environmental microbiology reports

Environmental Microbiology Reports (2013) **5**(2), 322–332 doi:10.1111/1758-2229.12023

New haptophyte lineages and multiple independent colonizations of freshwater ecosystems

Marianne Simon, Purificación López-García, David Moreira and Ludwig Jardillier*
Unité d'Ecologie, Systématique et Evolution, CNRS UMR 8079 Université Paris-Sud, 91405 Orsay, France.

Summary

The diversity and ecological relevance of small haptophytes in marine systems is increasingly recognized. Similar investigations in freshwater remain scarce, despite some recent studies showing the existence of divergent haptophyte lineages and indicating that these microalgae can occur at high abundance in lakes. We studied the diversity of haptophytes in a wide variety of marine, salty continental and, most particularly, freshwater environments by amplifying, cloning and sequencing 18S rRNA genes. For this purpose, we designed two sets of primers specific for the two recognized haptophyte classes, Prymnesiophyceae and Pavlovophyceae. We detected pavlovophyte sequences only in freshwater systems as well as several novel prymnesiophyte phylotypes in both freshwater and marine environments. In addition, we retrieved a cluster of sequences (HAP-3) from the Marmara Sea branching deeply in the haptophyte tree with no clear affiliation to either of the two recognized classes. Five of the freshwater prymnesiophyte phylotypes detected formed a divergent monophyletic group (EV) without close described representatives that branched within the Isochrysidales, a group of generally marine and most often calcifying coccolithophorids. The presence of several sequences of freshwater haptophytes scattered among marine taxa in phylogenetic trees confirms the occurrence of several independent haptophyte transitions between marine and freshwater environments.

Introduction

Haptophytes are unicellular aquatic, mostly marine, photosynthetic eukaryotes. They are broadly distributed and recognized as key players in biogeochemical cycles in marine ecosystems. Their wide distribution and relative high abundance in marine waters was initially shown by high-performance liquid chromatography analyses of its characteristic pigment 19'-hexanoyloxyfucoxanthin (19-Hex) (Andersen *et al*., 1996). According to recent molecular investigations, haptophytes contributed from 30% to 50% of the photosynthetic standing stocks in the photic layer across the oceans in the year 2000 (Liu *et al*., 2009). Similarly, targeted metagenomic analyses suggest that haptophytes constitute on average 25% of global eukaryotic picophytoplankton carbon biomass (Cuvelier *et al*., 2010). Haptophytes may play a significant role in primary production even when they are not dominant; for instance, a new group of haptophytes was recently shown to contribute significantly to CO_2 fixation despite its low relative abundance (Jardillier *et al*., 2010). This might be due to their higher growth rates and their bigger size as compared with other more abundant planktonic phototrophs (Cuvelier *et al*., 2010) and, perhaps, also to the presence of efficient carbon-concentrating mechanisms (Reinfelder, 2011). Haptophytes may also play additional roles in the C-cycle through heterotrophic pathways, as some haptophytes have been shown to be mixotrophic (Legrand *et al*., 2001; Frias-López *et al*., 2009).

In addition to their increasingly recognized ecological importance in oceans, haptophytes are at the heart of a phylogenetic debate, as their position in the eukaryotic tree remains unresolved. Haptophytes were initially affiliated to the chromalveolates, a eukaryotic super-group including also alveolates, stramenopiles (heterokonts) and cryptophytes, the ancestor of which was thought to have acquired a red algal plastid as secondary endosymbiont (for review, see Keeling, 2009). However, whereas the monophyly of alveolates and stramenopiles is easily retrieved, recent molecular phylogenetic analyses suggest that cryptophytes and haptophytes do not form a monophyletic group with them. They rather seem to be sister to other clades, such as the kathablepharids, telonemids and centrohelid heliozoa, which do not have known photosynthetic members, forming another supergroup recently named Hacrobia (Okamoto *et al*., 2009) or CCTH group (cryptophytes, centroheliozoa, telonemids, haptophytes) (Burki *et al*., 2009), even if its monophyly remains discussed (Burki *et al*., 2012). At a finer phylogenetic scale, several questions also remain open. Haptophytes currently encompass two classes, the

Received 16 August, 2012; revised 26 November, 2012; accepted 26 November, 2012. *For correspondence. E-mail ludwig.jardillier@u-psud.fr; Tel. (+33) 169154991; Fax (+33) 169154697.

© 2012 Society for Applied Microbiology and Blackwell Publishing Ltd

Pavlovophyceae and Prymnesiophyceae. The monophyly of those classes is supported by both morphology and molecular data, although the phylogenetic relationships within each class are often reshuffled (see Edvardsen and Medlin, 2007; Bendif *et al*., 2011; Edvardsen *et al*., 2011 for recent propositions). Pavlovophytes have mainly been observed in littoral and brackish environments, although they have also been detected in freshwater systems by amplification of their 18S rRNA genes using pavlovophyte-specific primers (Shalchian-Tabrizi *et al*., 2011). They are currently represented by only four genera (namely *Diacronema*, *Exanthemachrysis*, *Pavlova* and *Rebecca*) distributed in four clades according to phylogenies based on 18S rRNA genes (Bendif *et al*., 2011). Compared with pavlovophytes, prymnesiophytes display a higher diversity and abundance and, consequently, have been studied more thoroughly. They comprise the orders Coccolithales, Isochrysidales, Phaeocystales, Prymnesiales, Syracosphaerales and Crepidolithales (De Vargas *et al*., 2007) or Zygodiscales (Jordan *et al*., 2004; Edvardsen and Medlin, 2007) depending on the authors. The legitimacy of the orders Syracosphaerales and Crepidolithales or Zygodiscales is under discussion; they might be included within the Coccolithales (Edvardsen and Medlin, 2007). Several haptophyte genera based on cell morphology later revealed to be polyphyletic based on molecular analyses (Edvardsen *et al*., 2011) and, conversely, several morphological species were found indistinguishable based on molecular markers (Bendif *et al*., 2011), which urges for a revision of haptophyte systematics based on reliable molecular phylogenies.

In spite of their ecological and phylogenetic importance, the diversity of haptophytes is not fully explored. Since the beginning of the century, the use of molecular methods based on the amplification, cloning and sequencing of 18S rRNA genes has uncovered a vast diversity of marine protist lineages affiliating to known taxa and a few undescribed divergent lineages scattered in the eukaryotic tree (López-García *et al*., 2001; Moon-van der Staay *et al*., 2001; Massana and Pedros-Alio, 2008). However, in those studies using general eukaryotic primers, the diversity and novelty of haptophytes was rather low. In the particular case of haptophytes, Moon-van der Staay and colleagues (2000) showed a great discrepancy between the proportion of haptophytes in marine picoplankton based on their characteristic 19-Hex pigment and their relative representation in 18S rDNA libraries. It is well known that the use of lineage-specific primers may enhance the recovery of diverse environmental sequences within a given taxon, as has been shown, for example, for cercozoans (Bass and Cavalier-Smith, 2004) or diplonemids (Lara *et al*., 2009). Actually, the use of specific primers targeting the 28S rRNA gene allowed the discovery of an unsuspected diversity of non-calcifying prymnesiophytes in oceanic waters (Liu *et al*., 2009). The enrichment of haptophyte fractions by flow cytometry has been another alternative showing a hitherto unveiled diversity within the group. In this way, a novel clade of highly divergent haptophytes was discovered in South Pacific waters (Shi *et al*., 2009). Subsequent metagenomic studies on haptophyte fractions sorted by flow cytometry confirmed a wide prymnesiophyte diversity in oceans (Cuvelier *et al*., 2010).

Compared with marine ecosystems, haptophytes in freshwater systems have been less studied and seem less diverse. Only a dozen freshwater haptophyte species have been described, whereas over 400 species have been defined from marine environments (Preisig, 2002). Most research efforts have concentrated on toxic species, as they trigger massive fish kills (Hansen *et al*., 1994; Edvardsen and Imai, 2006). Described freshwater haptophytes belong to the orders Pavlovales, Coccolithales and Prymnesiales (Preisig, 2002), with *Chrysochromulina* and *Prymnesium* being the most common genera. Consistent with classical observations, molecular diversity studies of small eukaryotes in lakes (Richards *et al*., 2005; Lepère *et al*., 2008; Triadó-Margarit and Casamayor, 2012) revealed neither a broad diversity nor a high proportion of haptophyte sequences, most of the sequences retrieved being close to *Chrysochromulina parva*. Nonetheless, under certain conditions, lake haptophytes can reach relatively high proportions in the euphotic zone. For instance, haptophytes accounted for 7.7% of microbial eukaryotes in Lake Aydat and up to 62.8% of small planktonic protists in Lake Bourget (Lepère *et al*., 2010). Most studies on freshwater haptophytes have concentrated in lakes. However, freshwater environments are varied and highly heterogeneous, and the eukaryotic diversity in these ecosystems is far from being described, especially in small freshwater bodies. In one of the first protist molecular surveys carried out in freshwater systems, Šlapeta and colleagues (2005) studied two ponds with different redox status and retrieved 18S rDNA phylotypes that formed a monophyletic group with pavlovophytes and prymnesiophytes but branched deeply in the haptophyte clade. The occurrence of this group (HAP-1) in freshwater systems has been confirmed by subsequent studies in Scandinavian lakes using 454 pyrosequencing of the V4 region of 18S rRNA genes (Shalchian-Tabrizi *et al*., 2011). This same study also suggested, based on the premise of a marine origin for haptophytes, that several transitions between marine and freshwater systems have occurred along the evolutionary history of the group. However, as the diversity and distribution of haptophytes in different ecosystems across a salinity gradient is not well known, this hypothesis awaits confirmation.

To contribute to a comprehensive account of haptophyte diversity, we designed new specific 18S rDNA

primers covering the diversity of known prymnesiophytes, pavlovophytes and the rest of known haptophyte environmental lineages. We then carried out molecular surveys in a variety of ecosystems across a salinity gradient, including oceanic samples, a variety of freshwater lakes and ponds, as well as brackish and hypersaline shallow lakes. Our study confirms a wider diversity of haptophytes in marine, as compared with freshwater, ecosystems, and reveals the occurrence of novel divergent lineages in both marine and freshwater systems. The ecological distribution of the new haptophyte lineages further supports the hypothesis of multiple transitions from marine to freshwater systems.

