



Supporting Online Material for

Picobiliphytes: A Marine Picoplanktonic Algal Group with Unknown Affinities to Other Eukaryotes

Fabrice Not,* Klaus Valentin, Khadidja Romari, Connie Lovejoy, Ramon Massana, Kerstin Töbe, Daniel Vaultot, Linda Medlin*

*To whom correspondence should be addressed. E-mail: lkmedlin@awi-bremerhaven.de (L.M.); not@icm.csic.es (F.N.)

Published 12 January 2007, *Science* **315**, 253 (2007)
DOI: 10.1126/science.1136264

This PDF file includes:

Materials and Methods
SOM Text
Figs. S1 and S2
Tables S1 to S8
References

Supporting Online Material: Material and Methods

1. Sequence analysis

Full length sequences from the picobiliphytes were obtained following Bezsteri et al. 2005 (S4). These sequences were imported into the ARB database, and aligned with its secondary structure model (S1). A selection of species representing each of the six major groups of eukaryotes (table S1) were used to construct a molecular phylogeny to place the picobiliphytes phylogenetically. Positions that occurred in at least 50 % of the taxa were selected for phylogenetic analysis. This resulted in a database of 174 taxa and 1,825 positions (available upon request from LKM). This data set was subjected to the Modeltest program (Version 2.2, ref. S5) in which the AIC criterion selected the general time reversal model of evolution with the following rate parameters: Lset Base=(0.2450 0.124 0.2609), Nst=6, Rmat=(1.0000 2.2127 1.0000 1.0000 3.1084), Rates=gamma, Shape=0.6199, Pinvar=0. A Mr Bayes analysis (<http://morphbank.ebc.uu.se/mrbayes/>), version 3.1, was run in two parallel runs saving every 1000th tree. We increased the complexity of our MrBayes analysis, which initially used a gamma correction and 6 rate categories with one million generations with 4 chains; then 1.5 million generations with 4 chains; then one million generations with 6 chains; and finally 1.5 million generations with 6 chains with increased temperature to encourage more swapping between chains. A consensus tree was made from the last 100 trees and presented in Fig. 1. A weighted MP analysis was performed in PAUP* (S6). For this analysis, a maximum parsimony tree (MP) was obtained in the following sequential analyses. The data set was weighted with a rescaled consistency index and analyzed with heuristic search using 1000 random additions with a NNI branch swapping algorithm. The resulting suboptimal trees were used as input into a second analysis using a TBR branch swapping algorithm to obtain an optimal shortest tree. This tree was loaded into MacClade and the trees rearranged forcing our picobiliphytes into a polytomy with each major lineage of eukaryotes. The trees from each rearrangement were used to constrain another MP analysis and the resulting trees from the constrained analysis were then tested to determine if they were significantly different from the best tree

obtained in the MP analysis. In addition, all other eukaryotes were combined into a polytomy and using a reverse constraint analysis, we tested the non-monophyly of our picobiliphytes. The resulting trees were used as input for the Kishino-Hasegawa Test (table S3). Distance analysis was performed using PAUP*. Dissimilarity values, based on pairwise comparisons of sequences (S7), were transformed into distances using models determined from the Modeltest program. Branching order stability was estimated by bootstrap analysis as above. Stability of the branching order was estimated using bootstrap analysis (BT) (S8) for 100 replicates for both the distance and the weighted MP trees because the data set was so large.

2. Probe design and tests

The two probes PICOBIO1 and PICOBIO2 were designed to target environmental sequences using the ARB software package (S1, tables S4, S5). The 18S rDNA database used together with this software is currently maintained by the Oceanic Plankton team at the Station Biologique de Roscoff and contains over 30,000 aligned sequences from Eukaryotes, Bacteria and Archaea. Because the picobiliphyte isolates do not exist in culture, we were not able to perform any positive hybridization tests for these probes. However, using 40% formamide (based on the GC % and positions of mismatches) to adjust stringency, we tested the probes on a range of cultured species (table S6). Among them the species belonging to the divisions Rhodophyta and Cryptophyta were the closest relatives available in culture. This experiment aimed at unveiling some potential unspecific labelling. Results did not show any non-specific hybridization (table S6). However, we observed positive signals on natural samples. For these reasons we believe our probes are specific for the picobiliphytes.

3. Tyramide Signal Amplification - Fluorescent In Situ Hybridization (TSA-FISH)

Contributions of picobiliphytes to the total picoeukaryotic community are presented in table S8. These results were obtained by the application of the TSA-FISH technique using the two probes PICOBIO1 and PICOBIO2 on natural 3 μm filtered seawater samples harvested at different dates at the Roscoff ASTAN sampling site.

Abundance of cells (cells ml⁻¹) belonging to the picobiliphyte clades and to the total picoeukaryotic community were determined by TSA-FISH following Not et al. (S2), with the probes PICOBI01, PICOBI02, and a mix of three general probes (EUK1209R, CHLO01, and NCHLO02), respectively.

Because the picobiliphytes exhibited phycobilin-like pigments, we wanted to assess their contribution to the orange fluorescing cells present in the environment (table S7). Cells from the estuarine sampling station Roscoff Dourduff were simultaneously enumerated and sorted by flow cytometry based on their orange fluorescence (2,253 cells sorted in 3 hours). These cells were then concentrated on Anodisc filters by filtration and prepared for TSA-FISH (S2). Cell abundances for the picobiliphytes were estimated individually by TSA-FISH with the probes PICOBI01 and PICOBI02. Fluorescent *in situ* hybridizations have been done in replicate for both probes on the same filter (table S7).

4. Solid phase cytometry (ChemScan)

Helgoland surface samples were collected on a cruise with the RV Uthörn from 30/05-02/06/2006. 1 L samples were collected and subsequently fractionated with a 10 µm, 5 µm and 3 µm polycarbonate filter of 47 mm diameter (Millipore, USA) and finally filtered onto 0.2 µm polycarbonate filters (Millipore, USA) for TSA-FISH and solid phase cytometry (S3).

