., 2020; Read et al.,2013).

Subsequent haptophyte genome projects have included the oleaginous Prymnesialean species *Chrysochromulina tobinii* and *Chrysochromulina parva* (Hovde et al., 2015), the temperate and Arctic Pavlovophyceae *Diacronema lutheri* NIVA-4/92 and *Diacronema sp.* CCMP2436 (Dorrell et al., 2021a; Hulatt, Wijffels, & Posewitz, 2021); the alkenone-producing Isochrysidalean *Tisochrysis lutea;* and *Phaeocystis globosa* and *P. antarctica*, respectively, tropical and Antarctic members of the Phaeocystales (Fig 6a) (Grigoriev et al., 2021; Nelson et al., 2021).



Species	*Emiliania huxleyi* CCMP1516 (Read, 2013)	*Chrysochromulina tobin* (Hovde, 2015)	Tisochrysis lutea CCAP 927/14 (Carrier, 2018)	Diacronema lutheri NIVA-4/92 (Hulatt, 2021)	Diacronema sp. CCMP2436 (Dorrell, 2021)	Chrysochromulina parva (Grigoriev, 2021)	*Phaeocystis antarctica* CCMP1374 (Grigoriev, 2021)	*Phaeocystis globosa* Pg-G (Grigoriev, 2021)
Lineage	Isochrysidales	Prymnesiales	Isochrysidales	Pavlovophyceae	Pavlovophyceae	Prymnesiales	Phaeocystales	Phaeocystales
Assembly length, Mbp	141,7	59	57,7	43,5	165,4	65,76	198,94	155,85
Repeat space, Mbp	16,9	3,1	n.d.	10	16,9	n.d.	0,2	15,3
% repeat space	11,9	5,3	n.d.	22,9	10,2	n.d.	0,1	9,8
# scaffolds	6 995	n.d.	7,695	103	2 243	8 362	375	1284
Contig N50	109	n.d.	1,97	16	196	1243	42	135
Scaffold L50, Kbp	404,8	n.d.	8,07	852,3	252,3	20	1,56	0,36
% GC	64,5	63,4	n.d.	73,3	55,2	63,6	64,4	64,6
# haploid gene models	35 385	15 911	20, 582	14 446	26 034	28 185	41 088	32 196

Figure 6. Current and missing understanding in haptophyte genomic diversity. A: Tabulated features of completed haptophyte genomes, as described in Dorrell et al. (2021a), Grigoriev et al. (2021), Marinov et al. (2021), Nelson et al. (2021) and elsewhere. **B**: Biogeographical distribution of the isolation sites of haptophyte genomes and MMETSP transcriptomes (Keeling et al., 2014), shaded by haptophyte order, compared to cultured haptophyte strains available within the Roscoff Culture Collection and sites sampled within the *Tara* Oceans expedition (Ibarbalz et al., 2019; Vaulot et al., 2004). Although the majority ofcultured species have been isolated from sites around continental North America and Europe, cultured and environmental resourcesexist for profiling haptophyte genomic diversity on a global scale.

These targeted genomes have been supplemented by the inclusion of haptophytes in multi-species sequencing efforts, e.g., transcriptomes for 32 haptophyte species within the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP) (Keeling et al., 2014). Most recently, Delmont et al. have published assemblies for 92 haptophyte Metagenome-Assembled Genomes (MAGs) from *Tara* Oceans data, of uncertain taxonomy but presumably enriched in environmentally abundant genera (*Chrysochromulina, Phaeocystis*) (Delmont et al., 2022).

The rapid and exponential increase in genomic resources available, alongside quantitative estimates of gene expression obtained both from transcriptomes of model species (e.g., *Emiliania*) (Bochenek et al., 2013; Hernández Limón et al., 2020; Vincent et al., 2021) and from meta-genomic and meta-transcriptomic datasets such as *Tara* Oceans (Caputi et al., 2019; Carradec et al., 2018), provides an unprecedented opportunity for understanding fundamental principles in haptophyte biology and diversification. Here, we outline some of the questions that may now be addressed by haptophyte genomic and environmental resources and knowledge gaps in haptophyte genomic diversity that may be addressed by future studies.

Functional evolution of the chimeric haptophyte chloroplast

Although the haptophyte chloroplast genome is most closely related to those of cryptomonads (Rice & Palmer, 2006), it is supported by nucleus-encoded and chloroplast-targeted proteins most probably derived from a pelagophyte or dictyochophyte within the ochrophytes (Dorrell et al., 2017; Stiller et al., 2014) (Fig 2). Resolving the exact contributions of both these donors to the early evolution of haptophytes rests not only on denser genomic sampling of early-diverging haptophyte groups (e.g., pavlovophyceae and rappemonads) (Kawachi et al., 2021) but also on comparative genomic analyses of the immediate sister groups of haptophytes (centrohelids) (Burki et al., 2016), cryptomonads (*Goniomonas*) (Cenci et al., 2018) and pelagophytes and dictyochophytes (e.g., pinguiophytes) (Di Franco, Baurain, Glöckner, Melkonian, & Phillippe, 2022). This would help to identify exactly which sets of genes have been transferred between each group of organisms, at what time point and in what direction. Identifying the specific connection between these gene transfers and the haptophyte chloroplast may only be possible following experimental resolution of the haptophyte chloroplast proteome, as recently accomplished in other algal lineages such as euglenids (Vanclová et al., 2020) and diatoms (Schober, Flori, Finazzi, Kroth, & Bártulos, 2018) and/or environmental and experimental characterization of key chloroplast proteins using *Tara* Oceans data (Carradec et al., 2018) and emergent transformable haptophyte model systems (Faktorová et al., 2020).

Diversity of haptophyte cell biological systems and metabolism

As alluded to above, the plastid is far from the only interesting cell biological facet of haptophytes that can be informed by genomic data (Eikrem et al., 2017). The haptonema, for example, is an enigmatic structure associated with predation/phagocytosis and unique to haptophytes. Investigating the complement of cytoskeletal components or membrane-trafficking proteins uniquely encoded in haptophytes could well suggest candidate proteins to serve as handles for downstream molecular cell biological studies. The cell biology of coccolith formation and the haptophyte-specific coccolith vesicle is similarly mysterious (González & Read, 2005; Nam, Park, Lee, & Jin, 2019) but could be approached in the same manner, in this case by looking for components present in coccolith forming versus non-mineralizing species. Genes involved in the production of alkenones could likewise be identified by comparisons of genomic complements of Isochrysidales (which produce them) to other haptophyte orders (which typically do not) (Araie et al., 2018; Richter et al., 2019; Shi, 2019). Finally, comparative genomics of laboratory species with obligate phototrophic versus mixotrophic life strategies may be used to identify marker genes whose presence or absence is associated with phagotrophy, and to develop predictive models of haptophyte metabolism either in newly cultured species or in environmental datasets (Bock et al., 2021).

