The ISME Journal (2010) 4, 1044–1052

www.nature.com/ismej

ORIGINAL ARTICLE

Diversity of active marine picoeukaryotes in the Eastern Mediterranean Sea unveiled using photosystem-II *psbA* transcripts

Dikla Man-Aharonovich[1,5], Alon Philosof[1,5], Benjamin C Kirkup[2], Florence Le Gall[3], Tali Yogev[4], Ilana Berman-Frank[4], Martin F Polz[2], Daniel Vaulot[3] and Oded Béjà[1]
[1]*Faculty of Biology, Technion-Israel Institute of Technology, Haifa, Israel;* [2]*Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA;* [3]*UPMC (Paris-06) and CNRS, UMR7144, Station Biologique de Roscoff, Roscoff Cedex, France and* [4]*Mina and Everard Faculty of Life Sciences, Bar-Ilan University, Ramat Gan, Israel*

In vast areas of the oceans, most of the primary production is performed by cells smaller than 2–3 μm in diameter (picophytoplankton). In recent years, several *in situ* molecular studies showed a broad genetic diversity of small eukaryotes by sequencing 18S rRNA genes. Compared with photosynthetic cyanobacteria that are dominated by two genera, *Prochlorococcus* and *Synechococcus*, marine photosynthetic picoeukaryotes (PPEs) are much more diverse, with virtually every algal class being represented. However, the genetic diversity and ecology of PPEs are still poorly described. Here, we show using *in situ* molecular analyses of *psbA* transcripts that PPEs in the Eastern Mediterranean Sea are highly diverse, probably very active, and dominated by groups belonging to the red algal lineages, Haptophyta, Heterokontophyta (also called Stramenopiles), and Cryptophyta.
The ISME Journal (2010) **4,** 1044–1052; doi:10.1038/ismej.2010.25; published online 18 March 2010
Subject Category: microbial ecology and functional diversity of natural habitats
Keywords: picoeukaryotes; PPEs; psbA; transcripts; Mediterranean Sea; diversity

Introduction

Photosynthetic organisms have a crucial role in the marine environment and in vast areas of the oceans the majority of the primary production is performed by cells smaller than 2–3 μm (picophytoplankton) (Marañón *et al.*, 2001). Marine photosynthetic picoeukaryotes (PPEs) are found in the ocean photic zone (Sieburth *et al.*, 1978) at concentrations between 10^2 and 10^5 cells ml^{-1} (Li, 2009). PPEs contribute significantly to global biomass and primary productivity, in spite of their relatively low abundance as compared with prokaryotic picophytoplankton (Li, 1994; Worden *et al.*, 2004). Their distribution has been mainly estimated by flow cytometry, based on their size and natural pigment fluorescence (Li *et al.*, 1993), and by high-performance liquid chromatography analysis (Andersen *et al.*, 1996). In the past 10 years, molecular studies (Díez *et al.*, 2001; López-García *et al.*, 2001; Moon-van der Staay *et al.*, 2001) showed wide genetic diversity of picoplanktonic eukaryotes by sequencing 18S ribosomal RNA (rRNA) genes directly from natural samples. However, it was soon realized that this approach favors heterotrophic organisms such as those belonging to stramenopiles or alveolates (Vaulot *et al.*, 2002). Recently, approaches such as plastid 16S rRNA gene phylogenies (Fuller *et al.*, 2006) or flow cytometric sorting (Shi *et al.*, 2009) have allowed more direct targeting of PPE diversity. However, most studies have focused on the gene (DNA) rather than transcript (RNA) level, thus estimating abundance rather than biomass or activity of the cells. In the case of eukaryotes this issue might get very complicated, as in some cases they contain between one and a few thousand copies of the rRNA gene (Prokopowich *et al.*, 2003; Zhu *et al.*, 2005; Massana *et al.*, 2008). The few studies that have looked at rRNA instead of the rRNA genes (Stoeck *et al.*, 2007; Not *et al.*, 2009) have shown that the most abundant groups are in general not the most active ones. For example, in the North West Mediterranean Sea (Not *et al.*, 2009), though picoplanktonic rRNA gene sequences are dominated by alveolates and radiolarians, rRNA transcripts are dominated by marine heterotrophic stramenopiles, which are known to be very active predators (Massana *et al.*, 2009).

Correspondence: O Béjà, Faculty of Biology, Technion-Israel Institute of Technology, Haifa 32000, Israel.
E-mail: beja@tx.technion.ac.il
[5]These two authors contributed equally to this work.
Received 7 December 2009; revised 8 February 2010; accepted 11 February 2010; published online 18 March 2010

Compared with photosynthetic cyanobacteria that are dominated by just two genera (*Prochlorococcus* and *Synechococcus*), PPEs are more diverse with every algal class containing at least some organism of picoplanktonic size, although the number of described species remains very low (Thomsen, 1986; Vaulot *et al*., 2008). The most abundant PPE groups observed in the marine environment (Díez *et al*., 2001; Moon-van der Staay *et al*., 2001; Marie *et al*., 2006; Le Gall *et al*., 2008) are the Chlorophyta (Prasinophyceae), Stramenopiles (or Heterokontophyta, containing in particular diatoms but also numerous other classes such as the Pelagophyceae or the Chrysophyceae), Alveolates (Dinophyceae), Haptophyta (Prymnesiophyceae), and Cryptophyta (Cryptophyceae).

Although previous studies have focused on the survey of rRNA operons, we decided to use a functional, photosynthetic protein-coding gene. Not using rRNA genes enables focusing on photosynthetic microorganisms without the masking of high background data coming from heterotrophic microorganisms found in the same water column. We decided to focus on *psbA* genes and transcripts (coding for the protein D1 of photosystem-II reaction center), which proved previously to be good indicators for a variety of photosynthetic marine microbial groups: dinoflagellates on the DNA level (Zhang *et al*., 2000), marine picoeukarya on the DNA level (Zeidner *et al*., 2003; Zeidner and Béjà, 2004), and cyanobacteria on the DNA and RNA levels (Zeidner *et al*., 2003; Zeidner and Béjà, 2004; Sharon *et al*., 2007).

