

Genetic diversity of picoeukaryotes in a semi-enclosed harbour in the subtropical western Pacific Ocean

Man Kit Cheung, Ka Hou Chu, Chi Pang Li, Hoi Shan Kwan, Chong Kim Wong*

Department of Biology, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong SAR, China

ABSTRACT: Picoeukaryotes (<2 to 3 μm) are important components of aquatic ecosystems. The genetic diversity and seasonal variability of marine picoeukaryotes were compared between a semi-enclosed harbour and the adjacent open sea off the subtropical coast in the western Pacific Ocean based on 18S rRNA clone library analysis. Examination of 733 clones revealed 186 different restriction fragment length polymorphism (RFLP) patterns, representing 186 operational taxonomic units (OTUs). At least 19 high-level taxonomic groups of picoeukaryotes were recorded. Alveolates group II, ciliates and stramenopiles comprising 37, 17 and 11 % of the picoeukaryotes, respectively, were the most dominant groups. Phototrophs such as prasinophytes, cryptophytes and haptophytes were retrieved occasionally. Members from the 2 newly defined phyla, picobiliphytes and Telonemia, were also obtained. A differential spatial distribution of OTUs was observed between samples collected from the 2 sampling sites. Seasonal variations in picoeukaryote composition were more pronounced in the open sea libraries than in the semi-enclosed harbour libraries.

KEY WORDS: Picoplankton · Picophytoplankton · 18S · Urban harbour · Western Pacific Ocean · Spatial · Seasonal

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

Picoeukaryotes are eukaryotes smaller than 2 to 3 μm in diameter. They occur in aquatic environments worldwide and play fundamental roles in marine ecosystems (Sherr & Sherr 2000). Early studies were conducted to investigate their diversity in the major oceans (Díez et al. 2001, López-García et al. 2001, Moon-van der Staay et al. 2001). Worden (2006) studied the diversity of picoeukaryotes off the Californian coast, but no information is available on the diversity of these organisms in the subtropical coastal waters of the western Pacific Ocean or in semi-enclosed subtropical harbours.

This study was conducted in subtropical coastal seas in the northern part of the South China Sea, a large water mass in the western Pacific Ocean. The 2 sampling sites, Tolo Harbour and Mirs Bay, are located in the northeastern corner of Hong Kong. Tolo Harbour is

a landlocked bay with a long history of eutrophication (Wear et al. 1984). Because of its semi-enclosed topography, Tolo Harbour has poor water circulation. Mixing of water between the inner part of Tolo Harbour and the adjacent open sea is further restricted by the northeasterly wind, which blows against the direction of the outgoing tide. Mirs Bay, in contrast, is exposed to ocean currents from the South China Sea (Hong Kong Environmental Protection Department 2003). It is relatively unpolluted and contains 2 marine parks and most of the coral beds in Hong Kong. The availability of long-term data on phytoplankton composition and physio-chemical variables (Yung et al. 1997) makes the 2 sites good candidates for studying the diversity of picoeukaryotes.

Spatial and seasonal clone libraries from the 2 sampling sites were constructed and representative clones were selected for subsequent sequencing. The objectives are to (1) study the diversity of picoeukaryotes in

*Corresponding author. Email: chongkimwong@cuhk.edu.hk

a semi-enclosed subtropical harbour and the adjacent open sea off the coast in the western Pacific Ocean for the first time, (2) compare the compositional variations of picoeukaryotes between the semi-enclosed harbour and the adjacent open sea, and (3) investigate seasonal variations in picoeukaryote communities in the coastal waters of Hong Kong.

MATERIALS AND METHODS

Sample collection. Surface seawater samples (~2.5 l) were collected from stations in Tolo Harbour (22° 26' N, 114° 13' E) and Mirs Bay (22° 30' N, 114° 21' E) in January, April, July and October 2006 (Fig. 1). Immediately after collection, the water samples were filtered through a 200 µm mesh sieve to remove most of the mesozooplankton and large particles. Water temperatures, salinities and dissolved oxygen (DO) levels were measured on board ship using a HydroLab sensor. In the laboratory, chlorophyll *a* (chl *a*) concentrations were determined using a Turner Designs 10-AU fluorometer as described in Wong & Wong (2003). Diatom and dinoflagellate concentrations were determined by counting them under an inverted microscope. Two litres of water were prefiltered through 3 µm pore size Nuclepore® membranes (Whatman) and the microbial biomass was then collected onto a GF/F filter (Whatman). A gentle vacuum (<20 cm Hg) created by a hand pump was used to facilitate the filtration processes. The filter was then immersed in DNA lysis buffer (0.75 M sucrose, 40 mM EDTA, 50 mM Tris-HCl, pH 8), immediately frozen in liquid nitrogen and stored at -80°C until DNA extraction.

DNA extraction and PCR. DNA was extracted using a cetyltrimethylammonium bromide (CTAB) extraction procedure (Doyle & Doyle 1990). Extracts were purified with the GeneClean® II Kit (Q-BIOgene) and then stored at -20°C until use. The 18S rRNA gene was amplified by PCR using the eukaryotic primers

Euk328f (5'-ACC TGG TTG ATC CTG CCA G-3') and Euk329r (5'-TGA TCC TTC YGC AGG TTC AC-3') complementary to regions of conserved sequences proximal to the 5' and 3' termini of the 18S rRNA gene (Moon-van der Staay et al. 2000, Romari & Vaulot 2004). Each PCR mixture (50 µl) contained 2 µl of DNA template, 5 µl of 10× PCR buffer (50 mM KCl, 10 mM Tris-HCl and 1.5 mM MgCl₂), 200 µM of dNTP, 0.2 µM of each primer and 2.5 U *Taq* polymerase (Promega). The PCR thermal regime consisted of an initial denaturation of 3 min at 94°C, followed by 32 cycles of 45 s at 94°C, 30 s at 58°C, 2 min at 72°C and a final cycle of 10 min at 72°C. PCR products of the expected size (ca. 1800 bp) were purified with the QIAquick gel purification kit according to manufacturer's instructions (QIAGEN).

Cloning and screening. Purified PCR products were cloned using TA cloning kit (TaKaRa Biotechnology) following the manufacturer's recommendations. About 100 putative positive clones from each library were randomly selected, and the presence of the 18S rRNA gene insert was checked by PCR amplification with the same primers. Five microlitres of PCR product were digested for 2 h at 37°C with 2.5 U of restriction enzyme *Hae*III (Promega). Restriction fragments were separated by electrophoresis at 90 V for 50 min in a 2.5% agarose gel. Fragment patterns were visualised under UV lights after ethidium bromide staining. A 50 bp DNA ladder (Invitrogen) was included in each gel to aid the comparison of restriction fragment length polymorphism (RFLP) patterns of clones from different gels. Clones showing the same RFLP pattern were grouped and considered to belong to the same operational taxonomic unit (OTU). The average similarity (using partial sequences) among clones from the same OTU was 99.1% (range: 97.9 to 99.9%; data from 8 different RFLP pattern pairs of different high-level taxonomic groups).