A ChemScan RDI (Chemunex, France) was used for solid phase cytometry. An overlapping scan of the whole filter membrane surface was carried out with an argon laser at a wavelength of 488 nm to detect cells with FITC labelled tyramides. The computer software (MatLab, Matworks, Natick, Mass.), automatically applies different discrimination criteria based on optical characteristics like particle size and signal shape and therewith enables the differentiation between autofluorescent particles, unlabelled cells and labelled target cells.

The positive counted signals are shown as a representation of the filter on the computer screen, termed a scan map (fig. S2). The filters were validated

microscopically directly after the scan with an epifluorescence microscope, which is connected to the ChemScan and equipped with a motorized stage. After highlighting a signal with the cursor on the scan map, the motorized stage moves to the corresponding position on the filter and a validation of the counted signals is performed optically.

Supporting Online Material: Results of the phylogenetic analyses

Using a Bayesian analysis of the 18S rRNA sequences from organisms representing each of six major groups of eukaryotes (table S1) aligned by secondary structure in the ARB alignment program with increasing complexity of parallel runs of the MrBayes (MB) program, we found that the runs did not converge on the same tree. Initial analyses placed picobiliphytes sister to haptophytes or as an independent group. Complex analyses found picobiliphytes either sister to haptophytes (posterior probabilities or pp = 55) or to a cryptophyte/katablepharid clade (pp = 100) (Fig. 1). Similarities in the pigment composition and a DAPI staining organelle in the plastid may provide support for the latter sister relationship. Some sister relationships in our analysis, e.g., Heterokonta and Cercozoa, are likely artifacts because this is a single gene phylogeny. We do not recover all sister relationships found in concatenated phylogenies (S10), e.g., because we do not have living cells for additional genes, or similar sister relationships found in rate weighted phylogenies from rRNA genes, e.g., as in van de Peer et al. (S11). Therefore, we used the rRNA gene to test if picobiliphytes fall inside another major eukaryotic group. The independence of our lineage was assessed using the Kashino-Hasagawa test in PAUP (12) (table S3). All trees forcing picobiliphytes into other eukaryote groups were significantly different from the best tree (table S3) and the only group that could be interpreted as being a possible sister to our picobiliphytes is the rhodophytes because the number of steps from the best tree to this constrained tree is the shortest. Bootstrap analyses using a weighted Maximum Parsimony analysis and Neighbor-joining analysis with gamma corrections established from Modeltest found high support for all the terminal taxa but little or no support for sister relationships.

A consensus of the last ten trees in our most complex MB analysis showed a weak sister relationship with the rhodophytes (pp = 75).

Supporting Figures

Fig. S1. Use of probes PICOBIO1 and PICOBIO2 using TSA-FISH to detect marine picobiliphytes. **a.** Cells from Roscoff Astan (RA, September 26, 2001) and Roscoff Dourduff (RD, September 17, 2002) with overlaid epifluorescence pictures showing the nucleus stained with DAPI in blue (UV excitation), and probe fluorescence in green (blue excitation). The red fluorescence likely originated from the autofluorescence of a phycobilin-containing plastid under blue excitation, such as those from the red algae and the cryptophytes (see b). The paler yellow fluorescence in some of the pictures is residual chlorophyll. **b.** A cell of the cryptophyte *Rhodomonas salina* hybridized with the PICOBIO2 probe. The absence of a green color indicates that the probe did not label the cell. The plastid that contains phycobilins shows a clear red autofluorescence similar to that in the cells from the natural samples (a).