In this paper, we report a detailed analysis of the diversity of photosynthetically active PPEs in oligotrophic East Mediterranean waters at different seasons based on *psbA* transcripts.

Materials and methods

Cultures

All cultures were obtained from the Roscoff Culture Collection (http://www.sb-roscoff.fr/Phyto/RCC/; Vaulot *et al*., 2004) and are listed in Table 1.

Sample collection

Seawater samples were collected during four cruises (March, May, and October 2006 and January 2007) on board the R/V *Mediterranean Explorer*. Two locations were sampled: near-shore station (Tb200) and open sea station (Tb1000). Station Tb200 is located 20 km from the coast (32°09′N, 34°34′E) at ca. 200 m bottom depth and station Tb1000 is located 51 km offshore (32°10′N, 34°14′E) at ca. 1000 m bottom depth (Figure 1). The highest concentration of photosynthetic biomass (based on Chl *a* concentrations and variable fluorescence) was encountered at both stations during the winter mixing period. At that time of the year, the deep chlorophyll maximum (DCM) layer (~110–150 m) was not developed at Tb200, showing uniform Chl *a* distribution (~0.2–0.3 μg Chl $a\,l^{-1}$), whereas at Tb1000 it was widely spread between 50 and 130 m (~0.4 μg Chl $a\,l^{-1}$). During the stratified period (June–September) Chl *a* concentrations peaked to 0.45 $\mu g\,l^{-1}$ at the DCM. Additional description of these stations can be found in Bar-Zeev *et al*. (2008). Surface water samples were collected at both stations, whereas samples from the DCM were collected only at Tb1000. Twenty liters were pre-filtered through a 3 μm polycarbonate (GE Water & Process Technologies filter, Trevose, PA, USA) (samples from January 07) or a Whatman GF/A glass-fiber filter (1.6 μm nominal pore size) in all other samples, and the

Table 1 List of cultured strains for which the *psbA* gene was sequenced

RCC	*Class*	*Genus*	*Species*	*Strain*	*Ocean origin*
21	Chrysophyceae	*Ochromonas*	*distigma*	Caen	Atlantic Ocean
92	Eustigmatophyceae	*Nannochloropsis*	*salina*	CCMP527	Atlantic Ocean
96	Pelagophyceae	*Aureococcus*	*anophagefferens*	CCMP1784	Atlantic Ocean
97	Pelagophyceae	*Aureoumbra*	*lagunensis*	CCMP1681	Atlantic Ocean
100	Pelagophyceae	*Pelagomonas*	*calceolata*	CCMP1214	Pacific Ocean
113	Prasinophyceae	*Bathycoccus*	*prasinos*	CCMP1898	Mediterranean Sea
114	Prasinophyceae	*Micromonas*	*pusilla*	CCMP490	Atlantic Ocean
135	Prasinophyceae	*Pycnococcus*	*provasolii*	CCMP1199	Atlantic Ocean
205	Bolidophyceae	*Bolidomonas*	*pacifica*	OLI 31 SE3	Pacific Ocean
289	Trebouxiophyceae	*Picochlorum*	sp.	OLI 26 SA	Pacific Ocean
344	Prasinophyceae	*Ostreococcus*	sp.	PROSOPE_3	Atlantic Ocean
382	Dictyochophyceae	*Mesopedinella*	*arctica*	PROSOPE_2	Atlantic Ocean
406	Prymnesiophyceae	*Imantonia*	*rotunda*	RA000609–18–5	English Channel
417	Prasinophyceae	*Mantoniella*	*squamata*	CCMP480	North Sea
434	Prasinophyceae	*Micromonas*	*pusilla*	BL_122	Mediterranean Sea
438	Eustigmatophyceae	*Nannochloropsis*	*granulata*	BL_39	Mediterranean Sea
450	Prasinophyceae	*Micromonas*	*pusilla*	CCMP489	Atlantic Ocean
621	Pinguiophyceae	*Pinguiococcus*	*pyrenoidosus*	CCMP1144	Atlantic Ocean
626	Chlorarachniophyceae	*Gymnochlora*	*stellata*	CCMP2057	Pacific Ocean
789	Prasinophyceae	*Ostreococcus*	sp.	BL_82–7_clonal	Mediterranean Sea

They were all obtained from the Roscoff Culture Collection (RCC).

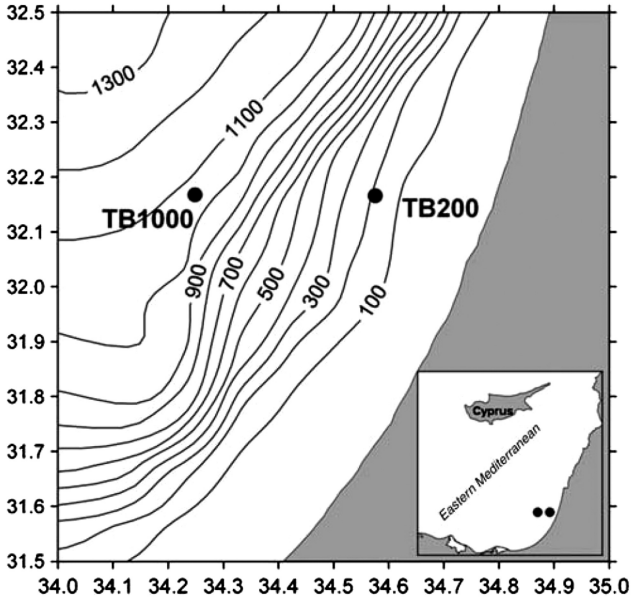

Figure 1 Map of station locations. Samples were taken at a near-shore station (Tb200), located 20 km from the coast (32°09′N, 34°34′E) at ca. 200 m bottom depth and at an open sea station (Tb1000), located 51 km offshore (32°10′N, 34°14′E) at ca. 1000 m bottom depth.

filtrate was collected on a 0.2 μm Sterivex filter (Millipore, Bedford, MA, USA) using a peristaltic pump (Cole Parmer Masterflex 5, channel hardware, Vernon Hills, IL, USA). Although the January samples were collected using a bigger-pore-size pre-filter, we see no significant bias compared with the other samples with regard to photosynthetic picoeukaryal proportions on the RNA level. After collection, the Sterivex filters were filled with 1 ml of lysis buffer (Massana *et al*., 1997) and stored at −80 °C.

