Vol. 42: 277–291, 2006 | AQUATIC MICROBIAL ECOLOGY
Aquat Microb Ecol | Published March 29

Genetic diversity and habitats of two enigmatic marine alveolate lineages

Agnès Groisillier[1], Ramon Massana[2], Klaus Valentin[3], Daniel Vaulot[1], Laure Guillou[1,*]

[1]Station Biologique, UMR 7144, CNRS, and Université Pierre & Marie Curie, BP74, 29682 Roscoff Cedex, France

[2]Department de Biologia Marina i Oceanografia, Institut de Ciències del Mar, CMIMA, CSIC. Passeig Marítim de la Barceloneta 37-49, 08003 Barcelona, Spain

[3]Alfred Wegener Institute for Polar Research, Am Handelshafen 12, 27570 Bremerhaven, Germany

ABSTRACT: Systematic sequencing of environmental SSU rDNA genes amplified from different marine ecosystems has uncovered novel eukaryotic lineages, in particular within the alveolate and stramenopile radiations. The ecological and geographic distribution of 2 novel alveolate lineages (called Group I and II in previous papers) is inferred from the analysis of 62 different environmental clone libraries from freshwater and marine habitats. These 2 lineages have been, up to now, retrieved exclusively from marine ecosystems, including oceanic and coastal waters, sediments, hydrothermal vents, and permanent anoxic deep waters and usually represent the most abundant eukaryotic lineages in environmental genetic libraries. While Group I is only composed of environmental sequences (118 clones), Group II contains, besides environmental sequences (158 clones), sequences from described genera (8) (*Hematodinium* and *Amoebophrya*) that belong to the Syndiniales, an atypical order of dinoflagellates exclusively composed of marine parasites. This suggests that Group II could correspond to Syndiniales, although this should be confirmed in the future by examining the morphology of cells from Group II. Group II appears to be abundant in coastal and oceanic ecosystems, whereas permanent anoxic waters and hydrothermal ecosystems are usually dominated by Group I. Based upon the similarity of partial sequences, we organized these 2 groups into clusters. The diversity of Group II (16 clusters) is wider than that of Group I (5 clusters). Two clusters from Group I have a widespread distribution and are found in all explored marine habitats. In contrast, all other clusters seem to be limited to specific marine habitats. For example, some clusters belonging to Group I and Group II are only detected in extreme environments (anoxic and hydrothermal vents), whereas many clusters from Group II have only been retrieved from coastal waters. We determined near-complete SSU rRNA gene sequences for 26 environmental clones, selected in order to obtain at least one complete sequence per cluster. Phylogenetic analyses (maximum likelihood, neighbor joining, maximum parsimony, and Bayesian reconstruction) based upon complete sequences all concurred to place both Group I and II as sister lineages of dinoflagellates. This result contradicts several published studies, which placed both groups within dinoflagellates.

KEY WORDS: Alveolates · rRNA environmental sequences · Group I and Group II · Dinoflagellates · Syndiniales · Biogeography

INTRODUCTION

In the past 15 yr, direct sequencing of genes such as the one coding for the small subunit of ribosomal RNA (SSU) in environmental samples has revealed an unexpected diversity within the smallest size fraction of the marine plankton. The ubiquitous presence of completely novel lineages, with no representatives in cultures, has been established for the 3 domains of life, Bacteria (Giovannoni et al. 1990), Archaea (Delong 1992, Fuhrman et al. 1992), and more recently Eukaryota (Díez et al. 2001, López-García et al. 2001, Moon-van der Staay et al. 2001). Two phyla dominate eukaryotic SSU rRNA gene sequences retrieved in the small size fractions from marine environmental samples: alveolates and stramenopiles (López-García et al. 2001, Moon-van der Staay et al. 2001). A phylogenetic analysis of environmental sequences belonging to the marine

*Corresponding author. Email: lguillou@sb-roscoff.fr

stramenopiles has been recently published (Massana et al. 2004b). In this paper, we identify clusters within each group, a necessary step prior to quantitative studies relying on specific oligonucleotide probes detected by fluorescent *in situ* hybridization (FISH) or quantitative PCR (qPCR). Our analysis permits assessment of the ecological distribution of the major clusters in marine waters. Our conclusions aid in choosing clones for which the SSU rRNA gene could be fully sequenced since long sequences are necessary to resolve phylogenetic relationships among related groups.

The objective of the present study parallels that of Massana et al. (2004b) but for 'novel' marine alveolates. Sequences belonging to this group often make up a significant fraction of clones retrieved from marine waters or hydrothermal vent environments (Moon-van der Staay et al. 2001, Edgcomb et al. 2002, Stoeck et al. 2003, Romari & Vaulot 2004, Medlin et al. 2006).

Alveolates are characterized by the presence of membrane-bound flattened vesicles named alveoli as a synapomorphic character (Cavalier-Smith 1993, Patterson 1999). Alveolates include 3 main groups of protists (ciliates, dinoflagellates, and apicomplexans), many of which are parasites, e.g. all apicomplexans including the causative agents of malaria, *Plasmodium*, and of toxoplasmosis, *Toxoplasma*. Alveolates also include important free-living organisms that can be active marine predators (ciliates) or develop all transitional feeding modes from phototrophy to heterotrophy. For example, dinoflagellates can be photosynthetic, phagotrophic, mixotrophic, or even parasitic. Other enigmatic protists have recently been related to these 3 main groups. This was the case for the parasitic Perkinsozoa, which have been considered as the earliest diverging sister lineage to the dinoflagellates (Norén et al. 1999, Saldarriaga et al. 2003, Leander & Keeling 2004) or the predatory flagellates belonging to the genus *Colpodella*, a sister lineage of apicomplexans (Kuvardina et al. 2002, Leander et al. 2003b).

In the present study, we undertake a complete analysis of the genetic diversity of environmental SSU rRNA gene sequences related to 2 novel alveolate lineages named Group I and Group II by several authors (López-García et al. 2001, Moreira & López-García 2002). These environmental sequences were obtained from 62 environmental genetic libraries (Table 1), which cover a range of aquatic ecosystems, in particular fresh, coastal, and oceanic waters, sediments, as well as deep-sea hydrothermal vents (Table 1). We first analyzed partial SSU rRNA sequences. This helped us to delineate clusters and to analyze their ecological distribution. We then determined complete SSU rRNA gene sequences for 26 environmental clones, and analyzed the position of Groups I and II within alveolates by different phylogenetic methods. We finally discuss the ecological distribution and genetic diversity of these 2 novel alveolate lineages.

MATERIALS AND METHODS

Genetic libraries. We analyzed 62 environmental genetic libraries of the 18S rRNA gene (Table 1). Sample filtration, primers, and number of clones analyzed for each library are listed in Table 1. All genetic libraries used in this study have been published, with the exception of OLI04660, for which water was collected from 60 m depth in the equatorial Pacific Ocean (1°S, 150°W) during the OLIPAC cruise in November 1994. Sample collection, DNA extraction, and amplification were as described in Moon-van der Staay et al. (2001). PCR products were cloned using the TOPO-TA cloning kit (Invitrogen) following the manufacturer's recommendations. The presence of 18S rRNA inserts in colonies was checked by PCR amplification. Positive clones were analyzed by RFLP using the restriction enzyme *Hae*III as explained in Romari & Vaulot (2004). Cloned fragments representative of each RFLP type were partially sequenced (550 bp) by Qiagen Genomics Sequencing Services using the internal primer Euk528f (5′-CCG CGG TAA TTC CAG CTC-3′).