Cross-referencing the prevalence of meta-genes encoding such components with the prevalence of environmental factors across the *Tara* Oceans dataset, for example, CO_3^{2-} abundance and pH in the case of coccolith-forming genes (Lohbeck, Riebesell, & Reusch, 2014), should also provide insight into the cellular modulation in response to the environment. Cross-referencing the expression levels of these genes to biological abundance data (e.g., V9 OTUs) might even allow the identification of possible biotic functions and constraints on haptophyte metabolic processes (Lima-Mendez et al., 2015; Vincent et al., 2021). For example, the expression of haptonema-associated genes (associated with phagocytosis) may show positive correlations with the abundance of potential prey species, whereas coccolith-associated gene expression may show correlations with the abundances of marine viruses (which have been shown in lab and in mesocosm experiments to modify coccolith formation dynamics) (Frada et al., 2008; Vincent et al., 2021). Understanding of the functional dynamics and regulation of each gene will be further aided by genomics-enabled exploration of non-coding and repetitive DNA in haptophyte genomes (Read et al., 2013), particularly with regard to epigenomic features such as histone and DNA modifications (Alaguraj Veluchamy et al., 2013, 2015) and long non-coding RNAs (de Carvalho, Sun, Bowler, & Chua, 2016), as have been explored in other marine algae such as diatoms.

Finally, it must be noted that a very large proportion of cell biological information comes from studies in animal and fungal model systems, and it is then assumed that the models of how cells work is widely applicable to all eukaryotes. However, emerging data show that while this is very often true, many aspects of cell biology and protein machinery, that are present in the broad diversity of eukaryotes, are not present in animal and fungal models (More, Klinger, Barlow, & Dacks, 2020). Several components of this evolutionary class (termed Jtnarlogs) have been identified in *E. huxleyi* (Hirst et al., 2014; Schlacht, Mowbrey, Elias, Kahn, & Dacks, 2013), raising questions as to their prevalence in haptophytes more generally. Having a more complete picture of what cellular machinery haptophytes possess jumpstarts downstream functional investigation not only in this lineage but also more broadly across the eukaryotic tree of life.

How does the environment structure haptophyte genome evolution?

An environmentally informed vision of haptophyte genomic content may reveal what micro-adaptations allow haptophytes to succeed in different marine environments on a global scale. Analysis of 18S rDNA V9 barcode data from *Tara* Oceans indicates that haptophytes show important spatial segregation and limitations to dispersal, principally defined by latitude, suggesting the probable occurrence of polar, temperate and subtropical-specific haptophyte assemblies and genomes (Sommeria-Klein et al., 2021). Indeed, a recent phylogenomic analysis of Arctic algal genome content, including the Pavlovophycean *Diacronema* sp. CCMP2436, identified the presence of substantial numbers of Arctic-specific genes unique to marine algae from this biome. These genes are shared between distantly related Arctic algal species by within-oceanic horizontal gene transfer to the exclusion of closely related species from lower latitudes (Dorrell et al., 2021a). These include ice-binding domain proteins, which allow psychrophilic algae to survive freezing temperatures and potentially to adhere to floating sea ice. The ice-binding protein genes in *Diacronema sp.* CCMP2436 have been acquired by horizontal transfer from an Arctic pelagophyte, whereas the Antarctic Phaeocystalean haptophyte *Phaeocystis antarctica* has independently acquired ice-binding proteins from Antarctic diatoms (Dorrell et al., 2021a; Raymond & Kim, 2012). Identifying what genes enable haptophyte adaptations to specific environmental habitats, and indeed under what conditions these genes are expressed, will be invaluable to understanding the pan-genomic variation of haptophytes with global environmental distributions (e.g., *Emiliania huxleyi*) (Bendif et al., 2016; von Dassow et al., 2015) and for identifying factors that may underpin responses of different haptophyte groups to rising sea temperatures (Chaffron et al., 2021).

Knowledge gaps and future directions

Despite the exponential increase in genomic resources available across the haptophytes at large, relatively little is still known about the genomic content of certain haptophyte groups. These include, necessarily, the nuclear genomes of uncultured (DPL, Prymnesiophyte Clades D and E) (Choi et al., 2017) or recently cultivated lineages (rappemonads) (Kawachi et al., 2021). At a functional level, no genomes have yet been sequenced for species known to engage in mixotrophy under laboratory conditions (e.g., *Chrysocampanula spinifera* and *Haptolina hirta*) (Anderson et al., 2018), toxic species (*Prymnesium parvum*), siliceous species (*Prymnesium neolepis*) (Yoshida et al., 2006) or species with a pulsatile vacuole, a specialized organelle which may allow adaptation to variable salinity environments (*Diacronema noctivaga, D. vlkianum* and Pavlova *granifera*) (Bendif et al., 2011).

At a global scale, existing haptophyte genomes and transcriptomes are necessarily biased towards temperate and coastal species, with much less known about the genomic content of tropical, polar and open ocean haptophyte groups (Fig 6b). This knowledge gap is nonetheless readily accessible, via the establishment of globally distributed haptophyte species in culture (e.g., the Roscoff Culture Collection, Fig 6b) (Vaulot et al., 2004) and indeed through metagenome-assembled genomes from the *Tara* Oceans dataset (Delmont et al., 2022). Additional challenges include the identification and sequencing of haptophytes from freshwater environments (Deodato et al., 2019; Singer et al., 2020), of which species are known but few are in culture; alongside species that are inaccessible, rare or fragile such as sea-ice, terrestrial and obligately non-photosynthetic haptophytes (Eikrem et al., 2017; Singer et al., 2020; Søgaard et al., 2021). Here, targeted sequencing of environmental samples, e.g., single-cell genomics, may be instrumental in closing the most recalcitrant gaps in our knowledge of haptophyte genomic diversity (Mangot et al., 2017; Seeleuthner et al., 2018).

Concluding remarks

Despite their great abundance within the oceans, and central position in eukaryotic algal evolutionary histories (Figs 1, 2), our understanding of haptophyte genomics, ecophysiology and diversity remains fragmented (Fig 6a). This reflects both the uncultivability of many ecologically prominent species, including entire haptophyte orders, alongside sampling biases both in terms of seasonality (relative to blooming events) and the geographical distribution of cultured species (biased towards continental oceans from temperate northern latitudes; Fig 6b).