Flow cytometry

Samples of 1.8 ml were taken directly from the Niskin bottles and were fixed immediately at room temperature with 23 μl of 25% glutaraldehyde (Sigma G-5882, St Louis, MO, USA) for 20 min after which they were frozen in liquid nitrogen. In the laboratory, the samples were kept at −80 °C until analysis. Samples were thawed at 37 °C and analyzed on a FACScan (Becton Dickinson, Franklin Lakes, NJ, USA) using excitation with an argon laser (488 nm), for either 10–15 min or until 10 000 cells were counted. Discrimination was based on the following parameters: forward and side scatter both related to cell size and pigment orange and red fluorescences for phycoerythrin and chlorophyll (585 and 630 nm, respectively). Beads (0.93 μm, Polysciences, Warrington, PA, USA) served as a standard.

Carbon biomass estimations

Carbon biomass was calculated according to Verity *et al*. (1992). Cell carbon content was calculated as carbon (pg) = 0.433 × (biovolume)$^{0.863}$ using the following biovolume parameters for *Prochlorococcus* (0.6)^3 μm^3, *Synechococcus* (1)^3 μm^3, and picoeukarya (2.5)^3 μm^3. Carbon biomass = carbon(pg) × cell concentration (cell ml^{-1}).

Nucleic acid extraction and reverse transcription

RNA and DNA were extracted from the samples according to Man-Aharonovich *et al*. (2007). Extracted DNA was stored frozen at −20 °C. The RNA was treated with RNase-Free DNase I (Ambion, Austin, TX, USA) for 30 min at 37 °C to remove DNA. DNase was inactivated by heat denaturation at 75 °C for 10 min and samples were stored at −80 °C until further use. For RT–PCR, total RNA (100–300 ng) was reverse transcribed with *psbA* degenerate reverse primer psbA-2R (Wang and Chen, 2008) using Bio-RT (Bio-Lab, Ipswich, MA, USA) according to the manufacturer's instructions. Reaction mixtures were incubated at 37 °C for 1 h.

Amplification of psbA genes

PsbA gene fragments (~750 bp) were amplified by PCR from cDNA and genomic DNA using the degenerated PCR primers designed by Wang and Chen (2008) that target the conserved YPIWEA and HNFPLD regions. PCR was performed in a total volume of 25 μl containing 10 ng of template DNA/cDNA, 2.5 μl of 10 X OptiBuffer, 2 μl of dNTP, 1.3 μl of 50 mM $MgCl_2$, 1 μl of 25 μM psbA-1F (TAYCCNATYTGGGAAGC), 1 μl of 25 μM psbA-2R (TCRAGDGGGAARTTRTG), and 1.2 U of BIO-X-ACT (Bioline, London, UK). Amplification conditions comprised steps at 95 °C for 2 min, and 30 cycles at 94 °C for 1 min, 55 °C for 1 min, and 68 °C for 1 min followed by one step of 7 min at 68 °C. We performed two tests for presence of contaminating DNA in the RNA sample: (1) PCR of the RNA samples without the reverse transcription step and (2) treating the RNA samples with RNase and subjecting them to RT–PCR. To test the reagents for DNA contamination, PCR reactions without template were performed. All tests gave the expected negative results.

Cloning of psbA genes, library construction, and sequencing

PCR products were cloned using the Qiagen PCR cloning kit according to the manufacturer's instruction. Clones were randomly picked up into 20 96-well plates. Each 96-well plate represented different date (March, May, and October 2006 and January 2007), different station and depth (Tb200 surface, Tb1000 surface, and Tb1000 DCM), and different source (DNA and RNA). All plates were sequenced at the MPI for molecular genetics in Berlin.

Rarefaction analysis and community structure analysis
All 1205 *psbA* sequences, which had previously been predicted to be either of bacteriophage, bacterial, or eukaryotic origin (Tzahor *et al*., 2009), were used to infer phylogenies using PhyML (Guindon and Gascuel, 2003). The data were analyzed both as DNA and (translated) amino acid sequences. To estimate total sequence diversity and efficiency of sampling, maximum likelihood (ML) trees were broken down into a distance matrix (using MATLAB), which served as input for the rarefaction analysis and the ACE non-parametric diversity estimator all implemented in the DOTUR (Schloss and Handelsman, 2005) and SONS (Schloss and Handelsman, 2006) packages. These ML distances differ from the more commonly used Hamming distances or percent sequence identity used in such analyses; however, when operational taxonomic units are calculated as sequence clusters at predetermined similarity cutoffs, ML distances provide a higher probability that operational taxonomic unit boundaries coincide with true phylogenetic clusters. To ensure comparison with more traditional distance measures, the range of percent sequence similarity corresponding to a specific ML-based distance was computed from several randomly selected sequences.

GenBank deposition
psbA sequences from RCC cultures were deposited in GenBank under accession #s EU851954-EU851972; environmental *psbA* sequences were deposited in GenBank under accession #s EU940373-EU940692.