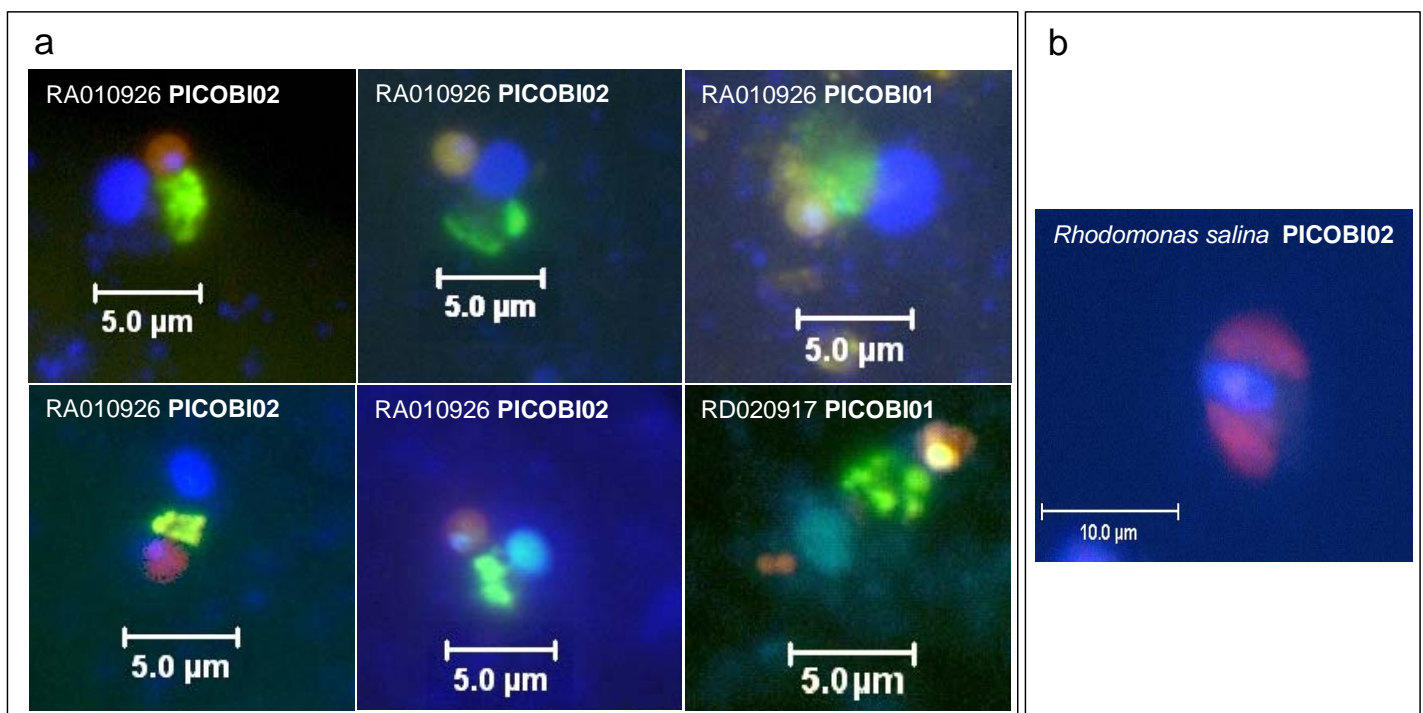
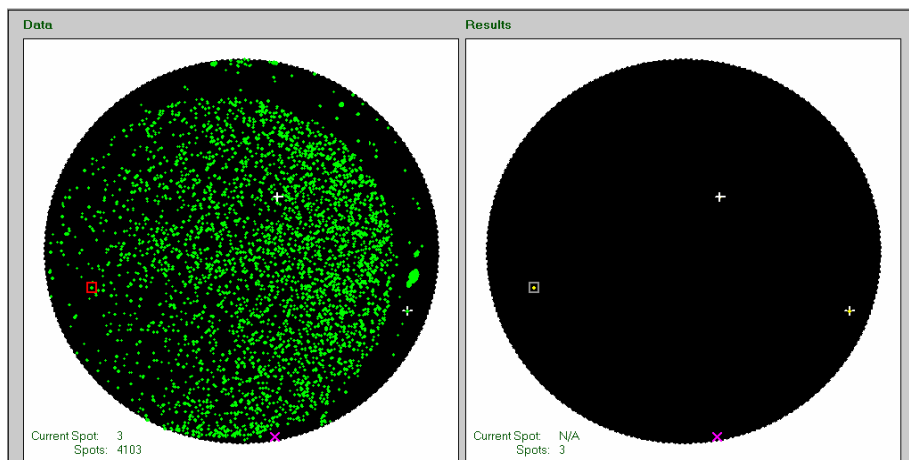
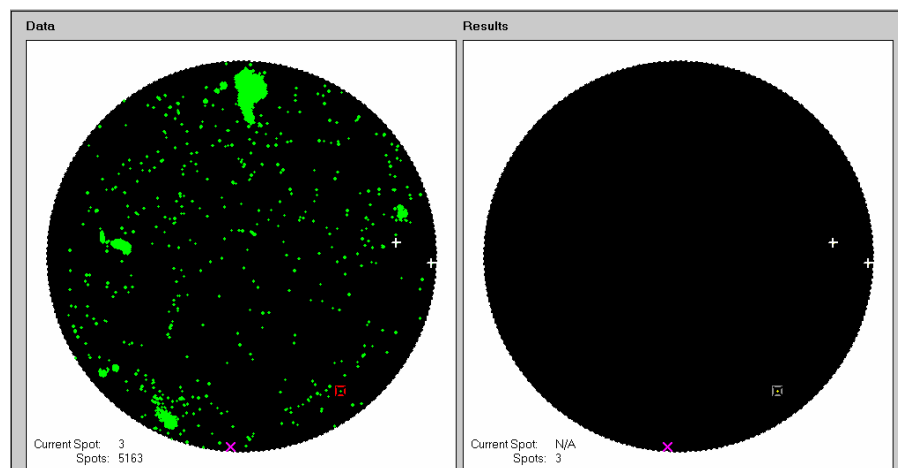


Fig S2. Application of probes PICOBI01 and PICOBI02 using TSA-FISH to detect marine picobiliphytes using the ChemScan machine from the less than 3 μm fraction sample from two locations near Helgoland in the German Bight. The sample is filtered, hybridized with the probe, and scanned by the ChemScan laser for fluorescent signals. The figure on the left represents all fluorescent signals on the filter, and the figure on the right displays only the cells recognized by the probe. A set of discriminate values provided by the ChemScan analysis package based on optical characteristics of the generated signals like wavelengths, signal shape and particle size eliminates all fluorescent signals that cannot be associated with a probe signal. + denotes positive cells subsequently checked in the microscope to verify the fluorescent signal of the cells. A more detailed use of the ChemScan machine can be found in ref S3.



Supporting Tables

Table S1. Species names, accession numbers, and taxonomic affiliations of full length sequences used for the phylogenetic analysis presented in Fig. 1. The taxonomic affiliation follows the revised eukaryotic classification in Adl et al. (S9).