Nonetheless, the increasing quantity of information produced from meta-genomic approaches, which can be correlated to quantitative environmental variables, may provide unprecedented opportunities to understand holistically the complexity and the diversity of haptophytes. For instance, the Tara Oceans expeditions may allow us to better understand the specificecological preferences of individual haptophyte orders, with some results questioning current knowledge of haptophyte ecology (Figs 3–5). Combining these meta-genomic data with both phylogenetic and molecular approaches, for example, emergent systems for haptophyte orders transformation and cell biology (Faktorová et al., 2020; Prasad et al., 2014), may allow us to unravel common features and specific adaptations between haptophyte orders, genus, species and ecotypes. These approaches may therefore cast insight into observed haptophyte geographical and seasonal patterns, alongside the real-world ecological significance of evolutionary and physiological phenomena observed within cultured species, such as chimeric chloroplast proteomes, flexible trophic life strategies and biogeographically structured genomes. Exploring the interplay between haptophyte genomes, meta-genomes and encoded functions may allow us to better understand the roles haptophytes play in oceanic biogeochemical dynamics, their centrality to marine ecosystems and food networks and their inferred sensitivity and resilience to environmental modifications induced by climate change.

Acknowledgments

The authors thank Margaux Carmichael, Priscillia Gourvil, Florence Le Gall, Ian Probert and Daniel Vaulot (Station Biologique de Roscoff), El Mahdi Bendif (University of Oxford) and Ryoma Kamikawa and Mami Nomura (University of Kyoto) for the provision of images used in Fig 1. This article is a contribution number 137 to *Tara* Oceans. The authors furthermore thank the groupof Chris Bowler (IBENS, Paris) for critical discussion of the interpretation of Tara data presented in this study.

Disclosure Statement

No potential competing interest was reported by the authors.

Funding

RGD acknowledges an ANR JCJC («PanArctica», ANR-21-CE02-0014-01, awarded 2021-2025) an ANR t-ERC (“ChloroMosaic”, awarded 2021-2023). RGD, BR and JBD further acknowledge a JGI CSP («The Haptophyte Pan-Genome», awarded 2021).

ORCID

Joel B. Dacks http://orcid.org/0000-0003-4520-5694
Richard G. Dorrell http://orcid.org/0000-0001-6263-9115

References

Absolon, D., Smith, A., & Helliwell, K. (2019). Shining a light on microbial dark matter: A role for the forgotten B vitamins in marine algal communities? *Access Microbiology*, *1*. doi:10.1099/acmi.ac2019.po0295

Al-Nahdi, M. S., & Sayegh, F. A. A. (2021). Feeding and bioremediation relationships between *Chrysochromulina* species and bacteria. *Advances in Environmental Biology*, *15*, 1–15.

Alexander, H., Rouco, M., Haley, S. T., Wilson, S. T., Karl, D. M., & Dyhrman, S. T. (2015). Functional group-specific traits drive phytoplankton dynamics in the oligotrophic ocean. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, E5972–5979. doi:10.1073/pnas.1518165112

Andersen, R. A., Saunders, G. W., Paskind, M. P., & Sexton, J. P. (1993). Ultrastructure and 18S ribosomal RNA gene sequence for *Pelagomonas calceolata* gen et sp. nov and the description of a new algal class, the Pelagophyceae. *Journal of Phycology*, *29*, 701–715. doi:10.1111/j.0022-3646.1993.00701.x

Andersen, R. A. (2004). Biology and systematics of heterokont and haptophyte algae. *American Journal of Botany*, *91*, 1508–1522. doi:10.3732/ajb.91.10.1508

Andersen, R. A. (2005). *Algal Culturing Techniques* (Vols. 0-12-088426-7). Amsterdam, NL.: Elsevier.

Anderson, R., Charvet, S., & Hansen, P. J. (2018). Mixotrophy in chlorophytes and haptophytes-effect of irradiance, macronutrient, micronutrient and vitamin limitation. *Frontiers in Microbiology*, *9*, 1704. doi:10.3389/fmicb.2018.01704

Araie, H., Nakamura, H., Toney, J. L., Haig, H. A., Plancq, J., Shiratori, T., and Shiraiwa, Y. (2018). Novel alkenone-producing strains of genus *Isochrysis* (haptophyta) isolated from Canadian saline lakes show temperature sensitivity of alkenones and alkenoates. *Organic Geochemistry*, *121*, 89–103. doi:10.1016/j.orggeochem.2018.04.008

Assmy, P., Fernández-Méndez, M., Duarte, P., Meyer, A., Randelhoff, A., Mundy, C. J., Granskog, M. A. (2017). Leads in Arctic pack ice enable early phytoplankton blooms below snow-covered sea ice. *Scientific Reports*, *7*, 40850. doi:10.1038/srep40850

Bendif, E. M., Probert, I., Hervé, A., Billard, C., Goux, D., Lelong, C., and Véron, B. (2011). Integrative taxonomy of the Pavlovophyceae (Haptophyta): a reassessment. *Protist*, *162*, 738–761. doi:10.1016/j.protis.2011.05.001

Bendif, E. M., Probert, I., Díaz-Rosas, F., Thomas, D., van den Engh, G., Young, J. R., & von Dassow, P. (2016). Recent reticulate evolution in the ecologically dominant lineage of coccolithophores. *Frontiers in Microbiology*, *7*, 784. doi:10.3389/fmicb.2016.00784

Bendif, E. M., Nevado, B., Wong, E. L. Y., Hagino, K., Probert, I., Young, J. R., Filatov, D. A. (2019). Repeated species radiations in the recent evolution of the key marine phytoplankton lineage *Gephyrocapsa*. *Nature Communications*, *10*, 4234. doi:10.1038/s41467-019-12169-7

Blanc-Mathieu, R., Krasovec, M., Hebrard, M., Yau, S., Desgranges, E., Martin, J., and Piganeau, G. (2017). Population genomics of picophytoplankton unveils novel chromosome hypervariability. *Science Advances*, *3*, e1700239. doi:10.1126/sciadv.1700239

Blossom, H. E., Rasmussen, S. A., Andersen, N. G., Larsen, T. O., Nielsen, K. F., & Hansen, P. J. (2014). *Prymnesium parvum* revisited: Relationship between allelopathy, ichthyotoxicity, and chemical profiles in 5 strains. *Aquatic Toxicology (Amsterdam, Netherlands)*, *157*, 159–166. doi:10.1016/j.aquatox.2014.10.006

Bochenek, M., Etherington, G. J., Koprivova, A., Mugford, S. T., Bell, T. G., Malin, G., & Kopriva, S. (2013). Transcriptome analysis of the sulfate deficiency response in the marine microalga *Emiliania huxleyi*. *The New Phytologist*, *199*, 650–662. doi:10.1111/nph.12303

Bock, N. A., Charvet, S., Burns, J., Gyaltshen, Y., Rozenberg, A., Duhamel, S., & Kim, E. (2021). Experimental identification and in silico prediction of bacterivory in green algae. *The ISME Journal*, *15*, 1987–2000. doi:10.1038/s41396-021-00899-w

Bork, P., Bowler, C., de Vargas, C., Gorsky, G., Karsenti, E., & Wincker, P. (2015). Tara oceans studies plankton at planetary scale. Introduction. *Science*, *348*, 873. doi:10.1126/science.aac5605

Bown, P. R. (1998). *Calcareous Nannofossil Biostratigraphy*. London: Kluwer Academic.