Results and discussion

psbA clone libraries
We amplified by PCR *psbA* genes and transcripts directly from DNA and RNA (cDNA) extracted simultaneously from the same samples (derived from a coastal station (Tb200) and a pelagic station (Tb1000) in the Eastern Mediterranean Sea) from mixed picoplankton assemblages using recently designed *psbA* primers (Wang and Chen, 2008). These primers amplify *psbA* from eukaryotes, cyanobacteria, and cyanophages. A total of 1205 randomly picked clones containing *psbA* inserts were sequenced; 618 were derived from DNA and 587 from RNA. We first differentiated *psbA* sequences of eukaryotic and bacterial origin. For rapid classification of the cyanobacterial and cyanophage *psbA* genes, we used the approach developed by Tzahor *et al*. (2009), which uses genomic signature and position-specific codon. This method successfully classified different *psbA* fragments into seven taxonomic groups (*Synechococcus*, HL-*Prochlorococcus*, LL-*Prochlorococcus*, *Synechococcus*-like Myovirus, *Synechococcus*-like Podovirus, *Prochlorococcus*-like Myovirus, and *Prochlorococcus*-like Podovirus), not including eukarya. Unexpectedly, though completely absent in the DNA extracts,

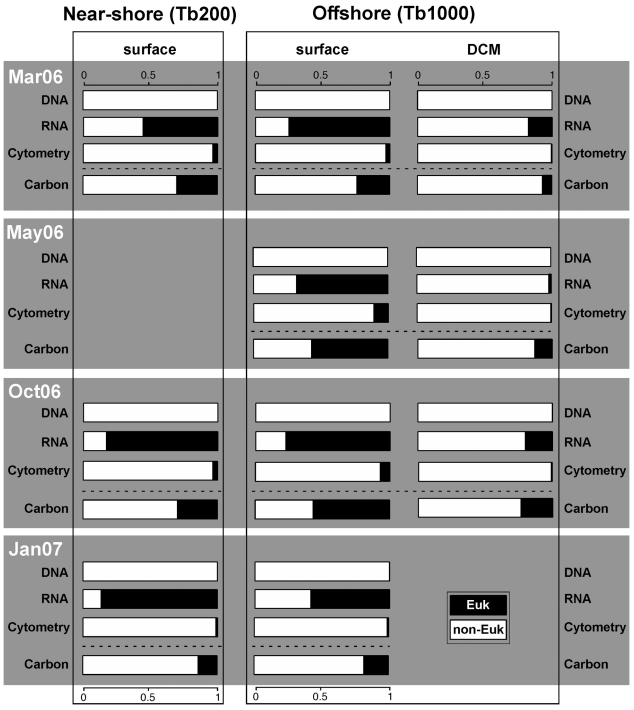

Figure 2 PPE abundance versus carbon biomass estimate. Relative *psbA* clone abundance (DNA and RNA), cell abundance (flow cytometry), and estimated carbon biomass, based on Verity *et al*. (1992), of PPE (black) compared with cyanobacteria and cyanophages (white) retrieved from station Tb1000 and Tb200 in the Eastern Mediterranean Sea.

eukaryotic *psbA* transcripts dominated surface water samples and could reach up to 88% of the total *psbA* RNA (Figure 2). This is not observed with the DCM samples, in which the eukaryotic *psbA* transcripts were found to be between 3% and 24% of the total *psbA* RNA.

Comparison of the observed sequence diversity with that estimated by the ACE non-parametric richness estimator (Supplementary Figure S1) suggests that a large fraction of the taxonomic diversity was sampled. Using an operational taxonomic unit definition of ~95–97% amino acid sequence similarity (or 0.02 ML-based distance), corresponding roughly to the order/class level (Supplementary Table S2), more than half of the diversity appears to be accounted for. Changing this cutoff to 80–85% amino acid identity (or 0.16 ML-based distance), corresponding to a level above the division, sampling was effectively saturated.

Photosynthetic picoplankton abundance and carbon biomass estimates
Flow cytometry analysis was used to provide information on abundance, cell size, and pigment content of the major photosynthetic picoplankton groups (cyanobacteria and PPEs) in our samples. Indeed, picophytoplankton cell counts (Supplementary Table S1) showed that *Prochlorococcus* and *Synechococcus* were numerically dominating at all stations and

depths (Figure 2). In most cases, PPEs represented between 1 and 11% of the total cell counts. When converting abundance to carbon biomass (Verity *et al.*, 1992), the relative part of PPEs increased, especially in surface waters (Figure 2). PPEs represented about 60% of the photosynthetic picoplankton carbon biomass in surface water at the pelagic station (Tb1000) in May and October 2006, and between 15 and 30% at the coastal station (Tb200). At the DCM depth, the PPEs carbon biomass fraction was lower (3–24% of the total). These observations agree with previous observations that although cyanobacteria are more abundant, PPEs sometimes dominate with respect to chlorophyll, carbon, and primary production, because of their higher carbon content (Li, 1994, 1995; Worden *et al.*, 2004).