For a subset of libraries (Roscoff in France, Blanes in Spain, Helgoland in Germany, and Olipac in the Pacific Ocean), we determined near-complete SSU rRNA sequences (about 1800 bp) for 26 clones for which only partial sequences were previously available (Massana et al. 2004a, Romari & Vaulot 2004, Medlin et al. 2006). For these sequences, the original names of the clones are conserved and new accession numbers are provided. For Roscoff and Olipac clones, sequencing reactions were performed with a Big Dye v.3 kit (Applied Biosystems) and an ABI PRISM model 3100 automated sequencer (Applied Biosystems). For Helgoland and Blanes clones, sequences were obtained from the Qiagen Genomics Sequencing Services. Sequences were submitted to the CHECK_CHIMERA program at the Ribosomal Database project (www.rdp.cme.msu.edu/html) and were subjected to a BLAST search on the National Center for Biotechnology Information web server (www.ncbi.nlm.nih.gov). New sequences were deposited in GenBank: DQ138055 to DQ138059 for partial sequences from Equatorial Pacific; DQ145103 to DQ145114 for complete sequences from Equatorial Pacific; DQ186525 to DQ186538 for complete sequences from coastal waters (Roscoff, Blanes, and Helgoland).

Phylogenetic analyses. Environmental sequences belonging to novel alveolates Group I and Group II from the genetic libraries listed in Table 1 and sequences from a selection of related organisms belong-

Table 1. Characteristics of the rRNA genetic clone libraries used in this study

No.	Code sample	Origin	Ecosystem/ Sample type	Coordinates	Date	Depth	Size fraction (µm)	No. of clones	Screening method	Primers (see also Fig. 1)	Groups I or II	Source
1	ME1	Mediterranean Alborán Sea	Oceanic surface waters	36°14′N, 04°15′W	9 Nov 1997	Surface	0.2 to 5	64	RFLP	EukA, EukB	YES	Díez et al. (2001)
2	ANT37	Antarctica Weddell Sea	Oceanic surface waters	60°32′S, 44°12′W	26 Jan 1998	Surface	0.2 to 1.6	58	RFLP	EUK309F, EukB	NO	Díez et al. (2001)
3	ANT12	Antarctica Scotia Sea	Oceanic surface waters	58°16′S, 44°27′W	23 Jan 1998	Surface	0.2 to 1.6	67	RFLP	EUK309F, EukB	NO	Díez et al. (2001)
4	NA11	North Atlantic	Oceanic surface waters	59°30′N, 21°08′W	14 Jun 1998	Surface	0.2 to 2	17	RFLP	EUK309F, EukB	NO	Díez et al. (2001)
5	NA37	North Atlantic	Oceanic surface waters	59°34′N, 21°03′W	21 Jun 1998	Surface	0.2 to 2	20	RFLP	EUK309F, EukB	YES	Díez et al. (2001)
6	DH144	Antarctic	Oceanic deep waters	59°19′S, 55°45′W	Dec 1998	250 m	0.2 to 5	24 to 104[a]	Sequencing	EK-1F, EK-82F, EK-1520R	YES	López-García et al. (2001)
7	DH145	Antarctic	Oceanic deep waters	59°19′S, 55°45′W	Dec 1998	500 m	0.2 to 5	24 to 104[a]	Sequencing	EK-1F, EK-82F, EK-1520R	YES	López-García et al. (2001)
8	DH147	Antarctic	Oceanic deep waters	59°19′S, 55°45′W	Dec 1998	2000 m	0.2 to 5	24 to 104[a]	Sequencing	EK-1F, EK-82F, EK-1520R	YES	López-García et al. (2001)
9	DH148	Antarctic	Oceanic deep waters	59°19′S, 55°45′W	Dec 1998	3000 m	0.2 to 5	24 to 104[a]	Sequencing	EK-1F, EK-82F, EK-1520R	YES	López-García et al. (2001)
10	DH148-5	Antarctic	Oceanic deep waters	59°19′S, 55°45′W	Dec 1998	3000 m	>5	24 to 104[a]	Sequencing	EK-1F, EK-82F, EK-1520R	YES	López-García et al. (2001)
11	OLI	Equatorial Pacific	Oceanic surface waters	150°W, 11°5S	1 Nov 1994	75 m	0.2 to 3	103	Sequencing	Euk328f, Euk329r	YES	Moon-van der Staay et al. (2001)
12	OLI04660	Equatorial Pacific	Oceanic surface waters	150°W, 1°S	1 Nov 1994	60 m	0.2 to 3	17	Sequencing	Euk328f, Euk329r	YES	This study
13	BAQ	Berkeley, CA	Anoxic marine sediment	NA	NA	1 to 3 cm below surface	Whole community	1000	RFLP	82FE/1391RE, and 360FE/1492R	NO	Dawson & Pace (2002)
14	BOL	Bolinas, CA	Anoxic marine sediment	NA	NA	1 to 3 cm below surface	Whole community	1000	RFLP	82FE/1391RE	NO	Dawson & Pace (2002)
15	LEM	Bloomington, IN	Anoxic freshwater sediment	NA	NA	3 to 5 cm below surface	Whole community	1000	RFLP	360FE/1492R	NO	Dawson & Pace (2002)
16	A1	Southern Guaymas vent field	Hydrothermal vent sediment	NA	1998	0 to 1 cm below surface	Whole community	45	Sequencing	Nested PCR U514F/ U1492R + E528F/U1492R	YES	Edgcomb et al. (2002)
17	A2	Southern Guaymas vent field	Hydrothermal vent sediment	NA	1998	1 to 2 cm below surface	Whole community	45	Sequencing	Nested PCR U514F/ U1492R + E528F/U1492R	YES	Edgcomb et al. (2002)
18	A3	Southern Guaymas vent field	Hydrothermal vent sediment	NA	1998	2 to 3 cm below surface	Whole community	45	Sequencing	Nested PCR U514F/ U1492R + E528F/U1492R	NO	Edgcomb et al. (2002)
19	C1	Southern Guaymas vent field	Hydrothermal vent sediment	NA	1998	0 to 1 cm below surface	Whole community	45	Sequencing	Nested PCR U514F/ U1492R + E528F/U1492R	YES	Edgcomb et al. (2002)
20	C2	Southern Guaymas vent field	Hydrothermal vent sediment	NA	1998	1 to 2 cm below surface	Whole community	45	Sequencing	Nested PCR U514F/ U1492R + E528F/U1492R	YES	Edgcomb et al. (2002)
21	C3	Southern Guaymas vent field	Hydrothermal vent sediment	NA	1998	2 to 3 cm below surface	Whole community	45	Sequencing	Nested PCR U514F/ U1492R + E528F/U1492R	YES	Edgcomb et al. (2002)
22	CS	Southern Guaymas vent field	Hydrothermal vent sediment	NA	1998	Surface of sediment	Whole community	45	Sequencing	Nested PCR U514F/ U1492R + E528F/U1492R	YES	Edgcomb et al. (2002)

Table continued on next 2 pages

Table 1. (continued)