Species	Accession Number	Super Group	First Rank if known
<i>Acanthamoeba castellanii</i>	AF114438	Amoebozoa	Amoebozoa
<i>Acanthamoeba pustulosa</i>	AF019050	Amoebozoa	Amoebozoa
<i>Hartmannella vermiformis</i>	X75513	Amoebozoa	Amoebozoa
<i>Entamoeba dispar</i>	Z49256	Amoebozoa	Amoebozoa
<i>Entamoeba histolytica</i>	X65163	Amoebozoa	Amoebozoa
<i>Naegleria gruberi</i>	M18732	Amoebozoa	Amoebozoa
<i>Neoparamoeba pemaquidensis</i>	AF371968	Amoebozoa	Amoebozoa
<i>Balamuthia mandrillaris</i>	AF019071	Amoebozoa	Amoebozoa
<i>Phreatamoeba balamuthi</i>	L23799	Amoebozoa	Amoebozoa
<i>Vahlkampfia lobospinosa</i>	M98052	Amoebozoa	Amoebozoa
<i>Dictyostelium discoideum</i>	K02641	Amoebozoa	Mycetozoa
<i>Physarum polycephalum</i>	X13160	Amoebozoa	Mycetozoa
<i>Mesostigma viride</i>	AJ250109	Archeplastida	Chlorophyta
<i>Trebouxia asymmetrica</i>	Z21553	Archeplastida	Chlorophyta
<i>Ulva rigida</i>	AJ005414	Archeplastida	Chlorophyta
<i>Chlorella minutissima</i>	AB006046	Archeplastida	Chlorophyta

<i>Oogamochlamys gigantea</i>	AJ410465	Archeplastida	Chlorophyta
<i>Chlamydomonas reinhardtii</i>	M32703	Archeplastida	Chlorophyta
<i>Cyanoptyche gloeocystis</i>	AJ007275	Archeplastida	Glaucocystophyta
<i>Glaucocystis nostochinearum</i>	X70803	Archeplastida	Glaucocystophyta
<i>Gloeochaete wittrockiana</i>	X81901	Archeplastida	Glaucocystophyta
<i>Bangia atropurpurea</i>	L36066	Archeplastida	Rhodophyta
<i>Chondrus crispus</i>	Z14140	Archeplastida	Rhodophyta
<i>Gracilaria lemaneiformis</i>	M54986	Archeplastida	Rhodophyta
<i>Porphyra suborbiculata</i>	AB013180	Archeplastida	Rhodophyta
<i>Porphyridium aerugineum</i>	L27635	Archeplastida	Rhodophyta
<i>Arabidopsis thaliana</i>	X52322	Archeplastida	Streptophyta
<i>Ginkgo biloba</i>	D16448	Archeplastida	Streptophyta
<i>Glycine max</i>	X02623	Archeplastida	Streptophyta
<i>Magnolia tripetala</i>	AF206956	Archeplastida	Streptophyta
<i>Marchantia polymorpha</i>	AB021684	Archeplastida	Streptophyta
<i>Zamia pumila</i>	M20017	Archeplastida	Streptophyta
<i>Ammonia beccarii</i>	U07937	Chromoalveolata	Alveolata/Ciliata
<i>Apusomonas proboscidea</i>	L37037	Chromoalveolata	Alveolata/Ciliata
<i>Colpoda inflata</i>	M97908	Chromoalveolata	Alveolata/Ciliata
<i>Paraurostyla weissei</i>	AJ310485	Chromoalveolata	Alveolata/Ciliata

<i>Stylonychia pustulata</i>	M14600	Chromoalveolata	Alveolata/Ciliata
<i>Tetrahymena nanneyi</i>	M98016	Chromoalveolata	Alveolata/Ciliata
<i>Trithymostoma steini</i>	X71134	Chromoalveolata	Alveolata/Ciliata
<i>Vorticella convallaria</i>	AF070700	Chromoalveolata	Alveolata/Ciliata
<i>Alexandrium fundyense</i>	U09048	Chromoalveolata	Alveolata/Dinoflagellata
<i>Amblyospora</i> sp.	U68474	Chromoalveolata	Alveolata/Dinoflagellata
<i>Amoebophrya</i> sp.	AF069516	Chromoalveolata	Alveolata/Dinoflagellata
<i>Amyloodinium ocellatum</i>	AF080096	Chromoalveolata	Alveolata/Dinoflagellata
<i>Cryptosporidium parvum</i>	L16996	Chromoalveolata	Alveolata/Dinoflagellata
<i>Dinophysis norvegica</i>	AF239261	Chromoalveolata	Alveolata/Dinoflagellata
<i>Gonyaulax spinifera</i>	AF022155	Chromoalveolata	Alveolata/Dinoflagellata
<i>Gymnodinium catenatum</i>	AY421785	Chromoalveolata	Alveolata/Dinoflagellata
<i>Gymnodinium mikimotoi</i>	AF009131	Chromoalveolata	Alveolata/Dinoflagellata
<i>Babesia bigemina</i>	X59607	Chromoalveolata	Apicomplexa
<i>Eimeria mitis</i>	U40262	Chromoalveolata	Apicomplexa
<i>Gregarina niphandrodes</i>	AF129882	Chromoalveolata	Apicomplexa
<i>Hepatozoon canis</i>	AF176835	Chromoalveolata	Apicomplexa
<i>Sarcocystis dispersa</i>	AF120115	Chromoalveolata	Apicomplexa
<i>Sarcocystis muris</i>	M34846/M64244	Chromoalveolata	Apicomplexa
<i>Theileria youngi</i>	AF245279	Chromoalveolata	Apicomplexa