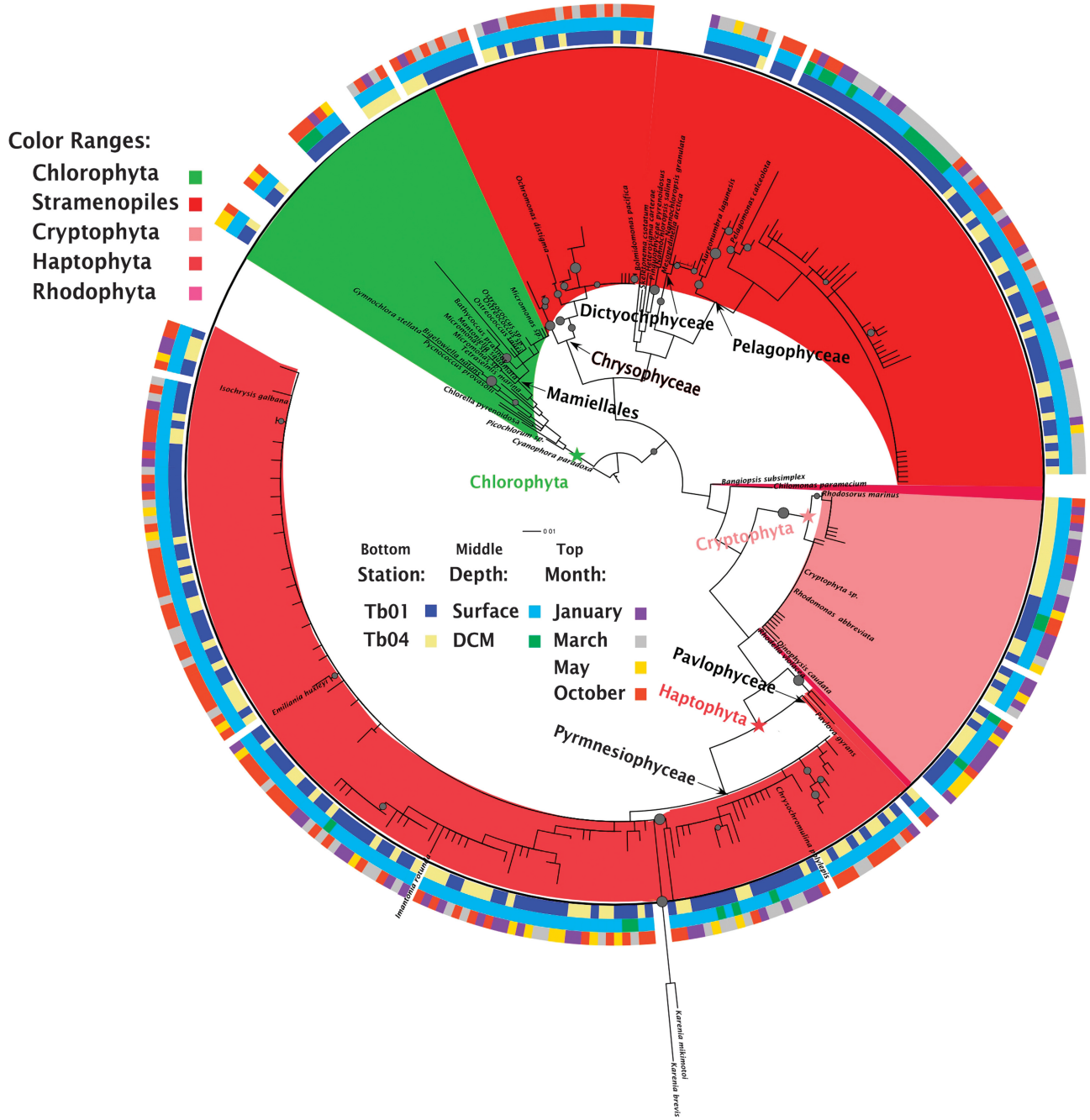

Figure 3 Phylogenetic relationships of Eastern Mediterranean picoeukaryotic *psbA* transcripts. An ML tree of eastern Mediterranean PPE and cultures *psbA* sequences (308 and 38 sequences, respectively). Only sequences from cultures are labeled whereas sequences from this study were left unlabeled for clarity. Coloring of the innermost circle indicates phyla affiliation whereas month, depth, and station attributes of each environmental sequence are marked on the three outer circles. Bootstrap analysis of 100 replicates was conducted, and gray circles indicate values above 50%. Scale bar represents 0.1 substitutions/site for each unit of branch length. Colored stars and black arrows indicate branching of major phylogenetic clades.

Diversity of active PPEs

Most of PPEs *psbA* sequences were clustered with the red algal lineage (94%), the rest being grouped with green algae (Figure 3). To better assign the different *psbA* transcript sequences retrieved, we included in the phylogenetic tree new *psbA* gene sequences corresponding to photosynthetic eukaryotes available in cultures (Table 1), trying to target phylogenetic groups containing picoeukaryotic species.

Haptophyta constitute the most numerous group in the *psbA* sequences (Figure 4), all of which seem to belong to the class Prymnesiophyceae and nearly all originating from surface waters. In addition, Haptophyta are also present in the sample with the lowest frequency of PPE *psbA* transcripts (3%, DCM of Station Tb1000 in May 2006). The high contribution of this group that we observe in surface waters is consistent with the dominance of the diagnostic pigment 19′hexanoylfucoxanthin in open oceanic waters where picoplankton is most important (Liu *et al.*, 2009) as well as in the picoplanktonic fraction itself (Moon-van der Staay *et al.*, 2001; Not *et al.*, 2008). However, despite the dominance of this carotenoid in the small size fractions, very few truly picoplanktonic Prymnesiophyceae species have been described (Vaulot *et al.*, 2008) and this class is in general quite under-represented in 18S clone libraries (Moon-van der Staay *et al.*, 2001; Vaulot *et al.*, 2008). However, several recent studies have shown that this under-representation is probably artefactual. First, the use of probes targeting 16S plastid rRNA has shown that Prymnesiophyceae can dominate the eukaryotic pico or ultra-phytoplankton fraction in environments as diverse as the Indian Ocean, the Mediterranean Sea, or the Pacific Ocean (Fuller *et al.*, 2006; McDonald *et al.*, 2007; Lepère *et al.*, 2009). Second, Liu *et al.* (2009) using primers targeting the LSU rRNA gene showed that a very large number of haptophyte sequences could be recovered from the picoplankton size fraction in a variety of oceanic waters. Third, Shi *et al.* (2009) obtained a large number of Haptophyta 18S rRNA gene sequences from flow cytometry sorted PPE populations from the South East Pacific, including some from a potentially new class intermediate between Prymnesiophyceae and Pavlophyceae. In this study, a large group of sequences were closely related to *Emiliania huxleyi* (Figure 3), an ubiquitous species that not only makes large scale blooms in mid- and high-latitude nutrient-rich waters but is also often isolated from oligotrophic waters (Le Gall *et al.*, 2008). The present data suggest that Haptophyta are also dominant members of the PPE community in the low-nutrient waters of the East Mediterranean Sea (Figure 4). In addition, two sequences were somewhat related to Haptophyta but fell outside its radiation and could belong a novel phylogenetic group (Figure 3).