No.	Code sample	Origin	Ecosystem/ Sample type	Coordinates	Date	Depth	Size fraction (µm)	No. of clones	Screening method	Primers (see also Fig. 1)	Groups I or II	Source
23	RA000412	English Channel	Coastal surface waters	48°45′N, 4°00′W	12 Apr 2000	Surface	0.2 to 3	100	RFLP	Euk328f, Euk329r	YES	Romari & Vaulot (2004)
24	RA000609	English Channel	Coastal surface waters	48°45′N, 4°00′W	9 Jun 2000	Surface	0.2 to 3	100	RFLP	Euk328f, Euk329r	YES	Romari & Vaulot (2004)
25	RA000907	English Channel	Coastal surface waters	48°45′N, 4°00′W	7 Sep 2000	Surface	0.2 to 3	100	RFLP	Euk328f, Euk329r	YES	Romari & Vaulot (2004)
26	RA001219	English Channel	Coastal surface waters	48°45′N, 4°00′W	19 Dec 2000	Surface	0.2 to 3	100	RFLP	Euk328f, Euk329r	YES	Romari & Vaulot (2004)
27	RA010412	English Channel	Coastal surface waters	48°45′N, 4°00′W	12 Apr 2001	Surface	0.2 to 3	100	RFLP	Euk328f, Euk329r	YES	Romari & Vaulot (2004)
28	RA010516	English Channel	Coastal surface waters	48°45′N, 4°00′W	16 May 2001	Surface	0.2 to 3	100	RFLP	Euk328f, Euk329r	YES	Romari & Vaulot (2004)
29	RA010613	English Channel	Coastal surface waters	48°45′N, 4°00′W	13 Jun 2001	Surface	0.2 to 3	100	RFLP	Euk328f, Euk329r	YES	Romari & Vaulot (2004)
30	RD010614	English Channel	Coastal surface waters	48°38′N, 3°51′W	14 Jun 2001	Surface	0.2 to 3	100	RFLP	Euk328f, Euk329r	YES	Romari & Vaulot (2004)
31	BL000921	NW Mediterranean Sea	Coastal surface waters	41°40′N, 2°48′E	21 Sep 2000	Surface	0.2 to 3	78	RFLP	EukA, EukB	YES	Massana et al. (2004a)
32	BL001221	NW Mediterranean Sea	Coastal surface waters	41°40′N, 2°48′E	21 Dec 2000	Surface	0.2 to 3	111	RFLP	EukA, EukB	YES	Massana et al. (2004a)
33	BL010320	NW Mediterranean Sea	Coastal surface waters	41°40′N, 2°48′E	20 Mar 2001	Surface	0.2 to 3	88	RFLP	EukA, EukB	YES	Massana et al. (2004a)
34	BL010625	NW Mediterranean Sea	Coastal surface waters	41°40′N, 2°48′E	25 Jun 2001	Surface	0.2 to 3	106	RFLP	EukA, EukB	YES	Massana et al. (2004a)
35	He000323	North Sea	Coastal surface waters	54°11′N, 7°54′E	23 Mar 2000	Surface	0.2 to 3	150-300[a]	RFLP	EukA, EukB	YES	Medlin et al. (2005)
36	He000427	North Sea	Coastal surface waters	54°11′N, 7°54′E	27 Apr 2000	Surface	0.2 to 3	150-300[a]	RFLP	EukA, EukB	YES	Medlin et al. (2005)
37	He000803	North Sea	Coastal surface waters	54°11′N, 7°54′E	3 Aug 2000	Surface	0.2 to 3	150-300[a]	RFLP	EukA, EukB	YES	Medlin et al. (2005)
38	He001005	North Sea	Coastal surface waters	54°11′N, 7°54′E	5 Oct 2000	Surface	0.2 to 3	150-300[a]	RFLP	EukA, EukB	YES	Medlin et al. (2005)
39	He001206	North Sea	Coastal surface waters	54°11′N, 7°54′E	6 Dec 2000	Surface	0.2 to 3	150-300[a]	RFLP	EukA, EukB	YES	Medlin et al. (2005)
40	He010218	North Sea	Coastal surface waters	54°11′N, 7°54′E	18 Feb 2001	Surface	0.2 to 3	150-300[a]	RFLP	EukA, EukB	YES	Medlin et al. (2005)
41	Or000415	North Atlantic	Orkney Island	NA	15 Apr 2000	Surface	0.2 to 3	150-300[a]	RFLP	EukA, EukB	YES	Medlin et al. (2005)
41	CAR_D	Caribbean Sea	Anoxic/oxic oceanic deep waters	10°50′N, 64°66′W	8 May 2002	270 m	Whole community	NA	RFLP	EukA, EukB	YES	Stoeck et al. (2003)
42	CAR_E	Caribbean Sea	Anoxic oceanic deep waters	10°50′N, 64°66′W	8 May 2002	340 m	Whole community	NA	RFLP	EukA, EukB	YES	Stoeck et al. (2003)
43	CAR_H	Caribbean Sea	Anoxic oceanic deep waters	10°50′N, 64°66′W	8 May 2002	900 m	Whole community	NA	RFLP	EukA, EukB	YES	Stoeck et al. (2003)
44	CCW	Cape Cod, MA	Salt marsh, water column	NA	NA	Water column	0.2 to 5	NA	RFLP	EukA, EukB	YES	Stoeck & Epstein (2003)
45	CCI	Cape Cod, MA	Salt marsh, Anoxic water-sediment interface	NA	NA	1 cm depth in sed.	Whole community	NA	RFLP	EukA, EukB	NO	Stoeck & Epstein (2003)
46	CCA	Cape Cod, MA	Salt marsh, anoxic sediment	NA	NA	20 cm depth in sed.	Whole community	NA	RFLP	EukA, EukB	NO	Stoeck & Epstein (2003)

Table 1. (continued)

No.	Code sample	Origin	Ecosystem/ Sample type	Coordinates	Date	Depth	Size fraction (µm)	No. of clones	Screening method	Primers (see also Fig. 1)	Groups I or II	Source
47	AT4	Mid-Atlantic Ridge Rainbow	Hydrothermal vent sediment	36°60′N, 33°11′W	NA	2264 m	Whole community	25 to 100[a]	Sequencing	18S-42F, 18S-82F, 18S-1498R, 18S-1520R	YES	López-García et al. (2003)
48	AT6	Mid-Atlantic Ridge Rainbow	Hydrothermal fluid	36°60′N, 33°11′W	NA	2264 m	Whole community	25 to 100[a]	Sequencing	18S-42F, 18S-82F, 18S-1498R, 18S-1520R	NO	López-García et al. (2003)
49	AT9	Mid-Atlantic Ridge Lucky strike	Hydrothermal fluid	37°17′N, 32°16′W	NA	1695 m	Whole community	25 to 100[a]	Sequencing	18S-42F, 18S-82F, 18S-1498R, 18S-1520R	NO	López-García et al. (2003)
50	AT2	Mid-Atlantic Ridge Lucky strike	Microcolonizers	37°17′N, 32°16′W	NA	1695 m	Whole community	25 to 100[a]	Sequencing	18S-42F, 18S-82F, 18S-1498R, 18S-1520R	NO	López-García et al. (2003)
51	AT5	Mid-Atlantic Ridge		NA	NA	NA	Whole community	25 to 100[a]	Sequencing	18S-42F, 18S-82F, 18S-1498R, 18S-1520R	NO	López-García et al. (2003)
52	AT8	Mid-Atlantic Ridge		NA	NA	NA	Whole community	25 to 100[a]	Sequencing	18S-42F, 18S-82F, 18S-1498R, 18S-1520R	YES	López-García et al. (2003)
53	SEY	The Seymas river (Geneva, Switzerland)	Fresh water sediment	NA	NA	NA	Whole community	81	Sequencing	s12.2, sB	NO	Berney et al. (2004)
54		Lake Godivelle	Oligotrophic freshwater	NA	Summer 2002	Euphotic zone	0.2 to 5	50	RFLP	EK-1F, EK-1520R	NO	Lefranc et al. (2005)
55		Lake Pavin	Oligo-mesotrophic fresh water	NA	Summer 2002	Euphotic zone	0.2 to 5	50	RFLP	EK-1F, EK-1520R	NO	Lefranc et al. (2005)
56		Lake Aydat	Eutrophic freshwater	NA	Summer 2002	Euphotic zone	0.2 to 5	50	RFLP	EK-1F, EK-1520R	NO	Lefranc et al. (2005)
57		Spain's river of fire	Acidic freshwater	NA	NA	NA	Whole community	NA	NA	NA	NO	Amaral-Zettler et al. (2002)
58	CH1	Haute Vallée de la Chevreuse	Sediment from oxic pool	NA	Early winter		Whole community	NA	Sequencing	18s-42F, 18S-82F, 18S-1498R, 18S-1520R	NO	Slapeta et al. (2005)
59	CH1	Haute Vallée de la Chevreuse	Freshwater from oxic pool	NA	Early winter		0.2 to 5 and >5	NA	Sequencing	18s-42F, 18S-82F, 18S-1498R, 18S-1520R	NO	Slapeta et al. (2005)
60	CV1	University Paris-Sud	Sediment from suboxic pond	NA	Early winter		Whole community	NA	Sequencing	18s-42F, 18S-82F, 18S-1498R, 18S-1520R	NO	Slapeta et al. (2005)
61	CV1	University Paris-Sud	Freshwater from suboxic pond	NA	Early winter		0.2 to 5 and >5	NA	Sequencing	18s-42F, 18S-82F, 18S-1498R, 18S-1520R	NO	Slapeta et al. (2005)
62	CV1	University Paris-Sud	Sand from suboxic pond	NA	Early winter		Whole community	NA	Sequencing	18s-42F, 18S-82F, 18S-1498R, 18S-1520R	NO	Slapeta et al. (2005)