Sequencing and phylogenetic analyses. One representative clone of each RFLP type was purified and then partially sequenced using the internal primer Euk528f (5'-GCG GTA ATT CCA GCT CCA A-3') (Elwood et al. 1985) and ABI PRISM® BigDye™ Terminators v3.1 Kit (Applied Biosystems) on an ABI PRISM® 3700 DNA Analyzer (Applied Biosystems). To determine the phylogenetic affiliation, all sequences were subjected to a BLAST search in GenBank (2 March 2008 database; Table A1, available in AME Supplementary Material online at www.int-res.com/articles/suppl/a053p295_app.pdf). Potential chimeric sequences were checked using the CHECK_CHIMERA program (Cole et al. 2003). Nucleotide sequences from the current study were deposited in GenBank under accession numbers EF538969 to EF539172 and FJ176218 to FJ176220. Library coverage

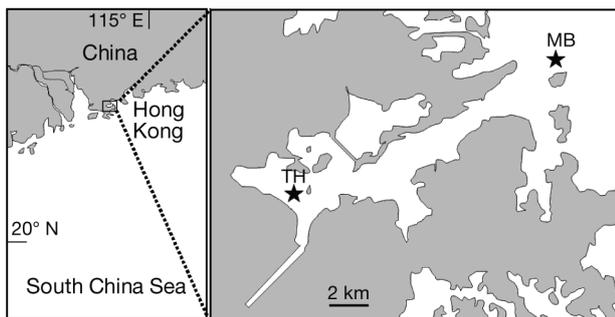


Fig. 1. Locations of sampling stations in Tolo Harbour (TH) and Mirs Bay (MB). Inset shows relative location along South China Sea coastline

Table 1. Collection information, hydrological variables and phytoplankton cell counts for water samples collected from January to October 2006

Sampling date	Temp. (°C)	DO (mg l ⁻¹)	Chl <i>a</i> (µg l ⁻¹)	Diatoms (cells ml ⁻¹)	Dinoflagellates (cells ml ⁻¹)
Tolo Harbour					
16 Jan	19.5	9.3	5.7	508	23
18 Apr	21.6	6.2	2.4	340	8
14 Jul	26.5	5.5	16.2	424	8
26 Oct	27.5	5.7	4.8	141	137
Mirs Bay					
16 Jan	16.8	9.4	0.8	27	9
18 Apr	20.8	7.1	2.0	262	5
14 Jul	27.0	6.2	2.8	122	12
26 Oct	27.3	7.3	1.5	6	4

Table 2. Summary of restriction fragment length polymorphism (RFLP) analysis of the 8 clone libraries. N: number of clones; OTUs: operational taxonomic units grouped by the same RFLP patterns; coverage value indicates how well the diversity within each sample was estimated by the library

Sampling date	Clone library	N	OTUs	Coverage (%)
Tolo Harbour				
16 Jan	TH01	100	34	84.0
18 Apr	TH04	96	32	79.2
14 Jul	TH07	87	36	78.2
26 Oct	TH10	99	41	79.8
Mirs Bay				
16 Jan	MB01	100	41	76.0
18 Apr	MB04	98	39	78.6
14 Jul	MB07	71	34	71.8
26 Oct	MB10	82	44	67.1

values were calculated using the software SPADE (Chao & Shen 2003).

High quality non-metazoan sequences that passed the chimera checking process, together with their closest GenBank relatives, were first aligned using the ClustalW multiple alignment tool integrated in MEGA 3.1 (Kumar et al. 2004) and the alignments were then manually improved. Ambiguously aligned regions were identified and removed by Gblocks v0.91b (Castresana 2000) (minimum block length = 5, allowed gap positions = with half). Maximum-likelihood (ML) phylogenetic trees were constructed using PHYML v2.4.4 (Guindon & Gascuel 2003) with the general time-reversible (GTR) nucleotide substitution model. Parameters for gamma distribution (G) and invariable sites (I) were estimated by the program. Statistical confidence on the inferred phylogenetic relationships was assessed by 100 ML bootstrap replicates. The inferred trees were displayed using MEGA.

RESULTS

Hydrological parameters

Water temperatures in both Tolo Harbour and Mirs Bay showed a similar seasonal pattern, marked by an increasing trend from January to October (Table 1). Surface DO levels varied between 5.5 and 9.4 mg l⁻¹ and were always lower in Tolo Harbour than in Mirs Bay. Mean chl *a* concentrations were 7.3 µg l⁻¹ in Tolo Harbour and 1.8 µg l⁻¹ in Mirs Bay. Concentrations of diatoms and dinoflagellates were generally higher in Tolo Harbour than in Mirs Bay.

Clone libraries

From 71 to 100 clones were examined in each clone library, revealing 32 to 44 different RFLP patterns or OTUs (Table 2). Coverage values calculated for most individual libraries were high (ca. 80%), except for libraries MB10 and TH10, which had values of around 70%. A total of 186 OTUs were observed in all 733 picoeukaryotic clones (Table 3). An additional 16 OTUs of metazoan origin (ca. 7.5% of total OTUs), including

Table 3. Operational taxonomic unit distribution of high-level taxonomic groups of picoeukaryotes in pooled Mirs Bay, pooled Tolo Harbour and total libraries. Number in parentheses represents the number of clones of a particular group in the library

Taxonomic group	Mirs Bay	Tolo Harbour	Total libraries
Alveolates group II	22 (125)	34 (146)	46 (271)
Alveolates group I	4 (16)	6 (49)	8 (65)
Dinophyceae	12 (40)	7 (14)	16 (53)
Ciliophora	23 (62)	20 (65)	34 (127)
Stramenopiles	11 (40)	25 (42)	31 (82)
Prasinophyceae	9 (28)	7 (13)	11 (41)
Picobiliphytes	2 (7)	4 (19)	4 (26)
Cercozoa	11 (22)	5 (9)	14 (31)
Haptophyceae	2 (4)	3 (3)	4 (7)
Cryptophyta	1 (1)	2 (5)	2 (6)
Cryptophyte nucleomorph	2 (2)	1 (2)	2 (4)
Katablepharidophyta	0 (0)	1 (2)	1 (2)
Telonemia	1 (1)	1 (1)	2 (2)
Fungi	2 (4)	0 (0)	2 (4)
Choanoflagellida	0 (0)	1 (1)	1 (1)
Radiozoa	1 (1)	2 (2)	3 (3)
Chlorarachniophyte nucleomorph	1 (1)	0 (0)	1 (1)
Mesomycetozoea	1 (1)	1 (1)	2 (2)
Ellobiopsids	1 (2)	1 (1)	1 (3)
Unknown group	1 (1)	1 (1)	1 (2)
Total	107 (358)	122 (375)	186 (733)

13 from Copepoda, 2 from Cnidaria, 1 from Ctenophora and 11 potential chimeric sequences, were detected and excluded from further analyses.