The second most abundant group of sequences was related to the Pelagophyceae. This is consistent with the carotenoid 19′butanoyloxyfucoxanthin, which is characteristics of this class, being important (after 19′hexanoylfucoxanthin) in oligotrophic waters (Not *et al.*, 2008). Surprisingly, these sequences diverged quite significantly (only 95–98% identity) from *Pelagomonas calceolata*, a picoplankton species

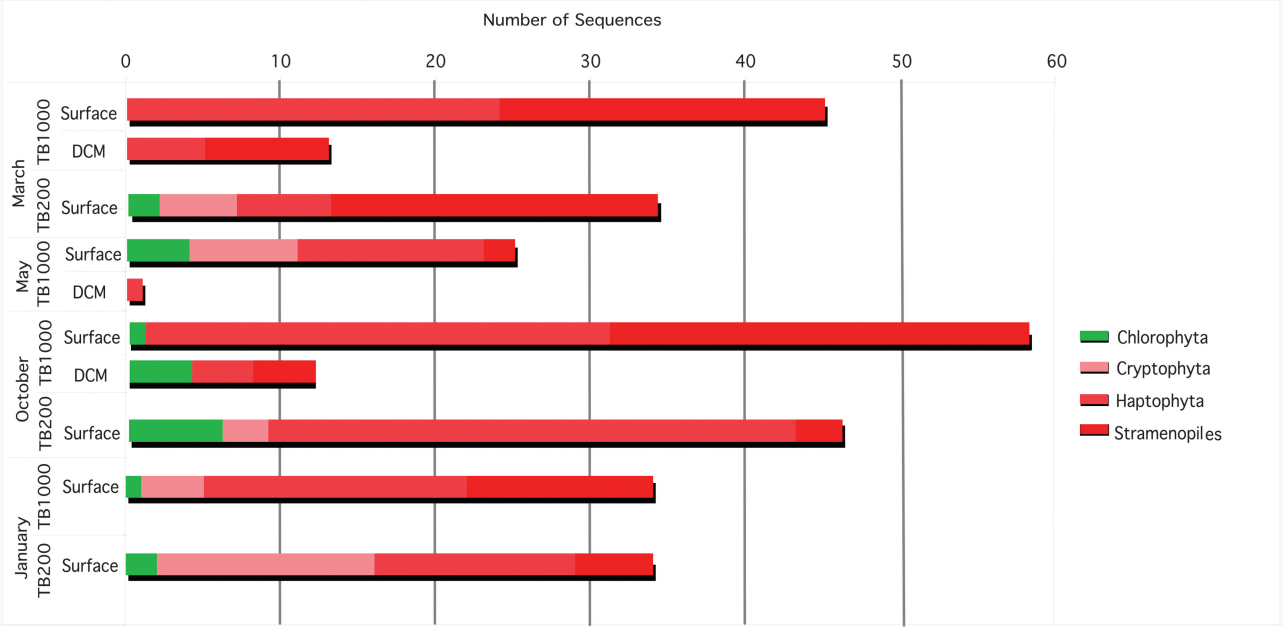

Figure 4 Taxonomic composition of each sample based on RNA-derived *psbA* sequences. Samples were taken on four different months, at two stations, TB200 (coastal) and TB1000 (Pelagic). At each station, samples were taken from the surface and, at TB1000, also from the DCM. Colors correspond to the coloring of phylogenetic affiliation in Figure 2.

(Andersen *et al*., 1993) that is very often isolated from oceanic waters (Le Gall *et al*., 2008) and whose sequences found in 18S rRNA gene clone libraries are generally highly conserved (Moon-van der Staay *et al*., 2001; Shi *et al*., 2009). It is therefore possible that these *psbA* sequences belong to a new class, not yet brought in culture.

Some *pbsA* sequences also fell into the closely related class of Dictyochophyceae of which a picoplanktonic species, *Florenciella parvula*, has been described recently (Eikrem *et al*., 2004).

Another large group of sequences belonged to Chrysophyceae (Figure 3). This is quite interesting because though Chrysophyceae are abundant in fresh water, most marine Chrysophyceae described to date are heterotrophic such as those belonging to the genus *Paraphysomonas* and they make only small contributions to picoplankton 18S rRNA gene clone libraries (Vaulot *et al*., 2008). However, many plastid 16S rRNA gene sequences related to Chrysophyceae have been recovered from the open ocean (Fuller *et al*., 2006; McDonald *et al*., 2007) and 16S rRNA probes hybridized on PCR products suggest that Chrysophyceae are important contributors in oligotrophic waters (Fuller *et al*., 2006; Lepère *et al*., 2009). More recently, Chrysophyceae 18S sequences have also been recovered from flow cytometry sorted PPE populations in the SE Pacific Ocean (Shi *et al*., 2009). However, the nature of these cells still escapes us as no marine photosynthetic Chrysophyceae cultures of picoplanktonic size have been isolated and characterized until now.

Quite a few Cryptophyceae sequences were also recovered. Sequences of the 18S rRNA gene from this group are abundant in fresh waters (Lepère *et al*., 2006) but in marine systems usually only found in coastal waters and absent in open ocean samples (Vaulot *et al*., 2008). Here, they were found in all near shore samples and offshore only in surface and not at the DCM (Figure 4). One possibility is that coastal populations may have been transported offshore, the other being that the ecology of Cryptophyceae is different in Mediterranean Sea waters and that this group is present in blue waters.