[a] Not detailed in the original publication

ing to alveolates were imported from GenBank. Three different alignments were built, i.e. partial sequences belonging to Group I, partial sequences belonging to Group II, and complete sequences of representative alveolates. Sequences were first aligned using the ClustalW multiple alignment tool integrated to the sequence editor BioEdit (Hall 1999), and alignments were then improved by hand. These 3 alignments are available from our web site: www.sb-roscoff.fr/Phyto/index.php?option=com_docman&task=doc_download&gid=216. Regions with significant variation were automatically removed using the Gblocks software (www1.imim.es/~castresa/Gblocks/Gblocks.html), with optimized parameters for rRNA alignments (minimum length of a block, 5; allowing gaps in half of the positions). Each alignment was analyzed by 3 different phylogenetic methods: maximum likelihood (ML), neighbor joining (NJ), and maximum parsimony (MP) using PAUP*4.0b10 (Swofford 2002). Different nested models of DNA substitution and associated parameters were estimated using Modeltest v.3.06 (Posada & Crandall 1998). These settings were used to perform ML and NJ analyses. A heuristic search procedure using the tree bisection/reconnection branch swapping algorithm (setting as in MP) was performed to find the optimal ML tree topology. In each case, we tried more than 70 000 rearrangements. For MP, the number of rearrangements was limited to 5000 for bootstrap replicates. Starting trees were obtained by randomized stepwise addition (number of replicates = 20). Bootstrap values for NJ and MP were estimated from 1000 replicates. We additionally used Bayesian reconstruction for the analysis of complete sequences with MrBayes, v.3.0b4 (Huelsenbeck & Ronquist 2001). The GTR model of substitution was used, taking into account a gamma-shaped distribution of the rates of substitution among sites. For each dataset, the chains were run for 2 000 000 generations. Trees were sampled every 100 generations. The first 5000 sampled trees, corresponding to the initial phase before the chains reach stationarity (burn-in), were discarded.

RESULTS

Overall, 38 of the 62 SSU rRNA gene (rDNA) environmental genetic libraries analyzed contained sequences belonging to the novel alveolate lineages called Group I and Group II (Table 1). Size fraction, primers used during the PCR amplification, and number of clones analyzed were important factors of variability (Table 1). For example, Group I and Group II were not detected for size fractions smaller than 1.6 µm (libraries ANT37 and ANT12, Table 1). Coverage, the number of clones actually sequenced, is also critical. This probably explains why no environmental novel alveolate sequences were detected in library NA11 for which coverage is very low (17 clones). In fact, alveolates were detected in library NA37, built by the same authors using a sample from the same location (Table 1). Primers may also play a role on the PCR amplification of such organisms. Primers used for the 62 libraries (Fig. 1) are quite similar. They may have identical nucleotide composition but different names (18S-82F and EK-82F), have different lengths but still be located in the same region (EukA/Euk328f/EK-1F and sB/ EukB), or differ by the addition of degenerate positions (82FE compared to 18S-82F, 1492R compared to U1492R, Euk329r compared to sB). All of these primers should in theory allow amplification of the majority of known environmental SSU rDNA sequences belonging to the 2 novel alveolate lineages (all the nucleotide positions have <10% of mismatches with the available complete SSU rDNA sequences), with 2 exceptions. Primer 360FE has a mismatch with almost all sequences belonging to Group II (the T at position 14 is usually a C in Group II), whereas primer s12.2 has a mismatch with almost all sequences belonging to Group I (the G at position 15 of the primer is usually a C in Group I). The 3′ end of reverse primers Euk329r, sB, EukB, and 18S-1520R does not target some sequences of *Amoebophrya* strains, which belong to Group II (AF472553, AF472554, AF472555, AY208892, AY208893, AY208894). Most environmental genetic libraries have been constructed using EukA/ EukB or Euk328f/Euk329r, which are quite similar in their position and base composition and therefore probably amplify similarly the sequences of these organisms.

Neither of the lineages was ever detected in freshwater samples (seawater or sediments), whatever the methodology used (Table 1). They were also absent from salt marshes (libraries 45, 46) and anoxic sediments (libraries 13, 14). They were not observed in some hydrothermal ecosystems, such as hydrothermal fluids or microcolonizers (library 18, 48 to 51) but present in others (libraries 16, 17, 19 to 22, 47). In the end, a total of 284 environmental partial sequences were considered (118 belonging to Group I and 158 to Group II).

Generally speaking (although there are some notable exceptions), Group II clones are abundant in coastal and oceanic ecosystems (Moon-van der Staay et al. 2001, Romari & Vaulot 2004, Medlin et al. 2006), whereas permanent anoxic waters and hydrothermal ecosystems are usually dominated by Group I (Edgcomb et al. 2002, Stoeck & Epstein 2003). The phylogeny of Group I and Group II was analyzed separately by ML, NJ, and MP. Clusters were defined within these 2 groups using the following set of criteria: (1) a cluster contains environmental sequences