Burki, F., Kaplan, M., Tikhonenkov, D. V., Zlatogursky, V., Minh, B. Q., Radaykina, L. V., and Keeling, P. J. (2016). Untangling the early diversification of eukaryotes: A phylogenomic study of the evolutionary origins of centrohelida, haptophyta and cryptista. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *283*. doi:10.1098/rspb.2015.2802

Caputi, L., Carradec, Q., Eveillard, D., Kirilovsky, A., Pelletier, E., Pierella Karlusich, J. J., and Wincker, P. (2019). Community-level responses to iron availability in open ocean planktonic ecosystems. *Global Biogeochemical Cycles*, *10*, 1029. doi:10.1029/2018GB006022

Carradec, Q., Pelletier, E., Da Silva, C., Alberti, A., Seeleuthner, Y., Blanc-Mathieu, R., Coordinators, T. O. (2018). A global ocean atlas of eukaryotic genes. *Nature Communications*, *9*, 373. doi:10.1038/s41467-017-02342-1

Cavalier-Smith, T. (1999). Principles of protein and lipid targeting in secondary symbiogenesis: Euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. *Journal of Eukaryotic Microbiology*, *46*, 347–366. doi:10.1111/j.1550-7408.1999.tb04614.x

Cavalier-Smith, T. (2002). Nucleomorphs: Enslaved algal nuclei. *Current Opinion Microbiology*, *5*, 612–619. doi:10.1016/s1369-5274(02)00373-9

Cavalier-Smith, T., & von der Heyden, S. (2007). Molecular phylogeny, scale evolution and taxonomy of centrohelid heliozoa. *Molecular Phylogenetics and Evolution*, *44*, 1186–1203. doi:10.1016/j.ympev.2007.04.019

Cenci, U., Sibbald, S. J., Curtis, B. A., Kamikawa, R., Eme, L., Moog, D., Archibald, J. M. (2018). Nuclear genome sequence of the plastid-lacking cryptomonad *Goniomonas avonlea* provides insights into the evolution of secondary plastids. *BMC Biology*, *16*, 137. doi:10.1186/s12915-018-0593-5

Chaffron, S., Delage, E., Budinich, M., Vintache, D., Henry, N., Nef, C., Eveillard, D. (2021). Environmental vulnerability of the global ocean epipelagic plankton community interactome. *Science Advanced*, *7*. doi:10.1126/sciadv.abg1921

Choi, C. J., Bachy, C., Jaeger, G. S., Poirier, C., Sudek, L., Sarma, V., Worden, A. Z. (2017). Newly discovered deep-branching marine plastid lineages are numerically rare but globally distributed. *Current Biology : CB*, *27*, 15–16. doi:10.1016/j.cub.2016.11.032

Cruz de Carvalho, M. H., Sun, H. X., Bowler, C., & Chua, N. H. (2016). Noncoding and coding transcriptome responses of a marine diatom to phosphate fluctuations. *The New Phytologist*, *210*, 497–510. doi:10.1111/nph.13787

De Vargas, C., Aubry, M.-P., Probert, I. A. N., & Young, J. (2007). Origin and evolution of coccolithophores: From coastal hunters to oceanic farmers. In P. G. Falkowski & A. H. Knoll (Eds.), *Evolution of Primary Producers in the Sea* (pp. 251–285). doi:10.1016/B978-012370518-1/50013-8. Burlington: Academic Press.

Decelle, J., Stryhanyuk, H., Gallet, B., Veronesi, G., Schmidt, M., Balzano, S., Lupette, J. (2019). Algal remodeling in a ubiquitous planktonic photosymbiosis. *Current Biology : CB*, *29*, 968–978. doi:10.1016/j.cub.2019.01.073

Delmont, T. O., Gaia, M., Hinsinger, D. D., Frémont, P., Vanni, C., Fernandez-Guerra, A., Pelletier, E. (2022). Functional repertoire convergence of distantly related eukaryotic plankton lineages abundant in the sunlit ocean. *Cell Genomics*, *2*, 100123. doi:10.1016/j.xgen.2022.100123

Deodato, C. R., Barlow, S. B., Hovde, B. T., & Cattolico, R. A. (2019). Naked *Chrysochromulina* (haptophyta) isolates from lake and river ecosystems: An electron microscopic comparison including new observations on the type species of this taxon. *Algal Research*, *40*, 101492. doi:10.1016/j.algal.2019.101492

Di Franco, A., Baurain, D., Glöckner, G., Melkonian, M., & Phillippe, H. (2022). Lower statistical support with larger datasets: Insights from the Ochrophyta radiation. *Molecular Biology and Evolution*, *39*, msab300. doi:10.1093/molbev/msab300

DiTullio, G., Grebmeier, J., Arrigo, K., Lizotte, M., Robinson, D., Leventer, A., and Dunbar, R. (2000). Rapid and early export of *Phaeocystis antarctica* blooms in the Ross Sea, Antarctica. *Nature*, *404*, 595–598. doi:10.1038/35007061

Dorrell, R. G., Gile, G., McCallum, G., Méheust, R., Bapteste, E. P., Klinger, C. M., and Bowler, C. (2017). Chimeric origins of ochrophytes and haptophytes revealed through an ancient plastid proteome. *eLife*, *6*, 23717. doi:10.7554/eLife.23717

Dorrell, R. G., Nisbet, R. E. R., Barbrook, A. C., Rowden, S. J. L., & Howe, C. J. (2019). Integrated Genomic and transcriptomic analysis of the peridinin dinoflagellate *Amphidinium carterae* plastid. *Protist*, *170*, 358–373. doi:10.1016/j.protis.2019.06.001

Dorrell, R. G., Kuo, A., Füssy, Z., Richardson, E., Salamov, A., Zarevski, N., Lovejoy, C. (2021a). Within-Arctic horizontal gene transfer as a driver of convergent evolution in distantly related microalgae. *bioRxiv*. doi:10.1101/2021.07.31.454568

Dorrell, R. G., Villain, A., Perez-Lamarque, B., Audren de Kerdrel, G., McCallum, G., Watson, A. K., and Blanc, G. (2021b). Phylogenomic fingerprinting of tempo and functions of horizontal gene transfer within ochrophytes. *Proceedings of the National Academy of Sciences of the United States of America*, *118*, 2009974118. doi:10.1073/pnas.2009974118

Edvardsen, B., Egge, E. S., & Vaulot, D. (2016). Diversity and distribution of haptophytes revealed by environmental sequencing and metabarcoding – A review. *Perspectives Phycology*, *2*, 12. doi:10.1127/pip/2016/0052

Egge, E. S., Johannessen, T. V., Andersen, T., Eikrem, W., Bittner, L., Larsen, A., Edvardsen, B. (2015). Seasonal diversity and dynamics of haptophytes in the Skagerrak, Norway, explored by high-throughput sequencing. *Molecular Ecology*, *24*, 3026–3042. doi:10.1111/mec.13160

Eikrem, W., Medlin, L. K., Henderiks, J., Rokitta, S., Rost, B., Probert, I., Edvardsen, B. (2017). Haptophyta. In J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits (Eds.), *Handbook of the Protists* (pp. 893–953). Cham: Springer International Publishing.

