environmental microbiology

Environmental Microbiology (2009) **11**(12), 3105–3117 doi:10.1111/j.1462-2920.2009.02015.x

Photosynthetic picoeukaryote community structure in the South East Pacific Ocean encompassing the most oligotrophic waters on Earth

Cécile Lepère,[1,2] Daniel Vaulot[2] and David J. Scanlan[1]*

[1]*Department of Biological Sciences, University of Warwick, Coventry, UK.*

[2]*UPMC (Paris 06) et CNRS, UMR7144, Station Biologique de Roscoff, Roscoff, France.*

Summary

Photosynthetic picoeukaryotes (PPEs), comprising organisms < 3 μm in size, are important primary producers in marine food webs and include representatives from all known algal lineages. Little is known, however, regarding the composition and distribution of PPE communities, particularly at large spatial scales, or in relation to the underlying biotic and abiotic factors that influence this structure. Here, we analysed PPE community structure along a transect in the South East Pacific Ocean (BIOSOPE cruise) that encompassed a large trophic gradient, including hyper-oligotrophic waters in the South Pacific Gyre (SPG), considered to be some of the 'clearest' natural waters on Earth. Using dot blot hybridizations with 16S rRNA oligonucleotide probes, we established that the PPE community was dominated by members of the classes Prymnesiophyceae and Chrysophyceae throughout the transect. Moreover, clone library construction followed by phylogenetic analysis of sequenced clones revealed several novel 16S rRNA gene lineages, including new clades of prymnesiophytes (designated Prym 16S-III) and prasinophytes (Pras 16S-VIII). Pras 16S-VIII was found at all five stations at which clone libraries were constructed, representing a range of trophic conditions, including the South Pacific Gyre, suggesting members of this clade have a broad distribution in this part of the South East Pacific at least. In contrast, Prym 16S-III sequences were largely restricted to oligotrophic stations of the SPG. Subsequent multivariate statistical analyses showed that, within the measured factors, chemical and biological factors seem to influence PPE community structure more than physical parameters. However, more than 50% of the variation in distribution of PPE classes remained unexplained.

Received 23 February, 2009; accepted 20 June, 2009. *For correspondence. E-mail d.j.scanlan@warwick.ac.uk; Tel. (+44) 24 76 522572; Fax (+44) 24 76 523701.

Introduction

The South East Pacific Ocean, which connects the tropical and high latitudes of the austral ocean, remains the most sparsely sampled region of the global ocean. The Biogeochemistry and Optics South Pacific Experiment (BIOSOPE) cruise (2004) consisted of an 8000 km long transect from subequatorial mesotrophic waters near the Marquesas Islands through the South Pacific Gyre (SPG) and extending to more eutrophic waters off the Chilean coastal upwelling near Concepción (Claustre *et al*., 2008). An intensive biogeochemical study of this poorly documented hyper-oligotrophic system was a predominant aim of the BIOSOPE cruise. The SPG is the most oligotrophic of all the subtropical gyres. The surface chlorophyll *a* (Chl-*a*) concentration is the lowest of the global open ocean with annual means as low as 0.019 mg Chl-*a* m^{-3} (Morel *et al*., 2007), corresponding to very clear waters with deep penetration not only of visible but also of UV radiation (Vasilkov *et al*., 2001). These hyper-oligotrophic characteristics are closely related with exceptional physical features (Claustre *et al*., 2008).

The very large trophic gradient that was extensively sampled during the BIOSOPE cruise offered a unique opportunity to better understand how the structure of biological communities and particularly photosynthetic picoeukaryotes (PPEs, cells < 3 μm) in the open ocean adapt to varying nutrient conditions, with a specific focus on the extremely oligotrophic conditions of the central SPG. Although marine PPEs are less numerous than their prokaryotic counterparts, because of their larger volume they can contribute greatly to global carbon cycling, biomass, and productivity in the sea (Li, 1994). Despite their ecological importance, picoeukaryotes have remained poorly described due to their small size and lack of morphological characteristics. However, in the last few years, thanks to the development of molecular techniques, diversity and spatial distribution analyses of marine picoeukaryotes have started to be performed (e.g.

see Lòpez-Garcia *et al*., 2001; Massana *et al*., 2004; Not *et al*., 2008). These studies, based on 18S rRNA gene sequence analyses, have shown a large diversity, including putative photosynthetic representatives from a wide range of classes e.g. Chrysophyceae, Cryptophyceae, Prasinophyceae and Prymnesiophyceae (see review in Vaulot *et al*., 2008). In coastal marine waters, work has shown prasinophytes and prymnesiophytes to be the dominant PPEs (Not *et al*., 2004). Recent work has complemented these 18S rDNA studies by focusing on the plastid 16S rRNA gene, a marker that specifically targets photosynthetic organisms, with the design of a marine algal-plastid biased PCR primer (Fuller *et al*., 2006a) as well as PPE class-specific oligonucleotide probes (Fuller *et al*., 2006b). Such tools have shown a high relative abundance of Chrysophyceae and Prymnesiophyceae in waters of the Arabian Sea, and in the Gulf of Naples (Fuller *et al*., 2006b; McDonald *et al*., 2007). However, there is still a relative dearth of information on the dominant PPE classes in open ocean regions, especially in oligotrophic waters which are estimated to contribute significantly to global CO_2 fixation (Field *et al*., 1998). Understanding the genetic structure of eukaryotic picophytoplankton communities *in situ* is important for accurately modeling growth and CO_2 fixation by these organisms but also for elucidating the biological, chemical and physical factors controlling the distribution of the different classes in the world ocean.

In this study, the *in situ* community structure of PPEs along the BIOSOPE transect was determined using both cloning-sequencing and dot blot hybridization approaches, revealing the complexity of this important community. Statistical approaches were subsequently used to assess the link between class distribution patterns to environmental variables, and various factors influencing the distribution of these classes are discussed.

Results

BIOSOPE transect: physical, chemical and flow cytometry profiles

The BIOSOPE transect was characterized by a clear salinity maximum extending from the surface down to 150 m between stations MAR1 and GYR (Fig. 1) with salinities of 35.5–36.3 PSU (Fig. 2). Temperature ranged from 16°C to 20°C at the centre of the gyre; colder and fresher waters were found at the Chilean coast. In the oligotrophic central SPG very little inorganic nitrogen was *a priori* available for biological production throughout the 0–150 m water column analysed (nitrate concentrations were typically < 3 nM, Raimbault *et al*., 2008; Fig. 2). Such waters, which are the clearest reported in the world ocean (Morel *et al*., 2007), were characterized by extremely low surface Chl-*a* concentrations (< 0.02 mg m^{-3}). In contrast, both phosphates and silicates (not shown) were always detectable (0.1 and 1 μM respec-

