ELSEVIER

FEMS Microbiology Ecology 52 (2005) 79–92

FEMS MICROBIOLOGY Ecology

www.fems-microbiology.org

# Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene

Fei Zhu [a,1], Ramon Massana [b], Fabrice Not [a], Dominique Marie [a], Daniel Vaulot [a,*]

[a] Station Biologique, UMR 7127 CNRS, INSU et Université Pierre et Marie Curie, Place George Teissier, 29680 Roscoff, France
[b] Department de Biologia Marina i Oceanografia, Institut de Ciències del Mar, CMIMA, CSIC, Passeig Marítim de la Barceloneta 37-49, 08003 Barcelona, Catalonia, Spain

Received 30 May 2004; received in revised form 18 October 2004; accepted 22 October 2004

First published online 4 November 2004

**Abstract**

A quantitative PCR (QPCR) assay based on the use of SYBR Green I was developed to assess the abundance of specific groups of picoeukaryotes in marine waters. Six primer sets were designed targeting four different taxonomic levels: domain (Eukaryota), division (Chlorophyta), order (Mamiellales) and genus (*Bathycoccus*, *Micromonas*, and *Ostreococcus*). Reaction conditions were optimized for each primer set which was validated in silico, on agarose gels, and by QPCR against a variety of target and non-target cultures. The approach was tested by estimating gene copy numbers for *Micromonas*, *Bathycoccus*, and *Ostreococcus* in seawater samples to which cultured cells were added in various concentrations. QPCR was then used to determine that rRNA gene (rDNA) copy number varied from one to more than 12,000 in 18 strains of phytoplankton. Finally, QPCR was applied to environmental samples from a Mediterranean Sea coastal site and the results were compared to those obtained by Fluorescent in situ hybridization (FISH). The data obtained demonstrate that Chlorophyta and more specifically Mamiellales were important in these waters, especially during the winter picoplankton bloom. The timing of major abundance peaks of the targeted species was similar by QPCR and FISH. When used in conjunction with other techniques such as FISH or gene clone libraries, QPCR appears as very promising to quickly obtain data on the ecological distribution of important phytoplankton groups. Data interpretation must take into account primer specificity and the varying rRNA gene copy number among eukaryotes.
© 2004 Federation of European Microbiological Societies. Published by Elsevier B.V. All rights reserved.

*Keywords:* Coastal ecosystems; Ecology; Fluorescent in situ hybridization; Picoplankton; Prasinophytes; Quantitative PCR

## 1. Introduction

Photosynthetic picoplankton (i.e. the fraction of the plankton composed of cells less than 2–3 μm in size) plays a critical role in marine ecosystems. Although its contribution to photosynthetic biomass (as estimated for example from chlorophyll *a*) is more important in open ocean oligotrophic waters (typically 80% [1]), it is far from negligible in coastal waters, outside of the bloom season (e.g. up to 87% in the English Channel off Roscoff [2]). While the prokaryotic component of picophytoplankton is reduced to two genera *Synechococcus* and *Prochlorococcus* [3], photosynthetic picoeukaryotes are much more diverse and can belong to a variety of divisions (classes): Chlorophyta (Prasinophyceae), Heterokontophyta (Pelagophyceae, Bolidophyceae), Haptophyta (Prymnesiophyceae) etc. The use of molecular methods and, in particular, the analysis of the 18S rRNA and *psba* genes has allowed to investigate in great details the composition of natural communities [4–6]. In particular, prasinophytes have been shown very

* Corresponding author. Tel.: +33 2 9829 2334; fax: +33 2 9829 2324.
*E-mail address:* vaulot@sb-roscoff.fr (D. Vaulot).
[1] Present address: Department of Pharmaceutical Sciences, University of Maryland, Baltimore, MD 21229, USA.

0168-6496/$22.00 © 2004 Federation of European Microbiological Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.femsec.2004.10.006

recently to be one of the key picophytoplankton group in marine waters [2]. They are composed of at least seven distinct clades, some of which correspond to established orders such as the Mamiellales or the Prasinococcales, while other do not contain any described species [7]. The importance of this group had already been suspected based for example on the relative abundance of chlorophyll *b* in some marine waters [8]. In coastal areas, the order Mamiellales that contains in particular the three typical picoplanktonic species *Micromonas pusilla* [9], *Bathycoccus prasinos* [10], and *Ostreococcus tauri* [11], appears specifically important as revealed by culture studies, optical and electron microscopy observations [12–14], as well as more recently by 18S rDNA analyses [2,7]. Therefore, it has become critical to assess the abundance of members of this group at different spatial and temporal scales in order to estimate their importance and to determine under which conditions they thrive. Since cells from genera such as *Bathycoccus* and *Ostreococcus* are too small to be identified by optical microscopy, the use of antibodies or of molecular markers constitute the only practical approach for their detection. Monoclonal antibodies provide a good choice to detect certain species or ecotypes [15] but they are expansive to produce and they cannot be applied to higher taxonomic levels since their specificity is always narrow. Probes targeted against 18S rRNA are much more flexible since they can be tailored (in theory at least) to target any taxonomic level from the ecotype to the division. Fluorescent in situ hybridization (FISH) with either mono-labeled probes or after amplification [16] can be used to count eukaryotic picoplankton by epifluorescence microscopy [2]. However, this method is quite tedious and, although automation of microscopy analysis [17] or detection by flow cytometry [18] can increase significantly sample throughput, alternative methods are required.

In this context, quantitative PCR (QPCR, see [19] for a review) appears as particularly promising. Its principle consists in amplifying a given gene (in the present case 18S rDNA) with primers specific of one taxonomic group and monitoring product formation in real-time by fluorescence. The later can be induced by a probe labeled with a fluorochrome and a quencher which is released as the probe binds to the product (Taqman approach) or more simply, by a dye, such as SYBR Green I, binding to double stranded DNA as it is formed during the PCR. The number of gene copies in the initial sample is deduced from the number of PCR cycles (cycle threshold or $C_T$) required to cross a certain fluorescence level. The major advantages of this approach are its linearity, its sensitivity, and the speed at which a large number of samples can be processed. Its use in environmental microbiology is rapidly expanding and it has been recently applied to aquatic environments to obtain large scale estimates of the abundance of major bacterial groups [20], of certain important bacterial groups such as *Pseudoalteromonas* [21], and of specific *Synechococcus* ecotypes [22].

In the present paper, we introduce six primer sets to be used with QPCR targeting four different eukaryotic taxonomic levels: domain (Eukaryota), division (Chlorophyta), order (Mamiellales) and genus (*Bathycoccus*, *Micromonas*, and *Ostreococcus*). The primers were first tested on cultures, then used to determine the number of rRNA gene copies in a variety of microalgal strains, and finally applied to coastal samples.

## 2. Materials and methods

### 2.1. Cultures

A variety of cultured strains were chosen to test primer specificity, to serve as templates for the optimization of QPCR assays, and to estimate the number of rDNA copies per genome (Table 1). Marine photosynthetic strains were grown in K media [23] in tissue culture flasks (Sarstedt, Orsay France) at 20 °C, under a light–dark cycle of 14–10 h, at 100 $\mu Ein\, m^{-2} s^{-1}$ light intensity (except for RCC 341 that was grown at 10 $\mu Ein\, m^{-2} s^{-1}$), and harvested in early stationary growth phase. Cells of *Escherichia coli* were cultivated in LB media (Sigma, L'Isle d'Abeau Chesnes France) at 37 °C for 12 h.

### 2.2. Marine samples

Water samples were taken from the Mediterranean Sea off the Spanish coast (Blanes: 41° 40′N, 2° 48′E) from 20 March 2001 to 23 October 2002. Up to 10 L of seawater were collected in sub-surface with a Niskin bottle and transferred through a 200 μm nylon mesh to a black bottle. Samples were pre-filtered through a 3 μm Nuclepore membrane (Whatman International Ltd, Maidstone, England) to separate picoplankton, and subsequently filtered through a Sterivex unit (Millipore, Billerica, MA, USA) with a peristaltic pump. The Sterivex units were directly filled with 1.8 ml lysis buffer (40 mM EDTA, 50 mM Tris–HCl, 0.75 M sucrose) and stored at −80 °C until DNA extraction.