Results and discussion

Prymnesiophyte and pavlovophyte diversity and distribution

In order to study haptophyte diversity in a variety of ecosystems, we first designed specific 18S rDNA primers targeting separately Prymnesiophyceae and Pavlovophyceae plus other divergent lineages (see supplementary *Materials and methods*). Specific haptophyte primers have been used in previous studies. However, they targeted either the 28S rRNA gene (Liu *et al*., 2009), for which a comprehensive species and environmental reference sampling is missing, or a particular haptophyte subset (pavlovophytes) (Shalchian-Tabrizi *et al*., 2011). Using our more inclusive haptophyte primer sets, we then amplified 18S rRNA genes from a variety of samples including freshwater, marine and saline systems (Table 1) (see *Supporting information* for methodological details). In all continental samples, haptophyte 18S rDNAs were amplified only after nested PCR reactions, suggesting that haptophytes were in low abundance in these ecosystems. Although we cannot totally discard a PCR bias, this is in agreement with the paucity of haptophyte sequences in most molecular surveys in lakes using general eukaryotic primers (Richards *et al*., 2005; Lepère *et al*., 2008; Triadó-Margarit and Casamayor, 2012). This might suggest that several haptophyte lineages were rare at the sampling time, forming part of a 'seed bank' that may bloom or become more prevalent at different seasons. We identified a total of 32 haptophyte operational taxonomic units (OTUs, defined as groups of 18S rDNA sequences being more than 98% identical), most of them belonging to the prymnesiophytes, which were detected in the majority of the ecosystems sampled (Table S1). Most of the haptophyte diversity was observed in marine systems where 21 OTUs were retrieved out of the total of 32 detected, while only nine and two OTUs were revealed in freshwater and salty (brackish or hypersaline) continental systems respectively (Figs 1–3, Table S2). The detected phylotypes affiliated to the five prymnesiophyte orders generally accepted (De Vargas *et al*., 2007), namely Phaeocystales, Prymnesiales, Coccolithales, Isochrysidales and Syracosphaerales (Fig. 2), confirming the validity of the new prymnesiophyte primer set. In turn, pavlovophytes could only be detected in four out of 15 freshwater systems and were found neither in the continental saline waters nor in the marine environments sampled (although we did amplify a divergent set of sequences from marine samples using this set of primers; Table S1). The two freshwater pavlovophyte OTUs retrieved in our study were affiliated to clade 4 (Bendif *et al*., 2011) (Fig. 3). Although the pavlovophyte diversity and distribution observed in our study were reduced, these results are consistent with current knowledge (Bendif *et al*., 2011).

New insights in the diversity of marine haptophytes

More than half the 32 OTUs detected in the environments sampled were not closely related to described species (Table S2). Among them, five marine OTUs were affiliated to previously described clades without any cultured representative species (Prym_14, Prym_20) or even to potential new clades (Prym_13 and Prym_29, Ma135_Pav3) (Figs 2 and 3, Table S2).

All sequences forming OTUs Prym_13 and Prym_29 were detected in the Sea of Marmara. They were grouped with the environmental sequence F01N5 that was retrieved at the epipelagic zone (0–200 m) in the Sargasso Sea (Not *et al*., 2007) and sequences SGZW1078 and FS04R14 from the euphotic zone (75 m depth) of Florida Straits and Sargasso Sea respectively (Figs 2 and S1) (Cuvelier *et al*., 2010). Interestingly, this group of sequences formed a new clade (that we call here clade B3) well nested within the Prymnesiales, along with clades B1 and B2 defined by Edvardsen and colleagues (2000; 2011). However, its position within the order is unstable; there is no significant statistical support for its sisterhood to either clade B1 or B2 (Fig. 2).

Operational taxonomic units Prym_14 and Prym_20 were affiliated to prymnesiophyte clade D and clade E respectively (Edvardsen *et al*., 2000; 2011), which are exclusively composed of marine environmental sequences from the South East Pacific Ocean and the Equatorial Pacific Ocean (Edvardsen *et al*., 2000; Moon-van der Staay *et al*., 2000; Shi *et al*., 2009). Over the last decade, the phylogenetic position of those two clades D and E changed with the increasing number of sequences retrieved in marine environments (Edvardsen *et al*., 2000; Moon-van der Staay *et al*., 2000; 2001). Recently, Shi and colleagues (2009) proposed clade D to form an independent lineage at the base of all other prymnesiophyte orders within the haptophytes. In addition to its widespread

Table 1. Major characteristics of the samples analysed in this study. Additional details are provided in Table S1.

Location	GPS coordinates	Ecosystem type	Sample name	Size fraction
Charca Verde, campus University Paris-Sud	48°42′02″N 2°10′28″E	Freshwater pond	CV1 MVSF	0.22–5 μm n.d.
Chevreuse, PNR Haute Vallée de Chevreuse	48°42′18.6″N 2°02′23.5″E	Freshwater urban pond	CH1 CH2	0.22–5 μm > 5 μm
Etang de la Tour, PNR Haute Vallée de Chevreuse	48°39′39.2″N 1°52′48.3″E	Freshwater pond	To1 To2	5–30 μm 0.22–5 μm
Etang des Vallées, PNR Haute Vallée de Chevreuse	48°41′23.0″N 1°54′59.2″E	Freshwater shallow lake	EV2 EV3 EV6 EV7 EV8 EV9 EV10	0.22–5 μm 5–30 μm 0.22–5 μm 5–30 μm 0.22–5 μm 5–30 μm 0.22–5 μm
Mare Gabard, PNR Haute Vallée de Chevreuse	48°39′15.83″N 1°55′20.26″E	Freshwater pond, forest	MG1 MG2	0.22–5 μm 5–30 μm
Saint Robert, PNR Haute Vallée de Chevreuse	48°39′54.82″N 1°56′45.28″E	Freshwater village pond	SR1 SR2	0.22–5 μm 5–30 μm
Ru Sainte Anne, PNR Haute Vallée de Chevreuse	48°36′45.91″N 1°58′16.61″E	Freshwater brook	RSA1 RSA2	0.22–5 μm 5–30 μm
La Claye, PNR Haute Vallée de Chevreuse	48°36′31.72″N 1°56′17.33″E	Freshwater pond, forest	LC1 LC2	0.22–5 μm 5–30 μm
Etang du Perray, PNR Haute Vallée de Chevreuse	48°41′49″N 1°51′37″E	Freshwater pond	Pe1	5–30 μm
Etang de Pourras, PNR Haute Vallée de Chevreuse	48°42′52″N 1°50′39″E	Freshwater pond	Po1	5–30 μm
Etang de Cernay, PNR Haute Vallée de Chevreuse	48°40′50″N 1°57′55″E	Freshwater pond	Ce1	5–30 μm
Lac du Bourget, Savoie	45°44′N 05°51′E	Freshwater lake	BG1 BG6	0.22–60 μm 0.22–60 μm
Lac d'Annecy Northern basin, Savoie	45°54′N 06°07′E	Freshwater lake	AN1 AN2 AN6	0.22–60 μm 0.22–60 μm 0.22–60 μm
Etang d'en haut, Paimpont, Britain	48°00′30.66″N 2°13′36.06″W	Freshwater pond	Ht1	0.22–5 μm
Etang du Châtenay, Paimpont, Britain	48°00′14.40″N 2°13′48.36″W	Freshwater pond	Châ1 Châ2	0.22–5 μm 5–30 μm
South Atlantic	56°18′57″S 57°39′45″E	Marine	DH122 DH123 DH125 DH129	0.22–5 μm 0.22–5 μm 0.22–5 μm 0.22–5 μm
Marmara Sea, Central basin	40°50′18.48″N 28°01′24.24″E	Marine	Ma101 Ma125 Ma130 Ma135	0.22–5 μm 0.22–5 μm 0.22–5 μm 0.22–5 μm
Salada Chiprana, Spain	41°14′30″N 0°10′50″W	Hypersaline	SCH1 SCH2 SCH3	n.a. 0.22–5 μm 0.22–5 μm
Ornithological Park of Teich, France	44°38′27.72″N 1°01′14.04″W	Brackish	CPT2 CPT3 CPT4	n.d. n.d. n.d.

n.d., not done; n.a., not applicable.

distribution in marine environments, our study confirms that these sequences form an independent clade among prymnesiophytes, and should therefore represent a new order (Fig. 2). Shi and colleagues (2009) also proposed clade E to be a sister group of the coccolithophorid orders (Coccolithales, Isochrysidales, Syracosphaerales). Our phylogenetic analysis confirms the placement of the clade E within the clade formed by coccolithophorid orders, but without any clear sisterhood to any of the described orders (Fig. 2). The inclusion of new coccolithophorid sequences should help place clade E within one of the described orders or support the erection of a novel order for organisms of this clade.

The use of pavlovophyte-specific primers revealed a divergent lineage, OTU Ma135_Pav3, grouping 60 clone sequences (Table S2), only detected at surface (15 m depth) in the Sea of Marmara. The position of this divergent OTU is unstable. It branched either as a sister group of pavlovophytes, although with very low support (Fig. S2), or at the base of both pavlovophytes and

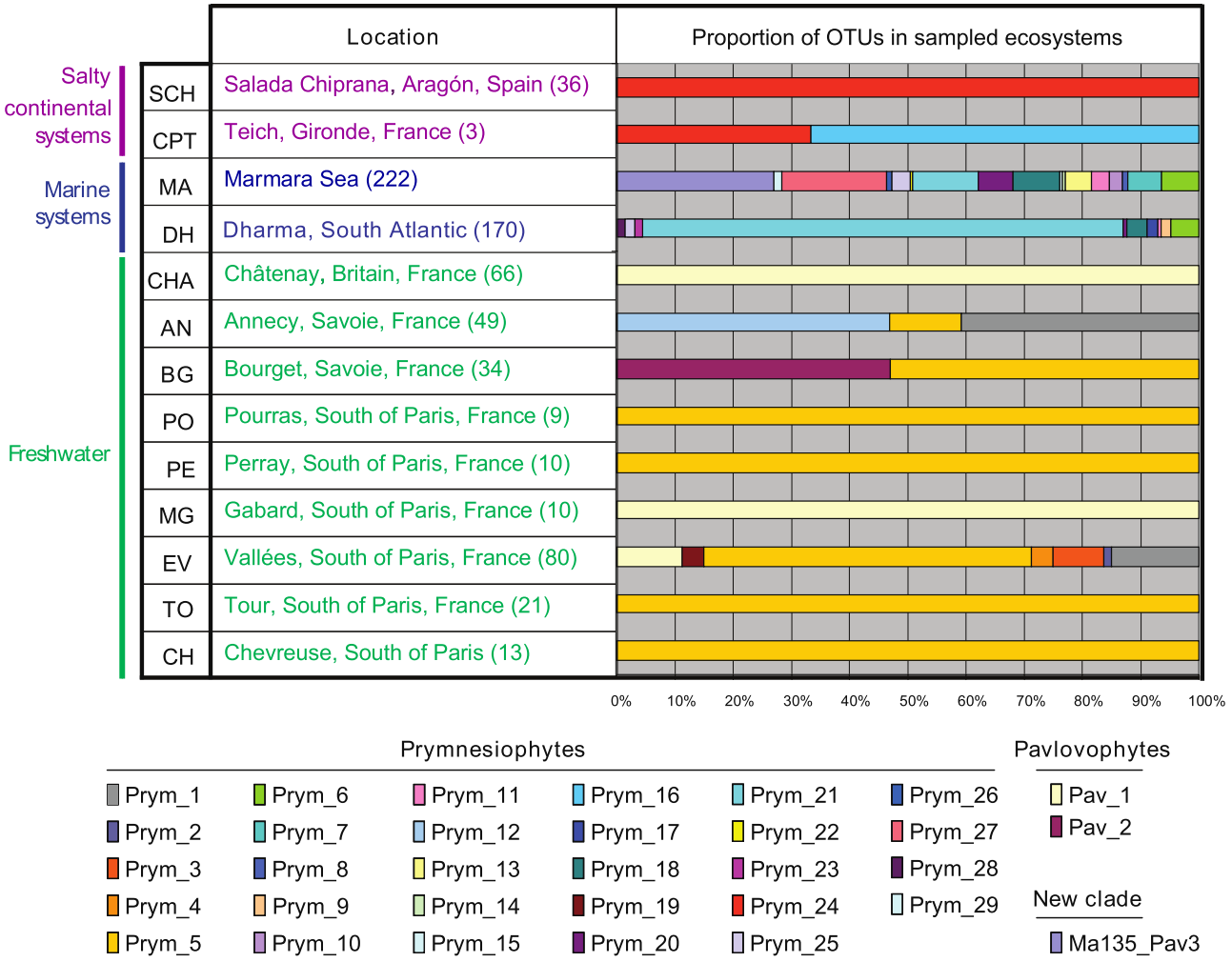

Fig. 1. Proportions of the OTUs detected in the different ecosystems investigated. The number of sequenced haptophyte clones from each ecosystem is shown in brackets; the colour code utilized for the different OTUs is indicated at the bottom.