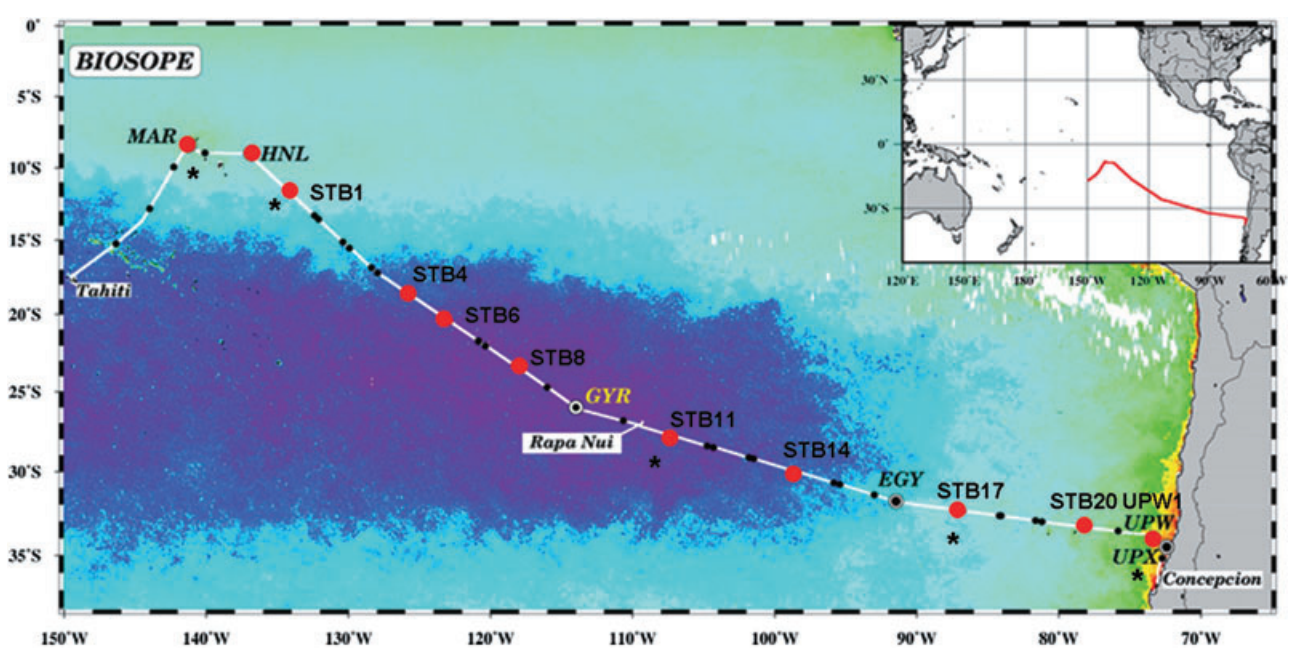

Fig. 1. Map of the BIOSOPE cruise track superimposed on a SeaWiFS ocean colour composite, the dark purple indicating extremely low values of total chlorophyll *a*. Red ellipses indicate the stations analysed by dot blot hybridization in this study. Asterisks indicate the stations at which the 16S rRNA gene clone libraries were constructed.

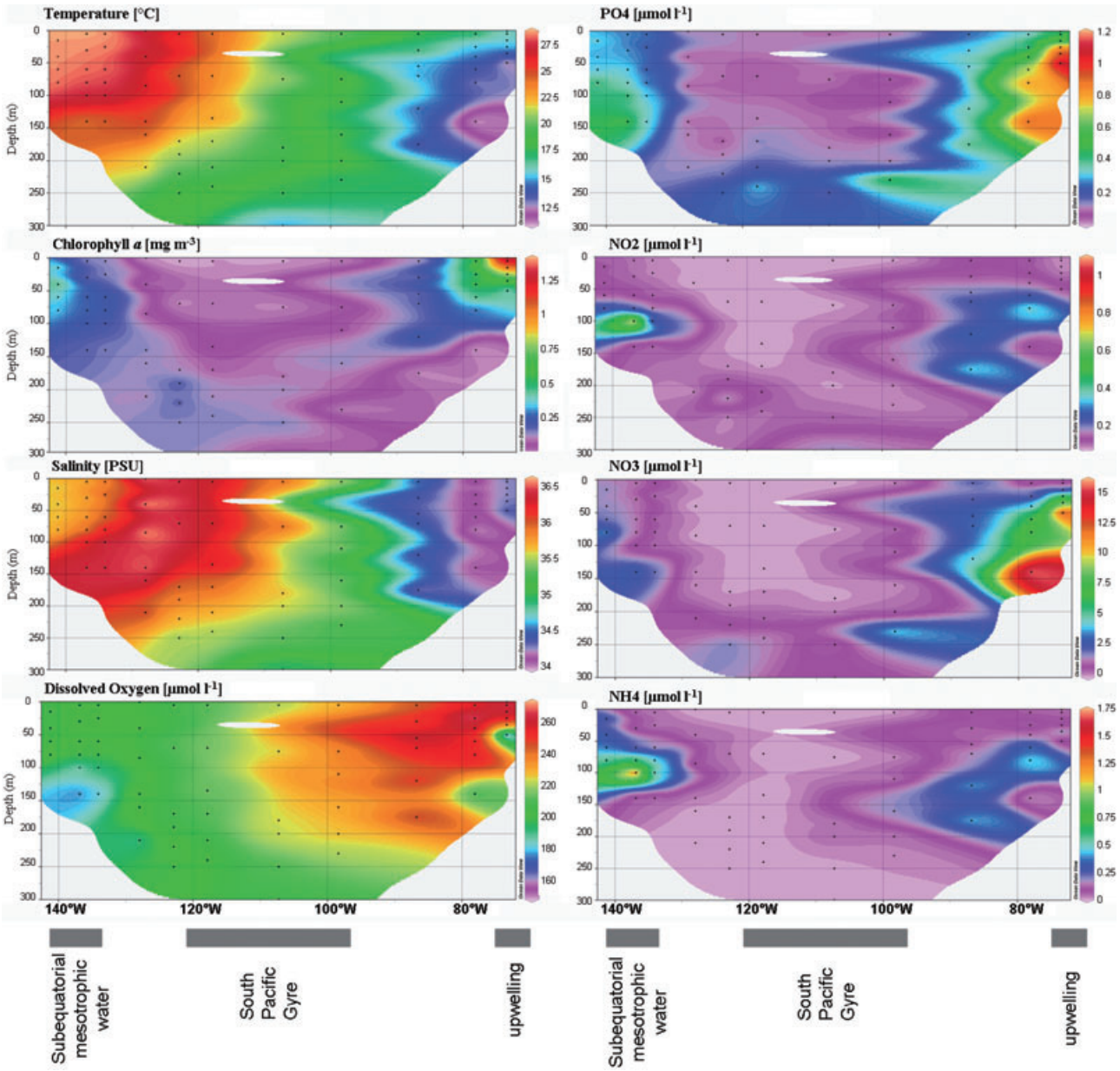

Fig. 2. Chemical and physical parameters along the BIOSOPE cruise track. Contour plots indicate nutrient or pigment concentrations (nitrite, nitrate, phosphate, ammonium, chlorophyll *a* or dissolved oxygen) or physical measurements (temperature, salinity) plotted as a function of depth along the cruise track (longitude degrees). Black dots represent sampling points for dot blot hybridization analysis.

tively, see Claustre *et al*., 2008), with concentrations highest near the Chilean coast. More detailed chemical and physical characteristics of the transect can also be found in Claustre and colleagues (2008).

Flow cytometry analysis (Grob *et al*., 2007) revealed peak abundances of PPEs in the Chilean upwelling with a maximum of 37×10^3 cells ml^{-1}. In contrast, the highly oligotrophic gyre supported average PPE cell numbers of 1.3×10^3 cells ml^{-1} (range 446–1508 cells ml^{-1} at the surface).

PPE community structure

Dot blot hybridization analyses using PPE class-specific oligonucleotide probes targeting plastid 16S rRNA revealed the presence of members of the Chrysophyceae, Prymnesiophyceae and Cryptophyceae at all stations along the BIOSOPE transect. In contrast, hybridization signals for the Chlorarachniophyceae, Prasinophyceae clade VI, Pelagophyceae, Eustigmatophyceae, Pinguiophyceae and Pavlovophyceae groups were generally low