### 2.3. DNA extraction

For cultured strains, DNA was extracted by the CTAB protocol [24] with some modifications. Cultures were harvested by centrifugation for 10 min at 8000*g*. 3% (w/v) CTAB (preheated to 60 °C) was added to the pellet and gently swirled to disperse cells. Filter samples were thawed and then treated. Samples were incubated at 60 °C for 30 min in a water bath with occasional vortexing. Nucleic acids were extracted once with an equal

Table 1
Strains used in this study

| Domain | Class | RCC[a] | Species | Strain name |
|---|---|---|---|---|
| Eukarya | Prasinophyceae | 113 | *Bathycoccus prasinos* | CCMP[b] 1898 |
| | Prasinophyceae | 116 | *Ostreococcus tauri* | OTTH0595 |
| | Prasinophyceae | 614 | *Ostreococcus tauri* | OTTH0595-Genome |
| | Prasinophyceae | 114 | *Micromonas pusilla* | CCMP490 |
| | Prasinophyceae | 450 | *Micromonas pusilla* | CCMP489 |
| | Prasinophyceae | 434 | *Micromonas pusilla* | BL122 |
| | Prasinophyceae | 136 | *Prasinococcus capsulatus* | CCMP1407 |
| | Prasinophyceae | 234 | *Tetraselmis* sp. | MIN 008-15m B |
| | Trebouxiophyceae | 347 | *Nannochloris* | OLI26SB |
| | Chlorophyceae | 1 | *Chlamydomonas concordia* | PLY491 |
| | Cryptophyceae | 20 | *Rhodomonas salina* | CCMP322 |
| | Prymnesiophyceae | 362 | *Emiliania huxleyi* | PROSOPE_115 |
| | Bacillariophyceae | 76 | *Thalassiosira weissflogii* | CCMP1336 |
| | Bacillariophyceae | 436 | *Thalassiosira* sp. | BL_77 |
| | Bacillariophyceae | | *Nitzschia closterium* | ROS97005 |
| | Eustigmatophyceae | 92 | *Nannochloropsis salina* | CCMP527 |
| | Chrysophyceae | 21 | *Ochromonas distigma* | Caen |
| | Pelagophyceae | 100 | *Pelagomonas calceolata* | CCMP1214 |
| | Pelagophyceae | 341 | *Pelagomonas calceolata* | PROSOPE_63 |
| | Dictyochophyceae | 382 | *Mesopedinella arctica* | PROSOPE-2 |
| | Dinophyceae | 89 | *Akashiwo sanguinea* | |
| | Dinophyceae | 88 | *Amphidinium carterae* | CCMP1314 |
| | Dinophyceae | 291 | *Prorocentrum minimum* | |
| | Dinophyceae | 303 | *Prorocentrum nux* | VillF I 50m |
| Bacteria | $\gamma$-Proteobacteria | | *Escherichia coli* | K12 |
| | Cyanophyceae | 27 | *Synechococcus* | WH 7803 |

[a] RCC: Roscoff Culture Collection.
[b] CCMP: Provasoli-Guillard National Center for Culture of Marine Phytoplankton.

volume of chloroform–isoamyl alcohol (48:2). After centrifugation for 10 min at 12,000*g*, the aqueous phase was transferred to a clean Corex tube and 2/3 volume of cold isopropanol was added to precipitate the nucleic acids at room temperature for several hours. Nucleic acids were recovered by centrifugation, washed once with 70% ethanol (v/v) and re-suspended in water prior to storage at −80 °C. For Mediterranean Sea natural samples, the DNA extraction was performed as described by Massana et al. [25], and genomic DNA extraction of *E. coli* was performed as described by Sambrook et al. [26].

### 2.4. Primer design

Eight individual primers (Table 2) were used to target different taxonomic levels within eukaryotes and used as six sets composed of a forward and a reverse primer (Table 3). Design strategy is detailed in Section 3. All primers were optimized using the Primer Express software (PE Biosystems) in order to obtain a theoretical melting temperature around 58–59 °C. The theoretical specificity of the primers (Tables 3 and 4) was verified with the ARB software [27] obtained from http://www.arb-home.de. Probes were tested against a database containing more than 28,000 complete or partial SSU rDNA sequences from both eukaryotes and prokaryotes. We started from the aligned SSU rDNA database released by the ARB team in June 2002 (http://www2.mikro.biologie.tu-muenchen.de/download/ARB/data/ssujun02.arb) and we added more than 1800 new sequences, either publicly available or unpublished.

### 2.5. QPCR assays

Reactions were performed in a final volume of 12.5 µl using SYBR PCR core reagents (Applied Biosystems, Courtaboeuf France) and containing 1.25 µl SYBR Green PCR buffer, 200 µM each of dATP, dCTP, and dGTP, 400 µM dUTP, 0.25 U of AmpErase uracyl *N*-Glycosylase (AmpErase UNG), and 0.05 U of Platinum Taq DNA polymerase (Life Technologies, Cergy Pontoise France). All reactions were performed with optical tubes (Applied Biosystems). First, 2.5 µl of template was delivered into the tubes followed by 10 µl of master mix. The tubes were sealed with optical caps (Applied Biosystems). All reactions were performed with an ABI 5700 sequence detection system (Applied Biosystems) programmed with a soak step of 2 min at 50 °C, allowing AmpErase UNG to hydrolyze PCR amplicons possibly carried over from previous reactions. An enzyme activation step (94 °C, 2 min) followed the initial soak step. Forty cycles of 15 s of denaturation at 94 °C, annealing-extension at 60 °C with the times listed in Table 3, and 25 s of data collection at 77 °C were performed. All data were analyzed using Sequence Detection System v 1.3 software (Applied Biosystems).

Table 2
Primers used in this study

| Name[a] | Sequence (5′- > 3′) | Length (bp) | GC (%) | $T_m$ (°C) | Minimum number of mismatches to non-target groups | References[b] |
|---|---|---|---|---|---|---|
| EUK345f | AAGGAAGGCAGCAGGCG | 17 | 65 | 59 | Four mismatches to prokaryotes | This work |
| EUK499r | CACCAGACTTGCCCTCYAAT | 20 | 55 | 58 | Five or more mismatches to prokaryotes | This work |
| EUK528f | CCGCGGTAATTCCAGCTC | 18 | 61 | 58 | Five or more mismatches to prokaryotes | [53] |
| CHLO02r | CTTCGAGCCCCCAACTTTC | 19 | 58 | 58 | One central mismatch to sequences of Chlorarachniophyceae, Cercozoa and some Apicomplexa | [40] |
| PRAS04r | CAAGCGTAAGCCCGCTTT | 18 | 56 | 58 | Two central mismatches to some Cryptophyceae | [2] |
| BATHY03r | ACCACGATGACTCCATGTCTCA | 22 | 50 | 59 | Five or more mismatches to nontarget groups | This work |
| MICRO04r | CGCGTCCTCTACAGGAAGTTG | 21 | 57 | 59 | Perfect match to *Micromonas* RCC 434. Two mismatches to *Micromonas* CCMP 489 and 490, *Mantionella* and *Mamiella* | [2] |
| OSTRE02r | AGTAACCACGGTGACTAAGTGGC | 23 | 52 | 58 | Four mismatches to *Bathycoccus* | [2] |

[a] f: Forward, r: Reverse.
[b] Reference in which the primer location appeared initially. Primers have been adapted for quantitative PCR constraints.

Table 3
Primer sets used in this study with optimal reaction conditions and linearized plasmid used as a standard

| Set | Target group | Primer forward | Primer reverse | Amplicon size (bp) | Standard | Amplicon $T_m$ (°C) | Primer-dimer $T_m$ (°C) | $MgCl_2$ (mM) | Annealing time (min) |
|---|---|---|---|---|---|---|---|---|---|
| EUK | Eucaryotes | Euk345f | Euk499r | 149 | pMIC | 83 | 80 | 3 | 1 |
| CHL | Chlorophyta | EUK528f | CHLO02r | 410 | pMIC | 81 | 76 | 5 | 2 |
| MAM | Mamiellales | EUK528f | PRAS04r | 205 | pMIC | 82 | 77 | 3 | 1 |
| BAT | *Bathycoccus* | EUK528f | BATHY03r | 171 | pBAT | 83 | 71 | 3 | 1 |
| MIC | *Micromonas* | EUK528f | MICRO04r | 135 | pMIC | 83 | 76 | 3 | 1 |
| OST | *Ostreococcus* | EUK528f | OSTRE02r | 167 | pOST | 85 | 76 | 3 | 1 |

The optimal concentrations for both forward and reverse primers were 400 nM. The annealing-extension and the detection temperatures for all primer sets were 60 and 77 °C, respectively.

### 2.6. Optimization of QPCR assays

#### 2.6.1. Annealing-extension time

The different primer sets amplified products with different size (Table 3) and therefore required different annealing-extension time to amplify the target sequences. We tested different annealing-extension times for each primer set and chose the minimal annealing-extension time that allowed samples to reach $C_T$ in a suitable cycle range (15–35).

#### 2.6.2. Melting curve analysis

Unlike Taqman, SYBR green I binds all double-stranded DNA without specificity. The dye can bind both PCR products and primer-dimers. The latter will affect the accuracy of the results if co-amplified with the PCR products. The dissociation curve from 65 to 95 °C was measured after the last QPCR cycle as detailed in the ABI 5700 software manual and the melting temperature ($T_m$) of both primer-dimers and specific PCR products was obtained (Table 3). In order to suppress fluorescence readings caused by the generation of primer-dimers, we set the temperature of the detection step above the $T_m$ of primer-dimers but approximately 3 °C below the $T_m$ of the specific PCR product.