prymnesiophytes (Fig. 3). The closest 18S rDNA sequence retrieved by a BLAST search against the Silva database (Pruesse *et al.*, 2007) was retrieved from the South East Pacific Ocean (Shi *et al.*, 2009) and shared only 93.9% identity with it (Table S2). Nonetheless, environmental sequences FS14K073 and EN360CTD001 (Cuvelier *et al.*, 2010) and SHAX 513 (Orsi *et al.*, 2012), which only partially overlapped our sequences, shared 97–98% identity with our sequence Ma135-Pav3-C1 and clustered together in phylogenetic trees (Fig. S2). This divergent OTU Ma135_Pav3 did not branch with any of the other deeply divergent haptophyte lineages detected so far: the freshwater clade HAP-1 formed by the environmental sequences CV1-B1-97 and CV1-B2-32 retrieved from a suboxic pond (Šlapeta *et al.*, 2005) and APB2H and AI9LL from Lake Finsevatn in Norway (Shalchian-Tabrizi *et al.*, 2011) (Figs 3 and S2) and the marine clade HAP-2 including sequences from the Biosope cruise T65.100 and T58.080 (Shi *et al.*, 2009). Therefore, the newly detected clade, which we name here HAP-3, together with HAP-1 and HAP-2, might represent new class-level groups of haptophytes along with pavlovophytes and prymnesiophytes.

Novel haptophytes in freshwater systems

We detected a group of five OTUs unrelated to any described species or to any environmental sequence (Group EV, Fig. 2) in two freshwater lakes, the shallow lake Etang des Vallées and Lake Annecy (Table 1). Interestingly, this phylogenetic group affiliates to Isochrysidales, a coccolithophorid order that so far was thought to be composed exclusively of marine haptophytes. This order of haptophytes is composed of non-calcifying (*Isochrysis*) and calcifying (*Chrysotila*, *Emiliania*, *Gephyrocapsa*) genera. To our knowledge, *Hymenomonas roseola* (Coccolithales) is the only calcifying freshwater haptophyte that has been described (Manton and Peterfi,

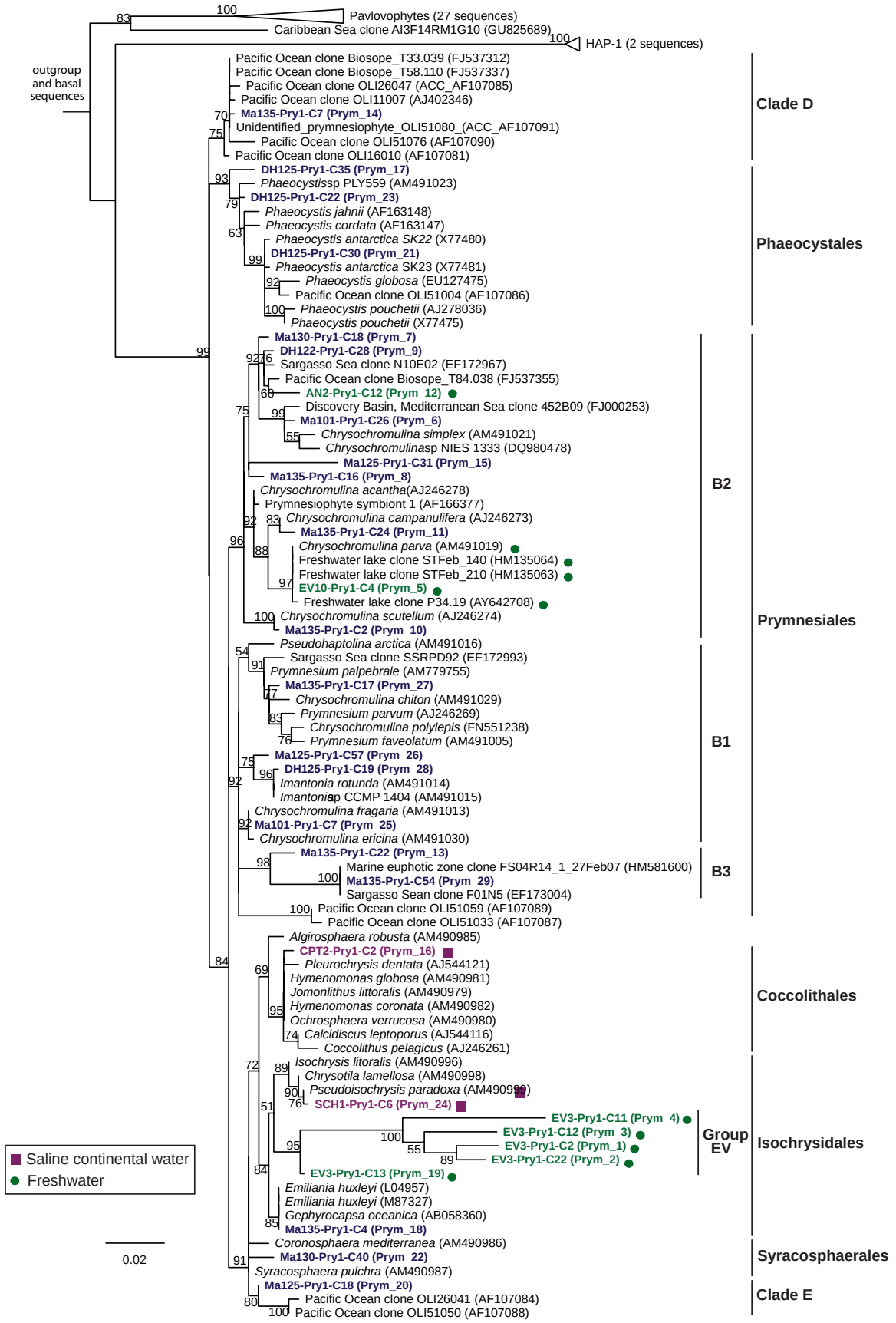
outgroup and basal sequences
Pavlovophytes (27 sequences)
Caribbean Sea clone AI3F14RM1G10 (GU825689)
HAP-1 (2 sequences)
Pacific Ocean clone Biosope_T33.039 (FJ537312)
Pacific Ocean clone Biosope_T58.110 (FJ537337)
Pacific Ocean clone OLI26047 (ACC_AF107085)
Pacific Ocean clone OLI11007 (AJ402346)
Ma135-Pry1-C7 (Prym_14)
Unidentified_prymnesiophyte_OLI51080_(ACC_AF107091)
Pacific Ocean clone OLI51076 (AF107090)
Pacific Ocean clone OLI16010 (AF107081)
Clade D
DH125-Pry1-C35 (Prym_17)
Phaeocystissp PLY559 (AM491023)
DH125-Pry1-C22 (Prym_23)
Phaeocystis jahnii (AF163148)
Phaeocystis cordata (AF163147)
Phaeocystis antarctica SK22 (X77480)
DH125-Pry1-C30 (Prym_21)
Phaeocystis antarctica SK23 (X77481)
Phaeocystis globosa (EU127475)
Pacific Ocean clone OLI51004 (AF107086)
Phaeocystis pouchetii (AJ278036)
Phaeocystis pouchetii (X77475)
Phaeocystales
Ma130-Pry1-C18 (Prym_7)
DH122-Pry1-C28 (Prym_9)
Sargasso Sea clone N10E02 (EF172967)
Pacific Ocean clone Biosope_T84.038 (FJ537355)
AN2-Pry1-C12 (Prym_12)
Discovery Basin, Mediterranean Sea clone 452B09 (FJ000253)
Ma101-Pry1-C26 (Prym_6)
Chrysochromulina simplex (AM491021)
Chrysochromulinasp NIES 1333 (DQ980478)
Ma125-Pry1-C31 (Prym_15)
Ma135-Pry1-C16 (Prym_8)
Chrysochromulina acantha(AJ246278)
Prymnesiophyte symbiont 1 (AF166377)
Chrysochromulina campanulifera (AJ246273)
Ma135-Pry1-C24 (Prym_11)
Chrysochromulina parva (AM491019)
Freshwater lake clone STFeb_140 (HM135064)
Freshwater lake clone STFeb_210 (HM135063)
EV10-Pry1-C4 (Prym_5)
Freshwater lake clone P34.19 (AY642708)
Chrysochromulina scutellum (AJ246274)
Ma135-Pry1-C2 (Prym_10)
B2
Pseudohaptolina arctica (AM491016)
Sargasso Sea clone SSRPD92 (EF172993)
Prymnesium palpebrale (AM779755)
Ma135-Pry1-C17 (Prym_27)
Chrysochromulina chiton (AM491029)
Prymnesium parvum (AJ246269)
Chrysochromulina polylepis (FN551238)
Prymnesium faveolatum (AM491005)
Ma125-Pry1-C57 (Prym_26)
DH125-Pry1-C19 (Prym_28)
Imantonia rotunda (AM491014)
Imantoniasp CCMP 1404 (AM491015)
Chrysochromulina fragaria (AM491013)
Ma101-Pry1-C7 (Prym_25)
Chrysochromulina ericina (AM491030)
B1
Ma135-Pry1-C22 (Prym_13)
Marine euphotic zone clone FS04R14_1_27Feb07 (HM581600)
Ma135-Pry1-C54 (Prym_29)
Sargasso Sean clone F01N5 (EF173004)
Pacific Ocean clone OLI51059 (AF107089)
Pacific Ocean clone OLI51033 (AF107087)
B3
Prymnesiales
Algirosphaera robusta (AM490985)
CPT2-Pry1-C2 (Prym_16)
Pleurochrysis dentata (AJ544121)
Hymenomonas globosa (AM490981)
Jomonlithus littoralis (AM490979)
Hymenomonas coronata (AM490982)
Ochrosphaera verrucosa (AM490980)
Calcidiscus leptoporus (AJ544116)
Coccolithus pelagicus (AJ246261)
Coccolithales
Isochrysis litoralis (AM490996)
Chrysotila lamellosa (AM490998)
Pseudoisochrysis paradoxa (AM490999)
SCH1-Pry1-C6 (Prym_24)
EV3-Pry1-C11 (Prym_4)
EV3-Pry1-C12 (Prym_3)
EV3-Pry1-C2 (Prym_1)
EV3-Pry1-C22 (Prym_2)
Group EV
EV3-Pry1-C13 (Prym_19)
Emiliania huxleyi (L04957)
Emiliania huxleyi (M87327)
Gephyrocapsa oceanica (AB058360)
Ma135-Pry1-C4 (Prym_18)
Isochrysidales
Coronosphaera mediterranea (AM490986)
Ma130-Pry1-C40 (Prym_22)
Syracosphaera pulchra (AM490987)
Syracosphaerales
Ma125-Pry1-C18 (Prym_20)
Pacific Ocean clone OLI26041 (AF107084)
Pacific Ocean clone OLI51050 (AF107088)
Clade E
Saline continental water
Freshwater
0.02