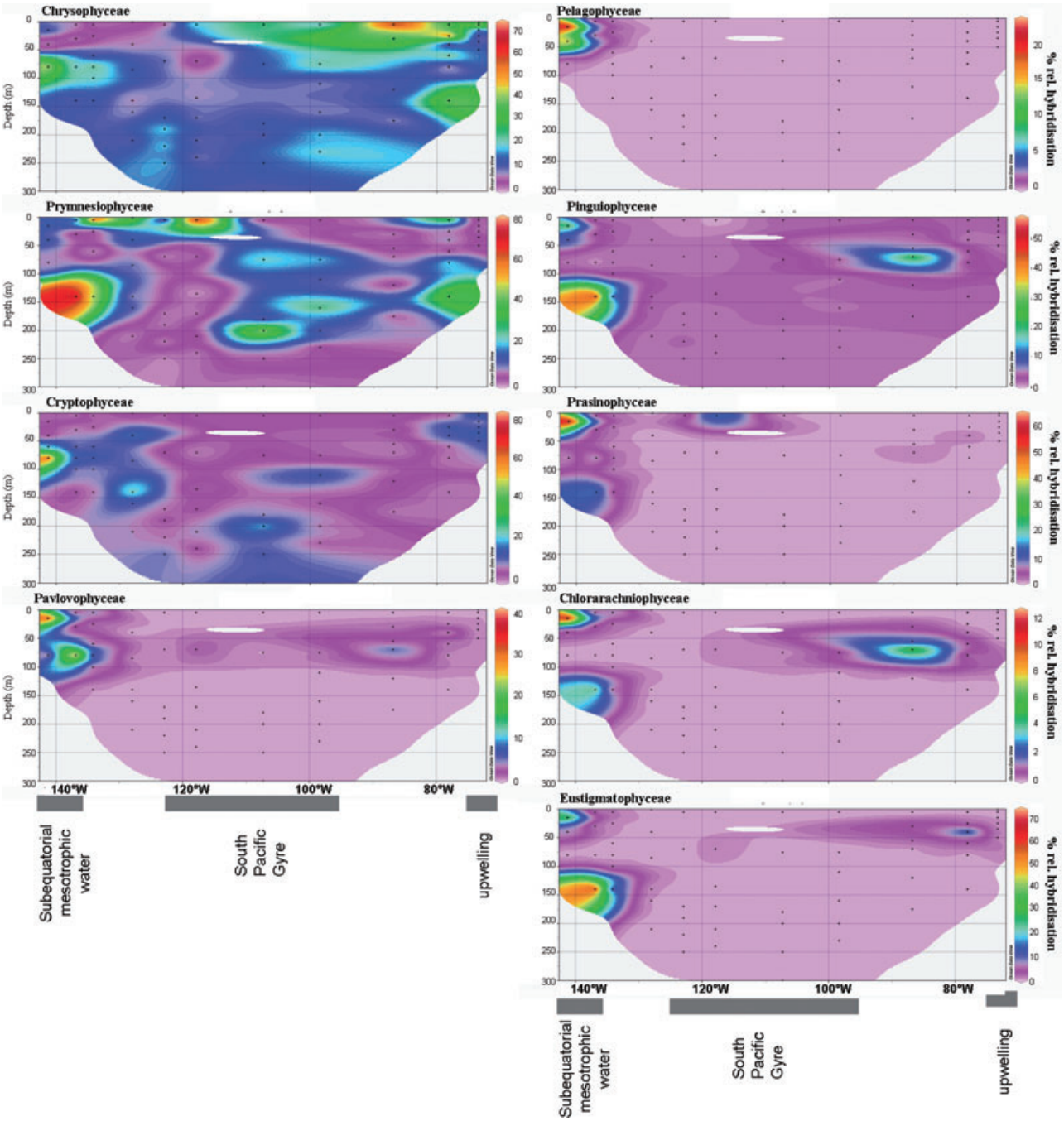

Fig. 3. Dot blot hybridization data describing PPE distribution patterns along the BIOSOPE cruise track. Contour plots indicate the percentage relative hybridization. Black dots represent sampling points.

(average 3%), although they became more significant at the mesotrophic stations MAR1 and HLN, where the PPE community appeared to be most diverse (Fig. 3, Table S1) (e.g. pelagophytes had a maximum signal of 24% at station MAR1 but were largely absent from the rest of the transect). Relative hybridization values for the Chrysophyceae showed this group to be the most important component of the PPE community along the transect, contributing on average 16% of the total signal at each station. Signals for this group were highest in surface waters, i.e. in the top 25 m of the water column, e.g. in the SPG, where the signal reached 25% (Fig. 3). They were also detected deep in the photic zone especially in more nutrient-rich areas of the transect (MAR1, STB20 and UPW1). Prymnesiophytes contributed on average 13% of the signal at each station, with a maximum relative hybrid-

ization signal of 80% at station HLN (150 m, Fig. 3, Table S1). In contrast to chrysophytes, however, prymnesiophytes were more abundant in deeper waters, especially in the SPG. Finally, cryptophytes represented on average 9% of the total signal at each station, the highest relative hybridization values being obtained in the more nutrient-rich waters of the transect.

Phylogenetic diversity of PPEs

Coverage values for the five clone libraries constructed from surface samples along the transect ranged from 86% to 96%. The S_{chao1} index ranged from 10.7 (STB11) to 92.5 (MAR1 station). Analysis of RFLP profiles from a total of 524 clones obtained from the five clone libraries revealed 89 operational taxonomic units (OTUs) (Table S2). To obtain an estimate of the total diversity of the PPE assemblage relative to sampling effort (i.e. number of clones sequenced), we also computed rarefaction curves from RFLP profiles (Fig. 4). These showed the PPE community in mesotrophic waters to be the most diverse while in the hyper-oligotrophic SPG and the Chile upwelling region rarefaction curves tending towards saturation implied a less complex community (Fig. 4).

Subsequent sequencing of OTUs revealed that the class Prymnesiophyceae showed the greatest diversity, with 41 distinct OTUs. Phylogenetic analysis of these sequences identified three distinct clades, one of which comprised sequences with no close cultured counterpart(s), suggesting it corresponds to a novel clade of the plastid 16S rRNA gene, designated Prym 16S-III in Fig. 5. The Prym16S-I clade does not include sequences from the central SPG whereas sequences forming the Prym16S-II clade were found at all five stations at which clone libraries were constructed, representing a range of trophic conditions. In contrast, Prym 16S-III comprised sequences mostly obtained from the libraries constructed from stations STB1, STB11 and STB17, corresponding to the more oligotrophic stations including the SPG.

Phylogenetic analysis of Prasinophyceae sequences (comprising 30 OTUs) revealed three clades, two of which have been previously described, Prasinophyceae clade VII and Prasinophyceae clade VI (Guillou *et al*., 2004). The third clade was not closely related to any other known taxa or environmental sequences (Fig. 6). Whereas sequences from Prasinophyceae clade VII and Prasinophyceae clade VI were confined to libraries constructed from stations with relatively high nutrient levels, 16S rRNA gene sequences forming this new group (designated Pras 16S-VIII in Fig. 6) were retrieved from all libraries representing a range of trophic conditions.

Sequences affiliated to chrysophytes obtained from the BIOSOPE cruise formed two phylogenetic clades (Chrys 16S-I and Chrys 16S-II in Fig. 7), and were well represented in library STB11 located in hyper-oligotrophic waters. However, they were also detected in more mesotrophic waters (stations STB1 and UPW1). Chrysophyceae OTUs from Chrys 16S-II, derived from the SPG, appeared to be closely related to previously described environmental sequences from the Arabian Sea (Fuller *et al*., 2006a), whereas sequences from Chrys 16S-II are related to Mediterranean Sea sequences (McDonald *et al*., 2007).

For other PPE classes further indication of trophic partitioning emerged. For example, trebouxiophyte sequences were only found at station MAR1 (Fig. 6), confirming the dot blot hybridization data (Fig. 4), while dictyochophyte sequences were only retrieved from the hyper-oligotrophic gyre waters at station STB11 (Fig. 7). For the cryptophytes, a sequence with relatively close affiliation to the cultured species *Guillardia theta* (formerly known as *Cryptomonas phi*) was only retrieved from the most eutrophic region of the transect (Fig. 6). Similar to the relative hybridization data, pelagophytes were confined to the STB1 and MAR1 libraries, where four sequences (Fig. 7) related to previously reported environmental sequences retrieved from the Arabian Sea were discernible (Fuller *et al*., 2006a). Finally, five OTUs could

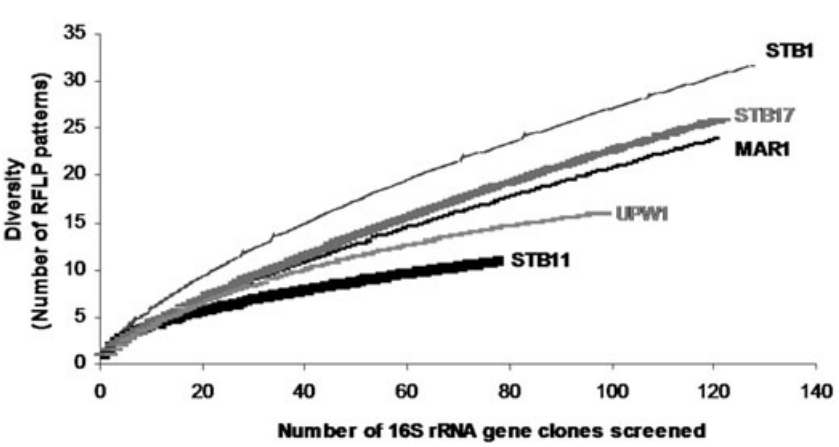

Fig. 4. Rarefaction curves determined for the five 16S rRNA gene libraries generated for the BIOSOPE transect. The number of different RFLP patterns was determined after digestion with restriction endonuclease *Hae*III.