#### 2.6.3. $MgCl_2$ and primer concentration

In order to obtain the highest amplification efficiency of QPCR assays, $MgCl_2$ concentrations (ranging from 1 to 9 mM) and primers concentrations (ranging from 200 to 1200 nM) were tested for all primer sets. We chose the minimal concentration of $MgCl_2$ and primers that allowed consistent amplification of the templates at different concentrations and reasonably high amplification efficiencies, estimated from the slopes of the standard curves empirically generated by QPCR assays (Table 3).

### 2.7. Primer specificity analysis

The specificity of the primer sets was first checked by standard PCR on a iCycler (Bio-Rad, Marnes-la-Coquette, France) using Promega (Madison, USA) PCR reagents. In a final 50 μl volume, reaction mixtures contained 5 μl of buffer, 3 mM of $MgCl_2$, 50 μM each of dATP, dCTP, dGTP, and dTTP, and 400 nM of primers. One nanogram of genomic DNA extracted from

Table 4
Specificity of primer sets

| | EUK | | | CHL | | | MAM | | | BAT | | | MIC | | | OST | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | Gel | $C_T$ | *n* | Gel | $C_T$ | *n* | Gel | $C_T$ | *n* | Gel | $C_T$ | *n* | Gel | $C_T$ | *n* | Gel | $C_T$ |
| *Bathycoccus prasinos* | 0 | + | 16 | 0 | + | 25 | 0 | + | 19 | 0 | + | 22 | 5+ | – | 40 | 5+ | – | 34 |
| *Micromonas* RCC114 | 0 | + | 16 | 0 | + | 22 | 0 | + | 19 | 5+ | – | 36 | 2 | + | 23 | 5+ | – | 35 |
| *Micromonas* RCC450 | 0 | + | 18 | 0 | + | 22 | 0 | + | 16 | 5+ | – | 35 | 2 | + | 25 | 5+ | – | 38 |
| *Micromonas* RCC434 | 0 | + | 21 | 0 | + | 25 | 0 | + | 17 | 5+ | – | 35 | 0 | + | 18 | 5+ | – | 40 |
| *Ostreococcus tauri* | 0 | + | 19 | 0 | + | 28 | 0 | + | 20 | 5+ | – | 36 | 5+ | – | 40 | 0 | + | 22 |
| *Prasinococcus capsulatus* | 0 | + | 15 | 0 | + | 27 | 5+ | – | 32 | 5+ | – | 35 | 5+ | – | 40 | 5+ | – | 40 |
| *Nannochloris* sp. | 0 | + | 16 | 0 | + | 25 | 4 | – | 40 | 5+ | – | 40 | 5+ | – | 40 | 5+ | – | 40 |
| *Rhodomonas salina* | 0 | + | 20 | 2 | – | 39 | 2 | +/– | 25 | 5+ | – | 40 | 5+ | – | 39 | 5+ | – | 40 |
| *Emiliania huxleyi* | 0 | + | 15 | 2 | – | 35 | 5+ | – | 32 | 5+ | – | 36 | 5+ | – | 40 | 5+ | – | 40 |
| *Thalassiosira weissflogii* | 0 | + | 19 | 2 | – | 40 | 5+ | – | 35 | 5+ | – | 35 | 5+ | – | 40 | 5+ | – | 40 |
| *Nannochloropsis salina* | 0 | + | 26 | 2 | – | 30 | 5+ | – | 31 | 5+ | – | 40 | 5+ | – | 40 | 5+ | – | 40 |
| *Amphidinium carterae* | 0 | + | 18 | 2 | – | 30 | 5+ | – | 35 | 5+ | – | 40 | 5+ | – | 40 | 5+ | – | 40 |
| *Ochromonas distigma* | 0 | + | 17 | 2 | – | 39 | 5+ | – | 33 | 5+ | – | 40 | 5+ | – | 40 | 5+ | – | 40 |
| *Pelagomonas calceolata* | 0 | + | 18 | 2 | – | 29 | 5+ | – | 33 | 5+ | – | 40 | 5+ | – | 40 | 5+ | – | 40 |
| *Escherichia coli* | 5+ | – | 40 | 5+ | – | 40 | 5+ | – | 27 | 5+ | – | 40 | 5+ | – | 40 | 5+ | – | 40 |
| *Synechococcus* sp. | 5+ | – | 40 | 5+ | – | 40 | 5+ | – | 28 | 5+ | – | 40 | 5+ | – | 40 | 5+ | – | 40 |
| Maximum $C_T$ for target organisms | | | 26 | | | 28 | | | 20 | | | 22 | | | 18 | | | 22 |
| Minimum $C_T$ for non-target organisms | | | 40 | | | 29 | | | 25 | | | 35 | | | 39 | | | 34 |

The six primer sets (Table 3) were tested against a range of cultures (Table 1). For each primer set, primer specificity was tested in silico using the ARB software and the number of mismatches to the 18S rDNA sequence of the species determined (first column, *n*; 5+: five or more mismatches). The second column (Gel) provides a binary result from conventional PCR based on whether a band was detected by agarose gel electrophoresis (+ means that a band is clearly present, – means that no band is observed and +/– means that a faint band is visible). The third column (CT) indicates the threshold cycle (fluorescence level = 0.3) for quantitative PCR using 1 ng genomic DNA of each organism (see Section 2). The difference between the maximum number of cycles necessary to detect a target organism and the minimum number of cycles to detect a non-target (last two lines) provides an indication of the specificity of the primer set.

the organisms listed in Table 4 was used as template. The following PCR steps were performed: a soak step of 5 min at 94 °C, 35 cycles of 30 s of denaturation at 94 °C, 30 s of annealing with gradient temperature from 55 to 65 °C, and 1 min of extension at 72 °C. Ten microlitres of the reaction product were run in 2% (w/v) agarose gel stained with ethidium bromide. Specificity was further verified by QPCR with 1 ng of template following the optimized protocol defined above. A fixed threshold (0.3) was used for $C_T$ computation in the Sequence Detection System software.

### 2.8. Calibration with plasmids

Three linear plasmids from cloned 18S rDNA for *Micromonas* RCC 114 (pMIC), *Bathycoccus* RCC 113 (pBAT), and *Ostreococcus* RCC 614 (pOST) were constructed to be used as standards for QPCR assays. The full 18S rRNA gene from representative strains for each genus was amplified with universal eukaryotic primers [28] and cloned using the TopoTA cloning kit (Invitrogen, Cergy Pontoise, France), according the instruction of the kit. Plasmids were extracted using Flexiprep kit (Amersham, Orsay, France). Linearized plasmids were produced from supercoiled plasmid by digestion with the restriction endonuclease *NotI* (New England Biolabs Beverly, MA, USA) according to the manufacturer's protocol, and purified with phenol–chloroform extraction. The concentration of genomic DNA from linear plasmids was measured fluorometrically with SYBR Green I (Molecular Probes, Leiden, The Netherlands) using $\lambda$ phage DNA as a standard according to the manufacturer's protocol. The number of copies in the standards was calculated using the following formula:

$$\text{molecules}/\mu\text{l} = a/(5681 \times 660) \times 6.022 \times 10^{23},$$

where $a$ is the plasmid DNA concentration (g/µl), 5681 is the plasmid length, including the vector (3931 bp) and inserted PCR fragment (average 1750 bp), 660 is the average molecular weight of one base pair, and $6.022 \times 10^{23}$ is the molar constant.

The pMIC plasmid was used for optimization of reaction conditions and calibration for all primer sets except for BAT and OST, for which pBAT and pOST were used respectively (Table 3). We observed that the use of pMIC, pBAT and pOST for calibration to analyze the same environmental sample with primer set MAM yielded different estimates. In order to be able to compare the concentrations of the different genera estimated by QPCR, we corrected the copy numbers of pBAT and pOST relative to pMIC (see Section 3 for details). This correction only concerned the BAT and OST primer sets.

### 2.9. Artificial mixture experiment

Cells of *M. pusilla* (RCC 114), *B. prasinos* (RCC 113), and *O. tauri* (RCC 614) were cultivated as described above for six days. Cell counts were obtained with a FACSsort flow cytometer (Becton Dickinson), following the protocol described by Marie et al. [29]. Known numbers of the cells from the three species were added to 1 L seawater samples collected off Roscoff on 1st July 2003. Samples filtration and DNA extraction were performed as described before. Estimates of 18S rRNA gene concentration of *M. pusilla*, *B. prasinos* and *O. tauri* were detected by primer set MIC, BAT, and OST, respectively, using pMIC, pBAT, and pOST as standards, respectively. Concentrations calibrated with the latter two were corrected as explained above.