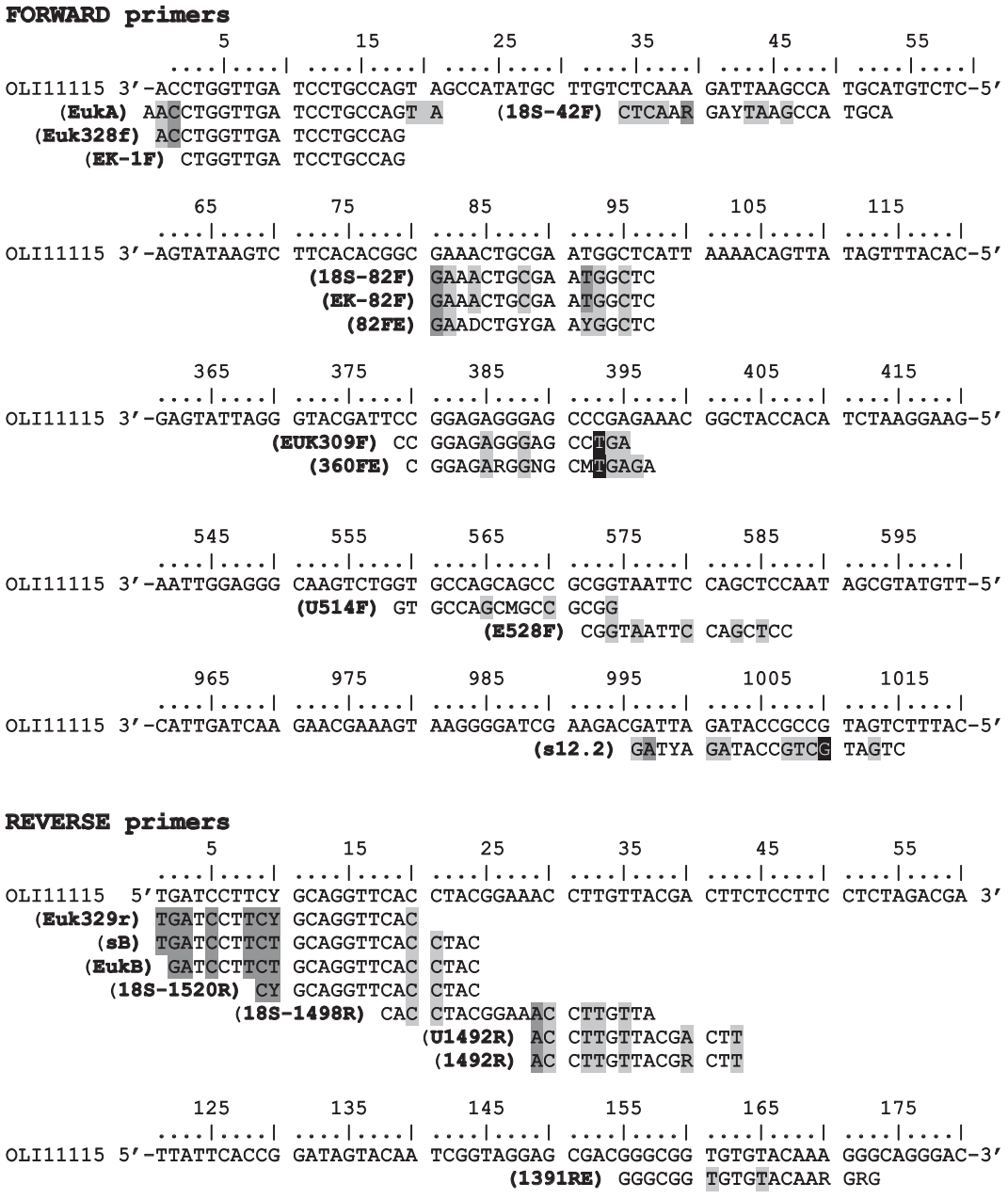

Fig. 1. Primers used to amplify SSU rDNA eukaryotic sequences for the 62 genetic libraries analyzed in this study. The number of complete environmental SSU rRNA sequences that have different nucleotide composition from primers is indicated by different tones. White: all sequences match the primer; light grey, dark grey, and black correspond, respectively, to 1 to 5%, 5 to 10%, and >10% of the sequences having a different nucleotide composition at this position. The environmental sequence named OLI11115 is used as reference for the localization of primers (3′ to 5′ direction for forward primers and 5′ to 3′ direction for reverse primers)

from at least 3 different sources (different genetic libraries or cultivated species), (2) a cluster is supported by the 3 tree construction methods (NJ, ML and MP), and (3) bootstrap values at the cluster node are higher than 80% with NJ and MP. For clusters fulfilling these criteria, the degree of identity percent among environmental sequences belonging to the same cluster is generally higher than 90%.

Within Group I, 5 clusters were detected (Fig. 2). Fifty percent of the environmental sequences belong to cluster I-1. This cluster contains sequences from all marine habitats considered (coastal waters, surface oceanic waters, deep waters, and hydrothermal vent sediments). This cluster is for example present in all libraries from Blanes (NW Mediterranean Sea), throughout the whole year. Its occurrence is restricted

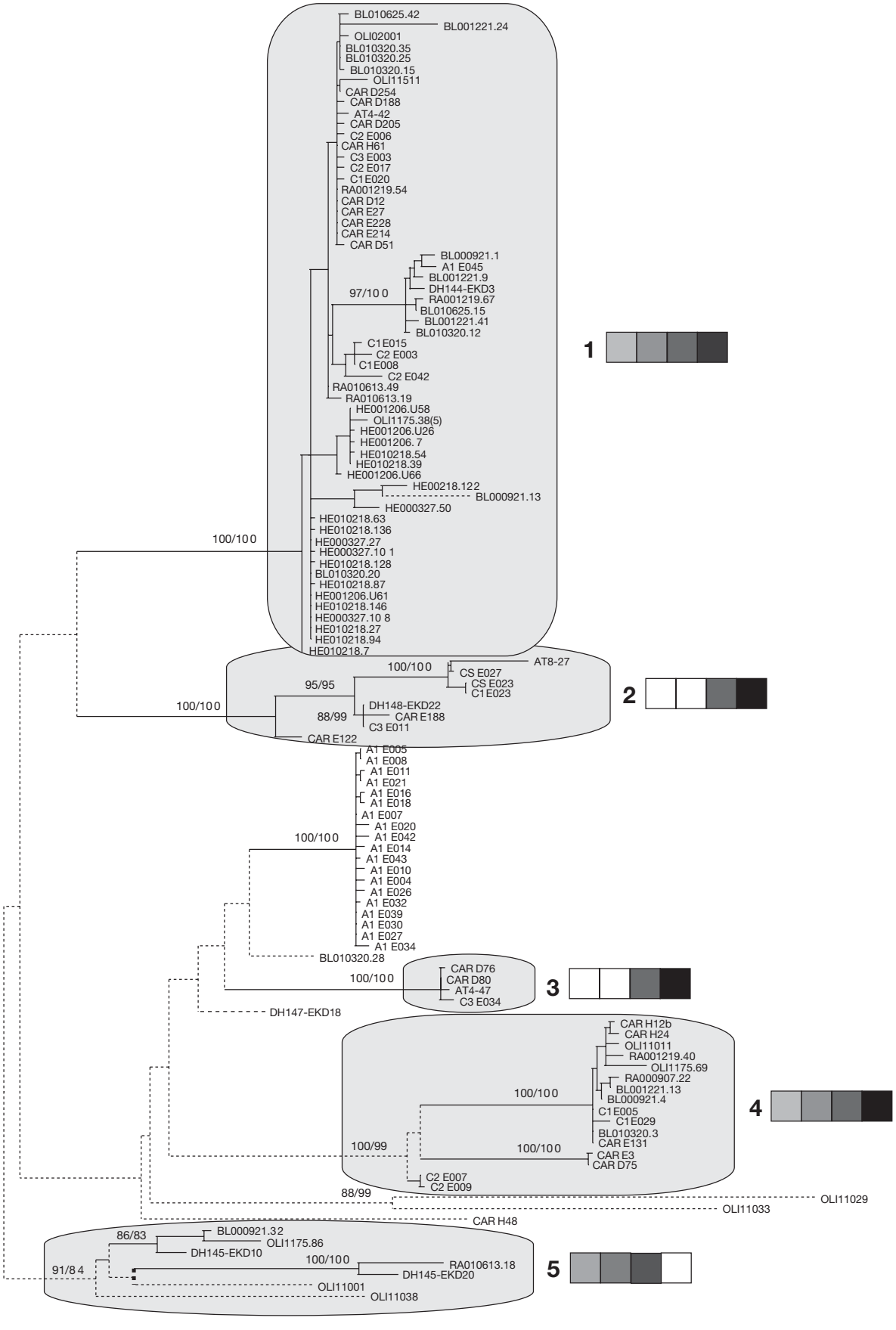
BL010625.42
BL001221.24
OLI02001
BL010320.35
BL010320.25
BL010320.15
OLI11511
CAR D254
CAR D188
AT4-42
CAR D205
C2 E006
CAR H61
C3 E003
C2 E017
C1 E020
RA001219.54
CAR D12
CAR E27
CAR E228
CAR E214
CAR D51
BL000921.1
A1 E045
BL001221.9
DH144-EKD3
RA001219.67
BL010625.15
BL001221.41
BL010320.12
97/100
C1 E015
C2 E003
C1 E008
C2 E042
RA010613.49
RA010613.19
HE001206.U58
OLI1175.38(5)
HE001206.U26
HE001206. 7
HE010218.54
HE010218.39
HE001206.U66
HE00218.122
BL000921.13
HE000327.50
HE010218.63
HE010218.136
HE000327.27
HE000327.10 1
HE010218.128
BL010320.20
HE010218.87
HE001206.U61
HE010218.146
HE000327.10 8
HE010218.27
HE010218.94
HE010218.7
100/100
1
AT8-27
100/100
CS E027
CS E023
C1 E023
95/95
DH148-EKD22
CAR E188
88/99
C3 E011
100/100
CAR E122
2
A1 E005
A1 E008
A1 E011
A1 E021
A1 E016
A1 E018
A1 E007
A1 E020
100/100
A1 E042
A1 E014
A1 E043
A1 E010
A1 E004
A1 E026
A1 E032
A1 E039
A1 E030
A1 E027
A1 E034
BL010320.28
CAR D76
CAR D80
100/100
AT4-47
C3 E034
3
DH147-EKD18
CAR H12b
CAR H24
OLI11011
RA001219.40
OLI1175.69
RA000907.22
100/100
BL001221.13
BL000921.4
C1 E005
C1 E029
BL010320.3
CAR E131
4
100/99
100/100
CAR E3
CAR D75
C2 E007
C2 E009
88/99
OLI11029
OLI11033
CAR H48
86/83
BL000921.32
OLI1175.86
DH145-EKD10
100/100
RA010613.18
91/84
DH145-EKD20
OLI11001
OLI11038
5
0.1