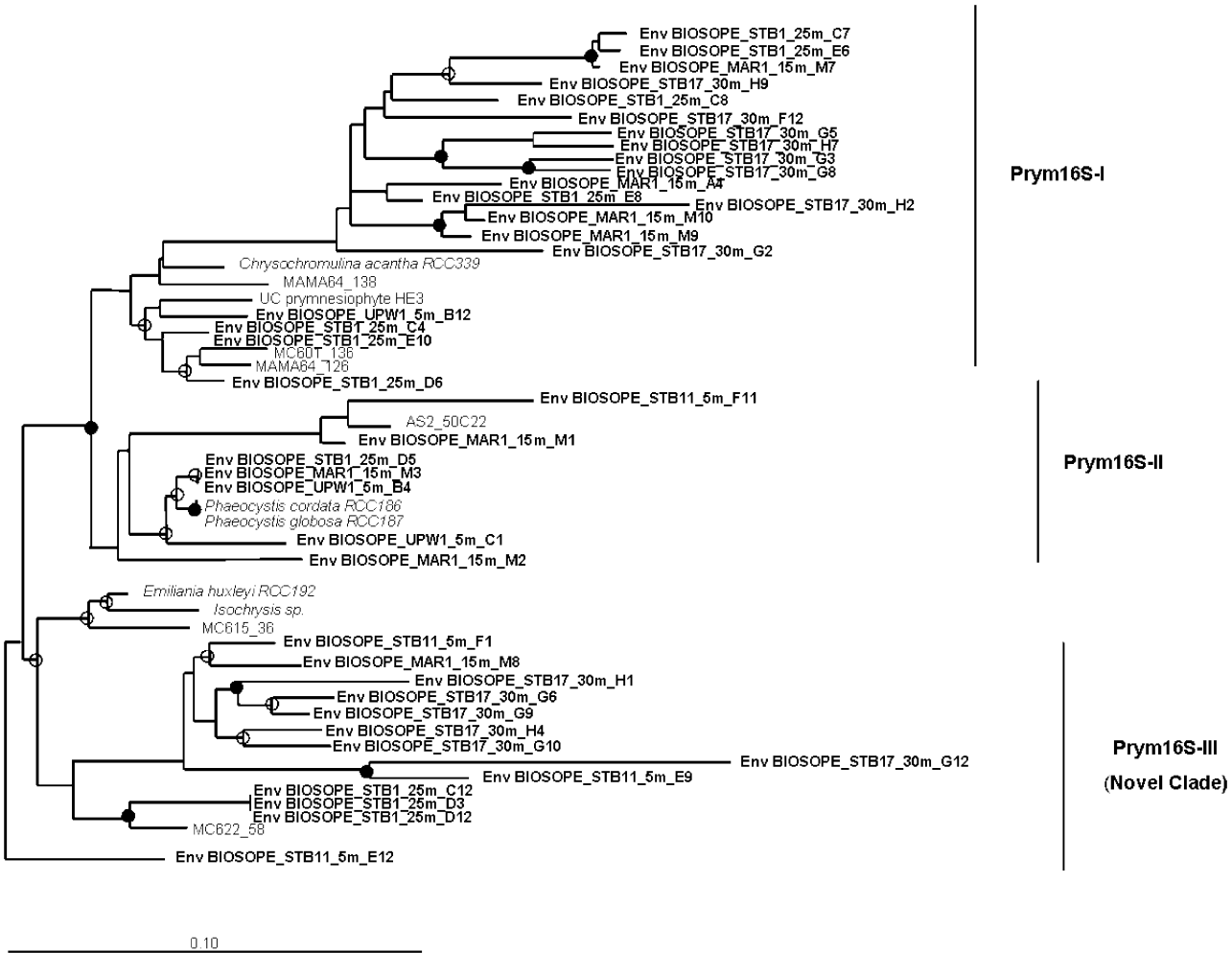

Fig. 5. Neighbour-joining phylogenetic tree of prymnesiophytes based on plastid 16S rRNA gene sequences. Bootstrap analysis was performed with ARB parsimony bootstrap. ●: values > 90%; ○: 70–90%; values < 70% are not shown. Sequences obtained during this study are highlighted in bold face. AS and MC sequences are derived from the Arabian Sea and Mediterranean Sea respectively (Fuller *et al.*, 2006a; McDonald *et al.*, 2007).

be phylogenetically affiliated with diatoms and were restricted to the mesotrophic stations MAR1 and UPW1.

PPE community composition in relation to environmental variables

A partial CCA (canonical correspondence analysis) was performed to help to determine the main parameters controlling the distribution of PPEs. Dot blot hybridization data were analysed in relation to the various physical, chemical and biological parameters measured along the transect. Such multivariate analysis hence allowed discrimination of the spatial distribution patterns of specific members of the PPE community as a function of environmental variables (Fig. 8). Members of the Cryptophyceae were linked to concentrations of Chl-*a*, phosphates and nitrites while Chrysophyceae appeared to be related to dissolved oxygen concentration, and Pinguiophyceae to salinity. A variation partitioning analysis was performed with environmental variables that independently explained a significant portion of the variation in CCA. This analysis showed that 13% of the variance in PPE community structure can be significantly explained by the physical parameters temperature and depth, while chemical and biological parameters were similarly important, 17 and 18% respectively. However, a large proportion of the variation remained unexplained by the variables measured as part of this study.

Discussion

The BIOSOPE transect encompassed a wide range of physical and chemical conditions, ranging from areas of high productivity in the Chile upwelling to ultra-

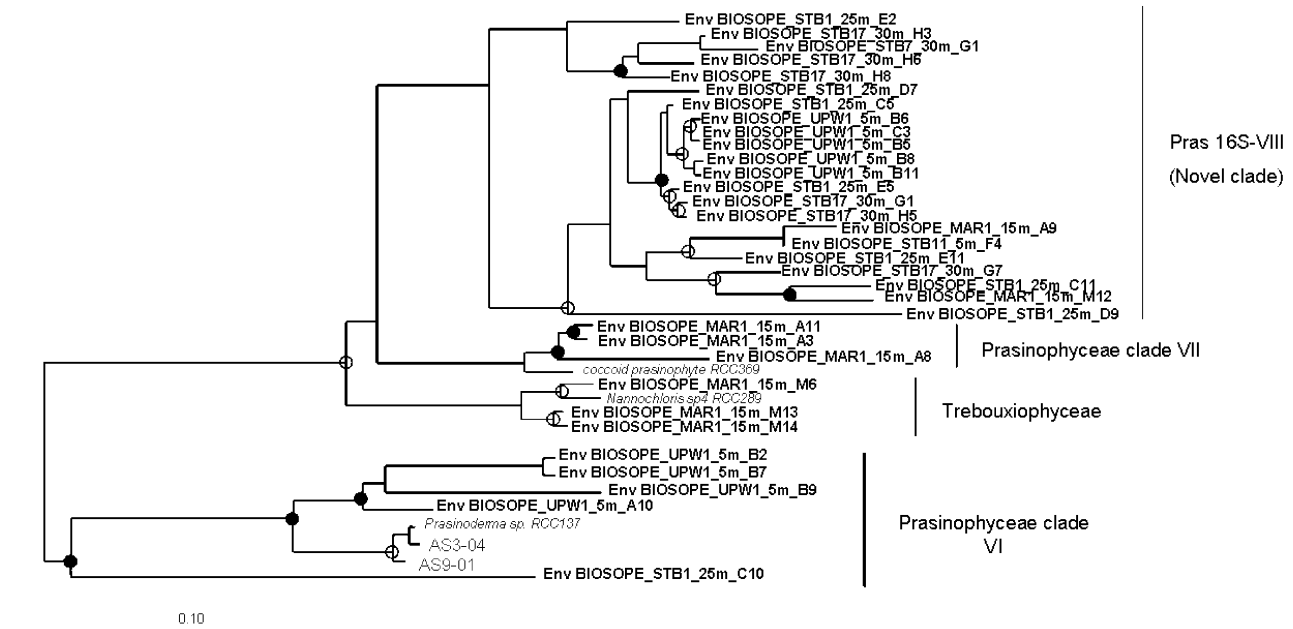

Fig. 6. Neighbour-joining phylogenetic tree of chlorophytes. See legend to Fig. 5 for details.

oligotrophic conditions in the central SPG. Concomitant with this, PPE abundance was lower in the nutrient-poor waters of the SPG, a region dominated numerically by *Prochlorococcus* (Grob *et al*., 2007), and greater in nutrient-rich coastal waters, a trend that is consistent with several previous data sets (e.g. Zubkov *et al*., 2000; Fuller *et al*., 2006b; Not *et al*., 2008).