### 2.10. Number of rRNA gene copies per genome

The number of rRNA gene copies per genome was estimated in 18 strains. Cells of the selected strains were cultivated for four days and cell numbers were counted by flow cytometry as described above. Two slightly different protocols were used. For a first set of strains, (RCC 1, 76, 92, 113, 114, 116, 362, 434, 436) different volumes of culture (1, 4, 7, 10, 20 ml, respectively) were filtered onto a 47 mm Supor-450 0.45 μm filter (Pall Gelman, New York, USA) and DNA was extracted from each filter using the CTAB protocol as described above. Genomic DNA was resuspended in 100 μl water and diluted 100-fold for QPCR assay. For a second group of strains (RCC 20, 81, 88, 89, 100, 234, 291, 303, 382), a more efficient protocol was used. A single volume of culture (8 ml) was filtered and extracted. Then four dilutions of the extracted DNA (corresponding to 1-, 10-, 100-, 1000-fold dilutions) were used for QPCR. The number of rDNA copies per sample were estimated with the EUK primer set using pMIC as a standard and plotted against the estimated cell number in each DNA sample. The slope of the regression provided an estimate of the number of rDNA gene copies for this strain.

### 2.11. QPCR analysis of natural samples

Four different plasmid concentrations (from 20 to 20,000 copies $\mu l^{-1}$) were used to construct a standard curve. Environmental samples were diluted 10-fold for QPCR assays and run in triplicate for each primer set. 2.5 μl DNA samples (corresponding to 200–1900 pg DNA per reaction tube) were used for each reaction. The original concentration of targeted 18S rDNA (copies $ml^{-1}$) was computed as follows:

$$((12.5 \times c/2.5) \times 10b)/a,$$

where 12.5 is the volume of QPCR reaction solution (μl), $c$ is the 18S rDNA concentration estimated by QPCR (copies $\mu l^{-1}$), 2.5 is the volume of the DNA sample in the reaction (μl), 10 corresponds to the sample dilution, $b$ is the volume into which the seawater DNA was resuspended initially (μl), and $a$ is the volume of seawater (ml) from which DNA was extracted.

### 2.12. FISH analysis of natural samples

FISH was used to determine the cell abundance of specific groups with the following probes: EUK1209R + CHLO01 + NCHLO01 (picoeucaryotes), CHLO02 (Chlorophyta), MICRO01 (*Micromonas*), BATHY01 (*Bathycoccus*), and OSTREO01 (*Ostreococcus*) as described by Not et al. [2].

## 3. Results

### 3.1. QPCR primer design, optimization and validation

In this work, we opted for the SYBR Green I approach [19], instead of the TaqMan approach, since the former requires only two primers for each target and therefore relies only on two conserved sites. For all targeted groups except eukaryotes (i.e. Chlorophyta, Mamiellales, and the three genera), we used the same primer on the 5′ side adapted from the internal sequencing primer Euk528f (Table 2). On the 3′ end, we modified slightly existing FISH probes (Table 2). For eukaryotes, no universal primer could be designed downward of Euk528f; therefore we designed two new primers upstream of the region targeted for the other groups (Table 2). In silico analysis of primer specificity (Tables 2 and 4) using the "probe match" function of the ARB software against a large database of SSU rDNA sequences (see Section 2.4), indicated that some primers were highly specific of the group they targeted and had at least four or five mismatches to non-target organisms (OSTRE02r or BATHY03r), while other displayed lower specificity (e.g. CHLO02r). Reaction conditions were then optimized for the six primer sets obtained by varying $MgCl_2$ and primer concentration, annealing time, and temperature for fluorescence reading (Table 3). The optimal $MgCl_2$ concentration fell into a range (3–5 mM) similar to what has been used in previous studies using QPCR to assess bacteria abundance in marine samples [30,31].

For all primer sets, except EUK, the $T_m$ of specific PCR products was 5 °C or more above that of primer dimers, and in excess of 80 °C (Table 3). When the primer sets were used on environmental samples (see below), the melting curves became complex, because of the sequence diversity of the amplicons induced complex melting curves, as observed previously [32]. For the EUK primer set, the difference between the $T_m$ of the primer dimers and that of the PCR products was only

3 °C. The best compromise appeared to perform detection for all primer sets after the annealing-extension step of the cycling program at 77 °C (25 s), although in the case of the EUK primer set, it was impossible to exclude completely primer dimer contribution from the PCR product fluorescence signal.

The next step was to test primer sets on a large number of cultured organisms representing a wide taxonomic range. Non-target cultures displayed from one to more than five mismatches to specific primer sets (Table 4). When tested by conventional PCR, primer sets EUK, CHL, MIC, BAT, and OST were specific of their target groups, without cross-amplification with any of the tested templates (Table 4). Primer set MAM targeting the order Mamiellales (prasinophytes) displayed faint cross-reactivity with one non-target strains, RCC 20 (*Rhodomonas salina*) which had only two mismatches to PRAS04r.

The use of QPCR allowed us to estimate more precisely the extent of cross-reaction between primer sets and cultured strains (Table 4). For a given primer set, there was considerable response variation for targeted strains. As an example, $C_T$ (the first PCR cycle for which fluorescence crosses a pre-determined level) varied from 16 to 26 for eukaryotes with primer set EUK. However, non-target strains, i.e. prokaryotes, displayed a much lower $C_T$ (40). The difference between the maximum $C_T$ for target strains and the minimum $C_T$ for non-target strains gives an idea of the specificity of the primer set (last two lines in Table 4). According to this criterion, primer set EUK and those targeting genera were clearly highly specific, in contrast to primer sets targeting intermediate taxonomic levels. Interestingly, the reverse primer MICRO04r matches perfectly the 18S rDNA sequence of RCC434 but has two mismatches with the two other *Micromonas* strains (RCC114, RCC450) that belong to different *Micromonas* clades [7]. This is evidenced by the slightly higher $C_T$ observed for the latter two. However, the MIC primer set still allowed to clearly distinguish between *Micromonas* and the other non-target species tested (Table 4).

### 3.2. Quantification aspects

In contrast to methods such as FISH that provide unequivocal data in terms of cell abundance, QPCR is inherently a relative method that requires internal references. Although it is possible to use genomic DNA from target organisms, a better choice appears to be linearized plasmids containing cloned 18S rRNA genes [31]. We used three different plasmids (pMIC, *Micromonas*; pBAT, *Bathycoccus*; pOST, *Ostreococcus*). pMIC was used to calibrate all primer sets except BAT and OST, for which we used pBAT and pOST respectively. The gene concentration in the plasmid solution was estimated by spectrofluorometry (see Section 2.8). We observed that identical estimated concentrations of pMIC, pBAT and pOST used with primer sets targeting the three plasmids (MAM and EUK) responded differently. Therefore, gene concentrations estimated based on pBAT (primer set BAT) and pOST (primer set OST) were corrected using pMIC as the reference, i.e. multiplied by 0.32 and 0.56, respectively.

Primer sets MIC, BAT, and OST were used to detect cells of cultured strains added to a natural seawater sample (Fig. 1). The number of estimated 18S rDNA copies was linearly related to the number of cells added to the sample, demonstrating that QPCR could be reliably used in natural samples. It should be noted that the slope of the relationship between gene copies and cell numbers is larger than one, reflecting the fact that eukaryote genomes contain in general more than one copy of the rDNA gene. In order to better interpret field data, it is critical to know how this number varies. This is difficult to achieve by conventional means and QPCR offers a practical solution [33]. We selected 18 algal strains representing 17 species belonging to different phylogenetic groups and ranging in size from 0.8 to 60 μm, for which we estimated the number of 18S rDNA copies using the EUK primer set. Copy number ranged from 1 in the picoplanktonic species *Nannochloropsis salina* (Eustigmatophyceae) to more than 12,000 for the large dinoflagellate *Akashiwo sanguinea* and was highly correlated with cell length (Fig. 2).

### 3.3. QPCR analysis of natural samples

The six primer sets were then used to estimate rDNA gene copy numbers in a temporal series of picoplankton (<3 μm) samples from the coastal Mediterranean Sea for which clone libraries data were available [34]. We also performed FISH analyses on the same samples. A clear annual cycle is observed with a picoplankton bloom taking place in late winter. This picoplankton bloom is much longer than typical diatom blooms that last only a couple of weeks [35]. This picoplankton bloom is dominated by Chlorophyta and more specifically by the two Mamiellales genera *Micromonas* and *Bathycoccus* (Fig. 3). Average eukaryotes 18S rRNA gene copy numbers from QPCR and cell abundance from FISH fell into very similar ranges. There was a particularly good match between average values for eukaryotes around 7000 cell $ml^{-1}$ for FISH or copies $ml^{-1}$ for QPCR (Table 5). A good correspondence was also observed for the contribution of each group to eukaryotes (%) except for *Micromonas*, for which QPCR provided 4-fold lower estimates of their contribution than FISH. A detailed analysis of the time series (Fig. 3) demonstrated that QPCR estimates of the different groups were very coherent with FISH