Endo, H., Ogata, H., & Suzuki, K. (2018). Contrasting biogeography and diversity patterns between diatoms and haptophytes in the central Pacific Ocean. *Scientific Reports*, *8*, 10916. doi:10.1038/s41598-018-29039-9

Faktorová, D., Nisbet, R. E. R., Fernández Robledo, J. A., Casacuberta, E., Sudek, L., Allen, A. E., Lukeš, J. (2020). Genetic tool development in marine protists: Emerging model organisms for experimental cell biology. *Nature Methods*, *17*, 481–494. doi:10.1038/s41592-020-0796-x

Foster, K. W., & Smyth, R. D. (1980). Light antennas in phototactic algae. *Microbiology reviews*, *44*, 572–630. doi:10.1128/mr.44.4.572-630.1980

Frada, M., Probert, I., Allen, M. J., Wilson, W. H., & de Vargas, C. (2008). The "Cheshire Cat" escape strategy of the coccolithophore *Emiliania huxleyi* in response to viral infection. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 15944–15949. doi:10.1073/pnas.0807707105

Fresnel, J., & Probert, I. (2005). The ultrastructure and life cycle of the coastal coccolithophorid *Ochrosphaera neapolitana* (prymnesiophyceae). *European Journal Phycology*, *40*, 105–122. doi:10.1080/09670260400024659

Fuller, N. J., Tarran, G. A., Cummings, D. G., Woodward, E. M. S., Orcutt, K. M., Yallop, M., Scanlan, D. J. (2006). Molecular analysis of photosynthetic picoeukaryote community structure along an Arabian Sea transect. *Limnology Oceanography*, *51*, 2502–2514. doi:10.4319/lo.2006.51.6.2502

González, E. L., & Read, B. A. (2005). Cell biology, biochemistry and genomics of coccolithophore biomineralization. *MRS Proceedings*, *873*, K4.4. doi:10.1557/PROC-873-K4.4

Granum, E., Roberts, K., Raven, J. A., & Leegood, R. C. (2009). Primary carbon and nitrogen metabolic gene expression in the diatom *Thalassiosira pseudonana* (bacillariophyceae): Diel periodicity and effects of inorganic carbon and nitrogen. *Journal of Phycology*, *45*, 1083–1092. doi:10.1111/j.1529-8817.2009.00728.x

Green, J., & Parke, M. (1974). A reinvestigation by light and electron microscopy of *Ruttnera spectabilis* geitler (haptophyceae), with special reference to the fine structure of the zoids. *Journal of the Marine Biological Association*, *54*, 539–550.

Green, J. C. (1980). The fine structure of *Pavlova pinguis* Green and a preliminary survey of the order Pavlovales (Prymnesiophyceae). *British Phycology Journal*, *15*, 151–191. doi:10.1080/00071618000650171

Grigoriev, I. V., Hayes, R. D., Calhoun, S., Kamel, B., Wang, A., Ahrendt, S., Kuo, A. (2021). PhycoCosm, a comparative algal genomics resource. *Nucleic Acids Research*, *49*, D1004–D1011. doi:10.1093/nar/gkaa898

Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., Coordinators, T. O. C. (2016). Plankton networks driving carbon export in the oligotrophic ocean. *Nature*, *532*, 465–470. doi:10.1038/nature16942

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Christen, R. (2013). The Protist Ribosomal reference database (PR2): A catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nuclear Acids Research*, *41*, D597–D604. doi:10.1093/nar/gks1160

Guiry, M. D., Guiry, G. M., Morrison, L., Rindi, F., Valenzuala Miranda, S., Mathieson, A. C., and Garbary, D. J. (2014). AlgaeBase: An on-line resource for algae *Cryptogamie Algologie*, *35*, 11. doi:10.7872/crya.v35.iss2.2014.105

Hehenberger, E., Gast, R. J., & Keeling, P. J. (2019). A kleptoplastidic dinoflagellate and the tipping point between transient and fully integrated plastid endosymbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 17934–17942. doi:10.1073/pnas.1910121116

Hernández Limón, M. D., Hennon, G. M., Harke, M. J., Frischkorn, K. R., Haley, S. T., & Dyhrman, S. T. (2020). Transcriptional patterns of *Emiliania huxleyi* in the North Pacific subtropical gyre reveal the daily rhythms of its metabolic potential. *Environmental Microbiology*, *22*, 381–396.

Hirst, J., Schlacht, A., Norcott, J. P., Traynor, D., Bloomfield, G., Antrobus, R., Robinson, M. S. (2014). Characterization of TSET, an ancient and widespread membrane trafficking complex. *eLife*, *3*, e02866. doi:10.7554/eLife.02866

Hovde, B. T., Deodato, C. R., Hunsperger, H. M., Ryken, S. A., Yost, W., Jha, R. K., Cattolico, R. A. (2015). Genome sequence and transcriptome analyses of *Chrysochromulina tobin*: Metabolic tools for enhanced algal fitness in the prominent order Prymnesiales (haptophyceae). *PLoS Genetics*, *11*. doi:10.1371/journal.pgen.1005469

Hulatt, C. J., Wijffels, R. H., & Posewitz, M. C. (2021). The genome of the haptophyte *Diacronema lutheri* (*Pavlova lutheri*, pavlovales): A model for lipid biosynthesis in eukaryotic algae. *Genome Biology and Evolution*, *13*, 178. doi:10.1093/gbe/evab178

Ibarbalz, F. M., Henry, N., Brandão, M. C., Martini, S., Busseni, G., Byrne, H., Zinger, L. (2019). Global trends in marine plankton diversity across kingdoms of life. *Cell*, *179*, 1084–1097.e1021. doi:10.1016/j.cell.2019.10.008

Inouye, I., & Kawachi, M. (1994). The Haptonema. In J. C. Green & B. S. C. Leadbeater (Eds.), The Haptophyte Algae. The Systematics Association Special Edition (Vol. 51, pp. 73–89). Oxford: Cirendon Press.

Johnsen, G., Dalløkken, R., Eikrem, W., Legrand, C., Aure, J., & Skjoldal, H. R. (1999). Eco-physiology, bio-optics and toxicity of the ichthyotoxic *Chrysochromulina leadbeateri* (Prymnesiophyceae). *Journal of Phycology*, *35*, 1465–1476.