Fig. 2. Phylogenetic analysis of partial SSU rRNA sequences belonging to novel alveolates Group I. Maximum likelihood tree based upon 118 environmental sequences. Three sequences of apicomplexans (*Sarcocystis muris*, *Toxoplasma gondii*, and *Cyclospora colobii*) were treated as outgroup (removed from the tree for clarity). The Gblock software retained 480 positions for phylogenetic analyses from the 516 initial nucleotides. A GTR+G model was selected by Modeltest 3.06 using the following parameters: Lset Base = (0.2627 0.1810 0.2622), Nst = 6, Rmat = (1.0000 2.8775 1.0000 1.0000 4.6908), Rates = gamma, Shape = 0.5006, Pinvar = 0 (–lnL = 6309.34567). The scale bar corresponds to 0.1% sequence divergence. Neighbor-joining and maximum parsimony bootstrap values are indicated at the nodes of major branches (1000 replicates in each case, values >70% shown). Dark lines join sequences that share more than 90% sequence identity. Clusters are indicated in grey (for the procedure of cluster definition see text). The marine habitats from which sequences have been retrieved are indicated for each cluster by tones of different % of black, from left to right. Left and 20% grey: coastal ecosystems; middle left and 40% grey: oceanic surface waters; middle right and 60% grey: deep waters (oxic or anoxic); right and 80% grey: hydrothermal vents. White indicates absence of environmental sequences belonging to novel alveolates Group I in corresponding habitats (result restricted to genetic libraries analyzed in this study)

to 2 genetic libraries from Roscoff in the English Channel (RA001219 and RA010613), and 3 genetic libraries from Helgoland in the North Sea (He000323, He001206 and He010218). Cluster I-4 is also widespread in marine ecosystems. In contrast, clusters I-2 and I-3 have been exclusively retrieved from deep marine waters or hydrothermal vent sediments. Both clusters include environmental sequences retrieved from the Atlantic (López-García et al. 2003) as well as from the Pacific (Edgcomb et al. 2002). Finally, cluster I-5 is only composed of sequences retrieved from 3 different genetic libraries, namely from Pacific surface waters, Antarctic deep waters, and NW Mediterranean coastal waters (Fig. 2).

In comparison, the genetic diversity of alveolate Group II is much wider (Fig. 3). We were able to define 16 clusters. However, 28 environmental sequences were too divergent to be included in any of these clusters; therefore, additional clusters probably exist and will have to wait for additional sequences in order to be further characterized. Most of the clusters were retrieved exclusively from surface marine waters (clusters II-1, 2, 3, 4, 5, 8, 10, 11, 12, 13, and 16), whereas 2 clusters (II-9 and 15) are restricted to extreme environments (anoxic or deep waters and hydrothermal sediments). In contrast to Group I, no cluster contains sequences retrieved from both surface waters and deep ecosystems. Sequences from the parasite *Amoebophrya* spp. are present in 4 independent clusters (II-1, 2, 3, and 4), that appear to be restricted to coastal and oceanic surface waters. Clusters II-1, 3, 4, 8, and 10 contain the largest number of sequences.

Based on this analysis of partial sequences, we selected 26 clones for complete sequencing of the rRNA gene. We aimed at obtaining at least one complete sequence per cluster. Analysis of these sequences revealed a probable chimera, the sequence RA010613.20, for which the beginning of the gene appear to be related to stramenopiles. Environmental sequence DH145-EKD20 was also confirmed to be a chimera (Berney et al. 2004). Both sequences were removed from the analyses of complete sequences. Full length sequences, together with those already available for other representative alveolates, were analyzed using NJ, MP, ML, and Bayesian methods (Fig. 4), the latter 2 providing similar tree topologies.

In agreement with published analyses, ciliates are the most basal lineage to emerge within alveolates according to all 4 methods. Apicomplexa and related flagellates (*Colpodella*, *Cryptosporidium*, *Perkinsus*, and gregarines) are grouped together by MP, whereas *Perkinsus* forms an independent lineage by NJ and Bayesian inference. Group I, Group II, and dinoflagellates are closely related in all our phylogenetic analyses. *Hematodinium* is always placed at the base of Group II, with the exception of MP (position not supported by the bootstrap analysis). The monophyly of Group I and Group II (with the exclusion of *Hematodinium*) is supported by relatively high bootstrap values for NJ and MP (better than 90%). The enigmatic environmental sequence OLI11005 was placed at the base of dinoflagellates with MP and NJ, and at the base of Group II (including *Hematodinium*) with Bayesian inference and ML. None of these placements were supported by bootstrap analyses. This sequence is closely related to the partial environmental sequences of BL010625.44 and RA000609.43.