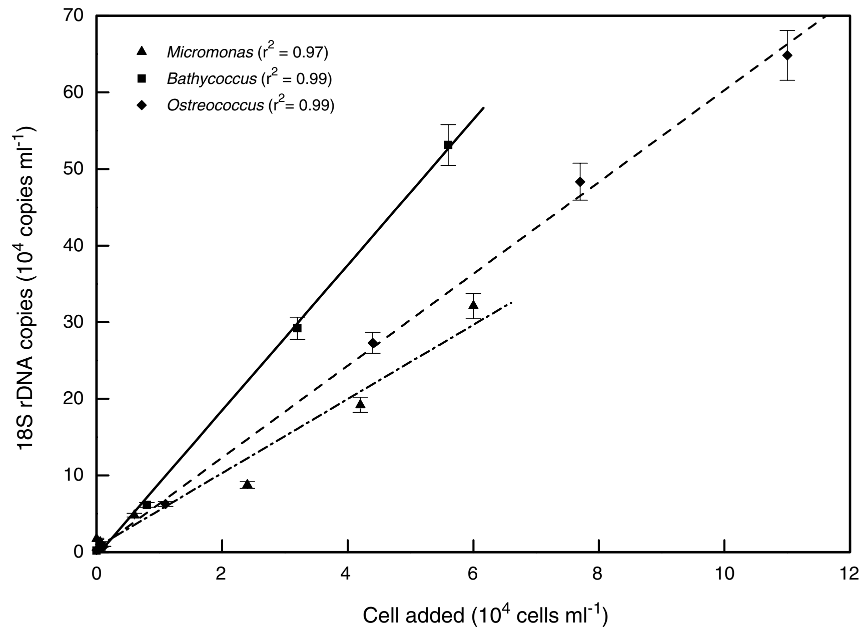


Fig. 1. Cell addition experiment. Known concentration of cultures from *Micromonas pusilla*, *Bathycoccus prasinos* and *Ostreococcus tauri* were added to Roscoff seawater. After DNA collection and extraction, 18S rDNA copy numbers were measured in each sample by QPCR using primer sets MIC, BAT, and OST, respectively. The $x$ axis corresponds to numbers of added cells estimated by flow cytometry and the $y$ axis to 18S rDNA copies measured by QPCR. Straight lines represent regression for each species.

data. For example, the picoeukaryote bloom occurring in February is clearly visible with QPCR. Three of the major peaks of Chlorophyta contribution (September 2001, April and September 2002) corresponded to peak contribution observed by FISH (Fig. 3(b)). Although for both Mamiellales and *Micromonas* estimates of QPCR and FISH fell into different ranges, seasonal variation was quite similar for both estimates (Fig. 3(c) and (e)). Finally, for the two primer sets targeting the genera *Bathycoccus* and *Ostreococcus*, the timing of the maximal contribution (March 2002 for *Bathycoccus* and October 2001 for *Ostreococcus*) were identical for QPCR and FISH (Fig. 3(d) and (f)).

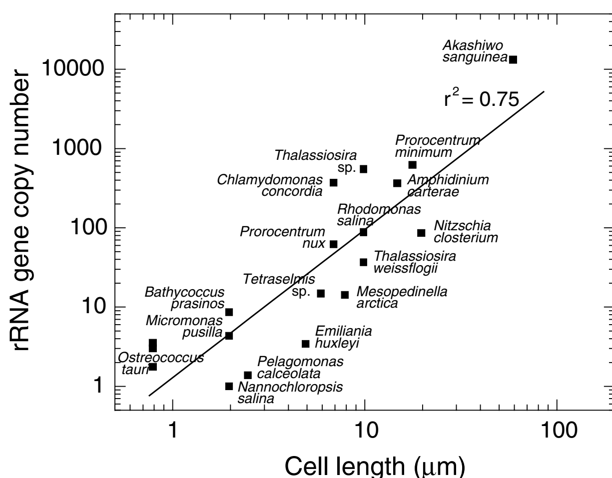


Fig. 2. Correlation between rDNA copy number estimated by QPCR and cell length from 18 strains of phytoplankton.

## 4. Discussion

### 4.1. Primers targeting taxonomic groups

In order to address the ecology of groups such as the prasinophytes, it is important to be able to determine the distribution of different taxonomic entities. At a broad level, one may want, for example, to distinguish between ecosystems dominated by green vs. brown algal classes (e.g. prasinophytes vs. prymnesiophytes), since the groups have deeply divergent phylogenetic histories as well as different physiological responses [36]. However, it is also critical to be able to analyze the distribution of certain genera or species. A typical example will be *Micromonas* which seems to dominate eukaryotic picoplankton in temperate nutrient-rich waters [2]. At an even finer level, one may want to analyze what are the ecological niches of specific ecotypes: recent physiological, genetic and genomic data on genera such as *Prochlorococcus* or *Ostreococcus* demonstrate that many ubiquitous picoplankton species display considerable ecotypic variability [37–39]. While it is relatively easy to design probes and primers targeting the domain level or the species level, this is considerably more tricky for intermediate levels such as those of the division (e.g. Chlorophyta) or class (e.g. Prymensiophyceae). Only a

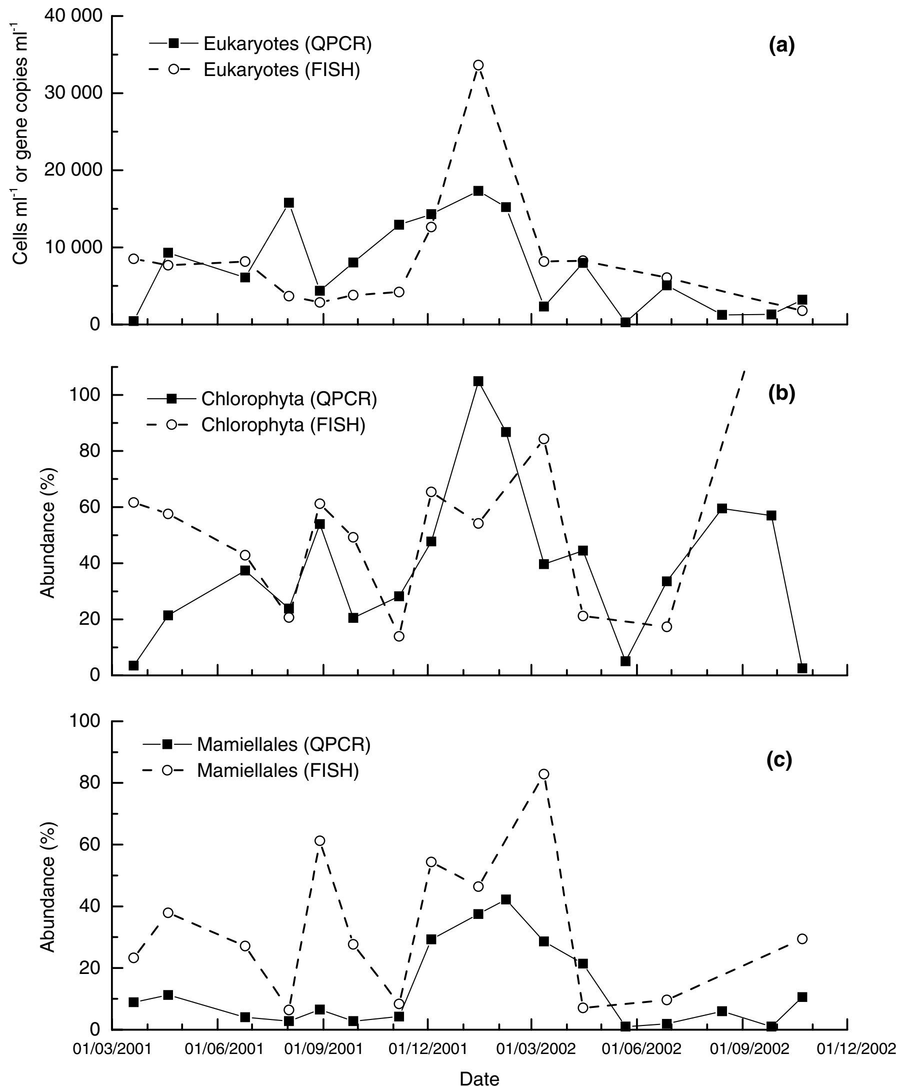


Fig. 3. Temporal change of picoplankton at Blanes (Mediterranean Sea) obtained by QPCR and FISH from 20 March 2001 to 23 November 2002. (a) Picoeucaryotes (cell $ml^{-1}$ for FISH or copies $ml^{-1}$ for QPCR); (b) Chlorophyta (percentage of picoeukaryotes); (c) Mamiellales (percentage of picoeukaryotes); (d) *Bathycoccus* (percentage of picoeukaryotes); (e) *Micromonas* (percentage of picoeukaryotes); (f) *Ostreococcus* (percentage of picoeukaryotes).

few regions in the 18S rRNA gene are suitable for such probes and these probes have only few mismatches (one or two) to many non-target sequences. For example, only two somewhat imperfect FISH probes have been designed for Chlorophyta: CHLO01 and CHLO02 [40,41]. QPCR adds more constraints than FISH since two primers are needed that must be located closely enough since PCR fragments in excess of 400 base pairs result in poor amplification efficiency [19]. Both primers must have nearly identical melting temperature ($T_m$) around 60 °C. In our specific case, it was impossible to find two specific primers meeting these conditions for some of the target groups. Therefore, one universal primer was used on the 5′ end and specificity was achieved through the reverse primer on the 3′ end. Among the six primer sets used in this study, three appear very satisfactory and highly specific of their target (EUK, BAT, OST), while those targeting intermediary levels (CHL, MAM) have more difficulty to distinguish targets from non-target groups and corresponding data should be evaluated more carefully. Primer set MIC is a bit special since the reverse primer does not match exactly two of the three existing *Micromonas* clades [7], while in practice it allows a clear distinction between target and non-target organisms (Table 4).