Regardless of low PPE cell abundance in the SPG, dot blot hybridization data (Fig. 4) as well as phylogenetic analysis of environmental sequences (Figs 5–7) showed a relative complex PPE community, but with specific classes clearly predominating. As such, the BIOSOPE data set adds further weight to support the suggestion that members of the Chrysophyceae are major components of the PPE community in pelagic marine ecosystems even though this class is generally regarded as being more important in freshwater environments and that very few marine photosynthetic species have been described to date, most marine species being heterotrophic (e.g. *Paraphysomonas*). This is consistent with recent plastid 16S rRNA data from the Arabian and Mediterranean Sea (Fuller *et al*., 2006b; McDonald *et al*., 2007). These studies show that Chrysophyceae are predominantly associated with warm surface waters. Phylogenetically however, the BIOSOPE sequences (Fig. 7) are only poorly related to the plastid sequences from the Mediterranean and Arabian Sea (Fuller *et al*., 2006a; McDonald *et al*., 2007). Interestingly, Chrysophyceae sequences formed two clades (Chrys 16S-I and Chrys 16S-II), the former comprising sequences solely from the SPG suggesting a distinct distribution of lineages. Unfortunately, because the corresponding cells have not yet been isolated into culture, their exact taxonomic affiliation remains uncertain. Hence, they cannot be related to nuclear phylogenies using the 18S rRNA gene, for which environmental sequences forming novel lineages have recently been obtained (X.L. Shi, D. Marie, L. Jardillier, D.J. Scanlan and D. Vaulot, unpubl. data).

Prymnesiophytes were also well represented along the BIOSOPE transect and particularly within the clearest waters of the SPG. This distribution pattern is consistent with studies from several other oceanic regions, e.g. in the subtropical Indian Ocean, equatorial Pacific Ocean and Arabian Sea (Moon-van der Staay *et al*., 2000; Fuller *et al*., 2006b; Not *et al*., 2008), where this group was recognized particularly by pigment analysis as a major component of the eukaryotic picoplankton and particularly in oligotrophic waters. This is interesting since described species of the Prymnesiophyceae are almost all larger than 3 μm (Vaulot *et al*., 2008). The picoplanktonic size of these cells has recently been confirmed using fluorescent *in situ* hybridization (S. Masquelier and D. Vaulot, unpubl. data; L. Jardillier, pers. com.). Even though prymnesiophytes represented nearly half of all the sequences obtained from the clone libraries, it is remarkable that most of these were not phylogenetically affiliated to cultured counterparts or to other environmental sequences in the databases, especially given the number of plastid prymnesiophyte sequences present in clone libraries constructed from waters in the Arabian (27 sequences, Fuller *et al*., 2006a) or Mediterranean Sea (114 sequences, McDonald *et al*., 2007). Furthermore, the fact that this

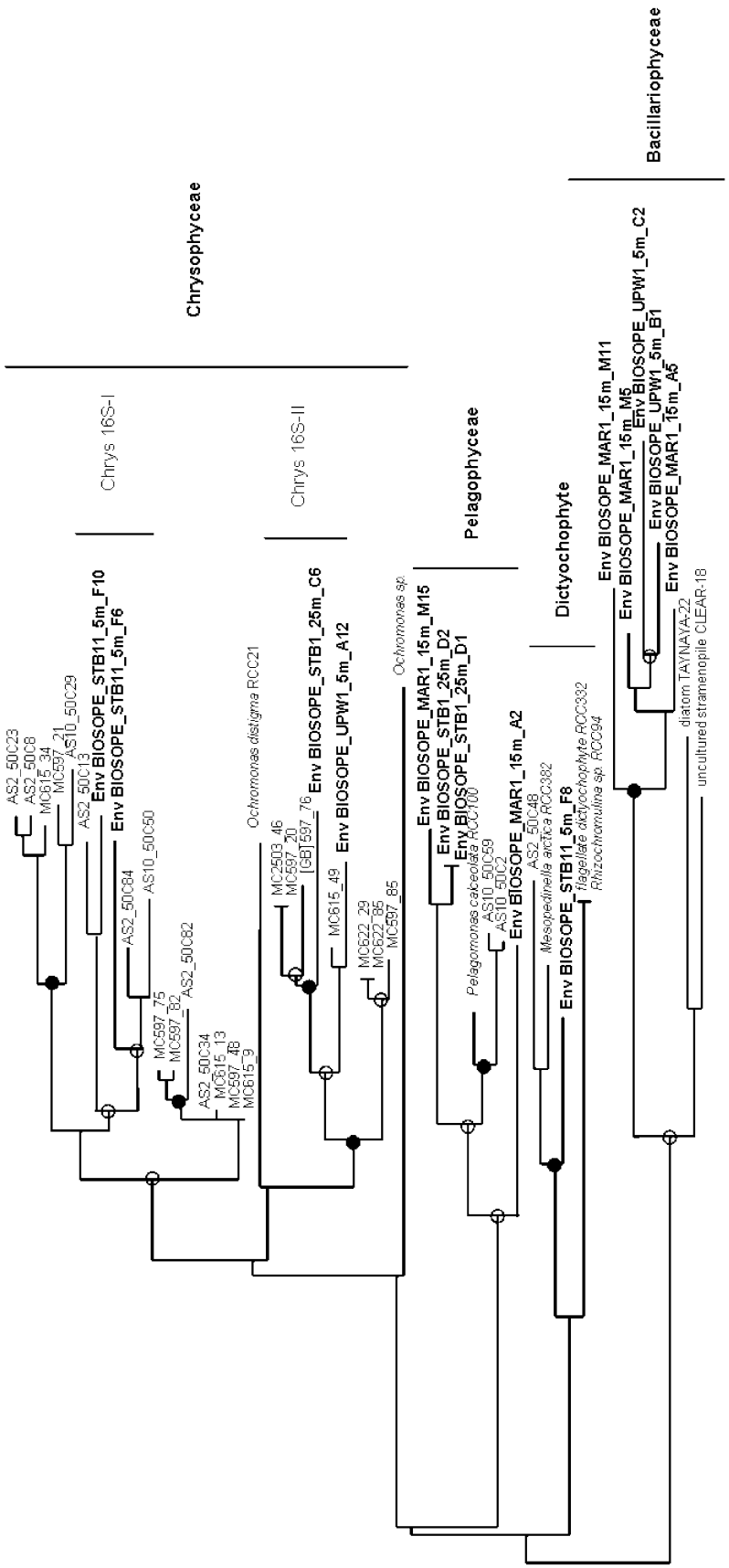

Fig. 7. Neighbour-joining phylogenetic tree of stramenopiles. See legend to Fig. 5 for details.

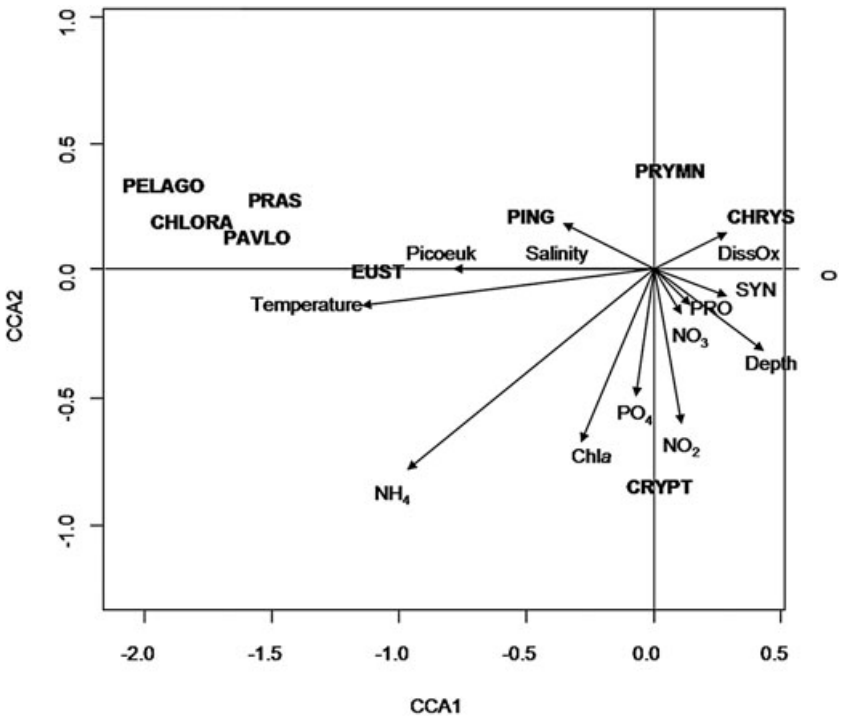

Fig. 8. Canonical correspondence analysis plot realized with percentage hybridization detected from dot blot hybridization analysis.

novel lineage was found in several clone libraries suggests it has a relatively broad distribution in this part of the South East Pacific at least.