Kawachi, M., Inouye, I., Maeda, O., & Chihara, M. (1991). The haptonema as a food-capturing device: Observations on *Chrysochromulina hirta* (Prymnesiophyceae). *Phycologia*, *30*, 563–573. doi:10.2216/i0031-8884-30-6-563.1

Kawachi, M., Nakayama, T., Kayama, M., Nomura, M., Miyashita, H., Bojo, O., and Kamikawa, R. (2021). Rappemonads are haptophyte phytoplankton. *Current Biology*, *31*, 2395–2403. doi:10.1016/j.cub.2021.03.012

Keeling, P. J., Burki, F., Wilcox, H. M., Allam, B., Allen, E. E., Amaral-Zettler, L. A., Worden, A. Z. (2014). The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biology*, *12*, e1001889. doi:10.1371/journal.pbio.1001889

Kegel, J. U., John, U., Valentin, K., & Frickenhaus, S. (2013). Genome variations associated with viral susceptibility and calcification in *Emiliania huxleyi*. *PLoS One*, *8*, e80684. doi:10.1371/journal.pone.0080684

Kim, E., Harrison, J. W., Sudek, S., Jones, M. D., Wilcox, H. M., Richards, T. A., and Archibald, J. M. (2011). Newly identified and diverse plastid-bearing branch on the eukaryotic tree of life. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 1496–1500. doi:10.1073/pnas.1013337108

Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., coordinators, T. O. (2015). Ocean plankton. Determinants of community structure in the global plankton interactome. *Science*, *348*, 1262073. doi:10.1126/science.1262073

Liu, H., Probert, I., Uitz, J., Claustre, H., Aris-Brosou, S., Frada, M., de Vargas, C. (2009). Extreme diversity in non-calcifying haptophytes explains a major pigment paradox in open oceans. *Proceedings National Academy of Sciences USA*, *106*, 12803–12808. doi:10.1073/pnas.0905841106

Liu, H., Aris-Brosou, S., Probert, I., & de Vargas, C. (2010). A time line of the environmental genetics of the haptophytes. *Molecular Biology and Evolution*, *27*, 161–176. doi:10.1093/molbev/msp222

Liu, H., Sun, J., Wang, D., Zhang, X., Zhang, C., Song, S., & Thangaraj, S. (2018). Distribution of living coccolithophores in eastern Indian Ocean during spring intermonsoon. *Scientific Reports*, *8*, 12488. doi:10.1038/s41598-018-29688-w

Lohbeck, K. T., Riebesell, U., & Reusch, T. B. (2014). Gene expression changes in the coccolithophore *Emiliania huxleyi* after 500 generations of selection to ocean acidification. *Proceedings R Social Biological Science*, *281*, 20140003. doi:10.1098/rspb.2014.0003

Longhurst, A. (2006). *Ecological Geography of the Sea* (2nd ed.). Cambridge, Massechussetts: Academic Press.

Malhi, Y., & Grace, J. (2000). Tropical forests and atmospheric carbon dioxide. *Trends in Ecology Evolution*, *15*, 332–337. doi:10.1016/S0169-5347(00)01906-6

Mangot, J. F., Logares, R., Sanchez, P., Latorre, F., Seeleuthner, Y., Mondy, S., Massana, R. (2017). Accessing the genomic information of unculturable oceanic picoeukaryotes by combining multiple single cells. *Scientific Reports*, *7*. doi:10.1038/srep41498

Manton, I. (1968). Observations on the microanatomy of the type species of *Pyramimonas* (*P. tetrarhynchus* Schmarda). *Proceedings of the the Linnean Society of London*, *179*, 147–152. doi:10.1111/j.1095-8312.1968.tb00974.x

Marinov, G. K., Trevino, A. E., Xiang, T., Kundaje, A., Grossman, A. R., & Greenleaf, W. J. (2021). Transcription-dependent domain-scale three-dimensional genome organization in the dinoflagellate *Breviolum minutum*. *Nature Genetics*, *53*, 613–617. doi:10.1038/s41588-021-00848-5

Medlin, L. K., Saez, A. G., & Young, J. R. (2008). A molecular clock for coccolithophores and implications for selectivity of phytoplankton extinctions across the K/T boundary. *Marine Micropaleontology*, *67*, 69–86. doi:10.1016/j.marmicro.2007.08.007

Mills, M. M., Turk-Kubo, K. A., van Dijken, G. L., Henke, B. A., Harding, K., Wilson, S. T., Zehr, J. P. (2020). Unusual marine cyanobacteria/haptophyte symbiosis relies on N. *The ISME Journal*, *14*, 2395–2406. doi:10.1038/s41396-020-0691-6

More, K., Klinger, C. M., Barlow, L. D., & Dacks, J. B. (2020). Evolution and natural history of membrane trafficking in eukaryotes. *Current Biology*, *30*, R553–R564. doi:10.1016/j.cub.2020.03.068

Nam, O., Park, J. M., Lee, H., & Jin, E. (2019). De novo transcriptome profile of coccolithophorid alga *Emiliania huxleyi* CCMP371 at different calcium concentrations with proteome analysis. *PLoS One*, *14*, e0221938. doi:10.1371/journal.pone.0221938

Nef, C., Jung, S., Mairet, F., Kaas, R., Grizeau, D., & Garnier, M. (2019). How haptophytes microalgae mitigate vitamin B12 limitation. *Scientific Reports*, *9*, 8417. doi:10.1038/s41598-019-44797-w

Nelson, D. R., Hazzouri, K. M., Lauersen, K. J., Jaiswal, A., Chaiboonchoe, A., Mystikoi, A., Salehi-Ashtiani, K. (2021). Large-scale genome sequencing reveals the driving forces of viruses in microalgal evolution. *Cell Host Microbe*, *29*, 250–266. doi:10.1016/j.chom.2020.12.005

Novák Vanclová, A. M., Zoltner, M., Kelly, S., Soukal, P., Záhonová, K., Füssy, Z., Lukeš, J. (2020). Metabolic quirks and the colourful history of the *Euglena gracilis* secondary plastid. *The New Phytologist*, *225*, 1578–1592. doi:10.1111/nph.16237

O'Dea, S. A., Gibbs, S. J., Bown, P. R., Young, J. R., Poulton, A. J., Newsam, C., & Wilson, P. A. (2014). Coccolithophore calcification response to past ocean acidification and climate change. *Nature Communications*, *5*, 5363. doi:10.1038/ncomms6363

Okamoto, N., Chantangsi, C., Horák, A., Leander, B. S., & Keeling, P. J. (2009). Molecular phylogeny and description of the novel katablepharid *Roombia truncata* gen. et sp. nov., and establishment of the Hacrobia taxon nov. *PLoS One*, *4*, e7080. doi:10.1371/journal.pone.0007080

Parfrey, L. W., Lahr, D. J., Knoll, A. H., & Katz, L. A. (2011). Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proceedings of the National Academy of Sciences United Staes of America*, *108*, 13624–13629. doi:10.1073/pnas.1110633108

Pesant, S., Not, F., Picheral, M., Kandels-Lewis, S., Le Bescot, N., & Gorsky, G. (2015). Tara oceans consortium. Open science resources for the discovery and analysis of *Tara* Oceans data. *Scientific Data*, *2*. doi:10.1038/sdata.2015.23