### 4.2. Quantification issues

With cell-based techniques such as FISH, quantification issues are straightforward since it is easy to determine whether a cell is labeled or not, especially when amplification techniques such as TSA are used [42].

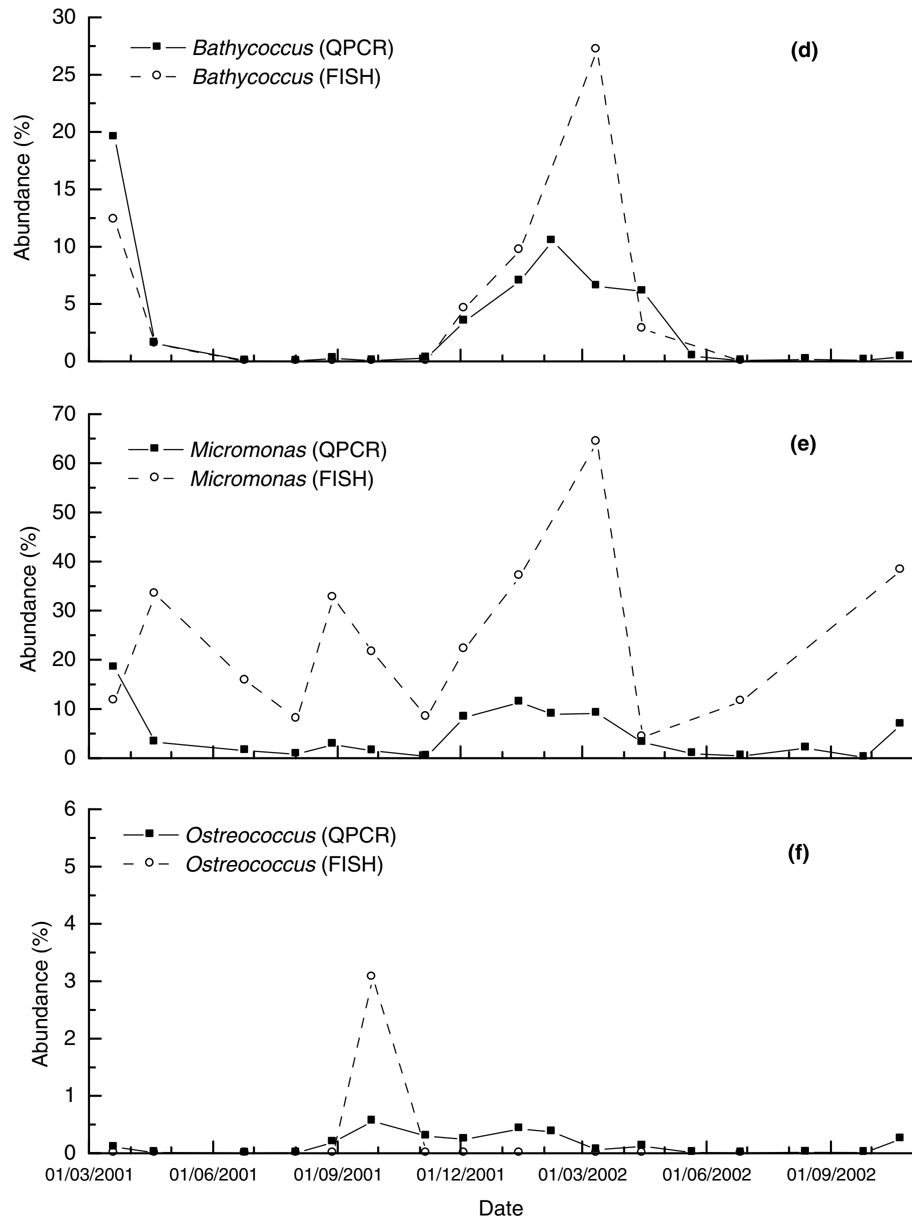


Fig. 3 (*continued*)

However, with QPCR a number of points must be considered when acquiring and interpreting the data.

QPCR is performed on DNA extracted from samples usually filtered on a porous membrane. The extraction procedure must be performed under very uniform conditions in order to achieve maximum efficiency and recovery. An alternative approach could be to concentrate of microbial cells by centrifugation and to resuspend them

Table 5
Comparison of average picoplankton abundance obtained by FISH and QPCR expressed as percentage of total eukaryotes at a Mediterranean Sea coastal site (Blanes) between 20 March 2001 and 23 October 2002

| | Eukaryotes (cell $ml^{-1}$) or (copies $ml^{-1}$)[a] | Chlorophyta (% of euks) | Mamiellales (% of euks) | *Bathycoccus* (% of euks) | *Micromonas* (% of euks) | *Ostreococcus* (% of euks) |
|---|---|---|---|---|---|---|
| FISH | 7714 | 52.0 | 28.0 | 4.2 | 21.0 | 0.4 |
| QPCR | 7357 | 49.2 | 18.2 | 3.4 | 4.6 | 0.2 |

[a] FISH data are expressed as cell $ml^{-1}$ and QPCR data as copies $ml^{-1}$.

in a lysis buffer [43]. However picoplankton can be difficult to centrifuge, although procedures reducing cell loss have been developed recently [18]. With respect to calibration, it would be best to use an internal standard (for example adding a known amount of target cells to the sample, see Fig. 1). However, this is not really feasible in practice when targeting very broad taxonomic levels and working on natural samples since standards would have to be added at sampling time. In our case, we followed the approach of Suzuki et al. [31] using linearized plasmids containing cloned 18S rDNA genes. The advantages of this approach are that standards are easily produced in unlimited amounts and can be generated from genetic library clones for uncultivated taxa, that are very common in picoplankton [4]. Using plasmids (which contain a single copy of the gene) also overcomes the problem of multiple rDNA copies per genomes that would occur if one were using genomic DNA as a standard. One problem we encountered was that plasmids containing genes from three different organisms (*Micromonas*, *Ostreococcus*, *Bathycoccus*) gave different results when used to calibrate data acquired with the same primer set (in a range from about 1–3) probably due to errors in plasmid quantification or maybe PCR efficiency. As a consequence, field data were corrected using the *Micromonas* plasmid as a reference.

A major complication with PCR compared to cell-based approaches is that the rDNA gene occurs in multiple copies. Although this problem is present with prokaryotes for which copy number varies between 1 (for example in the typical photosynthetic picoplankton genus *Prochlorococcus* [44]) and 13 (for example in *Bacillus cereus* [45]), this range is much wider in eukaryotes. For example, Prokopowich et al. [46] reports a range between 35 and 19,300 in animals and between 150 and more than 25,000 in plants. In protists and especially photosynthetic ones, the number of copies can be very small. For example in the unicellular alga *Cyanidioschyzon merolae* only three copies of the rrn operon are present on different loci with no repetition [47] while in the apicomplexan *Cryptosporidium parvum*, five copies are found [48]. The range of rDNA copies in 18 strains of microalgae that we estimated by QPCR (Fig. 2) agrees with previous data [46] but is wider since it expands all the way down to one copy for *Nannochloropsis*. The number of rDNA copies in *Ostreococcus*, which was used as a control in each PCR plate assayed, was found to vary between 1.7 and 3.4 (Fig. 2), which is coherent with an estimate of four copies arranged in a single cluster based on genome sequencing (Moreau, H., personal communication). Our estimates for nanoplanktonic dinoflagellates are also coherent with values around 1000 estimated for *Alexandrium minutum* [43]. Large dinoflagellates, such as *A. sanguinea*, however have rRNA gene copy numbers that reach the highest values observed for plants. The good correlation between rDNA copies and size (Fig. 2) is not surprising. On the one hand, it is well-known that genome size (or DNA content) is related to size (the so-called C-value paradox [49]) as demonstrated for phytoplankton [50]. On the other hand, rDNA copy number appears to be correlated to genome size [46]. The correlation we obtained is somewhat tighter than for plants and animals [46].

For the prasinophyte genera targeted in the present paper, copy numbers were equal to 8.27, $4.09 \pm 0.11$ ($n = 2$), $2.7 \pm 0.9$ ($n = 3$) for *Bathycoccus*, *Micromonas* and *Ostreococcus*, respectively, while for the other two picoplanktonic species examined, *Pelagomonas* and *Nannochloropsis*, it was equal to 1.33 and 0.96, respectively. Therefore, the range for picoplankton is quite small and similar to what is found in prokaryotes. In fact, the good correlation between rDNA gene copies and cell size suggests that QPCR data could be used as a better proxy for biomass than FISH data that only provides cell abundance estimates. Another way to circumvent the problem of multiple rRNA gene copies would be to start from rRNA and to perform a reverse-transcription step prior to QPCR [51], since rRNA levels are representative of cell activity [52].