DISCUSSION

'Novel' marine alveolates belong to 2 independent and well supported lineages, Group I and Group II. Group I is entirely composed of environmental sequences, whereas Group II also includes the parasite *Amoebophrya* (Syndiniales, dinoflagellates). This organism was described by Koeppen as *Hyalosaccus ceratii* more than a century ago (1903), which renders its 'novelty' quite relative. Furthermore, several phylogenetic analyses suggest that the SSU rRNA sequences of *Hematodinium* sp. (Syndiniales), a parasite of the blue crab (Gruebl et al. 2002), is closely related to those of Group II. Therefore the whole Group II may corre-

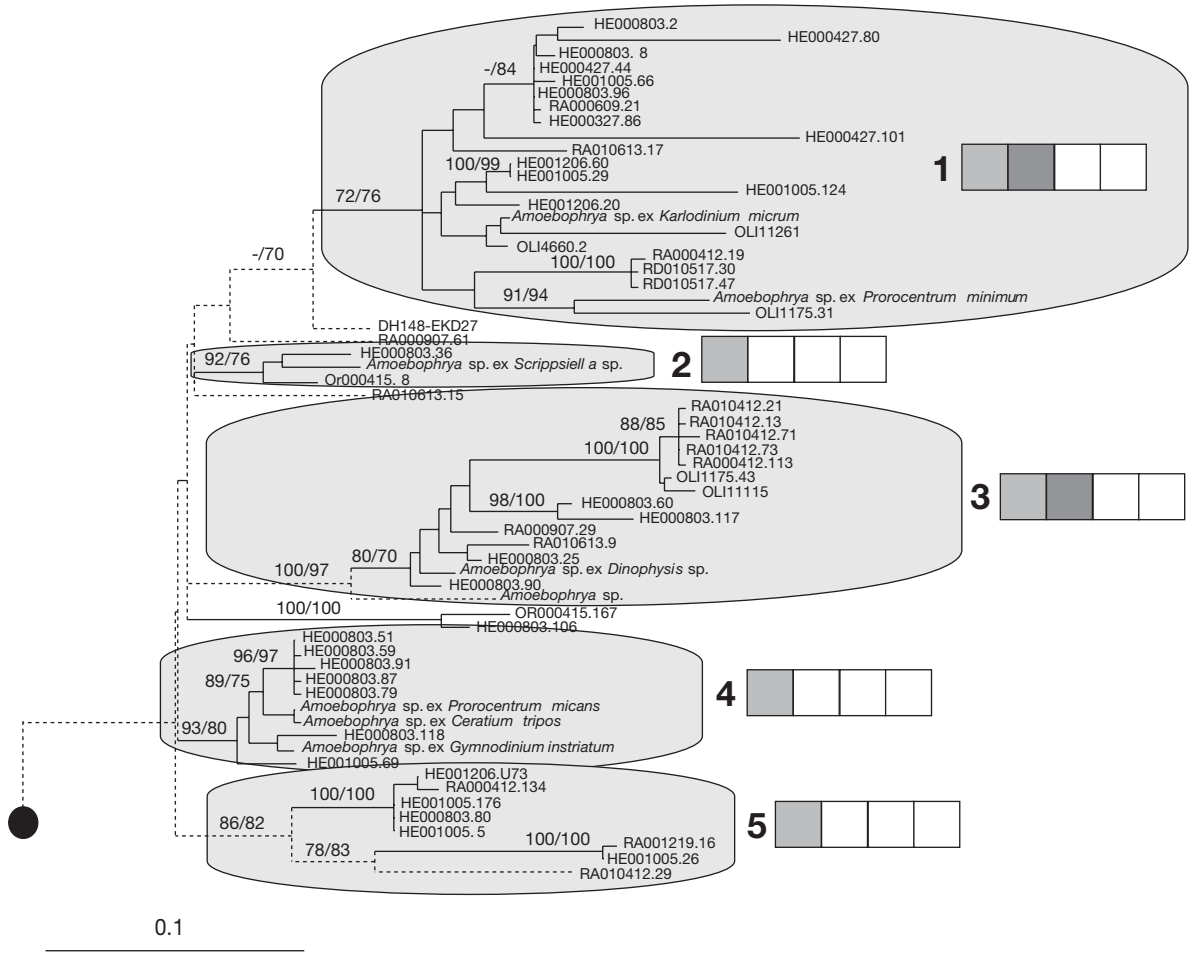

Fig. 3. Phylogenetic analysis of partial SSU rRNA sequences belonging to novel alveolates Group II. Maximum likelihood tree based upon the analysis of 164 sequences. Three sequences of apicomplexans were used as outgroup (removed from the tree for clarity). The Gblock software retained 474 positions for phylogenetic analyses from the initial 534 initial position. A GTR+G model was selected by Modeltest 3.06 using the following parameters: Lset Base = (0.2736 0.1700 0.2545), Nst = 6, Rmat = (1.0000 3.3780 1.2804 1.2804 4.9683), Rates = gamma, Shape = 0.6108, Pinvar = 0.1155 (–lnL 11989.71615). Other definitions as in Fig. 2

spond to the order Syndiniales (= Syndinia) (Corliss 1984), which are partially or entirely intracellular parasites and truly heterotrophic (they lack any chloroplast). They have been observed in a great variety of marine hosts, such as dinoflagellates, radiolarians, ciliates, crabs, or copepod eggs. All these parasites produce motile cells (dinospores) that have the general appearance and swimming behavior of other dinoflagellates. However, important cytological diagnostic features of dinoflagellates, such as the dinokaryon (nucleus with permanently condensed chromosome and without nucleosomal histones), are absent in Syndiniales. Our analysis, which includes 26 novel complete SSU rDNA sequences and relies on 4 different phylogenetic reconstruction methods, places these 2 alveolate lineages as independent groups closely related to dinoflagellates. These results contradict several previous studies that placed the enigmatic dinoflagellate *Noctiluca scintillans* at a more basal position than Groups I and II (Moon-van der Staay et al. 2001, Leander et al. 2003a, Berney et al. 2004, Saldarriaga et al. 2004, Taylor 2004). Nevertheless, even after the addition of a substantial number of novel complete sequences, we are not able to completely elucidate the phylogenetic placement of these lineages, since our analyses are not supported by high bootstrap values. More conserved eukaryotic genes, such as those coding for heat shock protein 90 (hsp90) or actin, that have provided more robust phylogenies of alveolates (Leander & Keeling 2004), are probably required to elucidate the position of these novel lineages relative to that of dinoflagellates.