Pelagophyte sequences were detected at stations MAR1 and STB1 (Fig. 7) and appear to be related to *Pelagomonas calceolata*, a small flagellate initially isolated from the North Pacific gyre (Andersen *et al*., 1993). Dot blot hybridization also confirmed the presence of this group near the Marquesas Island (Fig. 3). Several clone libraries from coastal regions do not include this group (Romari and Vaulot, 2004; McDonald *et al*., 2007), which suggests that this class might be restricted to more open ocean regions. Here, even though no sequences were detected in the gyre itself, this group was repeatedly isolated during the BIOSOPE cruise (Le Gall *et al*., 2008). Since only a single dictyochophyte sequence was detected, from the SPG station STB11, it is likely that this class is present only at low abundance in these South Pacific waters. Nonetheless, two dictyochophyte strains related to the picoplanktonic species *Florenciella parvula* (Eikrem *et al*., 2004) have been obtained from surface waters of the transect (Le Gall *et al*., 2008).

The total hybridization signal, i.e. the sum of all the class-specific relative hybridization signals, was often less than 100% (range 26–98%, Fig. 9). This likely reflects the presence of other algal classes for which plastid probes are not yet available, or as-yet, undiscovered classes. Indeed, sequencing of OTUs revealed both Prasinophyceae clade VII and dictyochophyte sequences, for neither of which specific probes exist. Prasinophytes

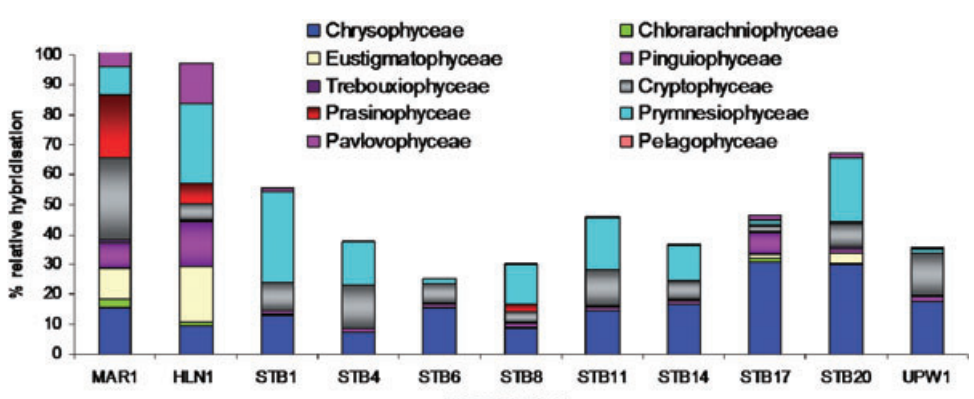

Fig. 9. Total percentage relative hybridization, i.e. the sum of all probes used, at each station along the BIOSOPE transect.

particularly are well represented among picoplankton 18S rRNA gene sequences, predominantly of the order Mamiellales. *Micromonas*, *Ostreococcus* and *Bathycoccus* appear to be the dominant genera especially in temperate coastal waters (Not *et al*., 2004). However, similar to an earlier study using the PLA491F-1313R PCR primer pair (McDonald *et al*., 2007), prasinophyte clade II (Mamiellales) sequences appear to be poorly detected using this approach and were not found in any of the clone libraries constructed along the BIOSOPE transect, despite the fact that they dominated 18S rRNA PPE clone libraries from the upwelling region (X.L. Shi, D. Marie, L. Jardillier, D.J. Scanlan and D. Vaulot, unpubl. data). Thus, as previously discussed (McDonald *et al*., 2007), there appears to be a partial trade-off between excluding cyanobacteria and bacteria and including all eukaryotes when using these plastid-targeted primers. Interestingly though, sequences affiliated with Prasinophyceae clade VI (Prasinococcales), VII and also a new clade (16S-VIII) were detected (Fig. 6). Sequences for the former two clades were restricted to upwelling and mesotrophic regions of the transect consistent with isolation studies (Le Gall *et al*., 2008) and pigment data (Ras *et al*., 2008) performed on the transect. Conversely, sequences from the potentially novel 16S rRNA gene clade Pras 16S-VIII were particularly well presented in libraries constructed from oligotrophic waters within and adjacent to the SPG (stations STB11 and STB17). It is possible that this clade may equate to clade IX of Viprey and colleagues (2008) observed initially in oligotrophic Mediterranean waters as well as in BIOSOPE picoeukaryote 18S rRNA clone libraries (X.L. Shi, D. Marie, L. Jardillier, D.J. Scanlan and D. Vaulot, unpubl. data) but for which, as yet, there are no plastid sequences.

In contrast to the Prasinophyceae, the other chlorophyte class Trebouxiophyceae was only retrieved at one station (MAR1) forming a single clade, loosely affiliated to *Nanochlorum*, a halotolerant genus that includes marine species as well as some isolated from hyper-saline waters (Henley *et al*., 2004). Indeed, culture representatives of Trebouxiophyceae, affiliated to the same *Nanochlorum* group, were isolated from the Chilean upwelling region (Le Gall *et al*., 2008).

While we obtained no sequences of Pavlovophyceae, Pinguiophyceae, Eustigmatophyceae or Chlorarachniophyceae in clone libraries, dot blot hybridization analysis revealed the presence of these groups along the transect, which exhibited very similar distributions consistent with their clustering in the CCA plot (Fig. 8). Although a few picoplanktonic species from some of these classes have been described (Kawachi *et al*., 2002; Suda *et al*., 2002), their oceanic distributions still remain poorly known.

Molecular analyses clearly provide much information on the diversity of picophytoplankton. However, we know much less of the factors involved in the regulation of these communities. The CCA analysis performed here showed a clear discrimination of groups on axis 1 or 2 of the plot indicative of a clear influence of specific environmental factors. Pelagophyceae, Chlorarachniophyceae, Prasinophyceae, Pavlovophyceae and Pinguiophyceae seemed to be more influenced by temperature and salinity, consistent with their higher relative hybridization values in western parts of the transect, which are generally warm, saltier waters (Fig. 8). Conversely, dissolved oxygen concentrations appear to explain chrysophyte distribution patterns (Fig. 8). The latter may be a function of the colder, highly oxygenated waters found in the eastern South Pacific.

According to variation partitioning analysis, the distribution of the different algal classes appears to be less influenced by the measured physical parameters of temperature and depth than to chemical and biological factors. It is significant that a large part of the variation (52%) remains unexplained by the parameters analysed here even though these are the parameters that are commonly considered when attempting to explain the spatial partitioning of marine microbes. Clearly other parameters must be taken into account when attempting to understand the distribution of this community. Missing from our data set are any mortality factors, e.g. predation (grazing), viral lysis, parasitism and/or competition, for which precise quantification is certainly non-trivial. Differential grazing rates on members of the PPE community have been reported, indicative of prey selection (Worden *et al*., 2004; Jezbera *et al*., 2006). These parameters require further inclusion in PPE community structure studies. The same can be said for the role of viruses in this process, which may be responsible in some cases for a considerable portion of total PPE mortality (Suttle *et al*., 1990; Evans *et al*., 2003; Baudoux *et al*., 2008). Clearly, further analyses *in situ* perhaps using long-term monitoring sites that encompass open ocean conditions, as well as using experimental approaches to analyse the influence of both biotic and abiotic factors on PPE community structure, are required.