<i>Theileria cervi</i>	AF086804	Chromoalveolata	Apicomplexa
<i>Chilomonas paramecium</i>	L28811	Chromoalveolata	Cryptophyta
<i>Cryptomonas pyrenoidifera</i>	AJ421147	Chromoalveolata	Cryptophyta
<i>Cryptomonas paramecium</i>	AJ715468	Chromoalveolata	Cryptophyta
<i>Cryptomonas pyrenoidifera</i> nucleomorph	AJ715473	Chromoalveolata	Cryptophyta
<i>Hanusia phi</i>	U53126	Chromoalveolata	Cryptophyta
<i>Geminigera cryophila</i>	AB058368	Chromoalveolata	Cryptophyta
<i>Geminigera cryophila</i> nucleomorph	U53123	Chromoalveolata	Cryptophyta
<i>Goniomonas truncata</i>	U03072	Chromoalveolata	Cryptophyta
<i>Guillardia theta</i> nucleomorph	AF165818	Chromoalveolata	Cryptophyta
<i>Chrysochromulina polylepis</i>	AJ004866	Chromoalveolata	Haptophyta
<i>Emiliana huxleyi</i>	X82156	Chromoalveolata	Haptophyta
<i>Pavlova virescens</i>	AJ515248	Chromoalveolata	Haptophyta
<i>Pavlova salina</i>	L34669	Chromoalveolata	Haptophyta
<i>Phaeocystis globosa</i>	X77476	Chromoalveolata	Haptophyta
<i>Achlya bisexualis</i>	M32705	Chromoalveolata	Heterokonta
<i>Allomyces macrogynus</i>	U23936	Chromoalveolata	Heterokonta
<i>Bacillaria paxillifer</i>	M87325	Chromoalveolata	Heterokonta
<i>Blastocystis hominis</i>	U51151	Chromoalveolata	Heterokonta
<i>Caecitellus parvulus</i>	AF174367	Chromoalveolata	Heterokonta

<i>Cafeteria roenbergensis</i>	AF174364	Chromoalveolata	Heterokonta
HE001005.33	<u>EF050072</u>	Chromoalveolata	Heterokonta
<i>Chattonella verruculosa</i>	AY788947	Chromoalveolata	Heterokonta
<i>Epipyxis pulchra</i>	AF123298	Chromoalveolata	Heterokonta
<i>Heterosigma akashiwo</i>	U41650	Chromoalveolata	Heterokonta
<i>Labyrinthuloides minuta</i>	L27634	Chromoalveolata	Heterokonta
<i>Mallomonas papillosa</i>	M55285	Chromoalveolata	Heterokonta
<i>Nannochloropsis granulata</i>	AF045041	Chromoalveolata	Heterokonta
<i>Paraphysomonas foraminifera</i>	AB022864	Chromoalveolata	Heterokonta
<i>Phytophthora megasperma</i>	X54265	Chromoalveolata	Heterokonta
<i>Proteromonas lacertae</i>	U37108	Chromoalveolata	Heterokonta
<i>Thraustochytrium kinnei</i>	L34668	Chromoalveolata	Heterokonta
<i>Tribonema aequale</i>	M55286	Chromoalveolata	Heterokonta
<i>Ulkenia profunda</i>	AB022114	Chromoalveolata	Heterokonta
<i>Uroglena americana</i>	AF123290	Chromoalveolata	Heterokonta
<i>Lagenidium giganteum</i>	X54266	Chromoalveolata	Hetrokonta
<i>Laminaria angustata</i>	AB022818	Chromoalveolata	Hetrokonta
<i>Mallomonas caudata</i>	U73228	Chromoalveolata	Hetrokonta
<i>Skeletonema pseudocostatum</i>	X85394	Chromoalveolata	Hetrokonta
<i>Katablepharis japonica</i>	AB231617	Chromoalveolata	Katablepharids

<i>Leucocryptos marina</i>	AB194980	Chromoalveolata	Katablepharids
<i>Giardia intestinalis</i> isolate BAG1	AF199448	Excavata	Diplomonadida
<i>Spironucleus muris</i>	X84231	Excavata	Diplomonadida
<i>Giardia intestinalis</i>	AF473852	Excavata	Dipolomonidae
<i>Astasia longa</i>	AF112871	Excavata	Euglenozoa
<i>Bodo saliens</i>	AF174379	Excavata	Euglenozoa
<i>Bodo caudatus</i>	X53910	Excavata	Euglenozoa
<i>Dimastigella trypaniformis</i>	X76495	Excavata	Euglenozoa
<i>Euglena gracilis</i>	M12677	Excavata	Euglenozoa
<i>Trypanosoma cruzi</i>	AF245381	Excavata	Euglenozoa
<i>Coronympha octonaria</i>	U17504	Excavata	Parabasalidea
<i>Amblyospora connecticus</i>	AF025685	Opisthokonta	Fungi
<i>Anurofeca richardsi</i>	AF070445	Opisthokonta	Fungi
<i>Aspergillus avenaceus</i>	AB008395	Opisthokonta	Fungi
<i>Basidiobolus ranarum</i>	D29946	Opisthokonta	Fungi
<i>Candida aaseri</i>	AB013564	Opisthokonta	Fungi
<i>Delitschia didyma</i>	AF242264	Opisthokonta	Fungi
<i>Dermocystidium salmonis</i>	U21337	Opisthokonta	Fungi
<i>Eupenicillium crustaceum</i>	D88324	Opisthokonta	Fungi
<i>Microsporidium prosopium</i>	AF151529	Opisthokonta	Fungi

<i>Neurospora crassa</i>	X04971	Opisthokonta	Fungi
<i>Psorospermium haeckelii</i>	U33180	Opisthokonta	Fungi
<i>Rhinosporidium seeberi</i>	AF158369	Opisthokonta	Fungi
<i>Septata intestinalis</i>	L19567	Opisthokonta	Fungi
<i>Sphaeroforma arctica</i>	Y16260	Opisthokonta	Fungi
<i>Thalassicolla nucleata</i>	AF057742	Opisthokonta	Fungi
<i>Trichosporon asteroides</i>	AB001729	Opisthokonta	Fungi
<i>Udeniomyces megalosporus</i>	D31657	Opisthokonta	Fungi
<i>Artemia salina</i>	X01723	Opisthokonta	Metazoa
<i>Diaphanoeca grandis</i>	L10824	Opisthokonta	Metazoa
<i>Drosophila melanogaster</i>	M21017/M29800	Opisthokonta	Metazoa
<i>Homo sapiens</i>	U13369	Opisthokonta	Metazoa
<i>Leucosolenia</i> sp.	AF100945	Opisthokonta	Metazoa
<i>Littorina obtusata</i>	X94274	Opisthokonta	Metazoa
<i>Mnemiopsis leidyi</i>	L10826	Opisthokonta	Metazoa
<i>Mus musculus</i>	X82564	Opisthokonta	Metazoa
<i>Obelia</i> sp.	Z86108	Opisthokonta	Metazoa
<i>Acanthoecopsis unguiculata</i>	L10823	Rhizaria	Acantharea
<i>Acanthometra</i> sp.	AF063240	Rhizaria	Acantharea
<i>Chaunacanthid</i> sp.	AF018158	Rhizaria	Acantharea