Taxonomic distribution

At least 19 high-level taxonomic groups of pico-eukaryotes were recorded (Table 3). In the combined dataset, alveolates group II, with 46 OTUs (ca. 25% of total) and 271 clones (ca. 37% of total), was the most abundant group. Ciliates and stramenopiles were also abundant, representing about 18 and 17%, respectively, of the total OTUs (ca. 17 and 11% of total clones). Members from Dinophyceae, Prasinophyceae and Cercozoa were also well represented, contributing 16 OTUs (53 clones), 11 OTUs (41 clones) and 14 OTUs (31 clones), respectively. Other groups were less common, each contributing fewer than 10 OTUs. Members from the 2 newly defined phyla, picobiliphytes (Not et al. 2007) and Telonemia (Shalchian-Tabrizi et al. 2006), were also recorded in this study.

When the datasets were grouped according to sampling sites, alveolates group II and stramenopiles harboured more OTUs in the Tolo Harbour libraries even though the total number of contributed clones was similar between the 2 sites (Table 3). While alveolates group I, picobiliphytes and cryptophytes were more abundant in Tolo Harbour, dinoflagellates and prasinophytes showed greater contribution in Mirs Bay. Cercozoa was both more diverse and abundant in Mirs Bay. The numbers of OTUs and clones for ciliates were similar for both sites. Katablepharidophytes and choanoflagellates were only found in Tolo Harbour

libraries. Fungi and chlorarachniophyte nucleomorphs were restricted to Mirs Bay libraries, and were represented by just a few clones.

All 6 major taxonomic groups—alveolates group I and group II, dinoflagellates, ciliates, stramenopiles and prasinophytes—appeared in all Tolo Harbour libraries throughout the year and varied seasonally only in relative clonal abundance (Fig. 2). In contrast, occurrence in the Mirs Bay libraries tended to be sporadic, with all 6 groups present in April, 5 in January and July, and only 4 in October.

Phylogenetic affiliations of OTUs

ML trees were constructed to investigate the phylogenetic positions of sequences obtained in the present study. Most of the high-level taxonomic groups were supported by high bootstrap values (Fig. 3). Alveolates, stramenopiles and Rhizaria were diverse and independent phylogenetic trees were drawn (Figs. 4, 5 & 6). Novel clades were defined here using the following set of criteria: (1) at least 3 sequences were involved; (2) sequences were retrieved from at least 3 independent environmental libraries; and (3) sequences in the group were at least 95% identical.

Alveolate sequences were divided mainly into 4 major groups: alveolates group I, alveolates group II, ciliates and dinoflagellates (Fig. 4). Within the most diverse alveolates group II, members from 13 of 16 previously defined clusters (Grosillier et al. 2006) were recovered in the current study (Fig. 4C). Within the *Amoebophrya* clade formed by cluster 1 to cluster 5, MB07.29 was >98% similar to AP-picoclone20 and TH10.4 was 93.7% similar to BB01-83. TH07.6, TH07.32 and RA010613.126 formed a distinct clade with 100% bootstrap support. Another clade was formed by TH07.2, TH07.12 and TH07.26 with 99% bootstrap support. Sequence identity between TH01.1 and UEPACCp3 and between TH01.8 and UEPACBp4 was >98%. Both UEPACCp3 and UEPACBp4 were from the eastern Pacific Ocean. TH10.1 was 95.7% similar to OLI11012 from the equatorial Pacific. The phylogenetic positions of MB07.15, MB01.42, TH10.49 and TH04.26 were unclear from the current analysis using partial sequences.

Of 5 previously defined alveolates group I clusters (Grosillier et al. 2006) 3 were recovered in the current study (Fig. 4B). MB04.17 had >98% identity with ENVP 21819.00079. TH07.4 represented a distant lineage of unclear phylogenetic position.

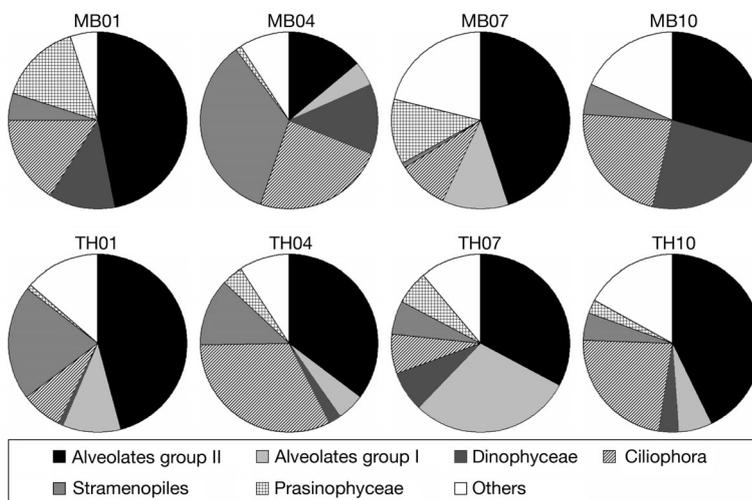
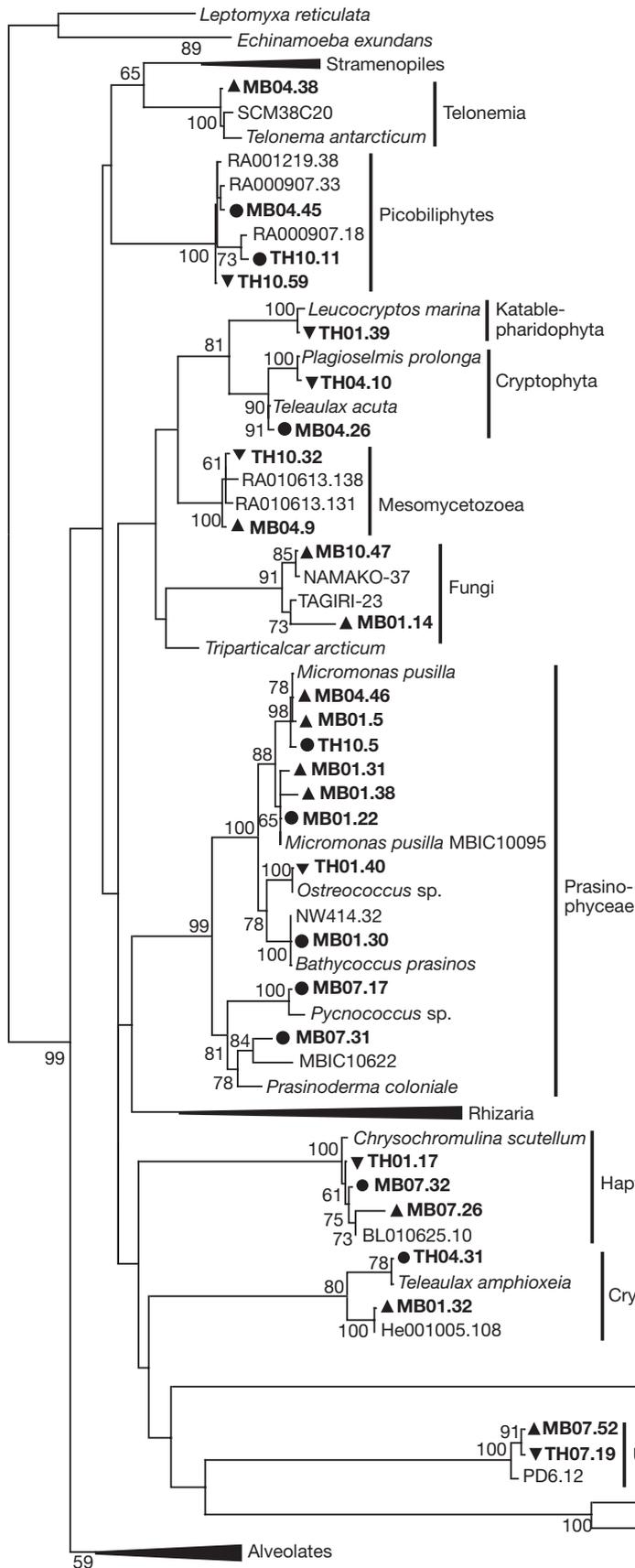