Finally, a few sequences were affiliated to Mamiellales, an order from the Prasinophyceae that contains three genera *Micromonas*, *Bathycoccus*, and *Ostreococcus* that can dominate PPEs in coastal waters, for example in the English Channel or in the Chile upwelling (Not *et al*., 2004; Shi *et al*., 2009), but are also found sporadically in open ocean waters, especially for the genus *Ostreococcus* (Marie *et al*., 2006). In this study, *Micromonas psbA* sequences were only found near shore whereas those related to *Ostreococcus* and *Bathycoccus* were found offshore.

Conclusions

The analysis of different genes (nuclear 18S rRNA, plastid 16S rRNA, and now *psbA*) begins to offer a better view of the most abundant and most active groups within the PPE community. The present data based on *psbA* transcripts bring two major conclusions:

(1) Eukaryotes account for a much higher fraction of *psbA* transcripts than of *psbA* genes, and therefore probably contribute significantly to primary production.
(2) Prymnesiophyceae, Pelagophyceae, Chrysophyceae, and Cryptophyceae appear as the most active PPEs in Eastern Mediterranean waters. This confirms previous estimates in the other regions of the Mediterranean Sea based on pigment signatures and on environmental plastid 16S rRNA gene sequences (McDonald *et al*., 2007). Nonetheless, many of the corresponding species remain to be isolated and described.

Acknowledgements

We thank the sequencing team in the MPI for Molecular Genetic in Berlin for technical support and the captain and crew of the R/V *Med Explorer* for their expert assistance at sea. We are grateful to Feng Chen and Kui Wang for allowing us to use their *psbA* primers prior to publication and to Ramon Massana for help during the initial stages of the project. This research was supported in part by an Israel Science Foundation Grant (no. 458/04, IB-F & OB), an EMBO YIP award (OB), a Marine Genomics Network Of Excellence EU grant (OB), and a grant by the Gordon and Betty Moore Foundation and US Department of Energy Genomes to Life (MFP). Strains were provided by the Roscoff Culture Collection supported by the Contrat de Plan Etat-Région Souchothèque de Bretagne and by the ASSEMBLE EU FP7 research infrastructure initiative. This research is part of the requirements for a PhD thesis for TY at Bar-Ilan University.

References

Andersen RA, Bidigare RR, Keller MD, 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 Part 2 Top Stud Oceanogr* **43**: 517–537.

Andersen RA, Saunders GW, Paskind MP, Sexton J. (1993). Ultrastructure and 18S rRNA gene sequence for *Pelagomonas calceolata* gen. and sp. nov. and the description of a new algal class, the Pelagophyceae classis nov. *J Phycol* **29**: 701–715.

Bar-Zeev E, Yogev T, Man-Aharonovich D, Kress N, Herut B, Béjà O *et al.* (2008). Seasonal dynamics of the endosymbiotic, nitrogen-fixing cyanobacterium *Richelia intracellularis* in the eastern Mediterranean Sea. *ISME J* **2**: 911–923.

Díez B, Pedrós-Alió C, Marsh TL, Massana R. (2001). Application of denaturing gradient gel electrophoresis (DGGE) to study the diversity of marine picoeukaryotic assemblages and comparison of DGGE with other molecular techniques. *Appl Environ Microbiol* **67**: 2942–2951.

Eikrem W, Romari K, Latasa M, Le Gall F, Throndsen J, Vaulot D. (2004). *Florenciella parvula* gen. and sp. nov. (Dictyochophyceae, Heterokontophyta) a small

flagellate isolated from the english channel. *Phycologia* **43**: 658–668.

Fuller NJ, Campbell C, Allen DJ, Pitt FD, Zwirglmaier K, Le Gall F *et al.* (2006). Analysis of photosynthetic picoeukaryote diversity at open ocean sites in the Arabian Sea using a PCR biased towards marine algal plastids. *Aquat Microb Ecol* **43**: 79–93.

Guindon S, Gascuel O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**: 696–704.

Le Gall F, Rigaut-Jalabert F, Marie D, Garczareck L, Viprey M, Godet A *et al.* (2008). Picoplankton diversity in the South-East Pacific Ocean from cultures. *Biogeosciences* **5**: 203–214.

Lepère C, Boucher D, Jardillier L, Domaizon I, Debroas D. (2006). Structure and regulation factors of eukaryotic picoplankton in lacustrine ecosystems. *Appl Environ Microbiol* **72**: 2971–2981.

Lepère C, Vaulot D, Scanlan DJ. (2009). Photosynthetic picoeukaryote community structure in the South East Pacific Ocean encompassing the most oligotrophic waters on earth. *Environ Microbiol* **11**: 3105–3117.

Li WKW. (1994). Primary production of prochlorophytes, cyanobacteria, and eucaryotic ultraphytoplankton: measurements from flow cytometric sorting. *Limnol Oceanogr* **39**: 169–175.

Li WKW. (1995). Composition of ultraphytoplankton in the central North Atlantic. *Mar Ecol Prog Ser* **122**: 1–8.

Li WKW. (2009). From cytometry to macroecology: a quarter century quest in microbial oceanography. *Aquat Microb Ecol* **57**: 239–251.

Li WKW, Zohary T, Yacobi YZ, Wood AM. (1993). Ultraphytoplankton in the eastern Mediterranean Sea—towards deriving phytoplankton biomass from flow cytometric measurements of abundance, fluorescence and light scatter. *Mar Ecol Prog Ser* **102**: 79–87.

Liu H, Probert I, Uitz J, Claustre H, Aris-Brossou S, Frada M *et al.* (2009). Haptophyta rule the waves: extreme oceanic biodiversity in non-calcifying haptophytes explains the 19-Hex paradox. *Proc Natl Acad Sci USA* **106**: 12803–12808.