<i>Symphyacanthid</i> sp.	AF063242	Rhizaria	Acantharea
<i>Cercomonas</i> ATCC50318	U42450	Rhizaria	Cercozoa
<i>Cercomonas longicauda</i>	AF101052	Rhizaria	Cercozoa
<i>Chlorarachnion reptans</i>	U03477	Rhizaria	Cercozoa
<i>Euglypha rotunda</i>	X77692	Rhizaria	Cercozoa
<i>Paulinella chromatophora</i>	X81811	Rhizaria	Cercozoa
<i>Chlorarachnion reptans</i>	X70809	Rhizaria	Chloroarachniophyta
<i>Chlorarachnion reptans</i> nucleomorph	U03275	Rhizaria	Chloroarachniophyta
<i>Gymnochlora stellata</i>	AF076171	Rhizaria	Chloroarachniophyta
<i>Chloraranion</i> sp. nucleomorph	U58510	Rhizaria	Chloroarachniophyta
<i>Sorites orbiculus</i>	AJ132369	Rhizaria	Foraminifera
<i>Acrosphaera</i> sp.	AF091148	Rhizaria	Polycystinea
<i>Siphonosphaera cyathina</i>	AF091145	Rhizaria	Polycystinea
Uncultured Polycystinea	AF382824	Rhizaria	Polycystinea
<i>Cryptotermes domesticus</i>	AB032215	unknown	Parabasalidea
BL000921.8	AY426835	unknown	Picobiliphytes
HE000427.214	DQ222872	unknown	Picobiliphytes
HE000803.72	DQ222873	unknown	Picobiliphytes
HE001005.148	DQ222874	unknown	Picobiliphytes
NW414.27	DQ060524	unknown	Picobiliphytes

NOR46.24	DQ060526	unknown	Picobiliphytes
NW617.02	DQ060525	unknown	Picobiliphytes
OR0004.159	DQ222875	unknown	Picobiliphytes
RA000907.33	DQ222876	unknown	Picobiliphytes
RA000907.54	DQ222877	unknown	Picobiliphytes
RA001219.38	DQ222878	unknown	Picobiliphytes
RA000907.18	DQ222879	unknown	Picobiliphytes
RA010613.144	DQ222880	unknown	Picobiliphytes

Table S2. Relative abundance of picobiliphyte sequences in clone libraries

Clone libraries Year, month, location	Eukaryotic clones	Picobiliphyte clones	% of picobiliphyte clones
2000, March, HE*	46	0	0
2000, April, HE*	94	1	1.1
2000, April, RA [♦]	82	1	1.2
2000, April, OR [#]	64	4	6.3
2000, June, RA [♦]	42	2	4.8
2000, August, HE*	103	1	1
2000, September, RA [♦]	40	7	17.5
2000, September, BL [◊]	71	1	1.4
2000, October, HE*	73	2	2.7
2000, December, HE*	36	0	0
2000, December, RA [♦]	34	2	5.9
2000, December, BL [◊]	106	0	0
2001, February/March, HE*	86	0	0
2001, February/March, BL [◊]	81	0	0
2001, April, RA [♦]	47	0	0
2001, May, RA [♦]	41	0	0
2001, June, RA [♦]	41	2	4.9
2001, June, BL [◊]	81	0	0
2002, August, NW01 [∨]	90	3	3.3
2002, August, NW08 [⊕]	70	1	1.4
2002, August, Z59 [^]	228	28	12.3

* HE = Helgoland, 54°11'N, 7°54'E. (5) [♦] RA = Roscoff ASTAN, 48°46'N, 3°56'E. (4) [◊] BL = Blanes Bay, 41°40'N, 2°48'E. (6) [#] OR = Orkney Islands. (5) [^] Z59 = Norwegian Sea, 76°19'N, 3°59'E. (7) [∨] NW01 = Canada Basin of the Arctic Ocean, 75°59'N, 156°52'W. (7) [⊕] NW08 = Canada Basin of the Arctic Ocean, 76°46'N, 148°57'W. (7)

Table S3. Results of Kishino-Hasegawa test where the length of a tree with the enforced polytomy of the picobiliphytes with each major eukaryotic group was tested against the best tree where the picobiliphytes was an independent lineage. The monophyly of the picobiliphytes was also tested using a reverse constraint analysis against all eukaryotes in a single clade.

Tree	Length	<u>Length</u> difference	s.d.(difference)	t	P*
Best Tree	40741				
Cryptophytes	41774	1033	69.24223	14.9186	<0.0001
Chlorophytes	41295	554	38.96773	14.2169	<0.0001
Glaucocystophytes	42346	1605	117.60304	13.6476	<0.0001
Discicristates	42349	1608	120.58334	13.3352	<0.0001
Haptophytes	42037	1296	98.73351	13.1262	<0.0001
Entamoebae	41506	765	58.75924	13.0192	<0.0001
Cercomonads	41348	607	47.62931	12.7443	<0.0001
Apicomplexa	41209	468	39.48443	11.8528	<0.0001
Stramenopiles	41644	903	78.57431	11.4923	<0.0001
Opisthokonts	41070	329	44.72071	7.3568	<0.0001
Rhodophytes	40775	34	5.77536	5.8871	<0.0001
Monophyly	49736	8995	346.98441	25.9234	<0.0001

* Probability of getting a more extreme T-value under the null hypothesis of no difference between the two trees (two-tailed test). All values were significantly different at $P < 0.05$.