Fig. 2. Relative clonal abundance of the 6 most represented picoeukaryote groups in the 8 clone libraries. See Table 2 for abbreviations for clone libraries



Most dinoflagellate sequences were closely related to known species including *Pyrodinium bahamense*, *Karlodinium micrum*, *Lepidodinium viride*, *Prorocentrum micans*, *Gyrodinium spirale*, *Ceratium longipes* and *Noctiluca scintillans* (Fig. 4B). MB01.33 and TH07.15 had 99 and 100% identity with SCM15C23 and CD8.10, respectively, forming a potentially novel dinoflagellate clade.

A diverse lineage of ciliates including members from Prostomatea, Colpodea and Spirotrichea was also recovered in the current study (Fig. 4A). Most of the sequences retrieved were closely related to environmental sequences, but distantly related to cultured organisms. In addition to the recovery of sequences belonging to the previously defined *Strombidium* clades *Strom A* and *Strom B* (Lovejoy et al. 2006), 2 potentially novel *Strombidium* clades, each supported with high bootstrap value, were also retrieved here.

Stramenopiles represented another diverse assemblage in the current study (Fig. 5). Most of the stramenopile sequences obtained (25 out of 31 OTUs) were from heterotrophic lineages, particularly from previously described marine stramenopile (MAST) clusters (Massana et al. 2004). Members from 7 of the 12 clusters were retrieved in the current study. For other heterotrophic members, MB04.21 had 99%

Fig. 3. Maximum-likelihood (ML) phylogenetic tree of 18S rDNA sequences from all major picoeukaryotic groups observed in the present study. The tree is based upon 79 partial sequences of 447 positions after Gblock processing. Sequences obtained in the current study are in bold. Symbol codes to the left of sequence names indicate the origin of sampling (▲ = Mirs Bay, MB; ▼ = Tolo harbour, TH; ● = both). Numbers at nodes are ML bootstrap percentages from 100 replicates. Values < 50% are not shown. The tree is rooted using 2 Amoebozoa sequences (*Leptomyxa reticulata* and *Echinamoeba exundans*). Scale bar = 0.1 substitution per base. Collapsed clusters (alveolates, stramenopiles and Rhizaria) are expanded in the next 3 figures

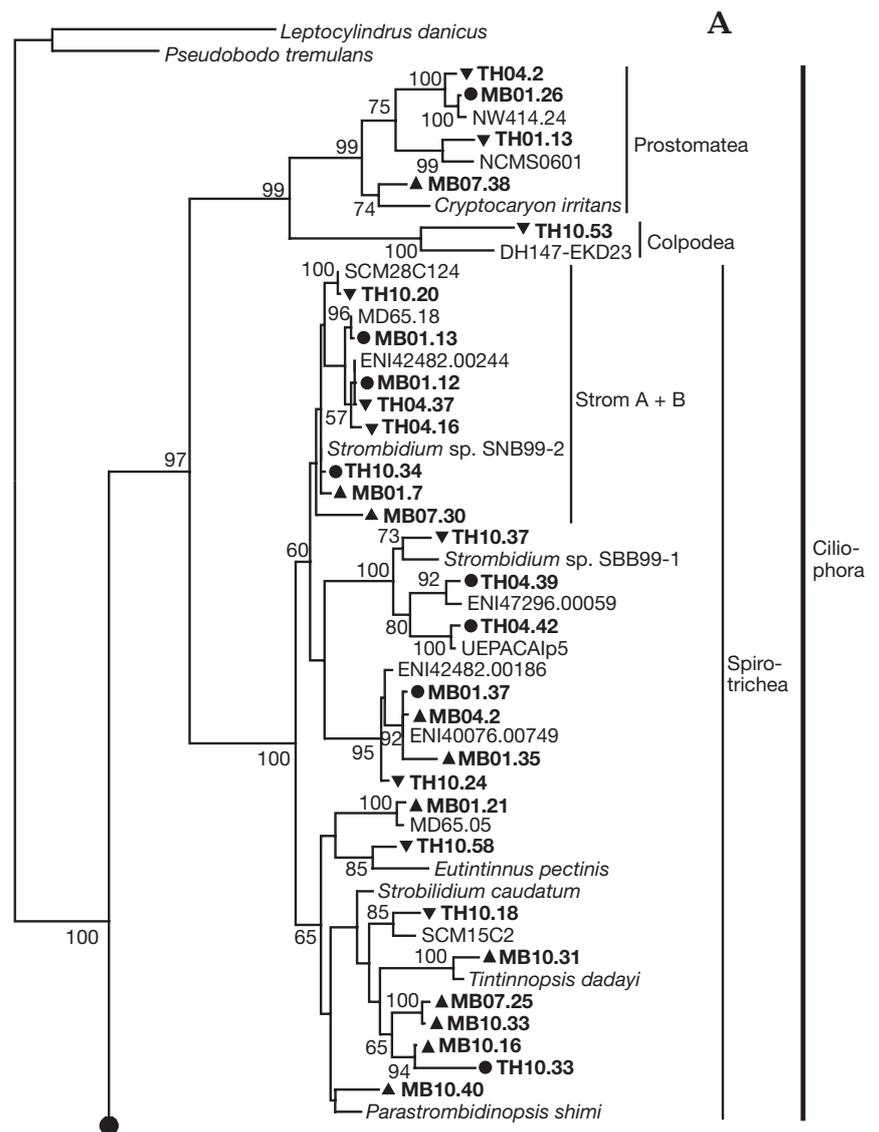
0.1

identity with oomycete clone C3-E035, TH04.36 was closely related (>98%) to *Pseudobodo tremulans* (Bicosoecida), and TH01.44 was loosely related to CAR_D107 and the closest known labyrinthulid relative *Labyrinthuloides haliotidis*. TH10.39 was >97% similar to ENVP10203.00002. MB01.34 and TH10.51 occupied unclear phylogenetic positions in the current analysis using partial sequences. Photosynthetic stramenopiles formed a monophyletic lineage. Six stramenopile sequences belonging to 3 typical algal classes (Bacillariophyceae, Bolidophyceae and Chrysophyceae) were retrieved in the current study.