### 4.3. Application of QPCR to natural samples

Three caveats must always be kept in mind when interpreting QPCR estimates of natural samples. First, QPCR is exponential by nature and therefore has an inherent precision of 1.4-fold ($2^{0.5}$, assuming that one can detect changes corresponding to half a PCR cycle). Second, as discussed above, the issue of rDNA copy number is also critical. For example, disrupted cells of larger protists or gametes of multicellular organisms passing through the pre-filter used to separate picoplankton (usually 3 μm) could seriously bias estimates, since a single cell would have as much impact as 1000 or 10,000 picoplankton cells. Third, the specificity of the primer sets might not be perfect and therefore, some non-target species may be included in a given estimate.

When applied to an annual series from a coastal site characterized by a strong seasonal cycle, some of the major features of the picoplankton community were very well recovered. Absolute QPCR estimates and FISH abundance agreed very well for eukaryotes (Table 5). This was in fact quite unexpected because of multiple bias possible in DNA extraction and QPCR calibration. The range observed of contributions for the other taxonomic groups also matched well (Table 5). In particular, taxonomic ranking was clearly respected: estimates of QPCR rDNA copy increased logically from genera to order to division and up to domain. Temporal variation was also very well recovered for most of the taxa. FISH and QPCR data agreed extremely well for *Bathycoccus*

with two recurrent peak contribution in March 2001 and February 2002. Similarly, the increased importance of Mamiellales in May and October of 2001 and in February and November of 2002 is also seen in the QPCR data.

However some discrepancies also appeared. The sum of the QPCR estimates for the three genera *Ostreococcus*, *Bathycoccus* and *Micromonas* only amounts to 40% of the estimated Mamiellales contribution (Table 5), suggesting the existence of other Mamiellales genera in the samples. However, no other Mamiellales sequences, besides those of these three genera, were recovered from four clone libraries analyzed at the Blanes site [7]. In the same way, the difference between the Mamiellales and Chlorophyta contributions (that should correspond to Chlorophyta belonging to other prasinophyte orders or to other Chlorophyta classes such as the Trebouxiophyceae) disagrees with the clone library data in which very few Chlorophyta besides the three genera mentioned above have been found [7]. These discrepancies could be explained by the lower specificity of the MAM and CHL primer sets, which may amplify some non-targeted groups and therefore overestimate the target groups. In fact, clone libraries constructed using primers EUK328f and CHLO02r (Viprey, M., personal communication) yielded besides Chlorophyta sequences, other sequences belonging to the Chrorarachniophyceae or even to the alveolates. It should be noted that the former group has only one mismatch with CHLO02r (Table 2). Another discrepancy is that the relative contribution of the *Micromonas* genus is much lower when estimated by QPCR than with FISH. In this case, primer specificity cannot be suspected, and the explanation is probably to seek in the effect of rDNA copy numbers. Clone libraries from the Blanes sites were rich, particularly, in sequences from uncultivated alveolate groups [34] which may correspond to organisms that have large genomes and therefore a high number of rDNA copies. Their strong contribution to the total eukaryotic rDNA pool would artificially decrease the relative the contribution of genera such as *Micromonas*.

Once primer and reaction conditions have been optimized, QPCR appears as a very promising technique to map the distribution of plankton taxonomic groups at large temporal and spatial scales, especially because many samples that can be processed in a short time. It will be best applied to monitor specific picoplankton species since the range of picoplankton rRNA gene copy number is quite restricted and it is relatively easy to design primer sets that are highly specific of narrow taxonomic groups. However in more difficult cases, i.e. when targeting broader taxonomic groups or when looking at a wider size range, the application of QPCR may still yield invaluable information that cannot be obtained by any other available technique. For example in the present study, we demonstrated the importance of Mamiellales during the winter Mediterranean Sea plankton bloom, extending the observations recently reported in the English Channel [2].

In conclusion, it should be emphasized that we will only gain a deep understanding of marine ecosystems by combining as many techniques as possible and QPCR should not be used alone but in conjunction with other molecular biology approaches such as, FISH, clone libraries, or DGGE, as well as with more classical tools such as pigment analyses.

## Acknowledgements

The work was supported by the following programs: PICODIV (EVK3-CT-1999-00021), BIOSOPE (CNRS INSUE), PICMANCH (Région Bretagne), GenoMer (Région Bretagne et Département du Finistère). FZ and FN benefited from fellowships from the French Ministry of Education and Research.

## References

[1] Ishizaka, J., Harada, K., Ishikawa, K., Kiyosawa, H., Furusawa, H., Watanabe, Y., Ishida, H., Suzuki, K., Handa, N. and Takahashi, M. (1997) Size and taxonomic plankton community structure and carbon flow at the equator, 175 °E during 1990–1994. Deep-Sea Res. II 44, 1927–1949.

[2] 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.

[3] Partensky, F., Blanchot, J. and Vaulot, D. (1999) Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: a review In: Marine Cyanobacteria (Charpy, L. and Larkum, A.W.D., Eds.), Numéro spécial, 19, pp. 457–475. Bulletin de l'Institut Océanographique de Monaco.

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

[5] Díez, B., Pedrós-Alió, C. and 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.

[6] Zeidner, G., Preston, C.M., Delong, E.F., Massana, R., Post, A.F., Scanlan, D.J. and Beja, O. (2003) Molecular diversity among marine picophytoplankton as revealed by *psbA* analyses. Environ. Microbiol. 5, 212–216.

[7] Guillou, L., Eikrem, W., Chrétiennot-Dinet, M.J., Le Gall, F., Massana, R., Romari, K., Pedrós-Alió, C. and Vaulot, D. (2004) Diversity of picoplanktonic prasinophytes assessed by direct nuclear SSU rDNA sequencing of environmental samples and novel isolates retrieved from oceanic and coastal marine ecosystems. Protist 155, 193–214.

[8] Michaels, A.F., Knap, A.H., Dow, R.L., Gundersen, K., Johnson, R.J., Sorensen, J., Close, A., Knauer, G.A., Lohrenz, S.E., Asper, V.A., Tuel, M. and Bidigare, R. (1994) Seasonal patterns of ocean biogeochemistry at the US JGOFS Bermuda Atlantic Time-series Study site. Deep-Sea Res. I 41, 1013–1038.

[9] Butcher, R. (1952) Contributions to our knowledge of the smaller marine algae. J. Mar. Biol. Assoc. UK 31, 175–191.

[10] Eikrem, W. and Throndsen, J. (1990) The ultrastructure of *Bathycoccus* gen. nov. and *B. prasinos* sp. nov., a non-motile picoplanktonic alga (Chlorophyta, Prasinophyceae) from the Mediterranean and Atlantic. Phycologia 29, 344–350.

[11] Chrétiennot-Dinet, M.J., Courties, C., Vaquer, A., Neveux, J., Claustre, H., Lautier, J. and Machado, M.C. (1995) A new marine picoeucaryote: *Ostreococcus tauri* gen. et sp. nov. (Chlorophyta, Prasinophyceae). Phycologia 34, 285–292.

[12] Throndsen, J. and Kristiansen, S. (1991) *Micromonas pusilla* (Prasinophyceae) as part of pico- and nanoplankton communities of the Barents Sea. Polar Res. 10, 201–207.

[13] Zingone, A., Sarno, D. and Forlani, G. (1999) Seasonal dynamics in the abundance of *Micromonas pusilla* (Prasinophyceae) and its viruses in the Gulf of Naples (Mediterranean Sea). J. Plankton Res. 21, 2143–2159.

[14] Johnson, P.W. and Sieburth, J.M. (1982) In situ morphology and occurence of eucaryotic phototrophs of bacterial size in the picoplankton of estuarine and oceanic waters. J. Phycol. 18, 318–327.

[15] Anderson, D.M., Kulis, D.M., Keafer, B.A. and Berdalet, E. (1999) Detection of the toxic dinoflagellate *Alexandrium fundyense* (Dinophyceae) with oligonucleotide and antibody probes: Variability in labeling intensity with physiological condition. J. Phycol. 35, 870–883.

[16] Speel, E.J. and Komminoth, P. (1999) CARD in situ hybridization: sights and signals. Endocr. Pathol. 10, 193–198.

[17] Pernthaler, J., Pernthaler, A. and Amann, R. (2003) Automated enumeration of groups of marine picoplankton after fluorescence in situ hybridization. Appl. Environ. Microbiol. 69, 2631–2637.

[18] Biegala, I., Not, F., Vaulot, D. and Simon, N. (2003) Quantitative assessment of picoeucaryotes in the natural environment using taxon specific oligonucleotide probes in association with TSA-FISH (Tyramide Signal Amplification – Fluorescent In Situ Hybridization) and flow cytometry. Appl. Environ. Microbiol. 69, 5519–5529.

[19] Bustin, S.A. (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. J. Mol. Endocr. 25, 169–193.