Table S4. *In silico* specificity of probe PICOBI01. Clone names, length, and taxonomic affiliation of sequences tested. In addition to the picobiliphyte full length sequences presented in this study, partial sequences available in GenBank, which are at present undetermined but likely represent picobiliphytes, are also shown. Nature and position of mismatches on the closest full-length 18S rRNA non-target sequence/species are also indicated..

PICOBI01				5' -	GCG	TGA	TGC	CAA	AAT	CCG	-3'
Target				3' -	CGC	ACU	ACG	GUU	UUA	GGC	-5'
Sequence name	length (bp)	Taxonomy	Acc number	Position of mismatches on closest sequences							
HE000803.72	1812	picobiliphytes	AY343928	3' -	---	---	---	---	---	---	-5'
HE000427.214	1732	picobiliphytes	DQ222872	3' -	---	---	---	---	---	---	-5'
NW414.27	1776	picobiliphytes	DQ060524	3' -	---	---	---	---	---	---	-5'
RA000907.18	1834	picobiliphytes	DQ222879	3' -	---	---	---	---	---	---	-5'
RA001219.38	1785	picobiliphytes	DQ222878	3' -	---	---	---	---	---	---	-5'
NW617.02	1779	picobiliphytes	DQ060525	3' -	---	---	---	---	---	---	-5'
RA000907.54	1784	picobiliphytes	DQ222877	3' -	---	---	---	---	---	---	-5'
ENI47296.00159	381	picobiliphytes	AY938310	3' -	---	---	---	---	---	---	-5'
ENI42482.00158	543	picobiliphytes	AY938048	3' -	---	---	---	---	---	---	-5'
ENI42482.00072	573	picobiliphytes	AY938005	3' -	---	---	---	---	---	---	-5'
ENI40076.00318	632	picobiliphytes	AY937616	3' -	---	---	---	---	---	---	-5'
BB01_42	593	picobiliphytes	AY885047	3' -	---	---	---	---	---	---	-5'
RA000907.60	546	picobiliphytes	AY295523	3' -	---	---	---	---	---	---	-5'
RA001219.38	543	picobiliphytes	AY295551	3' -	---	---	---	---	---	---	-5'
RA001219.56	543	picobiliphytes	AY295566	3' -	---	---	---	---	---	---	-5'
RA000907.6	548	picobiliphytes	AY295522	3' -	---	---	---	---	---	---	-5'
RA000907.54	546	picobiliphytes	AY295518	3' -	---	---	---	---	---	---	-5'
RA000907.23	405	picobiliphytes	AY295495	3' -	---	---	---	---	---	---	-5'
RA000907.21	548	picobiliphytes	AY295493	3' -	---	---	---	---	---	---	-5'
RA000907.18	547	picobiliphytes	AY295489	3' -	---	---	---	---	---	---	-5'
RA000609.19	548	picobiliphytes	AY295445	3' -	---	---	---	---	---	---	-5'
RA000609.13	547	picobiliphytes	AY295441	3' -	---	---	---	---	---	---	-5'
RA000412.151	546	picobiliphytes	AY295385	3' -	---	---	---	---	---	---	-5'
NOR46.29	1763	picobiliphytes	DQ060523	3' -	--T	---	---	---	---	---	-5'
RA000907.33	1840	picobiliphytes	DQ222876	3' -	---	--C	---	---	---	---	-5'
NOR50.52	1780	picobiliphytes	DQ060527	3' -	--T	--C	---	---	---	---	-5'
<i>Lophothalia hormocladus</i>	1677	Rhodophyta	AF373216	3' -	--A	G--	G--	---	---	---	-5'
<i>Clostridium cellulolyticum</i>	1642	Bacteria	X71847	3' -	---	---	---	T--	-G-	--T	-5'
<i>Linderina pennispora</i>	1753	Fungi	AF007538	3' -	---	--G	--C	-G-	---	---	-5'
<i>Palmaria palmata</i>	1771	Rhodophyta	Z14142	3' -	-C-	GG-	G--	---	---	---	-5'

Table S5. *In silico* specificity of probe PICOB102. Clone names, length, and taxonomic affiliation of sequences tested. In addition to the picobiliphytes' full length sequences presented in this study, partial sequences available in GenBank are also shown. Nature and position of mismatches on the closest full-length 18S rRNA non-target sequence/species are also indicated.

PICOB102				5' -	ATA TGC CCG TCA AAC CGT	-3'
Target				3' -	UAU ACG GGC AGU UUG GCA	-5'
Sequence name	length (bp)	Taxonomy	Acc number	Position of mismatches on closest sequences		
NOR46.24	1788	picobiliphytes	DQ06052	3' -	---	-5'
			6			
RA010613.144	1783	picobiliphytes	DQ22288	3' -	---	-5'
			0			
OR000415.9	1804	picobiliphytes	DQ22287	3' -	---	-5'
			5			
HE001005.148	1795	picobiliphytes	DQ22287	3' -	---	-5'
			4			
RA010613.40	550	picobiliphytes.	AY29570	3' -	---	-5'
			6			
BL000921.8	1803	picobiliphytes	AY42683	3' -	--C T--	-5'
			5			
<i>Trypanosoma congolense</i>	2217	Euglenozoa	AJ009145	3' -	--- -A- --- -A- --- -G	-5'
<i>Trypanosoma congolense</i>	2240	Euglenozoa	AJ223563	3' -	--- -A- --- -AA --- -G	-5'
<i>Trypanosoma</i> sp.	2229	Euglenozoa	AJ009169	3' -	--- -A- --- --- -T -GG	-5'
<i>Branchiostoma floridae</i>	1778	Metazoa	M97571	3' -	-T- --- T-- G-- C-- ---	-5'
<i>Philonema</i> sp.	1749	Metazoa	U81574	3' -	-CG --- --- --- -CC ---	-5'

Table S6. Roscoff Culture Collection (RCC: <http://www.sb-roscoff.fr/Phyto/RCC/>) strains that have been tested using Tyramide Signal Amplification-Fluorescent *In Situ* Hybridization with probes PICOBI01 and PICOBI02. Negative hybridizations for all strains used suggest no unspecific labeling from the probes. The number of mismatches to each probe is listed for each strain.