Sequences belonging to Cercozoa and Radiozoa were also retrieved from the current study (Fig. 6). Most of the cercozoan sequences belonged to the core Cercozoa cluster. Except for TH01.15, which had >98% identity with *Cryothecomonas aestivalis*, all other core cercozoan sequences were closely related to environmental sequences. MB07.6, TH10.54, MB07.23 and MB07.50 grouped with TAGIRI-3 and C1 to form a cluster previously referred to as NC2 (Bass & Cavalier-Smith 2004). MB10.21 and MB07.4 formed a highly supported group (100% bootstrap support) with NOR46.27 and NAMAKO-6, possibly representing a novel cercozoan clade. MB07.22 had >99% identity with SCM27C7 and TH07.17 was related to TH04.43. These 4 were grouped together with 82% bootstrap support. Besides the core cercozoan sequences, MB04.43 was closely related (>98%) to NOR26.21, an environmental sequence belonging to Chlorarachnea, while MB07.8 was loosely related to plasmodiophorid clone RD010517.43. For Radiozoa, one sequence closely related to *Pseudocubus obeliscus* (Polycystinea) and another to *Amphibelone anomala* (Acantharea) were recovered.

Prasinophyceae was the second most diverse photosynthetic assemblage in the current study (Fig. 3). Sequences highly similar (>98%) to cultured organisms belonging to the genera *Micromonas*, *Ostreococcus*, *Bathycoccus* and *Pycnococcus* were retrieved. The only exception was MB07.31, which was 95.8% similar to the coccoid prasinophyte clone MBIC10622. The 3 recorded haptophyte sequences were all closely related to the genus *Chrysochromulina* (Prymnesiophyceae). Two cryptophyte sequences, one closely related to *Plagioselmis prolonga* and the other to *Teleaulax acuta*, were also retrieved. One katablepharidophyte sequence shared >99% identity with *Leucocryptos marina*. Three sequences closely related (>98%) to environmental relatives belonging to picobiliphytes, a newly defined algal phylum (Not et al. 2007) were also recovered.

Fig. 4. (This and facing page.) ML phylogenetic tree of alveolate 18S rDNA sequences. The tree is based upon 162 partial sequences of 505 positions after Gblock processing. The tree is rooted using 2 stramenopile sequences (*Pseudobodo tremulans* and *Leptocylindrus danicus*). Other definitions as in Fig. 3. (A) Upper one-third of the tree, containing all ciliate sequences. (B) Middle one-third of the tree, containing all dinoflagellate and alveolates group I sequences. (C) Lower one-third of the tree, containing all alveolates group II sequences



Among the other heterotrophic lineages, one sequence belonged to the newly defined phylum *Telonemia* (Shalchian-Tabrizi et al. 2006). The sequences MB10.47 and MB01.14 appeared to be loosely related to the closest known fungus relative *Triparticalcar arcticum* and closely related to environmental clones NAMAko-37 and TAGIRI-23, respectively. Two sequences closely related to cryptophyte nucleomorph and one sequence loosely related to chlorarachniophyte nucleomorph were also found.

Seven OTUs could not be identified using analysis based on partial sequences. More detailed analyses based on full-length 18S sequences (GenBank accession numbers FJ176218 to FJ176220) revealed that 2 of them belonged to Mesomycetozoa, 3 belonged to Ellobiopsids and the remaining 2 belonged to a group with unknown identity (Fig. 3).

DISCUSSION

Picoeukaryotic diversity

In this study, the diversity of picoeukaryotes in a semi-enclosed harbour and the adjacent open sea area of the western Pacific Ocean was studied. Alveolates group II, the most diverse and abundant group, was first discovered in an 18S rRNA-based environmental

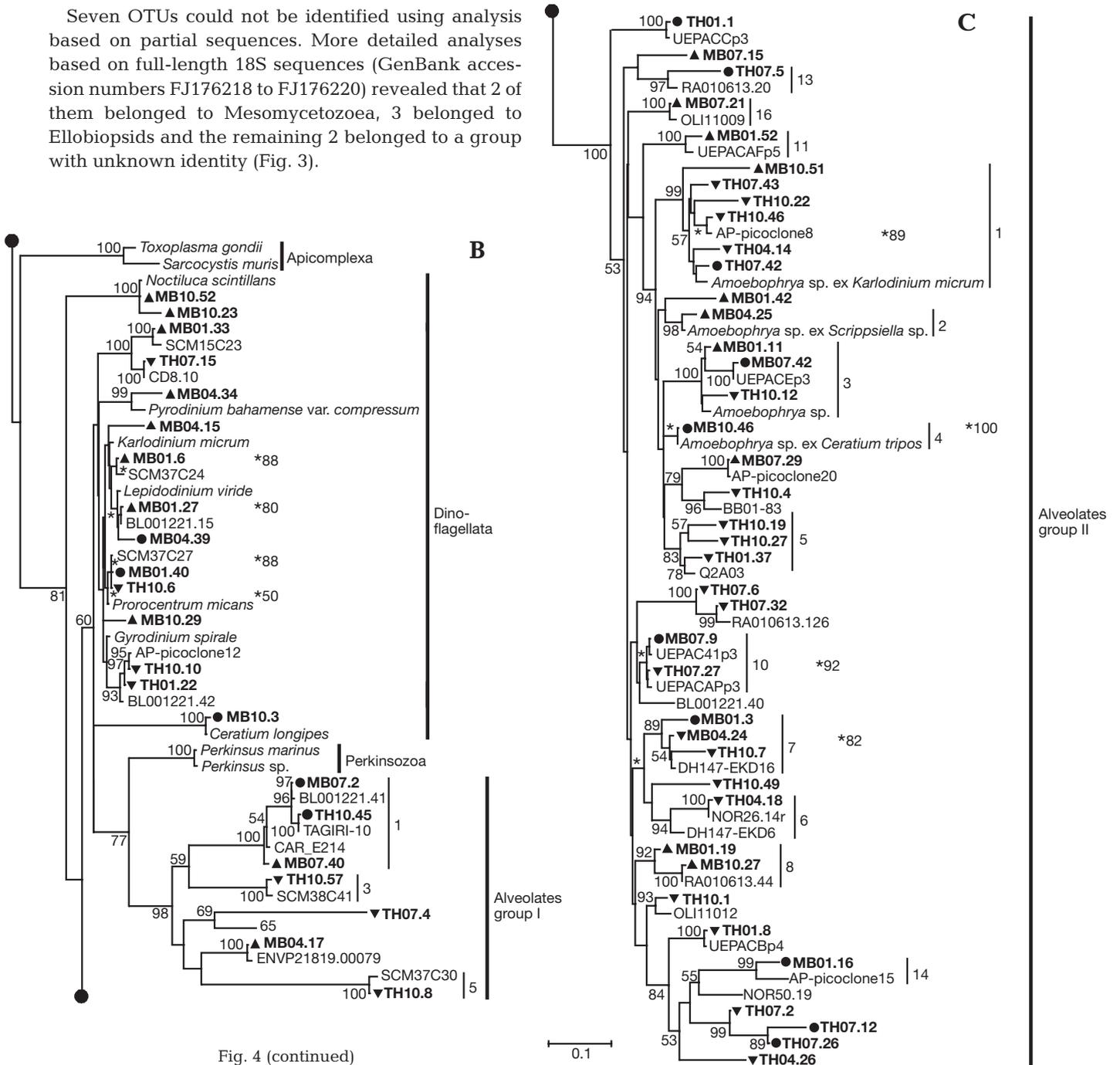


Fig. 4 (continued)