Fig. 2. Maximum likelihood phylogenetic tree of 18S rDNA haptophyte sequences of marine, freshwater and salty continental habitats showing prymnesiophyte clades. A total of 758 non-ambiguously aligned positions were used to reconstruct the tree; gaps were excluded. Two cryptophyte sequences were used as outgroup. 18S rRNA gene sequences from this work are shown in bold. Pavlovophyte and HAP-1 branches are shown collapsed. Bootstrap values greater than 50% are shown at nodes. The scale bar represents the estimated number of substitutions per 100 positions per a unit branch length.

1969). The lack of 18S rRNA gene sequences of *H. roseola* prevents us from establishing the phylogenetic relationship between this species and the sequences detected in our freshwater systems. Nonetheless, there is little chance that the EV group, affiliated to Isochrysidales, turns out to be closely related to *H. roseola*, which has been classified within the Coccolithales based on classical morphological description. Future morphological identification of members of the group EV should show whether the former are actually calcifying or not.

Even if our phylogenetic analysis strongly supports its monophyly, the group EV is highly diverse. Indeed, sequences forming this group share 95% of identity only (average value calculated on 16 complete sequences). For comparison, 18S rRNA sequences of the polyphyletic genus *Chrysochromulina*, which is scattered in the order Prymnesiales (Fig. 2), share 96% identity [average value calculated on full-length 18S rDNA genes from eight *Chrysochromulina* species by Caron *et al*. (2009)]. Sequences belonging to group EV were found in both 0.22–5 μm and 5–30 μm fractions at the Etang des Vallées (Table 1), with OTUs Prym_1 and Prym_3 being shared by both the size fractions, although more diversity was retrieved in the biggest cell-size fractions. There are three possible explanations for the observation of members of this group in different fractions. One explanation would be imperfect size fractionation that could happen if, for instance, the cells are fragile and lyse during

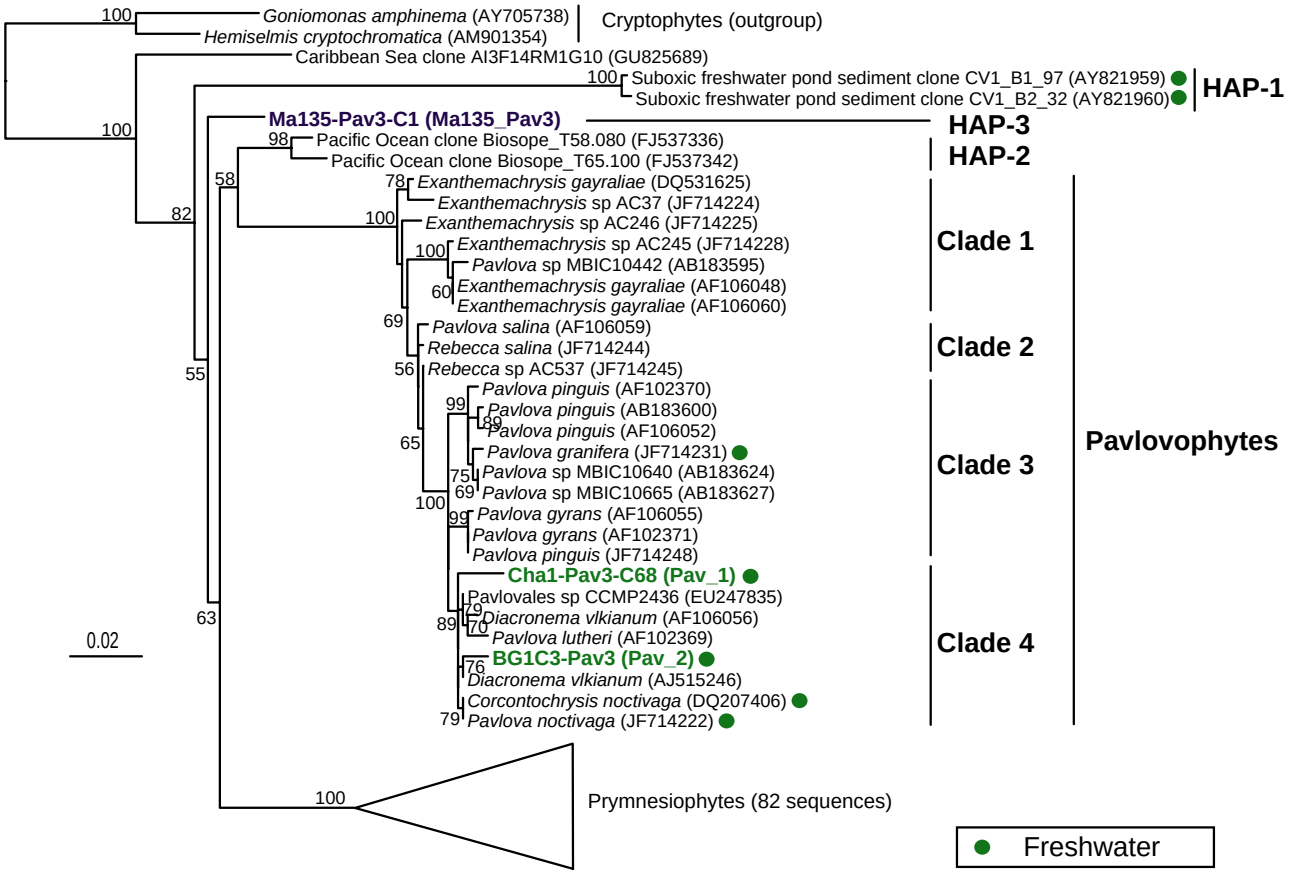

Fig. 3. Maximum likelihood phylogenetic tree of 18S rDNA haptophyte sequences of marine, freshwater and salty continental habitats showing the diversity of pavlovophytes and basal haptophyte lineages. A total of 851 non-ambiguously aligned positions were used to reconstruct the tree; gaps were excluded. 18S rRNA gene sequences from this work are shown in bold. The prymnesiophyte branch is showed collapsed. Bootstrap values greater than 50% are shown at nodes. The scale bar represents the estimated number of substitutions per 100 positions per unit branch length.

the filtration process, or can be deformed and pass through filters under the applied filtration pressure. Second, this could be explained by a diversity of cell sizes within members of group EV due, for instance, to the presence of smaller gametes or different cell sizes in different life cycle phases (e.g. haploid–diploid transitions). Finally, a third possibility would be the non-obligate physical association of those organisms with bigger ones (e.g. symbiosis). The fact that OTUs Prym_1, Prym_2, Prym_3 and Prym_4 form a tight and diverse subgroup with much longer branches than most other prymnesiophyte sequences in our phylogenetic tree (Fig. 2) might suggest that they are mutualistic or parasitic haptophytes, as symbiotic lineages tend to accelerate their evolutionary rate (Wernegreen, 2002; Bromham, 2009). Symbiotic prymnesiophytes have already been observed, for instance in association with the planktonic foraminifer *Globigerinella siphonifera* (Gast *et al*., 2000), which can either live with its host or be free-living.

This new group EV has been detected in two geographically distant and ecologically different freshwater ecosystems: the shallow and oligo-mesotrophic shallow lake Etang des Vallées and the deep and oligotrophic Lake Annecy (Fig. 1). Similarly, the divergent freshwater clade HAP-1 has also been detected in two different and geographically distant environments, first in the sediment of a suboxic pond in France (Šlapeta *et al*., 2005), then in the sediment of a high oligo- to mesotrophic alpine lake in Norway (Shalchian-Tabrizi *et al*., 2011). This suggests that, although not necessarily abundant, very divergent haptophyte lineages may still be found in the understudied freshwater systems. Investigating these lineages might help to reconstruct the evolutionary history of the group and to understand its ecology.

Multiple independent marine–freshwater transitions

The two pavlovophyte OTUs detected in this work were composed of sequences retrieved only from freshwater systems and close to the emended genus *Diacronema* (clade 4) (Bendif *et al*., 2011). OTU Pav_1 was detected in three ponds (Fig. 1, Table S2) and had as closest BLAST hit the sequence of *Corcontochrysis noctivaga* [synonym of *Diacronema noctivaga* (Bendif *et al*., 2011)] strain AC88 (Table S2), isolated from freshwater. It also seems close to sequences Finsevatn 89.12, Svaersvann14 and Svaersvann16, which were retrieved from freshwater (Shalchian-Tabrizi *et al*., 2011). OTU Pav_2 was only encountered in one lake (Fig. 1, Table S2) and affiliated to *Diacronema vlkanium*, which has been visually recorded in freshwater, brackish and marine habitats (Preisig, 2002). Among prymnesiophytes, OTUs Prym_5 and Prym_12 (Prymnesiales, clade B2) were only recorded in freshwater environments. OTU Prym_5 was affiliated to *C. parva*, a well-known and widely distributed freshwater toxic species (Hansen *et al*., 1994; Nicholls, 2003; Edvardsen and Imai, 2006; Luo *et al*., 2011). We detected *C. parva* sequences in seven out of the 15 freshwater ecosystems studied, thus confirming its broad distribution (Fig. 1, Table S2). Within the clade B2 of Prymnesiales, the OTU Prym_12, which was detected in the oligotrophic lake Annecy, clustered with environmental marine sequences (Fig. 2). The remaining freshwater OTUs Prym_1, Prym_2, Prym_3, Prym_4 and Prym_19 formed the group EV within the Isochrysidales, a prymnesiophyte order with no known freshwater representative. These freshwater haptophyte lineages may then represent five distinct transitions from marine to freshwater environments.