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

Man-Aharonovich D, Kress N, Bar Zeev E, Berman-Frank I, Béjà O. (2007). Molecular ecology of *nifH* genes and transcripts in Eastern Mediterranean Sea. *Environ Microbiol* **9**: 2354–2363.

Marañón E, Holligan PM, Barciela R, Gonzalez N, Mourino B, Pazo MJ *et al.* (2001). Patterns of phytoplankton size structure and productivity in contrasting open-ocean environments. *Mar Ecol Prog Ser* **216**: 43–56.

Marie D, Zhu F, Balagué V, Ras J, Vaulot D. (2006). Eukaryotic picoplankton communities of the Mediterranean Sea in summer assessed by molecular approaches (DGGE, TTGE, QPCR). *FEMS Microbiol Ecol* **55**: 403–415.

Massana R, Karniol B, Pommier T, Bodaker I, Béjà O. (2008). Metagenomic retrieval of a ribosomal DNA repeat array from an uncultured marine alveolate. *Environ Microbiol* **10**: 1335–1343.

Massana R, Murray AE, Preston CM, DeLong ED. (1997). Vertical distribution and phylogenetic characterization of marine planktonic *Archaea* in the Santa Barbara channel. *Appl Environ Microbiol* **63**: 50–56.

Massana R, Unrein F, Rodriguez-Martinez R, Forn I, Lefort T, Pinhassi J *et al.* (2009). Grazing rates and functional diversity of uncultured heterotrophic flagellates. *ISME J* **3**: 588–596.

McDonald SM, Sarno D, Scanlan DJ, Zingone A. (2007). Genetic diversity of eukaryotic ultraphytoplankton in the Gulf of Naples during an annual cycle. *Aquat Microb Ecol* **50**: 75–89.

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

Not F, del Campo J, Balagué V, de Vargas C, Massana R. (2009). New insights into the diversity of marine picoeukaryotes. *PLoS ONE* **4**: e7143.

Not F, Latasa M, Marie D, Cariou T, Vaulot D, Simon N. (2004). A single species *Micromonas pusilla* (Prasinophyceae) dominates the eukaryotic picoplankton in the western english channel. *Appl Environ Microbiol* **70**: 4064–4072.

Not F, Latasa M, Scharek R, Viprey M, Karleskind P, Balagué V *et al.* (2008). Protistan assemblages across the Indian Ocean, with a specific emphasis on the picoeukaryotes. *Deep Sea Res I* **55**: 1456–1473.

Prokopowich CD, Gregory TR, Crease TJ. (2003). The correlation between rDNA copy number and genome size in eukaryotes. *Genome* **46**: 48–50.

Schloss PD, Handelsman J. (2005). Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* **71**: 1501–1506.

Schloss PD, Handelsman J. (2006). Introducing SONS, a tool for operational taxonomic unit-based comparisons of microbial community memberships and structures. *Appl Environ Microbiol* **72**: 6773–6779.

Sharon I, Tzahor S, Williamson S, Shmoish M, Man-Aharonovich D, Rusch DB *et al.* (2007). Viral photosynthetic reaction center genes and transcripts in the marine environment. *ISME J* **1**: 492–501.

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

Sieburth JM, Smetacek V, Lenz J. (1978). Pelagic ecosystem structure: heterotrophic components of the plankton and their relationship to plankton size fractions. *Limnol Oceanogr* **23**: 1256–1263.

Stoeck T, Zuendorf A, Breiner HW, Behnke A. (2007). A molecular approach to identify active microbes in environmental eukaryote clone libraries. *Microb Ecol* **53**: 328–339.

Thomsen HA. (1986). A survey of the smallest eukaryotic organisms of the marine phytoplankton. *Can Bull Fish Aquat Sci* **214**: 121–158.

Tzahor S, Man-Aharonovich D, Kirkup BC, Yogev T, Berman-Frank I, Polz M *et al.* (2009). A supervised learning approach for taxonomic classification of core-photosystem-II genes and transcripts in the marine environment. *BMC Genomics* **10**: 229.

Vaulot D, Eikrem W, Viprey M, Moreau H. (2008). The diversity of small eukaryotic phytoplankton ($\leqslant 3\,\mu m$) in marine ecosystems. *FEMS Microbiol Rev* **32**: 795–820.

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**: 49–70.

Vaulot D, Romari K, Not F. (2002). Are autotrophs less diverse than heterotrophs in marine picoplankton? *Trends Microbiol* **10**: 266–267.

Verity PG, Robertson CY, Tronzo CR, Andrews MG, Nelson JR, Sieracki ME. (1992). Relationships between cell volume and the carbon and nitrogen content of marine photosynthetic nanoplankton. *Limnol Oceanogr* **37**: 1434–1446.

Wang K, Chen F. (2008). Prevalence of highly host-specific cyanophages in the estuarine environment. *Environ Microbiol* **10**: 300–312.

Worden AZ, Nolan JK, Palenik B. (2004). Assessing the dynamics and ecology of marine picophytoplankton: The importance of the eukaryotic component. *Limnol Oceanogr* **49**: 168–179.

Zeidner G, Béjà O. (2004). The use of DGGE analyses to explore eastern Mediterranean and Red Sea marine picophytoplankton assemblages. *Environ Microbiol* **6**: 528–534.

Zeidner G, Preston CM, Delong EF, Massana R, Post AF, Scanlan DJ *et al.* (2003). Molecular diversity among marine picophytoplankton as revealed by *psbA* analyses. *Environ Microbiol* **5**: 212–216.

Zhang Z, Green BR, Cavalier-Smith T. (2000). Phylogeny of ultra-rapidly evolving dinoflagellate chloroplast genes: a possible common origin for sporozoan and dinoflagellate plastids. *J Mol Evol* **51**: 26–40.

Zhu F, Massana R, Not F, Marie D, Vaulot D. (2005). Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol Ecol* **52**: 79–92.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)