Conclusion

The very large trophic gradient that was extensively sampled during the BIOSOPE cruise offered a unique opportunity to further increase our understanding of PPEs, particularly in the extremely oligotrophic conditions of the central SPG. The use of a targeted plastid clone library approach revealed several novel lineages, including new prasinophyte and prymnesiophyte clades. This suggests that the diversity of PPEs is still far from being completely elucidated. This is essential to identify the major PPE players, organisms that should not be under-

estimated as to their contribution and role in global C cycling. Moreover, the fact that a range of commonly measured physical, chemical and biological variables explained < 50% of the variation in distribution of specific PPE classes indicates we still have some way to go to identify the factors influencing PPE community structure, and ultimately controls on marine photosynthesis.

Experimental procedures

Sampling and oceanographic context

The BIOSOPE cruise took place on board the French research vessel l'Atalante in the South East Pacific Ocean from 26 October to 11 December 2004 (Claustre *et al*., 2008). Seawater samples were collected from 11 stations, 6 depths (0–300 m) along the BIOSOPE transect (Fig. 1) using Niskin bottles mounted on a CTD frame. Samples were subsequently filtered through a 47 mm diameter, 3 μm pore size polycarbonate filter (Nuclepore) and then onto a 0.22 μm pore size Sterivex filter (Millipore) using a peristaltic pump. The Sterivex filters were filled with DNA lysis buffer (0.75 M sucrose, 400 mM NaCl, 20 mM EDTA, 50 mM Tris-HCl pH 9.0), both ends obturated and then stored at −80°C until DNA extraction. Nutrients and Chl-*a* concentrations were determined as described in Claustre and colleagues (2008), with data being obtained from the BIOSOPE database (http://www.obs-vlfr.fr/proof/php/bio_open_access_data.php).

PCR amplification

Amplification of 16S rRNA genes from control strains and environmental samples for dot blot hybridization analysis and clone library construction used a marine algal plastid-biased primer, PLA491F (5′-GAGGAATAAGCATCGGCTAA-3′) (Fuller *et al*., 2006a), in conjunction with OXY1313R (5′-CTTCAYGYAGGCGAGTTGCAGC-3′) (Fuller *et al*., 2003), giving rise to an approximately 830 base pair PCR product. PCR amplification was carried out in a total reaction volume of 100 μl containing 200 μM deoxynucleotide triphosphates, 1.2 mM $MgCl_2$, 1.2 mM primers, 2.5 U of Taq polymerase in 1× enzyme buffer (Invitrogen), with 1 mg ml^{-1} bovine serum albumin (Roche). Amplification conditions comprised 95°C for 5 min, 80°C for 1 min, at which time the Taq polymerase was added, followed by 30 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 40 s, with a final extension at 72°C for 6 min

Dot blot hybridization

Dot blot hybridization conditions for the various marine algal class-specific oligonucleotides were optimised previously (Fuller *et al*., 2006b). Algal cultures used as controls were the same as those used by Fuller and colleagues (2006b) and were obtained from the Roscoff Culture Collection (RCC, http://www.sb-roscoff.fr/Phyto/RCC/) and the Provasoli-Guillard National Center for Culture of Marine Phytoplankton (CCMP, https://ccmp.bigelow.org/). 16S rDNA amplicons from BIOSOPE environmental DNAs and control strains were purified, blotted onto nylon membranes and hybridized to algal class-specific oligonucleotide probes, following the method of Fuller and colleagues (2003). The oligonucleotide probes used were: CHLA768, CHRY1037, CRYP862, EUST985, PAVL665, PELA1035, PING1024, PRAS826, PRYM666 and TREB708 targeting the plastids of Chlorarachniophyceae, Chrysophyceae, Cryptophyceae, Eustigmatophyceae, Pavlovophyceae, Pelagophyceae, Pinguiophyceae, Prasinophyceae clade VI (Prasinococcales), Prymnesiophyceae and Trebouxiophyceae respectively (Fuller *et al*., 2006b). Final wash (or dissociation) temperatures (T_d) for each probe were determined empirically (Fuller *et al*., 2006b), following a previously described method (Fuller *et al*., 2003). Hybridization was quantified by using a Fujifilm FLA-5000 phosphorimager and Total Laboratory software (Phoretix). Relative hybridization of the PPE class-specific probes to total oxygenic phototroph 16S rDNA sequences amplified by the PLA491F-OXY1313R primer pair was calculated according to the equation below:

$$\text{Relative hybridization}(\%) = \left[\left(\frac{S_{env}}{E_{env}}\right) \bullet \left(\frac{S_{con}}{E_{con}}\right)^{-1}\right] \times 100$$

where S*env* and E*env* represent hybridization to environmental DNA of the specific and eubacterial probes, respectively, and S*con* and E*con* are the slopes of the specific and eubacterial probe-binding curves, respectively, calculated by hybridizing each probe to dilution series of homogenous control DNAs. The relative hybridization of a given specific probe compared with that of the eubacterial probe to the control DNAs was averaged where more than one control DNA was used. Any sample giving a signal above 2% was considered above background.

Construction of plastid clone libraries

PCR products obtained from station MAR1 (15 m), STB1 (25 m), STB11 (5 m), STB17 (30 m) and UPW1 (5 m) using the PLA491F-OXY1313R primer pair, were cloned into vector pCR2.1, using the TOPO TA Cloning Kit (Invitrogen, Carlsbad, California, United States) according to the manufacturer's instructions. PCR amplicons derived from positive clones were subsequently digested with *Hae*III and the resulting restriction fragment length polymorphism (RFLP) products separated by agarose gel electrophoresis. Clones that produced the same RFLP pattern were considered members of the same OTU. At least one clone from each OTU was subsequently sequenced. Sequencing reactions were performed at the NERC Molecular Genetics Facility (Gene Pool, Ashworth Laboratories, Edinburgh).

Phylogenetic analysis

To determine phylogenetic affiliation, each sequence was compared with sequences available in databases using BLAST analysis (Altschul *et al*., 1997). We identified six cyanobacterial sequences in the clone libraries constructed here. The sequences were then aligned within ARB using the latter's automatic alignment tool (http://www.arb-home.de/) (Ludwig *et al*., 2004). The resulting alignments were checked and corrected manually, taking into account the secondary structure of the rRNA molecule. Sequences were screened for

potential chimeras using the Ribosomal Database project II program CHECK_CHIMERA and by the construction of alternative phylogenetic trees using 350 bp from either the 5′ or 3′ end of the sequenced fragment. The phylogenetic tree was constructed from sequences > 1200 nucleotides in length using Jukes-Cantor correction and a maximum frequency filter for plastids. Shorter sequences were added by parsimony using the same filter. Bootstrap analysis was performed using the ARB parsimony bootstrap algorithm. Nucleotide sequences determined in this study have been deposited in GenBank under accession numbers FJ649229–FJ649313.

Rarefaction analysis was performed using analytic rarefaction software (v 1.3) (http://www.uga.edu/~strata/software/Software.html) based on the analytical solution presented by Raup (1975) and Tipper (1979). The relative distribution of OTUs in the library was used to calculate coverage values (Good's coverage) (Hughes *et al*., 2001).

Statistical analyses

To explain the distribution of PPEs measured by dot blot hybridization (% relative hybridization), CCA was used (Ter Braak, 1986). Forward selection was performed to select the environmental variables that explained a significant part of the variation in PPE distribution pattern ($P < 0.05$). Variables included NH_4-N, NO_2-N, NO_3-N, PO_4-P, temperature, salinity, dissolved oxygen, depth, *Synechococcus* and *Prochlorococcus* abundance, Chl-*a* and PPE abundance.

To evaluate the effects of physical, biological and chemical parameters on PPE distributions a Variation Partitioning Analysis (VPA) was performed (Bocard *et al*., 1992; Lepère *et al*., 2006). All explanatory variables were divided into three groups: chemical variables (NH_4-N, NO_2-N, NO_3-N, PO_4-P, salinity and dissolved oxygen), physical variables (temperature, depth) and biological variables (*Synechococcus* and *Prochlorococcus* abundance, Chl-*a* and PPE abundance). For each experiment, we selected only variables that independently explained a significant proportion ($P < 0.05$) of the variation in the PPE classes by forward CCA. The VPA distinguished between pure chemical, physical and biological effects on PPEs and the proportion explained by interactions between these effects. These statistics were computed with R software using the Vegan package for the CCA and related methods (http://cran.r-project.org/).