In addition to OTUs clustering within known haptophyte orders, the divergent lineage HAP-1 (Šlapeta *et al*., 2005; Shalchian-Tabrizi *et al*., 2011), together with the species *Pavlova granifera* (Green, 1973) and six OTUs affiliated to the clade B1 of Prymnesiales (Finsevatn AI7UI, Finsevatn AKXPZ, Finsevatn AYOY0H) or to the pavlovophyte clade 4 (Finsevatn 8912, Svaersvann 14 and 16) (Shalchian-Tabrizi *et al*., 2011), have only been found in freshwater systems and might represent other examples of putative marine–freshwater transitions. Altogether, there might have been at least nine freshwater colonization events from marine waters. In addition, a few species, for which the 18S rRNA gene sequences are not yet available, have been visually observed in freshwater systems such as *Chrysochromulina laurentiana*, *Chrysochromulina inormamenta*, *Chrysochromulina breviturrita* (Hansen *et al*., 1994; Nicholls, 2003) and *H. roseola* (Manton and Peterfi, 1969).

Continental salty ecosystems also harbour particular haptophytes as shown by OTUs Prym_16 and Prym_24 (Fig. 2) that were isolated in a brackish pond (France) and from Chiprana, belonging to the hypersaline lake complex in the central Ebro basin (Spain) (Jonkers *et al*., 2003). OTU Prym_16 was affiliated to *Jomonlithus littoralis* ALGO Je5, a coastal marine species, while OTU Prym_24 was related to *Pseudoisochrysis paradoxa* CCAP949/1, isolated from the brackish York River Estuary in Virginia (USA) and *Chrysotila lamellosa*, a species often isolated from coastal marine regions (such as the strain ALGO HAP17) or brackish continental environments (such as CCAP 918/1 isolated in London, UK). Haptophytes recorded in salty continental systems thus appear to be phylogenetically close to species of either coastal habitats or continental brackish systems. This finding is in agreement with the importance of salinity as a barrier for marine–continental environment transitions, (Lozupone and Knight, 2007; Logares *et al*., 2009), in spite of the fact that haptophytes seem to have crossed that barrier several times in the course of their evolution.

Acknowledgements

We are most grateful to Stéphan Jacquet from the Institut National de la Recherche Agronomique, UMR42, CARRTEL, Thonon-les-Bains (France) and Anabel López-Archilla from the Universidad Autónoma de Madrid for generously providing samples of, respectively, Lakes Bourget and Annecy (France) and Salada de Chiprana (Spain). We also thank P. Henry and N. Cagatay, chief scientists of the oceanographic cruise MARNAUT (2007), for providing P.L.G. the opportunity to collect samples in the Marmara Sea, as well as the Captain, the crew and the scientific team onboard of the R/V L'Atalante. We acknowledge the help of the Turkish Navy to protect our vessel in the zones of heavy ship traffic. We thank INSU for providing a CTD-rosette for water sampling and L. Fichen, who operated it. This work was financed by a CNRS-INSU programme Ecosphère Continentale et Cotière (EC2CO) and a University Paris-Sud Appel à Projet Attractivité. M.S. is recipient of a fellowship from the Ecole Normale Supérieure.

References

Andersen, R.A., Bidigare, R.R., Keller, M.D., and Latasa, M. (1996) A comparison of HPLC pigment signatures and electron microscopic observations for oligotrophic waters of the North Atlantic and Pacific Oceans. *Deep Sea Res II* **43:** 517–537.

Bass, D., and Cavalier-Smith, T. (2004) Phylum-specific environmental DNA analysis reveals remarkably high global biodiversity of Cercozoa (Protozoa). *Int J Syst Evol Microbiol* **54:** 2393–2404.

Bendif, E.M., Probert, I., Hervé, A., Billard, C., Goux, D., Lelong, C., *et al*. (2011) Integrative taxonomy of the Pavlovophyceae (Haptophyta): a reassessment. *Protist* **162:** 738–761.

Bromham, L. (2009) Why do species vary in their rate of molecular evolution? *Biol Lett* **5:** 401–404.

Burki, F., Inagaki, Y., Bråte, J., Archibald, J.M., Keeling, P.J., Cavalier-Smith, T., *et al*. (2009) Large-scale phylogenomic analyses reveal that two enigmatic protist lineages, telonemia and centroheliozoa, are related to photosynthetic chromalveolates. *Genome Biol Evol* **1:** 231–238.

Burki, F., Okamoto, N., Pombert, J.-F., and Keeling, P.J. (2012) The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. *Proc Biol Sci* **279:** 2246–2254.

Caron, D.A., Countway, P.D., Savai, P., Gast, R.J., Schnetzer, A., Moorthi, S.D., *et al*. (2009) Defining DNA-based operational taxonomic units for microbial-eukaryote ecology. *Appl Environ Microbiol* **75:** 5797–5808.

Cuvelier, M.L., Allen, A.E., Monier, A., McCrow, J.P., Messié, M., Tringe, S.G., *et al*. (2010) Targeted metagenomics and ecology of globally important uncultured eukaryotic phytoplankton. *Proc Natl Acad Sci USA* **107:** 14679–14684.

De Vargas, C., Aubry, M.-P., Probert, I., and Young, J. (2007) Origin and evolution of coccolithophores: from coastal hunters to oceanic farmers. In *Evolution of Primary Producers in the Sea*. Falkowski, P.G., and Knoll, A.H. (eds). London, UK: Elsevier, pp. 251–285.

Edvardsen, B., and Imai, I. (2006) The ecology of harmful flagellates within Prymnesiophyceae and Raphidophyceae. In *Ecology of Harmful Algae*. Granéli, E., and Turner, J.T. (eds). Berlin, Germany: Springer, pp. 67–78.

Edvardsen, B., and Medlin, L.K. (2007) Molecular systematics of Haptophyta. In *Unravelling the Algae: The past, Present and Future of Algal Systematics*. Brodie, J., and Lewis, J. (eds). Oxford, UK: CRC Press, Taylor and Francis Group, pp. 183–196.

Edvardsen, B., Eikrem, W., Green, J.C., Andersen, R.A., Moon-van der Staay, S.Y., and Medlin, L.K. (2000) Phylogenetic reconstructions of the Haptophyta inferred from 18S ribosomal DNA sequences and available morphological data. *Phycologia* **39:** 19–35.

Edvardsen, B., Eikrem, W., Throndsen, J., Sáez, A.G., Probert, I., and Medlin, L.K. (2011) Ribosomal DNA phylogenies and a morphological revision provide the basis for a revised taxonomy of the Prymnesiales (Haptophyta). *Eur J Phycol* **46:** 202–228.

Frias-López, J., Thompson, A., Waldbauer, J., and Chisholm, S.W. (2009) Use of stable isotope-labelled cells to identify active grazers of picocyanobacteria in ocean surface waters. *Environ Microbiol* **11:** 512–525.

Gast, R.J., McDonnell, T.A., and Caron, D.A. (2000) srDna-based taxonomic affinities of algal symbionts from a planktonic foraminifer and a solitary radiolarian. *J Phycol* **36:** 172–177.

Green, J.C. (1973) Studies in the fine structure and taxonomy of flagellates in the genus *Pavlova*. II. A freshwater representative, *Pavlova granifera* (Mack) comb. *Brit Phycol J* **8:** 1–12.

Hansen, L., Kristiansen, J., and Rasmussen, J. (1994) Potential toxicity of the freshwater *Chrysochromulina* species *C. parva* (Prymnesiophyceae). *Hydrobiologia* **287:** 157–159.

Jardillier, L., Zubkov, M.V., Pearman, J., and Scanlan, D.J. (2010) Significant CO2 fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. *ISME J* **4:** 1180–1192.

Jonkers, H.M., Ludwig, R., De Wit, R., Pringault, O., Muyzer, G., Niemann, H., *et al*. (2003) Structural and functional analysis of a microbial mat ecosystem from a unique permanent hypersaline inland lake: 'La Salada de Chiprana' (NE Spain). *FEMS Microbiol Ecol* **44:** 175–189.

Jordan, R.W., Cros, L., and Young, J.R. (2004) A revised classification scheme for living haptophytes. *Micropaleontology* **50:** 55–79.

Keeling, P.J. (2009) Chromalveolates and the evolution of plastids by secondary endosymbiosis. *J Eukaryot Microbiol* **56:** 1–8.

Lara, E., Moreira, D., Vereshchaka, A., and López-García, P. (2009) Pan-oceanic distribution of new highly diverse clades of deep-sea diplonemids. *Environ Microbiol* **11:** 47–55.

Legrand, C., Johansson, N., Johnsen, G., Borsheim, K.Y., and Granéli, E. (2001) Phagotrophy and toxicity variation in the mixotrophic *Prymnesium patelliferum* (Haptophyceae). *Limnol Oceanogr* **46:** 1208–1214.

Lepère, C., Domaizon, I., and Debroas, D. (2008) Unexpected importance of potential parasites in the composition of the freshwater small-eukaryote community. *Appl Environ Microbiol* **74:** 2940–2949.

Lepère, C., Masquelier, S., Mangot, J.-F., Debroas, D., and Domaizon, I. (2010) Vertical structure of small eukaryotes in three lakes that differ by their trophic status: a quantitative approach. *ISME J* **4:** 1509–1519.

Liu, H., Probert, I., Uitz, J., Claustre, H., Aris-Brosou, S., Frada, M., *et al*. (2009) Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. *Proc Natl Acad Sci USA* **106:** 12803–12808.

Logares, R., Bråte, J., Bertilsson, S., Clasen, J.L., Shalchian-Tabrizi, K., and Rengefors, K. (2009) Infrequent marine-freshwater transitions in the microbial world. *Trends Microbiol* **17:** 414–422.

López-García, P., Rodriguez-Valera, F., Pedros-Alió, C., and Moreira, D. (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* **409:** 603–607.

Lozupone, C.A., and Knight, R. (2007) Global patterns in bacterial diversity. *Proc Natl Acad Sci USA* **104:** 11436–11440.

Luo, W., Bock, C., Li, H., Padisák, J., and Krienitz, L. (2011) Molecular and microscopic diversity of planktonic eukaryotes in the oligotrophic Lake Stechlin (Germany). *Hydrobiologia* **661:** 133–143.

Manton, I., and Peterfi, L.S. (1969) Observations of the fine structure of coccoliths, scales and the protoplast of a freshwater coccolithorid, *Hymenomonas roseola* Stein, with supplementary observations of the protoplast of *Cricosphaera carterae*. *Proc R Soc Lond B Biol Sci* **172:** 1–15.

Massana, R., and Pedros-Alio, C. (2008) Unveiling new microbial eukaryotes in the surface ocean. *Curr Opin Microbiol* **11:** 213–218.

Moon-van der Staay, S.Y., van der Staay, G.W.M., Guillou, L., Claustre, H., Medlin, L.K., and Vaulot, D. (2000) Abundance and diversity of prymnesiophytes in the picoplankton community from the equatorial Pacific Ocean inferred from 18S rDNA sequences. *Limnol Oceanogr* **45:** 98–109.