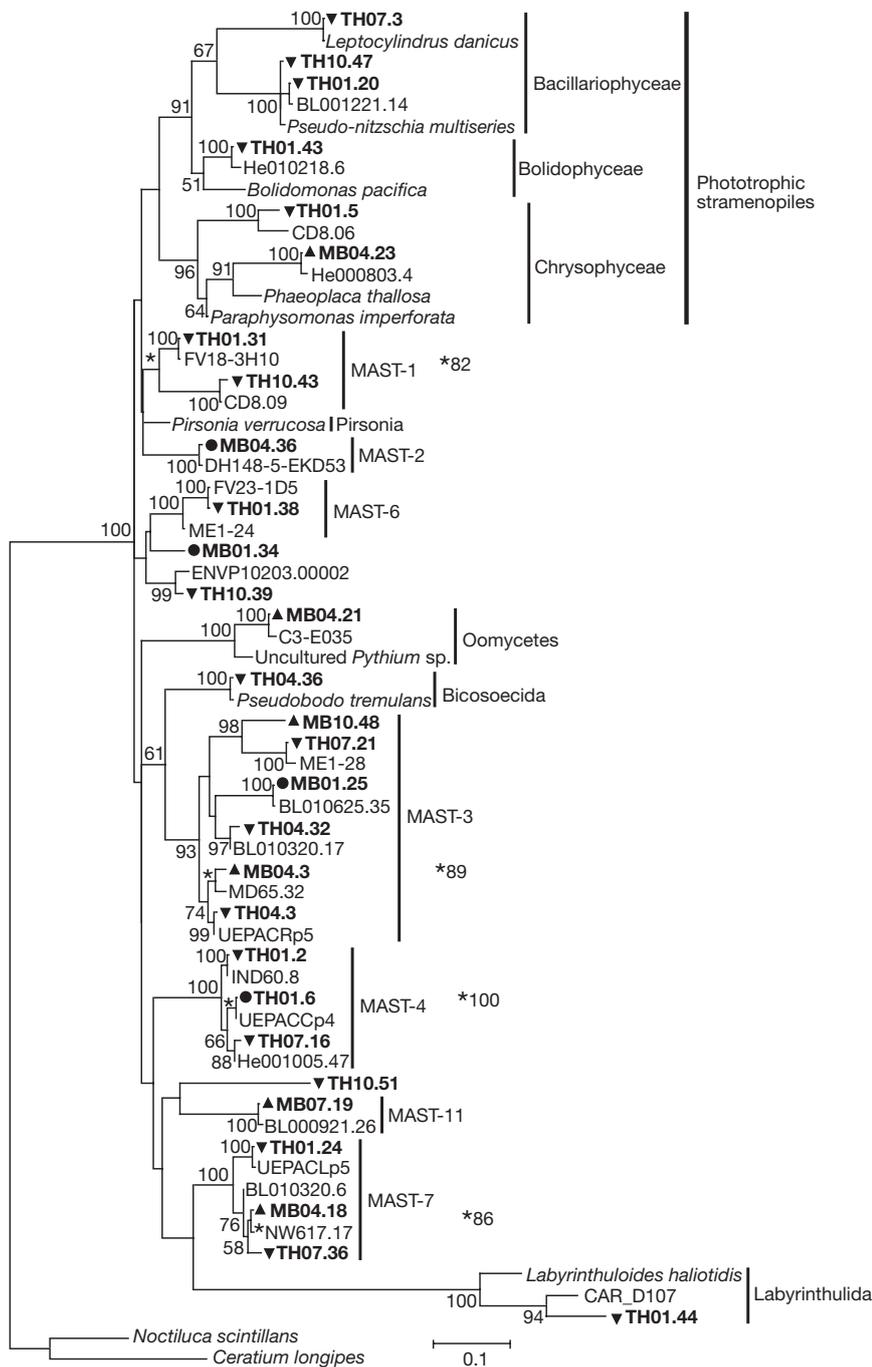


Fig. 5. ML phylogenetic tree of stramenopile 18S rDNA sequences. The tree is based upon 64 partial sequences of 517 positions after Gblock processing. The tree is rooted using 2 dinoflagellate sequences (*Noctiluca scintillans* and *Ceratium longipes*). Other definitions as in Fig. 3

diversity survey of small eukaryotes (<5 μm) in the Antarctic polar front (López-García et al. 2001). The parasitic nature of this group is suggested by its close phylogenetic relationship with *Amoebophrya* spp., known parasites of dinoflagellates (Grosillier et al. 2006). The high diversity and clonal abundance of the

group recovered here, together with its ubiquitous distribution in other marine ecosystems, support the suggestion that these organisms are important components of marine microbial ecosystems (Grosillier et al. 2006).

Grosillier et al. (2006) performed a comprehensive analysis on the genetic diversity and ecological distributions of the 2 novel groups of alveolates from 62 environmental clone libraries and identified 5 clusters from group I and 16 clusters from group II. However, some group II sequences obtained in our study could not be clearly assigned to previously defined clusters, raising the possibility of additional rare clusters that would need further analysis as the sequence database of the group expands (Fig. 4C). Specific examples included a moderately supported (79%) group composed of MB07.29, TH10.4, AP-picoclone 20 and BB01-83 within the *Amoebophrya* clade, and 2 highly supported ($\geq 99\%$) groups composed of sequences retrieved from less than 3 libraries (TH07.6 + TH07.32 + RA 010613.126 and TH07.2 + TH07.12 + TH07.26). Some of these sequences are highly similar to sequences retrieved from other environmental studies conducted in the Pacific Ocean (i.e. UEPACCp3, AP-picoclone20, OLI 11012 and UEPACBp4), suggesting that they might represent phylogenotypes restricted to the Pacific Ocean. Also, 39% of the current group II sequences and 25% of the current group I sequences are <95% similar to GenBank sequences (Table A1), suggesting that the vast diversity of these 2 novel alveolate groups deserves further examination.

Stramenopiles form an extremely diverse phylogenetic group that includes both phototrophs and heterotrophs. Phototrophic stramenopiles

formed a deeply branched monophyletic group in the current study (Fig. 5). This, together with the fact that heterotrophic stramenopiles are paraphyletic and occupy a basal position in the phylogenetic tree, is consistent with the hypothesis that stramenopiles are primarily heterotrophic and phototrophic groups origi-