Pierella Karlusich, J. J., Pelletier, E., Zinger, L., Lombard, F., Zingone, A., Colin, S., and Bowler, C. (2022). A robust approach to estimate relative phytoplankton cell abundances from metagenomes. *Molecular Ecological Research*. doi:10.1111/1755-0998.13592

Plancq, J., Couto, J. M., Ijaz, U. Z., Leavitt, P. R., & Toney, J. L. (2019). Next-generation sequencing to identify lacustrine haptophytes in the Canadian prairies: Significance for temperature proxy applications. *Journal of Geophysical Research*, *124*, 2144–2158. doi:10.1029/2018JG004954

Planes, S., Allemand, D., Agostini, S., Banaigs, B., Boissin, E., & Boss, E. (2019). The *Tara* Pacific Expedition-A pan-ecosystemic approach of the "-omics" complexity of coral reef holobionts across the Pacific Ocean. *PLoS Biology*, *17*, e3000483. 10.1371/journal.pbio.3000483

Prasad, B., Vadakedath, N., Jeong, H. J., General, T., Cho, M. G., & Lein, W. (2014). Agrobacterium tumefaciens-mediated genetic transformation of haptophytes (*Isochrysis* species). *Applied Microbiology and Biotechnology*, *98*, 8629–8639. doi:10.1007/s00253-014-5900-7

Rappé, M. S., Suzuki, M. T., Vergin, K. L., & 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. *Environmental Microbiology Reports*, *64*, 294–303. doi:10.1128/AEM.64.1.294-303.1998

Rastogi, A., Vieira, F. R. J., Deton-Cabanillas, A. F., Veluchamy, A., Cantrel, C., Wang, G., Tirichine, L. (2020). A genomics approach reveals the global genetic polymorphism, structure, and functional diversity of ten accessions of the marine model diatom *Phaeodactylum tricornutum*. *The ISME Journal*, *14*, 347–363. doi:10.1038/s41396-019-0528-3

Raymond, J. A., & Kim, H. J. (2012). Possible role of horizontal gene transfer in the colonization of sea ice by algae. *PLoS One*, *7*, e35968. doi:10.1371/journal.pone.0035968

Read, B. A., Kegel, J., Klute, M. J., Kuo, A., Lefebvre, S. C., & Maumus, F. (2013). *Emiliania huxleyi* annotation consortium. Pan genome of the phytoplankton *Emiliania* underpins its global distribution. *Nature*, *499*, 209–213. 10.1038/nature12221

Rice, D. W., & 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 Biology*, 31. doi:10.1186/1741-7007-4-31

Richter, N., Longo, W. M., George, S., Shipunova, A., Huang, Y., & Amaral-Zettler, L. (2019). Phylogenetic diversity in freshwater-dwelling Isochrysidales haptophytes with implications for alkenone production. *Geobiol*, *17*, 272–280. doi:10.1111/gbi.12330

Schlacht, A., Mowbrey, K., Elias, M., Kahn, R. A., & Dacks, J. B. (2013). Ancient complexity, opisthokont plasticity, and discovery of the 11th subfamily of Arf GAP proteins. *Traffic*, *14*, 636–649. doi:10.1111/tra.12063

Schober, A. F., Flori, S., Finazzi, G., Kroth, P. G., & Bártulos, C. R. (2018). Isolation of plastid fractions from the diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*. *Methods in Molecular Biology (Clifton, N.J.)*, *1829*, 189–203. doi:10.1007/978-1-4939-8654-5_13

Schoemann, V., Becquevort, S., Stefels, J., Rousseau, V., & Lancelot, C. (2005). *Phaeocystis* blooms in the global ocean and their controlling mechanisms: A review. *Journal of Sea Research*, *53*, 43–66. doi:10.1016/j.seares.2004.01.008

Seeleuthner, Y., Mondy, S., Lombard, V., Carradec, Q., Pelletier, E., Wessner, M., Coordinators, T. O. (2018). Single-cell genomics of multiple uncultured stramenopiles reveals underestimated functional diversity across oceans. *Nature Communications*, *9*, 310. doi:10.1038/s41467-017-02235-3

Shi, X. L., Marie, D., Jardillier, L., Scanlan, D. J., & Vaulot, D. (2009). Groups without cultured representatives dominate eukaryotic picophytoplankton in the oligotrophic South East Pacific Ocean. *PLoS One*, *4*, e7657. doi:10.1371/journal.pone.0007657

Shi, Q. (2019). Expression profiling of genes coding for abundant proteins in the alkenone body of marine haptophyte alga *Tisochrysis lutea*. *BMC Microbiology*, *19*, 56. doi:10.1186/s12866-019-1430-x

Simon, M., Lopez-Garcia, P., Moreira, D., & Jardillier, L. (2013). New haptophyte lineages and multiple independent colonizations of freshwater ecosystems. *Env Microbiology Reports*, *5*, 322–332. doi:10.1111/1758-2229.12023

Singer, D., Seppey, C. V. W., Lentendu, G., Dunthorn, M., Bass, D., Belbahri, L., Lara, E. (2020). Protist taxonomic and functional diversity in soil, freshwater and marine ecosystems. *Environment International*, *146*, 106262. doi:10.1016/j.envint.2020.106262

Søgaard, D. H., Sorrell, B. K., Sejr, M. K., Andersen, P., Rysgaard, S., Hansen, P. J., and Lund-Hansen, L. C. (2021). An under-ice bloom of mixotrophic haptophytes in low nutrient and freshwater-influenced Arctic waters. *Scientific Reports*, *11*, 2915. doi:10.1038/s41598-021-82413-y

Sommeria-Klein, G., Watteaux, R., Ibarbalz, F. M., Pierella Karlusich, J. J., Iudicone, D., Bowler, C., & Morlon, H. (2021). Global drivers of eukaryotic plankton biogeography in the sunlit ocean. *Science*, *374*, 594–599. doi:10.1126/science.abb3717

Song, H., Chen, Y., Liu, F., & Chen, N. (2021). Large differences in the haptophyte *Phaeocystis globosa* mitochondrial genomes driven by repeat amplifications. *Frontiers in Microbiology*, *12*. doi:10.3389/fmicb.2021.676447

Stiller, J. W., Schreiber, J., Yue, J., Guo, H., Ding, Q., & Huang, J. (2014). The evolution of photosynthesis in chromist algae through serial endosymbioses. *Nature Communications*, *5*, 5764. doi:10.1038/ncomms6764

Strassert, J. F. H., Irisarri, I., Williams, T. A., & Burki, F. (2021). A molecular timescale for eukaryote evolution with implications for the origin of red algal-derived plastids. *Nature Communications*, *12*, 1879. doi:10.1038/s41467-021-22044-z