Moon-van der Staay, S.Y., De Wachter, R., and Vaulot, D. (2001) Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* **409:** 607–610.

Nicholls, K.H. (2003) Haptophyte algae. In *Freshwater Algae of North America: Ecology and Classification*. Wehr, J.G., and Sheath, R.G. (eds). San Diego, CA, USA: Elsevier, pp. 511–521.

Not, F., Gausling, R., Azam, F., Heidelberg, J.F., and Worden, A.Z. (2007) Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environ Microbiol* **9:** 1233–1252.

Okamoto, N., Chantangsi, C., Horák, A., Leander, B.S., and 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.

Orsi, W., Song, Y.C., Hallam, S., and Edgcomb, V. (2012) Effect of oxygen minimum zone formation on communities of marine protists. *ISME J* **6:** 1586–1601.

Preisig, H.R. (2002) Phylum Haptophyta (Prymnesiophyta). In *The Freshwater Algal Flora of the British Isles: An Identification Guide to Freshwater and Terrestrial Algae*. John, D.M., Whitton, B.A., and Brook, A.J. (eds). Cambridge, UK: The press syndicate of the University of Cambridge, pp. 211–213.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., and Glöckner, F.O. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35:** 7188–7196.

Reinfelder, J.R. (2011) Carbon concentrating mechanisms in eukaryotic marine phytoplankton. *Ann Rev Mar Sci* **3:** 291–315.

Richards, T.A., Vepritskiy, A.A., Gouliamova, D.E., and Nierzwicki-Bauer, S.A. (2005) The molecular diversity of freshwater picoeukaryotes from an oligotrophic lake reveals diverse, distinctive and globally dispersed lineages. *Environ Microbiol* **7:** 1413–1425.

Shalchian-Tabrizi, K., Reier-Røberg, K., Ree, D.K., Klaveness, D., and Bråte, J. (2011) Marine–freshwater colonizations of haptophytes inferred from phylogeny of environmental 18S rDNA sequences. *J Eukaryot Microbiol* **58:** 315–318.

Shi, X.L., Marie, D., Jardillier, L., Scanlan, D.J., and Vaulot, D. (2009) Groups without cultured representatives dominate eukaryotic picophytoplankton in the oligotrophic South East Pacific Ocean. *PLoS ONE* **4:** e7657.

Šlapeta, J., Moreira, D., and López-García, P. (2005) The extent of protist diversity: insights from molecular ecology of freshwater eukaryotes. *Proc Biol Sci* **272:** 2073–2081.

Triadó-Margarit, X., and Casamayor, E.O. (2012) Genetic diversity of planktonic eukaryotes in high mountain lakes (Central Pyrenees, Spain). *Environ Microbiol* **14:** 2445–2456.

Wernegreen, J.J. (2002) Genome evolution in bacterial endosymbionts of insects. *Nat Rev Genet* **3:** 850–861.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Materials and methods

Fig. S1. Maximum likelihood phylogenetic tree of 18S rDNA haptophyte sequences showing all prymnesiophyte OTUs retrieved in this study, their first hit by blast against the SILVA database SSU104 and sequences representing known orders and environmental lineages, including partially overlapping sequences to our sequences, especially including sequences retrieved from freshwaters. Two cryptophyte sequences were used as outgroup. The alignment contained 1682 selected positions. Positions on a 2147 bp alignment having less than 50% gaps were retained to reconstruct the tree using BMGE. 18S rRNA gene sequences from this work are shown in bold. Full circles indicate sequences of freshwater and salty continental habitats; other sequences are from marine ecosystems. Bootstrap values greater than 50% are shown at nodes (1000 replicates). The scale bar represents the number of substitutions per 100 positions per a unit branch length.

Fig. S2. Maximum likelihood phylogenetic tree of 18S rDNA haptophyte sequences of marine, freshwater and salty continental habitats. This tree was built using the same sequences as in Fig. 3 plus shorter and/or partially overlapping environmental sequences related to our sequences. The

82 prymnesiophyte sequences used to construct the tree are shown collapsed. The alignment contained 1678 selected positions (positions with less than 50% gaps were selected using BMGE on a 2131 bp alignment). 18S rRNA gene sequences from this work are shown in bold. Full circles indicate freshwater sequences; other sequences are from marine ecosystems. Bootstrap values greater than 50% are shown at nodes (1000 replicates). The scale bar represents the number of substitutions per 100 positions per unit branch length.

Table S1. Major characteristics of the samples analysed in this study. The positive or negative amplification of 18S rRNA genes with prymnesiophyte (Prym.) or pavlovophyte (Pav.)-specific primers is indicated with '+' or '–' signs. n.a., not applicable; n.d., not done.

Table S2. OTUs identified in this work. The name of representative sequences for each OTU, their first BLAST hit in the Silva Database SSU104, their percentage of similarity as well as the total number of sequences retrieved in each system are given.

Environmental Microbiology Reports

New haptophyte lineages and multiple independent colonisations of freshwater ecosystems

Marianne Simon, Purificación López-García , David Moreira, and Ludwig Jardillier

Supplementary Information

Materials and methods

Study sites and sampling

To look for potential new lineages of marine haptophytes, we used DNA of marine plankton from two very distant water columns in the South Atlantic and the Sea of Marmara collected, respectively, in December 1998 (DHARMA cruise) (López-García et al., 2001) and June 2007 (MARNAUT cruise (Lara et al., 2009) (Table 1 and Table S1). To explore a variety of ecosystems across a salinity gradient, sediment and water samples from a hypersaline lake (La Salada de Chiprana, in Spain, see (Jonkers et al., 2003) for complete description) were also studied. We also used DNA extracted from culture enrichments from a brackish pond in the ornithological park of Teich (France, close to the Arcachon laguna shores). In addition, fifteen geographically distant and close freshwater ponds, lakes and brooks were sampled in France from 2003 to 2011 at surface (Table 1 and Table S2).

Plankton cells of different size ranges (0.22-60 µm, 5-30 µm and 0.22-5 µm; see Table 1) were collected and concentrated either by filtering onto polycarbonate filters of different pore-sizes (Millipore), using Cell Traps (Mem-Teq) retaining cells ranging from 0.22 to 60 µm in diameter thanks to a prefiltration through a 60 µm pore-size filter, or by centrifugation after culture enrichment with K-medium (Keller et al., 1987) and Volvic mineral water. DNA was than

extracted from the concentrated cells in a way depending on the collecting method used. Total DNA was extracted from filters stored at -20°C using the PowerSoil DNA extraction kit (MoBio) according to the manufacturer's instructions. DNA was eluted in 50-80µL of Tris-HCl (10 mM pH 8, stock). Total DNA extraction from samples concentrated in CellTraps, flash frozen and stored at -80°C, was performed according to a protocol adapted from Jardillier et al. (Jardillier et al., 2010). Briefly, 1 µL SDS 10% and 1 µL Proteinase K were added to 100 µL culture. After 45 min incubation at 50°C, cells were incubated 10 min at room temperature with 4 µL GenElute and 16 µL Sodium Acetate. DNA was then precipitated and washed in ethanol before being resuspended in 10 µL 10 mM Tris. Finally, when culture enrichments were performed, PCR were performed directly on culture pellets washed and resuspended on 10 mM Tris, or after DNA extraction as for DNA extraction from CellTraps.

18S rRNA gene amplification, cloning and sequencing

18S rRNA gene sequences were amplified using two sets of primers designed to target specifically Prymnesiophyceae (Pry421F: 5'-AGCAGGCGCGTAAATTGCCCG-3' + Pry1572R: 5'-TCAACGYRCGCTGATGACA-3') and Pavlovophyceae (Hap220F: 5'-ACCGGTCTCCGGTTGCGTGC-3' + Pav1702R: 5'-TAGATGATAAGGTTTGGGTG-3') respectively. PCR amplifications were performed in a total reaction volume of 25µl using 1µL of DMSO (Sigma-Aldrich), 1.5 mM $MgCl_2$, 0.2 mM dNTP, 0.2 µM of primers, 1 × PCR buffer and 0.5 U HotStart Taq polymerase (Taq Platinum, Invitrogen). TouchDown amplification conditions consisted of 10 cycles of 94°C for 15 s, 60 to 51°C for 30 s, 72°C for 2 min, followed by 25 cycles of 94°C for 15 s, 55°C for 30 to 54 s and 72°C for 2 min and a final extension step of 72°C for 7 min. 18S rRNA gene sequences from continental samples were amplified using a nested

PCR, first using general Eukaryote primers 82F and 1498R as described in López-García et al. (López-García et al., 2003) and then using either prymnesiophytes or pavlovophytes specific primers as described above. PCR gene amplicons were cloned using the TOPO TA cloning kit (Invitrogen) according to the supplier's instructions. 20-40 clones per sample (Fig. 1 and Table S2) were chosen randomly and the 18S rDNAs contained in those clones were sequenced using the forward primer previously used for the DNA amplification (Beckman Coulter Genomics, Takeley, IK).

Sequence and molecular phylogenetic analyses

Sequences were manually screened for the presence of chimeric artefacts by independent BLAST searches of the 5' and the 3' halves as well as by using the software package KeyDNATool (http://KeyDNAtools.com). The phylogenetic positions of the sequences were initially assessed by blasting them against the Silva SSU104-parc1000 database (Pruesse et al., 2007). Sequences affiliated to haptophytes were aligned with the ProbCons software (Do et al., 2005). Distance matrices calculated using ClustalX (Larkin et al., 2007) were used to construct Operational Taxonomic Units (OTUs) using a 98% sequence similarity cut-off with Mothur software (Schloss et al., 2009). Representative sequences of all OTUs were then sequenced using the reverse amplification primer in order to obtain complete clone sequences of about 1400 bp for pavlovophytes and 1070 bp for prymnesiophytes. For the final phylogenetic trees, we incorporated representative sequences of all cultured haptophyte taxa, plus a wide representation of environmental sequences that were nearly full-length. We also reconstructed trees containing short environmental haptophyte sequences (<500 bp) with pivotal positions allowing for an important proportion of gaps (see Figs. S1 and S2 below). However, due to the limited number of non-ambiguously aligned positions available for phylogenetic analysis, short haptophyte

sequences were not included in final phylogenetic trees where the position of divergent lineages was to be ascertained (Figs. 2 and 3). The number of positions selected from the multiple alignments was selected using Gblocks (Castresana, 2000) or BMGE (Criscuolo and Gribaldo, 2010). Phylogenetic trees were reconstructed by maximum likelihood (ML) using Treefinder (Jobb et al., 2004) applying a GTR model of sequence evolution considering 4 rate categories and taking among-site rate variation into account by using a four-category discrete approximation of a Γ distribution. ML bootstrap proportions were inferred using 1,000 replicates. Phylogenetic trees were visualized using the program FIGTREE (http://tree.bio.ed.ac.uk/software/figtree/). Complete sequences affiliated to haptophytes have been deposited in GenBank under accession numbers JX680338 - JX680446.