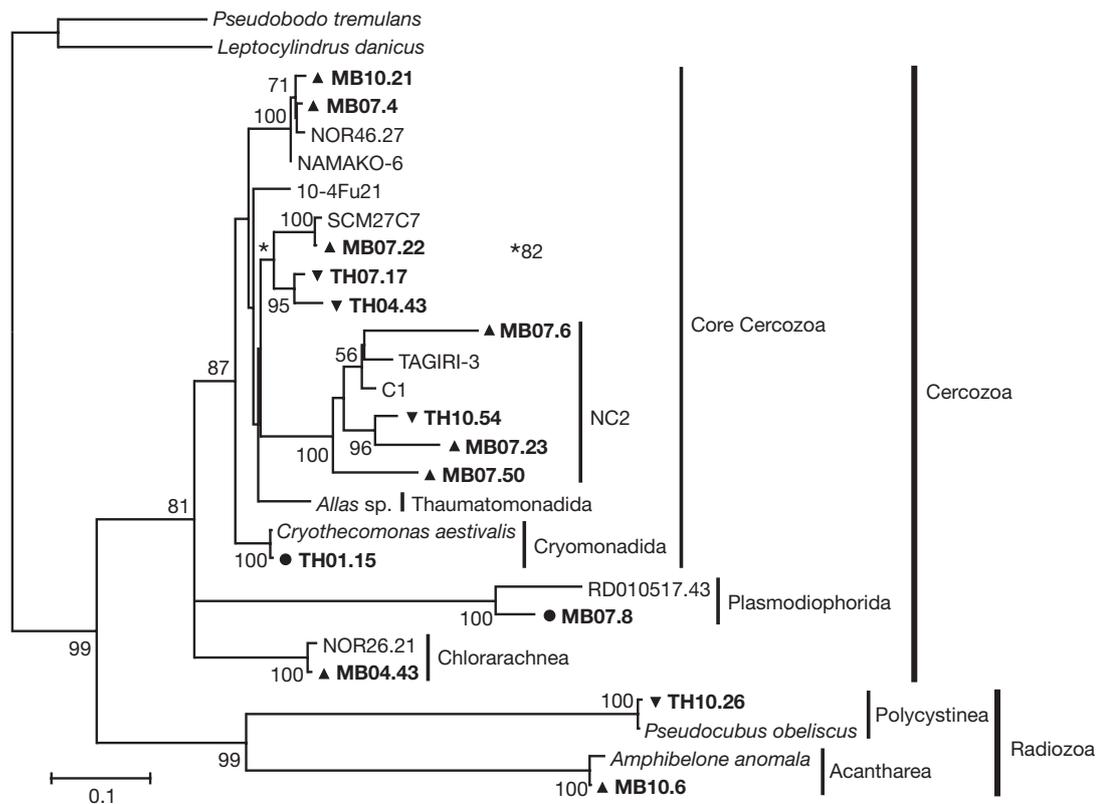


Fig. 6. ML phylogenetic tree of 18S rDNA sequences of Rhizaria, based upon 28 partial sequences of 513 positions after Gblock processing and rooted using 2 stramenopile sequences (*Pseudobodo tremulans* and *Leptocylindrus danicus*). Other definitions as in Fig. 3

nating from secondary endosymbiosis (Bhattacharya & Medlin 1995).

Massana et al. (2004) performed a comparative analysis on the genetic diversity and ecological distributions of marine stramenopiles (MAST) from 37 marine environmental clone libraries and defined 12 clusters. Of the 7 clusters recovered here, MAST-1, -3, -4 and -7 are large clusters commonly found in coastal systems and open seas, MAST-2 appears to be rare but widespread, and MAST-6 and -11 are minor components of marine picoeukaryotic assemblages (Massana et al. 2004, 2006a). MAST-1 has been found to be more pronounced in open oceans than in coastal seas (Massana et al. 2004), but TH01.31 and TH10.43 retrieved from the current study, together with CD8.09 retrieved from Norwegian Sea coast, indicate the potential importance of the cluster in coastal ecosystems. The vast diversity of heterotrophic stramenopiles recovered here suggests their importance in coastal ecosystems.

Rhizaria is a supergroup comprising the phyla Cercozoa, Radiozoa and Foraminifera (Nikolaev et al. 2004). Cercozoa is one of the most diverse and species-rich groups of flagellated protists (Bass & Cavalier-

Smith 2004). After the establishment of the phylum Cercozoa based on molecular data alone (Cavalier-Smith 1998), its composition has been continuously expanded, and its phylogeny and classification have been periodically updated (Bass & Cavalier-Smith 2004). A comprehensive phylum-specific environmental 18S rRNA survey of Cercozoa from 41 libraries revealed at least 9 novel cercozoan clades, several possibly at the level of order or above, suggesting that the diversity of this important group of protists could be much higher than anticipated previously (Bass & Cavalier-Smith 2004). In our study, additional environmental representatives from the previously defined NC2 clade were retrieved (Fig. 6). A potentially novel cluster was also recovered. These findings, together with the fact that 50% of the current cercozoan sequences are <95% similar to GenBank sequences (Table A1), suggest that diversity of Cercozoa could be further expanded.

Members from the 2 newly defined phyla, picobiliphytes (Not et al. 2007) and Telonemia (Shalchian-Tabrizi et al. 2006), were also identified in this study. Picobiliphytes represent a novel picoplanktonic algal group with a phycobiliprotein-containing plastid, pos-

sibly arisen through secondary endosymbiosis (Not et al. 2007). Picobiliphyte sequences have been retrieved from a variety of marine systems, including the European coast, the North Atlantic Ocean and the Arctic Ocean (Not et al. 2007, Cuvelier et al. 2008). To the best of our knowledge, this is the first record of picobiliphyte sequences from the Pacific coast. Clone library and FISH data indicate that picobiliphytes are well represented in polar and cold temperate coastal marine ecosystems (Not et al. 2007, Cuvelier et al. 2008). However, two-thirds of the picobiliphyte clones obtained in this study were present in libraries from July and October when water temperature was $\sim 27^{\circ}\text{C}$ (Table 1). This observation indicates that picobiliphytes may have cosmopolitan distribution in both warm and cold marine waters (Cuvelier et al. 2008).

Telonemia is another new protist phylum of close affinity to chromist lineages (Shalchian-Tabrizi et al. 2006). While early microscopic studies have suggested a worldwide distribution for *Telonema*-like organisms (Lee & Patterson 1998), this study represents the first record of *Telonemia* sequences in the Pacific Ocean.

Ciliates and dinoflagellates are generally considered to be larger than 2 to 3 μm . Their presence in the current clone libraries might be due to flexible cells that could be forced through the 3 μm filter pores, cell breakage during sample collection, sloppy feeding by zooplankton and life history stages of small size or undescribed picoplanktonic species. Nevertheless, reports of these organisms in pico-sized fractions, particularly from coastal environments (e.g. Romari & Vaulot 2004, Medlin et al. 2006, Worden 2006), appear to suggest the potential presence of pico-ciliates and pico-dinoflagellates.

Three OTUs retrieved in this study belonged to ellobiopsids, a group of multinucleated protist parasites of aquatic crustaceans. Ellobiopsids variously have been found to be affiliated with fungi, 'colourless algae' and dinoflagellates, and recently suggested to be within the alveolates (Silberman et al. 2004). However, our analysis using partial sequences shows that they occupy a basal position in the phylogenetic tree. MB07.52 and TH07.19 form an unknown group that is most closely related to clone PD6.12 ($\geq 97\%$) from a polar site in the Norwegian Sea (Massana et al. 2006b). Additional environmental sequences closely affiliated with this group include RA000609.30 from a coastal site in the English Channel in France (Romari & Vaulot 2004) and the Antarctica sequence ANT12.13 (Díez et al. 2001). These 3 GenBank sequences, together with MB07.52 and TH07.19, suggest that the group is robust and widely distributed. However, more detailed studies have to be performed to validate their phylogenetic status.