Takahashi, K., Benico, G., Lum, W. M., & Iwataki, M. (2019). *Gertia stigmatica* gen. et sp. nov. (kareniaceae, dinophyceae), a new marine unarmored dinoflagellate possessing the peridinin-type chloroplast with an eyespot. *Protist*, *170*, 125680. doi:10.1016/j.protis.2019.125680

Takishita, K., Nakano, K., & Uchida, A. (1999). Preliminary phylogenetic analysis of plastid-encoded genes from an anomalously pigmented dinoflagellate *Gymnodinium mikimotoi* (gymnodiniales, dinophyta). *Phycology Research*, *47*, 257–262. doi:10.1111/j.1440-1835.1999.tb00306.x

Thompson, A. W., Foster, R. A., Krupke, A., Carter, B. J., Musat, N., Vaulot, D., Zehr, J. P. (2012). Unicellular cyanobacterium symbiotic with a single-celled eukaryotic alga. *Science*, *337*, 1546–1550. doi:10.1126/science.1222700

Unrein, F., Gasol, J. M., Not, F., Forn, I., & Massana, R. (2014). Mixotrophic haptophytes are key bacterial grazers in oligotrophic coastal waters. *The ISME Journal*, *8*, 164–176. doi:10.1038/ismej.2013.132

Ustick, L. J., Larkin, A. A., Garcia, C. A., Garcia, N. S., Brock, M. L., Lee, J. A., Martiny, A. C. (2021). Metagenomic analysis reveals global-scale patterns of ocean nutrient limitation. *Science*, *372*, 287–291. doi:10.1126/science.abe6301

Valenti, T. W., James, S. V., Lahousse, M. J., Schug, K. A., Roelke, D. L., Grover, J. P., & Brooks, B. W. (2010). A mechanistic explanation for pH-dependent ambient aquatic toxicity of *Prymnesium parvum* carter. *Toxicon*, *55*, 990–998. doi:10.1016/j.toxicon.2009.09.014

Vaulot, D., Le Gall, F., Marie, D., Guillou, L., & Partensky, F. (2004). The Roscoff Culture Collection (RCC): A collection dedicated to marine picoplankton. *Nova Hedwigia*, *79*, 32. doi:10.1127/0029-5035/2004/0079-0049

Veluchamy, A., Lin, X., Maumus, F., Rivarola, M., Bhavsar, J., Creasy, T., Tirichine, L. (2013). Insights into the role of DNA methylation in diatoms by genome-wide profiling in *Phaeodactylum tricornutum*. *Nature Communications*, *4*, 2901. doi:10.1038/ncomms3091

Veluchamy, A., Rastogi, A., Lin, X., Lombard, B., Murik, O., Thomas, Y., Tirichine, L. (2015). An integrative analysis of post-translational histone modifications in the marine diatom *Phaeodactylum tricornutum*. *Genome Biology*, *16*, 102. doi:10.1186/s13059-015-0671-8

Vernette, C., Henry, N., Lecubin, J., de Vargas, C., Hingamp, P., & Lescot, M. (2021). The Ocean Barcode Atlas: A web service to explore the biodiversity and biogeography of marine organisms. *Molecular Ecology Resources*, *21*, 1347–1358. doi:10.1111/1755-0998.13322

Vincent, F., & Bowler, C. (2020). Diatoms are selective segregators in global ocean planktonic communities. *mSystems*, *5*. doi:10.1128/mSystems.00444-19

Vincent, F., Sheyn, U., Porat, Z., Schatz, D., & Vardi, A. (2021). Visualizing active viral infection reveals diverse cell fates in synchronized algal bloom demise. *Proceedings of the National Academy of Sciences of the United States of America*, *118*. doi:10.1073/pnas.2021586118

Vogt, M., O'Brien, C., Peloquin, J., Schoemann, V., Breton, E., Estrada, M., Stefels, J. (2012). Global marine plankton functional type biomass distributions: *Phaeocystis* spp. *Earth System Science Data*, *4*, 107–120.

von Dassow, P., John, U., Ogata, H., Probert, I., Bendif, E. M., Kegel, J. U., de Vargas, C. (2015). Life-cycle modification in open oceans accounts for genome variability in a cosmopolitan phytoplankton. *The ISME Journal*, *9*, 1365–1377. doi:10.1038/ismej.2014.221

Walker, G., Dorrell, R. G., Schlacht, A., & Dacks, J. B. (2011). Eukaryotic systematics: A user's guide for cell biologists and parasitologists. *Parasitol*, *138*, 1638–1663. doi:10.1017/S0031182010001708

Wang, X., Song, H., Wang, Y., & Chen, N. (2021). Research on the biology and ecology of the harmful algal bloom species *Phaeocystis globosa* in China: Progresses in the last 20 years. *Harmful Algae*, *107*, 102057. doi:10.1016/j.hal.2021.102057

Williams, D. M. (2021). Names, taxa and databases: Some aspects of diatom nomenclature. *Diatom Research*, *36*, 101–106. doi:10.1080/0269249X.2021.1873194

Wu, W., Huang, B., Liao, Y., & Sun, P. (2014). Picoeukaryotic diversity and distribution in the subtropical–tropical South China Sea. *FEMS Microbiology Ecology*, *89*, 563–579. doi:10.1111/1574-6941.12357

Wu, W., Wang, L., Liao, Y., & Huang, B. (2015). Microbial eukaryotic diversity and distribution in a river plume and cyclonic eddy-influenced ecosystem in the South China Sea. *Microbiology Open*, *4*, 826–840. doi:10.1002/mbo3.282

Yoon, H. S., Hackett, J. D., & Bhattacharya, D. (2002). A single origin of the peridinin- and fucoxanthin-containing plastids in dinoflagellates through tertiary endosymbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 11724–11729. doi:10.1073/pnas.172234799

Yoon, H. S., Hackett, J. D., Ciniglia, C., Pinto, G., & Bhattacharya, D. (2004). A molecular timeline for the origin of photosynthetic eukaryotes. *Molecular Biology and Evolution*, *21*, 809–818. doi:10.1093/molbev/msh075

Yoshida, M., Noël, M. H., Nakayama, T., Naganuma, T., & Inouye, I. (2006). A haptophyte bearing siliceous scales: Ultrastructure and phylogenetic position of *Hyalolithus neolepis* gen. et sp. nov. (prymnesiophyceae, haptophyta). *Protist*, *157*, 213–234. doi:10.1016/j.protis.2006.02.004

Zondervan, I. (2007). The effects of light, macronutrients, trace metals and CO2 on the production of calcium carbonate and organic carbon in coccolithophores - A review. *Deep-Sea Research*, *54*, 521–537. doi:10.1016/j.dsr2.2006.12.004

Zubkov, M. V., & Tarran, G. A. (2008). High bacterivory by the smallest phytoplankton in the North Atlantic Ocean. *Nature*, *455*, 224–226. doi:10.1038/nature07236