References

Castresana, J. (2000) Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Molecular Biology and Evolution* **17**: 540-552.

Criscuolo, A., and Gribaldo, S. (2010) BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evolutionary Biology* **10**: 210.

Do, C.B., Mahabhashyam, M.S.P., Brudno, M., and Batzoglou, S. (2005) ProbCons: Probabilistic consistency-based multiple sequence alignment. *Genome Res* **15**: 330-340.

Jardillier, L., Zubkov, M.V., Pearman, J., and Scanlan, D.J. (2010) Significant CO2 fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. *ISME J* **4**: 1180-1192.

Jobb, G., von Haeseler, A., and Strimmer, K. (2004) TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. In: BioMed Central Ltd.

Jonkers, H.M., Ludwig, R., De Wit, R., Pringault, O., Muyzer, G., Niemann, H. et al. (2003) Structural and functional analysis of a microbial mat ecosystem from a unique permanent hypersaline inland lake: 'La Salada de Chiprana' (NE Spain). *FEMS Microbiol Ecol* **44**: 175-189.

Keller, M.D., Selvin, R.C., Claus, W., and Guillard, R.R.L. (1987) Media for the culture of oceanic ultraphytoplankton. *J Phycol* **23**: 633-638.

Lara, E., Moreira, D., Vereshchaka, A., and López-García, P. (2009) Pan-oceanic distribution of new highly diverse clades of deep-sea diplonemids. *Environ Microbiol* **11**: 47-55.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H. et al. (2007) Clustal W and clustal X version 2.0. *Bioinformatics* **23**: 2947-2948.

López-García, P., Moreira, D., López-López, A., and Rodríguez-Valera, F. (2001) A novel haloarchaeal-related lineage is widely distributed in deep oceanic regions. *Environ Microbiol* **3**: 72-78.

López-García, P., Philippe, H., Gail, F., and Moreira, D. (2003) Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the Mid-Atlantic Ridge. *Proc Natl Acad Sci* **100**: 697-702.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., and Glöckner, F.O. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**: 7188-7196.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B. et al. (2009) Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Appl Environ Microbiol* **75**: 7537-7541.

Table S1. Major characteristics of the samples analysed in this study. The positive or negative amplification of 18S rRNA genes with prymnesiophyte (Prym.) or pavlovophyte (Pav.)-specific primers is indicated with '+' or '-' signs. n.a., not applicable; n.d., not done.

Location	GPS coordinates	Ecosystem type	Sample name	Sampling date	Depth	Size fraction	Volume filtered	Prym.	Pav.
Charca Verde, campus University Paris-Sud	48°42'02"N 2°10'28"E	Freshwater pond	CV1	November 2003	Surface	0.22 - 5 µm	150 ml	-	-
			MVSF	February 2005	Sediment surface	n.d.	n.a.	-	-
Chevreuse, PNR Haute Vallée de Chevreuse	48°42'18.6"N 2°02'23.5"E	Freshwater urban pond	CH1	December 2007	Surface	0.22 - 5 µm	500 ml	+	-
			CH2	December 2007	Surface	>5 µm	500 ml	+	-
Etang de la Tour, PNR Haute Vallée de Chevreuse	48°39'39.2"N 1°52'48.3"E	Freshwater pond	To1	June 2010	Surface	5 – 30 µm	500 ml	+	-
			To2	June 2010	Surface	0.22 - 5 µm	200 ml	+	-
Etang des Vallées, PNR Haute Vallée de Chevreuse	48°41'23.0"N 1°54'59.2"E	Freshwater shallow lake	EV2	February 2010	Surface	0.22 - 5 µm	300 ml	+	-
			EV3	February 2010	Surface	5 – 30 µm	300 ml	+	-
			EV6	June 2010	Surface	0.22 - 5 µm	100 ml	-	-
			EV7	June 2010	Surface	5 – 30 µm	500 ml	-	-
			EV8	September 2010	Surface	0.22 - 5 µm	140 ml	+	-
			EV9	September 2010	Surface	5 – 30 µm	200 ml	+	-
			EV10	March 2011	Surface	0.22 - 5 µm	100 ml	+	+
Mare Gabard, PNR Haute Vallée de Chevreuse	48°39'15.83"N 1°55'20.26"E	Freshwater pond, forest	MG1	March 2011	Surface	0.22 - 5 µm	100 ml	-	+
			MG2	March 2011	Surface	5 – 30 µm	200 ml	-	-
Saint Robert, PNR Haute Vallée de Chevreuse	48°39'54.82"N 1°56'45.28"E	Freshwater village pond	SR1	March 2011	Surface	0.22 - 5 µm	100 ml	-	-
			SR2	March 2011	Surface	5 – 30 µm	200 ml	-	-
Ru Sainte Anne, PNR Haute Vallée de Chevreuse	48°36'45.91"N 1°58'16.61"E	Freshwater brook	RSA1	March 2011	Surface	0.22 - 5 µm	100 ml	-	-
			RSA2	March 2011	Surface	5 – 30 µm	500 mL	-	-
La Claye, PNR Haute Vallée de Chevreuse	48°36'31.72"N 1°56'17.33"E	Freshwater pond, forest	LC1	March 2011	Surface	0.22 - 5 µm	100 ml	-	-
			LC2	March 2011	Surface	5 – 30 µm	500 ml	-	-
Etang du Perray, PNR Haute Vallée de Chevreuse	48°41'49"N 1°51'37"E	Freshwater pond	Pe1	June 2010	Surface	5 – 30 µm	200 ml	+	-
Etang de Pourras, PNR Haute Vallée de Chevreuse	48°42'52"N 1°50'39"E	Freshwater pond	Po1	June 2010	Surface	5 – 30 µm	200 ml	+	-
Etang de Cernay, PNR Haute Vallée de Chevreuse	48°40'50"N 1°57'55"E	Freshwater pond	Ce1	June 2010	Surface	5 – 30 µm	200 ml	-	-
Lac du Bourget, Savoie	45°44'N 05°51'E	Freshwater lake	BG1	May 2009	Surface	0.22 - 60 µm	10 l	n.d.	+
			BG6	April 2009	15 m	0.22 - 60 µm	10 l	+	-
Lac d'Annecy Northern basin, Savoie	45°54'N 06°07'E	Freshwater lake	AN1	May 2009	Surface	0.22 - 60 µm	10 l	-	-
			AN2	May 2009	15 m	0.22 - 60 µm	10 l	+	n.d.
			AN6	April 2009	15 m	0.22 - 60 µm	10 l	+	-
Etang d'en haut, Paimpont, Britain	48°00'30.66"N 2°13'36.06"W	Freshwater pond	Ht1	April 2010	Surface	0.22 - 5 µm	500 ml	-	-
Etang du Châtenay, Paimpont, Britain	48°00'14.40"N 2°13'48.36"W	Freshwater pond	Châ1	April 2010	Surface	0.22 - 5 µm	500 ml	-	+
			Châ2	April 2010	Surface	5 - 30 µm	200 ml	-	-
South Atlantic	56°18'57"S 57°39'45"E	Marine	DH122	December 1998	10 m	0.22 - 5 µm	23 l	+	-
			DH123	December 1998	25 m	0.22 - 5 µm	22 l	+	-
			DH125	December 1998	100 m	0.22 - 5 µm	20.5 l	+	-
			DH129	December 1998	2000 m	0.22 - 5 µm	32.5 l	+	-
Marmara Sea, Central basin	40°50'18.48"N 28°01'24.24"E	Marine	Ma101	June 2007	997 m	0.22 - 5 µm	10-15 l	+	-
			Ma125	June 2007	100 m	0.22 - 5 µm	5 l	+	-
			Ma130	June 2007	25 m	0.22 - 5 µm	4 l	+	-
			Ma135	June 2007	15 m	0.22 - 5 µm	2.5 l	+	+
Salada Chiprana, Spain	41°14' 30" N 0°10' 50" W	Hypersaline	SCH1	February 2010	Sediment surface	n.a.	n.a.	+	-
			SCH2	February 2010	0.5 m	0.22 - 5 µm	100 ml	+	-
			SCH3	February 2010	3.5 m	0.22 - 5 µm	25 ml	+	-
Ornithological Park of Teich, France	44°38'27.72"N 1°01'14.04"W	Brackish	CPT2	February 2010	Culture	n.d.	n.a.	+	-
			CPT3	February 2010	Culture	n.d.	n.a.	+	-
			CPT4	February 2011	Culture	n.d.	n.a.	+	-

Table S2. OTUs identified in this work. The name of representative sequences for each OTU, their first BLAST hit in the Silva Database SSU104, their percentage of similarity as well as the total number of sequences retrieved in each system are given.

		Freshwater									Sea		Salt lakes			
OTU	Representative Sequence	Chevreuse	Tour	Vallées	Gabard	Perray	Pourras	Bourget	Annecy	Châtenay	Dharma	Marmara	Teich	Salada	Chiprana	First hit (% identity)
Ma135-Pav3	Ma135-Pav3-C1											60				Unc marine_haptophyte_Biosope_T58_080 FJ537336 (93.9%)
Pav_1	Cha1-Pav3-C68			9	10					66						Corcontochrysis noctivaga DQ207406 (97.0%)
Pav_2	BG1C3-Pav3							16								Diacronema vlkianum AJ515246 (98.4%)
Prym_1	EV3-Pry1-C2			12					20							Emiliania huxleyi M87327 (92.4%)
Prym_2	EV3-Pry1-C22			1												Emiliania huxleyi M87327 (97.2%)
Prym_3	EV3-Pry1-C12			7												Emiliania huxleyi M87327 (95.2%)
Prym_4	EV3-Pry1-C11			3												Emiliania huxleyi M87327 (92.6%)
Prym_5	EV10-Pry1-C4	13	21	45		10	9	18	6							Chrysochromulina parva AM491019 (99.5%)
Prym_6	Ma101-Pry1-C26										8	14				Unc eukaryote_452B09 FJ000253 (98.9%)
Prym_7	Ma130-Pry1-C18											13				Unc eukaryote_N10E02 EF172967 (98.7%)
Prym_8	Ma135-Pry1-C16											2				Prymnesiophyte symbiont_1 AF166377 (98.5%)
Prym_9	DH122-Pry1-C28										3					Unc eukaryote_N10E02 EF172967 (98.7)
Prym_10	Ma135-Pry1-C2											5				Chrysochromulina scutellum AJ246274 (99.9%)
Prym_11	Ma135-Pry1-C24										1	7				Chrysochromulina campanulifera AJ246273 (99.2%)
Prym_12	AN2-Pry1-C12								23							Unc eukaryote_N10E02 EF172967 (97.9%)
Prym_13	Ma135-Pry1-C22											10				Prymnesiophyte symbiont_1 AF166377 (96.7%)
Prym_14	Ma135-Pry1-C7											1				Unc marine_haptophyte_Biosope_T58_110 FJ537337 (99.9%)
Prym_15	Ma125-Pry1-C31											1				Chrysochromulina scutellum AJ246274 (97.2%)
Prym_16	CPT2-Pry1-C2												2			Jomonlithus littoralis AM490979 (99.8%)
Prym_17	DH125-Pry1-C35										3					P_antarctica Karsten_SK22 X77480 (97.9%)
Prym_18	Ma135-Pry1-C4										6	18				Emiliania huxleyi M87327 (100%)
Prym_19	EV3-Pry1-C13			3												Emiliania huxleyi M87327 (98.4%)
Prym_20	Ma125-Pry1-C18										1	13				Syracosphaera pulchra AM490987 (99.0%)
Prym_21	DH125-Pry1-C30										145	25				P_antarctica Karsten_SK23 X77481 (99.9%)
Prym_22	Ma130-Pry1-C40											1				Syracosphaera pulchra AM490987 (98.7%)
Prym_23	DH125-Pry1-C22										1					Phaeocystis sp_PLY559 AM491023 (97.5%)
Prym_24	SCH-Pry1-C6												1	36		Pseudoisochrysis paradoxa AM490999 (99.6%)
Prym_25	Ma101-Pry1-C7										1	7				Chrysochromulina fragaria AM491013 (99.9%)
Prym_26	Ma125-Pry1-C57											2				Chrysochromulina ericina AM491030 (97.3%)
Prym_27	Ma135-Pry1-C17											40				Unc eukaryote_SSRPD92 EF172993 (99.1%)
Prym_28	DH125-Pry1-C19										1					Imantonia rotunda AM491014 (99.9%)
Prym_29	Ma135-Pry1-C54											3				Unc eukaryote_F01N5 EF173004 (99.8%)