[20] Suzuki, M.T., Preston, C.M., Chavez, F.P. and DeLong, E.F. (2001) Quantitative mapping of bacterioplankton populations in seawater: field tests across an upwelling plume in Monterey Bay. Aquat. Microb. Ecol. 24, 117–127.

[21] Skovhus, T.L., Ramsing, N.B., Holmstrom, C., Kjelleberg, S. and Dahllof, I. (2004) Real-time quantitative PCR for assessment of abundance of *Pseudoalteromonas* species in marine samples. Appl. Environ. Microbiol. 70, 2373–2382.

[22] Becker, S., Fahrbach, M., Boger, P. and Ernst, A. (2002) Quantitative tracing, by Taq nuclease assays, of a *Synechococcus* ecotype in a highly diversified natural population. Appl. Environ. Microbiol. 68, 4486–4494.

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

[24] Doyle, J.J. and Doyle, J.L. (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19, 11–15.

[25] Massana, R., Murray, A.E., Preston, C.M. and DeLong, E.F. (1997) Vertical distribution and phylogenetic characterization of marine planktonic *Archaea* in the Santa Barbara Channel. Appl. Environ. Microbiol. 63, 50–56.

[26] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[27] Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, A., Buchner, A., Lai, T., Steppi, S., Jobb, G., Forster, W., Brettske, I., Gerber, S., Ginhart, A.W., Gross, O., Grumann, S., Hermann, S., Jost, R., Konig, A., Liss, T., Lussmann, A., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A. and Schleifer, K.H. (2004) ARB: a software environment for sequence data. Nucl. Acids Res. 32, 1363–1371.

[28] Moon-van der Staay, S.Y., van der Staay, G.W.M., Guillou, L., Vaulot, D., Claustre, H. and Medlin, L.K. (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.

[29] Marie, D., Brussaard, C.P.D., Partensky, F. and Vaulot, D. (1999) Flow cytometric analysis of phytoplankton, bacteria and viruses (Robinson, J.P., Ed.), Current Protocols in Cytometry, vol. 11.11, pp. 1–15. John Wiley & Sons, New York.

[30] Campbell, M.S. and Wright, A.C. (2003) Real-time PCR analysis of *Vibrio vulnificus* from oysters. Appl. Environ. Microbiol. 69, 7137–7144.

[31] Suzuki, M.T., Taylor, L.T. and DeLong, E.F. (2000) Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5′-nuclease assays. Appl. Environ. Microbiol. 66, 4605–4614.

[32] Holtzendorff, J., Marie, D., Post, A.F., Partensky, F., Rivlin, A. and Hess, W.R. (2002) Synchronized expression of *ftsZ* in natural *Prochlorococcus* populations of the Red Sea. Environ. Microbiol. 4, 644–653.

[33] Zhang, H. and Lin, S.J. (2003) Complex gene structure of the form II RUBISCO in the dinoflagellate *Prorocentrum minimum* (Dinophyceae). J. Phycol. 39, 1160–1171.

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

[35] Sournia, A., Birrien, J.-L., Douville, J.-L., Klein, B. and Viollier, M. (1987) A daily study of the diatom spring bloom at Roscoff. I. The spring bloom within the annual cycle (France) in 1985. Est. Coast. Shelf. Sci. 25, 355–367.

[36] Falkowski, P.G., Schofield, O., Katz, M.E., van de Schootbrugge, B. and Knoll, A. (2004) Why is the land green and the ocean red In: Coccolithophorids (Thierstein, H. and Young, J., Eds.), pp. 429–453. Springer-Verlag, Berlin.

[37] Moore, L.R., Rocap, G. and Chisholm, S.W. (1998) Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. Nature 393, 464–467.

[38] Rocap, G., Distel, D.L., Waterbury, J.B. and Chisholm, S.W. (2002) Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S–23S ribosomal DNA internal transcribed spacer sequences. Appl. Environ. Microbiol. 68, 1180–1191.

[39] Rodríguez, F., Derelle, E., Guillou, L., Le Gall, F., Vaulot, D., and Moreau, H. (2004) Ecotype diversity in the marine picoeukaryote *Ostreococcus* (Chlorophyta, Prasinophyceae). Environ. Microbiol., in press.

[40] Simon, N., Campbell, L., Ornolfsdottir, E., Groben, R., Guillou, L., Lange, M. and Medlin, L.K. (2000) Oligonucleotide probes for the identification of three algal groups by dot blot and fluorescent whole-cell hybridization. J. Euk. Microbiol. 47, 76–84.

[41] Simon, N., Lebot, N., Marie, D., Partensky, F. and Vaulot, D. (1995) Fluorescent *in situ* hybridization with rRNA-targeted oligonucleotide probes to identify small phytoplankton by flow cytometry. Appl. Environ. Microbiol. 61, 2506–2513.

[42] Not, F., Simon, N., Biegala, I.C. and Vaulot, D. (2002) Application of fluorescent *in situ* hybridization coupled with tyramide signal amplification (FISH-TSA) to assess eukaryotic picoplankton composition. Aquat. Microb. Ecol. 28, 157–166.

[43] Galluzzi, L., Penna, A., Bertozzini, E., Vila, M., Garces, E. and Magnani, M. (2004) Development of a real-time PCR assay for rapid detection and quantification of *Alexandrium minutum* (a dinoflagellate). Appl. Environ. Microbiol. 70, 1199–1206.

[44] Dufresne, A., Salanoubat, M., Partensky, F., Artiguenave, F., Axmann, I.M., Barbe, V., Duprat, S., Galperin, M.Y., Koonin, E.V., Le Gall, F., Makarova, K.S., Ostrowski, M., Oztas, S., Robert, C., Rogozin, I.B., Scanlan, D.J., de, M.N.T., Weissenbach, J., Wincker, P., Wolf, Y.I. and Hess, W.R. (2003) Genome sequence of the cyanobacterium *Prochlorococcus marinus* SS120, a nearly minimal oxyphototrophic genome. Proc. Natl. Acad. Sci. USA 100, 10020–10025.

[45] Acinas, S.G., Marcelino, L.A., Klepac-Ceraj, V. and Polz, M.F. (2004) Divergence and redundancy of 16S rRNA sequences in genomes with multiple rrn operons. J. Bacteriol. 186, 2629–2635.

[46] Prokopowich, C.D., Gregory, T.R. and Crease, T.J. (2003) The correlation between rDNA copy number and genome size in eukaryotes. Genome 46, 48–50.

[47] Matsuzaki, M., Misumi, O., Shin, I.T., Maruyama, S., Takahara, M., Miyagishima, S.Y., Mori, T., Nishida, K., Yagisawa, F., Yoshida, Y., Nishimura, Y., Nakao, S., Kobayashi, T., Momoyama, Y., Higashiyama, T., Minoda, A., Sano, M., Nomoto, H., Oishi, K., Hayashi, H., Ohta, F., Nishizaka, S., Haga, S., Miura, S., Morishita, T., Kabeya, Y., Terasawa, K., Suzuki, Y., Ishii, Y., Asakawa, S., Takano, H., Ohta, N., Kuroiwa, H., Tanaka, K., Shimizu, N., Sugano, S., Sato, N., Nozaki, H., Ogasawara, N., Kohara, Y. and Kuroiwa, T. (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. Nature 428, 653–657.

[48] Abrahamsen, M.S., Templeton, T.J., Enomoto, S., Abrahante, J.E., Zhu, G., Lancto, C.A., Deng, M.Q., Liu, C., Widmer, G., Tzipori, S., Buck, G.A., Xu, P., Bankier, A.T., Dear, P.H., Konfortov, B.A., Spriggs, H.F., Iyer, L., Anantharaman, V., Aravind, L. and Kapur, V. (2004) Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. Science 304, 441–445.

[49] Cavalier-Smith, T. (1985) Eukaryote gene numbers, non-coding DNA and genome size In: The Evolution of Genome Size (Cavalier-Smith, T., Ed.), pp. 69–103. John Wiley & Sons Ltd, New York.

[50] Boucher, N., Vaulot, D. and Partensky, F. (1991) Flow cytometric determination of phytoplankton DNA in cultures and oceanic populations. Mar. Ecol. Prog. Ser. 71, 75–84.

[51] Labrenz, M., Brettar, I., Christen, R., Flavier, S., Botel, J. and Hofle, M.G. (2004) Development and application of a real-time PCR approach for quantification of uncultured bacteria in the central Baltic Sea. Appl. Environ. Microbiol. 70, 4971–4979.

[52] Binder, B.J. and Liu, Y.C. (1998) Growth rate regulation of rRNA content of a marine *Synechococcus* (Cyanobacterium) strain. Appl. Environ. Microbiol. 64, 3346–3351.

[53] Elwood, H.J., Olsen, G.J. and Sogin, M.L. (1985) The small-subunit ribosomal RNA gene sequences from the hypotrichous ciliates *Oxytricha nova* and *Stylonychia pustulata*. Mol. Biol. Evol. 2, 399–410.