Class	Species	RCC strain number	PICOBI01 mismatches	PICOBI02 mismatches	Hybridization results
Chlorophyceae	<i>Dunaliella tertiolecta</i>	6*	10	13	-
Prasinophyceae	<i>Micromonas pusilla</i>	114	8	12	-
Prasinophyceae	<i>Micromonas pusilla</i>	451	8	10	-
Prasinophyceae	<i>Ostreococcus tauri</i>	116	5	12	-
Prasinophyceae	<i>Pseudoscourfieldia cf. marina</i>	261	9	10	-
Cryptophyceae	<i>Rhodomonas salina</i>	20*	7	8	-
Cryptophyceae	<i>Rhodomonas baltica</i>	350*	7	10	-
Cryptophyceae	<i>Hemiselmis</i> sp.	439	7	10	-
Cryptophyceae	<i>Hemiselmis</i> sp.	660*	6	10	-
Bangiophyceae	<i>Porphyridium aerugineum</i>	652	9	10	-
Bangiophyceae	<i>Rhodella maculata</i>	655*	8	12	-
Bolidophyceae	<i>Bolidomonas pacifica</i>	205*	10	13	-
Pelagophyceae	<i>Pelagomonas calceolata</i>	100*	11	12	-

* Accession numbers DQ009772, AF508274, U53128, AJ007284, AB045608, AF123595 and U14389, respectively, were used to represent RCC strains for which no 18S rDNA sequence is available.

Table S7. Contribution of picobiliphytes to orange fluorescing cells sorted from a sample taken on September 22, 2004 at the estuarine Dourduff station (close to the Roscoff Astan sampling site, 48°38'N, 3°51'W) in the English Channel. The original concentration of orange fluorescing cells was 322 cells ml⁻¹.

Probes	Probe positive cells (cells ml ⁻¹) *	% of orange fluorescing cells (cells ml ⁻¹) *
PICOBI01	59-54	18-17
PICOBI02	96-142	30-44
PICOBI01 + PICOBI02	155-196	48-61

*Cell abundances for picobiliphytes estimated by TSA-FISH, values for the two replicates are given (replicate 1 - replicate 2).

Table S8. Contribution of picobiliphytes to the 3 μm fractionated picoeukaryotic community at the Roscoff Astan, (RA) sampling station in the English Channel as estimated with TSA-FISH and flow cytometry. Samples were filtered, and each filter was cut into sectors and two were hybridized to each probe. Positive controls using universal probes, EUK1209R, CHLO01, and NCHLO02 and no probe controls were made on different sectors of the filter.

Sample	Total picoeukaryotes (cells ml ⁻¹) #	Orange fluorescing cells (cells ml ⁻¹) ♦	Probe		Sum of Probes PICOBI01 & PICOBI02 (cells ml ⁻¹)	% of pico- eukaryotes	% of orange fluorescing cells		
			PICOBI01 (cells ml ⁻¹)	PICOBI02 (cells ml ⁻¹)					
			Piece	Piece	Piece	Piece			
			1	2	1	2			
RA010305*	4,804	n.a.	21	20	0	0	21	0.4	n.a.
RA010926*	6,693	98	42	29	63	22	79	1.2	80.6
RA011207*	4,224	209	42	37	42	14	68	1.6	32.5
RA020122*	4,590	187	43	n.a.	22	n.a.	65	1.4	34.8
RA020307*	3,927	119	21	n.a.	21	n.a.	42	1.1	35.3
Average	4,848	153	34	29	30	12	55	1.1	45.8

* Sampling date (year/month/day). Hybridizations showing no positive results (for both probes) were performed on the following summer samples: RA010412, RA010530, RA010628, RA010731, RA010814.. # Cell abundances for picobiliphytes and for the total picoeukaryotes estimated by probes.

♦ Abundances of cells with orange fluorescence estimated by flow cytometry. n.a. data not available..

Supporting references

1. W. Ludwig *et al.*, *Nucleic Acids Res.* **32**, 1363 (2004).
2. F. Not, N. Simon, I. C. Biegala, D. Vaultot, *Aquat. Microb. Ecol.* **28**, 157 (2002).
3. K. Töbe, G. Eller, L. K. Medin, *J. Plankton Res.* **7**, 643 (2006).
4. B. Beszteri, E. Acs, L. K. Medlin, *Protist* **156**, 317 (2005).
5. D. Posada, K. A. Crandall, *Bioinformatics* **14**, 817 (1998).
6. W. P. Maddison, D.R. Maddison, *Analysis of phylogeny and character evolution*. Version 3. Sinauer Associates, Sunderland, Massachusetts (1992).
7. W. M. Fitch, E. Margoliash, *Science* **155**, 279 (1967)
8. J. Felsenstein, *Evolution* **39**, 783 (1985)
9. S. M. Adl *et al.*, *J. Eukaryot. Microbiol.* **52**, 399 (2005)
10. D. Moreira , H. Le Guyader, H. Phillippe, *Nature* **405**, 69 (2000).
11. Y. Van de Peer, *et al.*, *J. Mol. Evol.* **51**, 565 (2000).