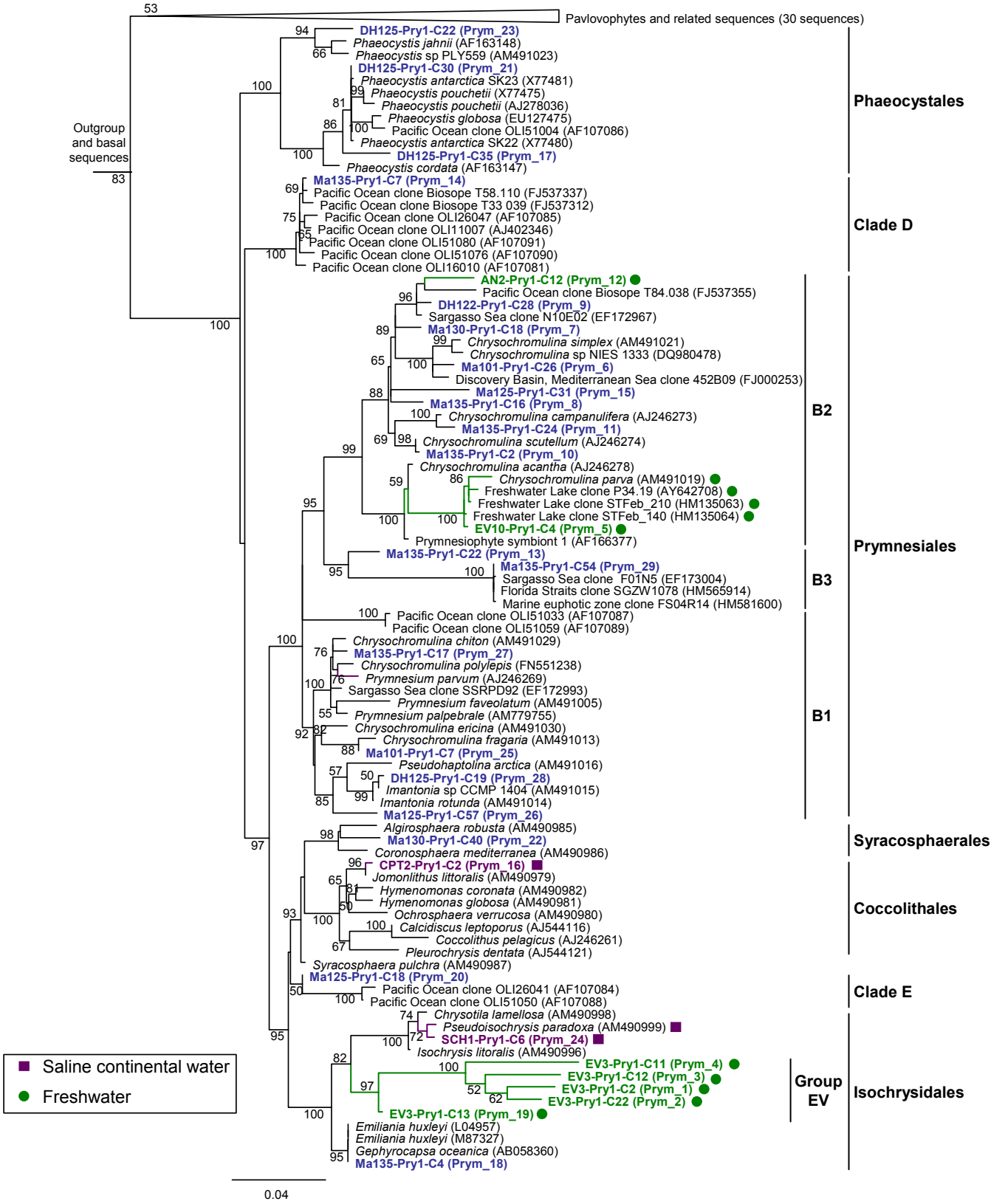

Figure. S1. Maximum likelihood phylogenetic tree of 18S rDNA haptophyte sequences showing all prymnesiophyte OTUs retrieved in this study, their first hit by blast against the SILVA database SSU104 and sequences representing known orders and environmental lineages, including partially overlapping sequences to our sequences, especially including sequences retrieved from freshwaters. Two cryptophyte sequences were used as outgroup. The alignment contained 1682 selected position. Positions on a 2,147bp alignment having less than 50% gaps were retained to reconstruct the tree using BMGE. 18S rRNA gene sequences from this work are shown in bold; colour codes correspond to marine (blue), freshwater (green) or salty continental (purple) systems. Full circles indicate sequences of freshwater and salty continental habitats, other sequences are from marine ecosystems. Bootstrap values greater than 50% are shown at nodes (1000 replicates). The scale bar represents the number of substitutions per 100 positions per a unit branch length.

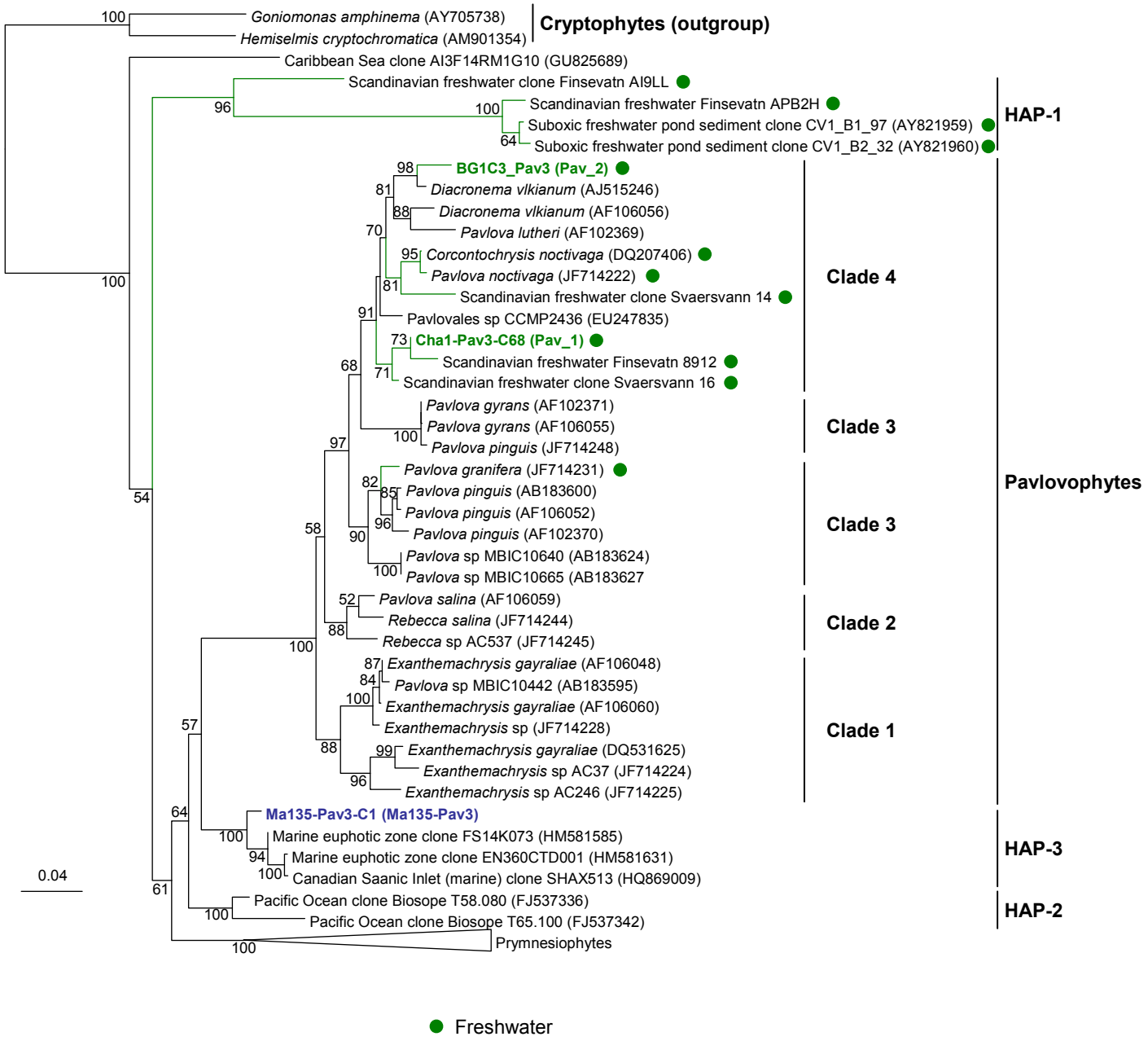

Figure S2. Maximum likelihood phylogenetic tree of 18S rDNA haptophyte sequences of marine, freshwater and salty continental habitats. This tree was built using the same sequences as in Fig. 3. plus shorter and/or partially overlapping environmental sequences related to our sequences. The 82 prymnesiophyte sequences used to construct the tree are shown collapsed. The alignment contained 1,678 selected positions (positions with less than 50% gaps were selected using BMGE on a 2,131 bp alignment). 18S rRNA gene sequences from this work are shown in bold. Full circles indicate freshwater sequences, other sequences are from marine ecosystems. Bootstrap values greater than 50% are shown at nodes (1000 replicates). The scale bar represents the number of substitutions per 100 positions per unit branch length.