Spatial and seasonal variations

To study if any spatial variations of picoeukaryotes existed between water samples collected from a semi-enclosed harbour and the adjacent open sea, species composition was compared between water samples collected from Tolo Harbour and Mirs Bay. While alveolates group I, picobiliphytes and cryptophytes were more abundant in the harbour libraries than in the open sea libraries, contribution by dinoflagellates, prasinophytes and cercozoans was higher in the Mirs Bay libraries (Table 3).

The abundance of alveolates groups II and stramenopiles was comparable between both sampling sites, but there seemed to be a differential distribution of some OTUs/clusters constituting the 2 sites. For instance, most current OTUs composing alveolates group II cluster 1 and MAST-4 were collected from the harbour libraries (Figs. 4C & 5). All 3 OTUs composing alveolates group II cluster 5 in this study were also from the harbour libraries (Fig. 4C). These findings suggest that these clusters may be more abundant in enclosed harbours with high nutrient levels than in the open ocean.

Seasonal clone libraries are valuable for formulating and testing hypotheses (Worden 2006). This is particularly applicable to coastal ecosystems, which are prone to large temporal fluctuations induced by episodic events such as tidal currents, intrusion of freshwater and algal blooms. All 6 of the most dominant picoeukaryotic groups recovered in this study appeared in all Tolo Harbour libraries throughout the year, varying only in relative clonal abundance. In contrast, occurrence in Mirs Bay libraries tended to be more sporadic, with all 6 groups present in April, 5 in January and July, and 4 in October (Fig. 2). These observations suggest that seasonal variations in picoeukaryotic composition may be more pronounced in open sea waters than in the more enclosed harbour waters.

Acknowledgements. We thank Y. H. Yung for sample collection and K. C. Cheung for technical assistance. The work was supported by research grants from the Research Committee, The Chinese University of Hong Kong.

LITERATURE CITED

- Bass D, 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
- Bhattacharya D, Medlin L (1995) The phylogeny of plastids: a review based on comparisons of small-subunit ribosomal RNA coding regions. *J Phycol* 31:489–498
- Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17:540–552

- Cavalier-Smith T (1998) A revised six-kingdom system of life. *Biol Rev Camb Philos Soc* 73:203–266
- Chao A, Shen TJ (2003) Program SPADE (Species Prediction And Diversity Estimation). Program and user's guide. <http://chao.stat.nthu.edu.tw>
- Cole JR, Chai B, Marsh TL, Farris RJ, Wang Q, Kulam SA (2003) The ribosomal database project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res* 31: 442–443
- Cuvelier ML, Ortiz A, Kim E, Moehlig H and others (2008) Widespread distribution of a unique marine protistan lineage. *Environ Microbiol* 10:1621–1634
- Diez B, Pedrós-Alió C, Massana R (2001) Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. *Appl Environ Microbiol* 67:2932–2941
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13–15
- Elwood HJ, Olsen GJ, Sogin ML (1985) The small-subunit ribosomal RNA gene sequences from the hypotrichous ciliates *Oxytricha nova* and *Stylonychia pustulata*. *Mol Biol Evol* 2:399–410
- Groissillier A, Massana R, Valentin K, Vaultot D, Guillou L (2006) Genetic diversity and habitats of two enigmatic marine alveolate lineages. *Aquat Microb Ecol* 42: 277–291
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696–704
- Hong Kong Environmental Protection Department (2003) Marine water quality in Hong Kong for 2002. Government Printer, Hong Kong
- Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* 5:150–163
- Lee WJ, Patterson DJ (1998) Diversity and geographic distribution of free-living heterotrophic flagellates — analysis by PRIMER. *Protist* 149:229–243
- López-García P, Rodríguez-Valera F, Pedrós-Alió C, Moreira D (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* 409:603–606
- Lovejoy C, Massana R, Pedrós-Alió C (2006) Diversity and distribution of marine microbial eukaryotes in the Arctic Ocean and adjacent sea. *Appl Environ Microbiol* 72: 3085–3095
- Massana R, Castresana J, Balagué V, Guillou L and others (2004) Phylogenetic and ecological analysis of novel marine stramenopiles. *Appl Environ Microbiol* 70: 3528–3534
- Massana R, Terrado R, Forn I, Lovejoy C, Pedrós-Alió C (2006a) Distribution and abundance of uncultured heterotrophic flagellates in the world oceans. *Environ Microbiol* 8:1515–1522
- Massana R, Guillou L, Terrado R, Forn I, Pedrós-Alió C (2006b) Growth of uncultured heterotrophic flagellates in unamended seawater incubations. *Aquat Microb Ecol* 45: 171–180
- Medlin LK, Metfies K, Mehl H, Wiltshire K, Valentin K (2006) Picoeukaryotic plankton diversity at the Helgoland time series site as assessed by three molecular methods. *Microb Ecol* 52:53–71
- Moon-van der Staay SY, Van der Staay GWM, Guillou L, Vaultot D, Claustre H, Medlin LK (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 SY, De Wachter R, Vaultot D (2001) Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* 409:607–610
- Nikolaev SI, Berney C, Fahrni JF, Bolivar I and others (2004) The twilight of Heliozoa and rise of Rhizaria, an emerging supergroup of amoeboid eukaryotes. *Proc Natl Acad Sci USA* 101:8066–8071
- Not F, Valentin K, Romari K, Lovejoy C and others (2007) Picobiliophytes: a marine picoplanktonic algal group with unknown affinities to other eukaryotes. *Science* 315: 253–254
- Romari K, Vaultot 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
- Shalchian-Tabrizi K, Eikrem W, Klaveness D, Vaultot D and others (2006) Telonemia, a new protist phylum with affinity to chromist lineages. *Proc R Soc Lond B Biol Sci* 273: 1833–1842
- Sherr EB, Sherr BF (2000) Marine microbes. An overview. In: Kirchman DL (ed) *Microbial ecology of the oceans*. Wiley-Liss, New York, p 13–46
- Silberman JD, Collins AG, Gershwin LA, Johnson PJ, Roger AJ (2004) Ellobiopsids of the genus *Thalassomyces* are alveolates. *J Eukaryot Microbiol* 51:246–252
- Wear RG, Thompson GB, Stirling HP (1984) Hydrography, nutrients and plankton in Tolo Harbour, Hong Kong. *Asian Mar Biol* 1:59–75
- Wong CK, Wong CK (2003) HPLC pigment analysis of marine phytoplankton during a red tide occurrence in Tolo Harbour, Hong Kong. *Chemosphere* 52:1633–1640
- Worden AZ (2006) Picoeukaryote diversity in coastal waters of the Pacific Ocean. *Aquat Microb Ecol* 43:165–175
- Yung YK, Wong CK, Bloom MJ, Ogden JA, Chan SCM, Leung Y (1997) Long-term changes in hydrography, nutrients and phytoplankton in Tolo Harbour, Hong Kong. *Hydrobiologia* 352:107–115

Editorial responsibility: Daniel Vaultot, Roscoff, France

*Submitted: September 11, 2007; Accepted: September 9, 2008
Proofs received from author(s): December 9, 2008*