Acknowledgements

We wish to thank all scientists, engineers and crew from the BIOSOPE cruise, and particularly H. Claustre who coordinated the project as well as M. Viprey and L. Garczarek for help in sampling. We also thank K. Zwirglmaier for extraction of DNA samples. This research was supported by the following programs: BIOSOPE cruise (CNRS-INSU LEFE) and PICOFUNPAC (ANR Biodiversité 06-BDIV-013).

References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25:** 3389–3402.

Andersen, R.A., Saunders, G.W., Paskind, M.P., and 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.

Baudoux, A.C., Veldhuis, M.J.W., Noordeloos, A.A.M., Vann Noort, G., and Brussaard, C.P.D. (2008) Estimates of virus- vs. grazing induced mortality of picophytoplankton in the North Sea during summer. *Aquat Microb Ecol* **52:** 69–82.

Bocard, D., Legendre, P., and Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology* **73:** 1045–1055.

Claustre, H., Sciandra, A., and Vaulot, D. (2008) Introduction to the special section bio-optical and biogeochemical conditions in the South East Pacific in late 2004: the BIOSOPE program. *Biogeosciences* **5:** 679–691.

Eikrem, W., Romari, K., Latasa, M., Le Gall, F., Throndsen, J., and Vaulot, D. (2004) *Florenciella parvula* gen. and sp. nov. (Dictyochophyceae, Heterokontophyta) a small flagellate isolated from the English Channel. *Phycologia* **43:** 658–668.

Evans, C., Archer, S.D., Jacquet, S., and Wilson, W.H. (2003) Direct estimates of the contribution of viral lysis and microzooplankton grazing to the decline of a *Micromonas spp*. population. *Aquat Microb Ecol* **30:** 207–219.

Field, C., Behrenfeld, M.J., Randerson, J.T., and Falkowski, P. (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281:** 237–240.

Fuller, N., Marie, D., Partensky, F., Vaulot, D., Post, A.F., and Scanlan, D.J. (2003) Clade-specific 16S ribosomal DNA oligonucleotides reveal the predominance of a single marine *Synechococcus* clade throughout a stratified water column in the Red Sea. *Appl Environ Microbiol* **69:** 2430–2443.

Fuller, N., Campbell, C., Allen, D.J., Pitt, F., Le Gall, F., Vaulot, D., and Scanlan, D.J. (2006a) 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.

Fuller, N.J., Tarran, G.A., Cummings, D., Woodward, E.M.S., Orcutt, K.M., Yallop, M., *et al.* (2006b) Molecular analysis of photosynthetic picoeukaryote community structure along an Arabian Sea transect. *Limnol Oceanogr* **51:** 2502–2514.

Grob, C., Ulloa, O., Claustre, H., Huot, Y., Alarcon, G., and Marie, D. (2007) Contribution of picoplankton to the total particulate organic carbon concentration in the eastern South Pacific. *Biogeosciences* **4:** 837–852.

Guillou, L., Eikrem, W., Chrétiennot-Dinet, M.J., Le Gall, F., Massana, R., Romari, K., *et al.* (2004) Diversity of picoplanktonic prasinophytes assessed by direct SSU rDNA sequencing of environmental samples and novel isolates retrieved from oceanic and coastal marine ecosystems. *Protist* **155:** 193–214.

Henley, W.J., Hironaka, J.L., Guillou, L., Buchheim, M.A., Buchheim, J.A., Fawley, M.W., and Fawley, K.P. (2004) Phylogenetic analysis of the '*Nannochloris*-like' algae and diagnoses of *Picochlorum oklahomensis* gen. et sp. nov. (Trebouxiophyceae, Chlorophyta). *Phycologia* **43:** 641–652.

Hughes, J.B., Hellmann, J.J., Ricketts, T.H., and Bohannan, B.J.M. (2001) Counting the uncountable: statistical approaches to estimating microbial diversity. *Appl Environ Microbiol* **67:** 4399–4406.

Jezbera, J., Horňák, K., and Šimek. K. (2006) Prey selectivity of bacterivorous protists in different size fractions of reservoir water amended with nutrients. *Environ Microb* **8:** 1330–1339.

Kawachi, M., Atsumi, M., Ikemoto, H., and Miyachi, S. (2002) *Pinguiochrysis pyriformis* gen. et sp. nov. (Pinguiophyceae), a new picoplanktonic alga isolated from the Pacific Ocean. *Phycol Res* **50:** 49–56.

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

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

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

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

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, et al. (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* **32:** 1363–1371.

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

Massana, R., Balagué, V., Guillou, L., and Pedrós-Alió, C. (2004) Picoeukaryotic diversity in a oligotrophic coastal site studied by molecular and culturing approaches. *FEMS Microbiol Ecol* **50:** 231–243.

Moon-van der Staay, S.Y., Van der Staay, G.W.M., Guillou, L., 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.

Morel, A., Gentili, B., Claustre, H., Babin, M., Bricaud, A., Ras, J., and Tieche, F. (2007) Optical properties of the 'clearest' natural waters. *Limnol Oceanogr* **52:** 217–229.

Not, F., Latasa, M., Marie, D., Cariou, T., Vaulot, D., and 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., Sharek, 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 Part I* **55:** 1456–1473.

Raimbault, P., Garcia, N., and Cerutti, F. (2008) Distribution of inorganic and organic nutrients in the South Pacific Ocean – evidence for a long-term accumulation of organic matter in nitrogen-depleted waters. *Biogeosciences* **5:** 281–298.

Ras, J., Claustre, H., and Uitz, J. (2008) Spatial variability of phytoplankton pigment distributions in the Subtropical South Pacific Ocean: comparison between *in situ* and predicted data. *Biogeosciences* **5:** 353–369.

Raup, D.M. (1975) Taxonomic diversity estimation using rarefaction. *Paleobiology* **1:** 333–342.

Romari, K., and Vaulot, D. (2004) Composition and temporal variability of picoeukaryote communities at a coastal site of the English Channel from 18S rDNA sequences. *Limnol Oceanogr* **49:** 784–798.

Suda, S., Atsumi, M., and Miyashita, H. (2002) Taxonomic characterization of a marine *Nannochloropsis* species, *N. oceanica* sp. nov. (Eustigmatophyceae). *Phycologia* **41:** 273–279.

Suttle, C.A., Chan, A.M., and Cottrell, M.T. (1990) Infection of phytoplankton by viruses and reduction of primary productivity. *Nature* **347:** 467–469.

Ter Braak, C.J.F. (1986) Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* **67:** 1167–1179.

Tipper, J.C. (1979) Rarefaction and rarefiction – the use and abuse of a method in paleontology. *Paleobiology* **5:** 423–434.

Vasilkov, A., Krotkov, P., Herman, N., McClain, J.C., Arrigo, K.R., and Robinson, W. (2001) Global mapping of underwater UV irradiances and DNA-weighted exposures using total ozone mapping spectrometer and sea-viewing wide filed-of-view sensor data products. *J Geophy Res* **106:** 27205–27219.

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

Viprey, M., Guillou, L., Ferréol, M., and Vaulot, D. (2008) Wide genetic diversity of picoplanktonic green algae (Chloroplastida) in the Mediterranean Sea uncovered by a phylum-biased PCR approach. *Environ Microbiol* **10:** 1804–1822.

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

Zubkov, M.V., Sleigh, M.A., Burkill, P.H., and Leakey, R.J.G. (2000) Picoplankton community structure on the Atlantic Meridional Transect: a comparison between seasons. *Prog Oceanog* **45:** 369–386.

Supporting information

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

Table S1. Dot blot hybridization data for each station along the BIOSOPE cruise track. Dot blot hybridization values are given as percentage relative hybridization (e.g. the Chrysophyceae specific probe signal is as a proportion of the total oxygenic phototroph 16S rRNA gene amplicons produced by the PLA491F-OXY1313R primer pair). *Not determined.

Table S2. Number of OTUs, and clones in brackets, identified within each algal class in the five genetic libraries.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.