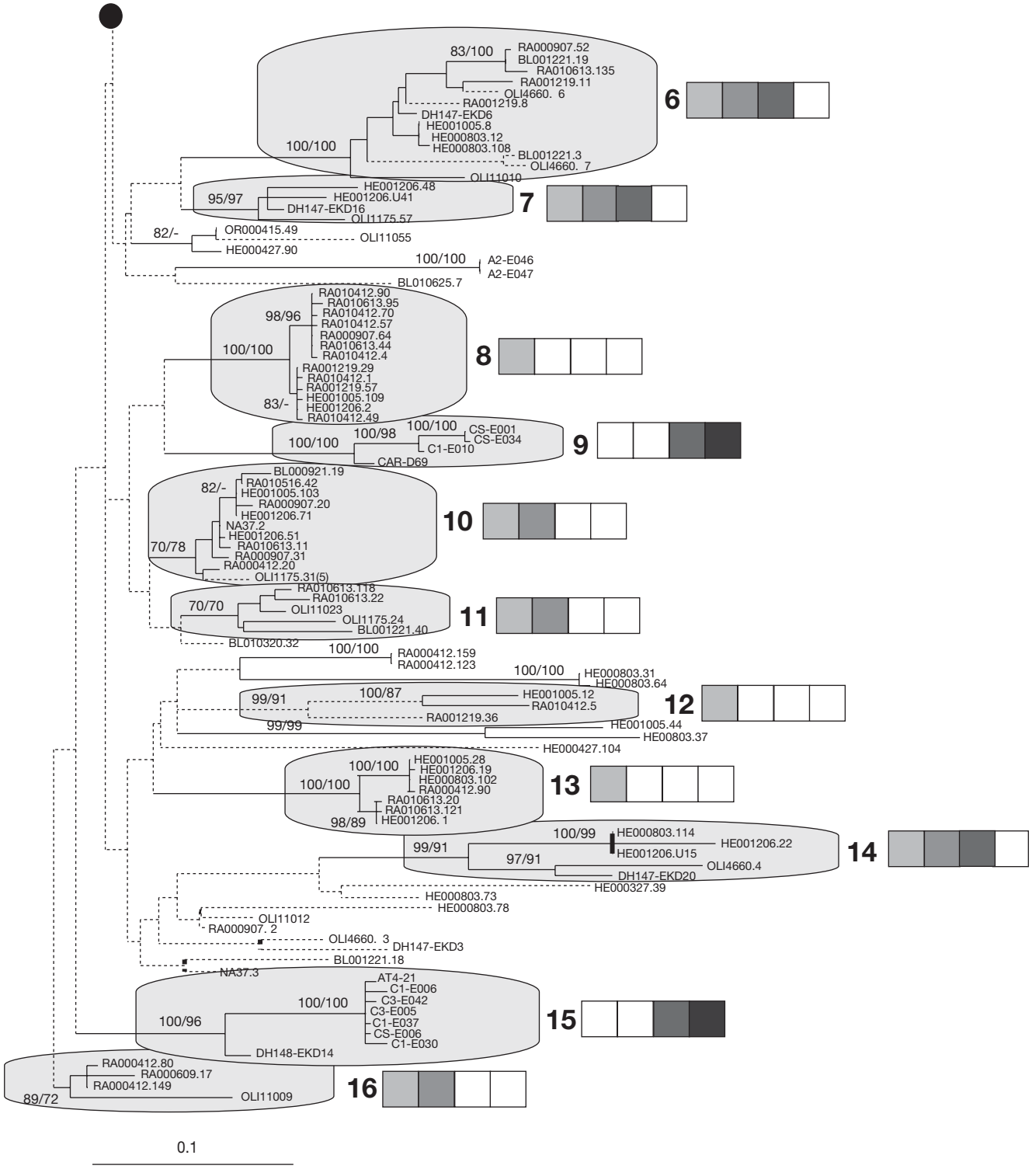

Fig. 3 (continued)

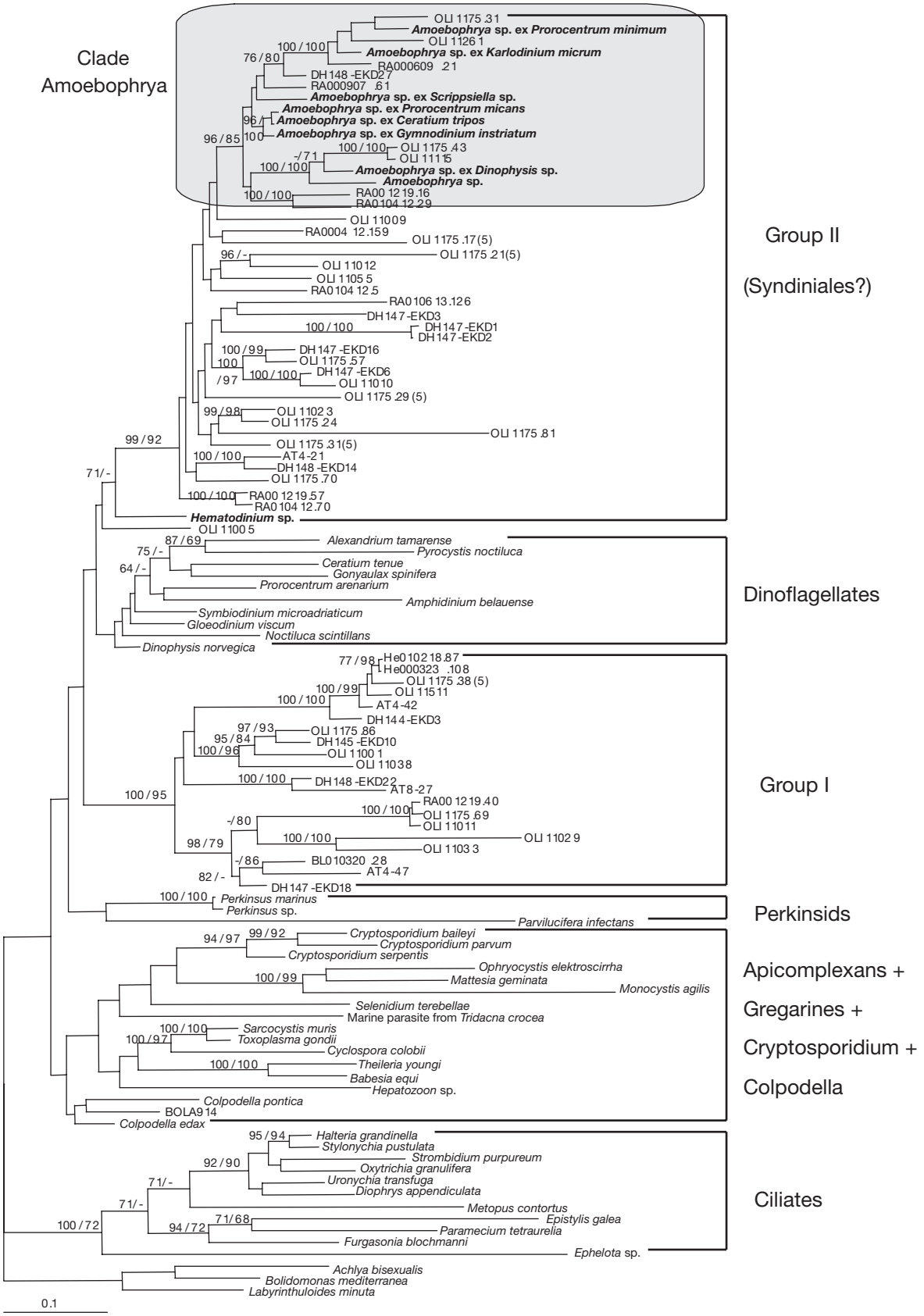
Clade Amoebophrya
OLI 1175.31
Amoebophrya sp. ex Prorocentrum minimum
OLI 11261
Amoebophrya sp. ex Karlodinium micrum
RA000609.21
DH148-EKD27
RA000907.61
Amoebophrya sp. ex Scrippsiella sp.
Amoebophrya sp. ex Prorocentrum micans
Amoebophrya sp. ex Ceratium tripos
Amoebophrya sp. ex Gymnodinium instriatum
OLI 1175.43
OLI 11115
Amoebophrya sp. ex Dinophysis sp.
Amoebophrya sp.
RA0012 19.16
RA0104 12.29
OLI 11009
RA0004 12.159
OLI 1175.17(5)
OLI 1175.21(5)
OLI 11012
OLI 11055
RA0104 12.5
RA0106 13.126
DH147-EKD3
DH147-EKD1
DH147-EKD2
DH147-EKD16
OLI 1175.57
DH147-EKD6
OLI 11010
OLI 1175.29(5)
OLI 11023
OLI 1175.24
OLI 1175.81
OLI 1175.31(5)
AT4-21
DH148-EKD14
OLI 1175.70
RA0012 19.57
RA0104 12.70
Hematodinium sp.
OLI 11005
Alexandrium tamarense
Pyrocystis noctiluca
Ceratium tenue
Gonyaulax spinifera
Prorocentrum arenarium
Amphidinium belauense
Symbiodinium microadriaticum
Gloeodinium viscum
Noctiluca scintillans
Dinophysis norvegica
He010218.87
He000323.108
OLI 1175.38(5)
OLI 11511
AT4-42
DH144-EKD3
OLI 1175.86
DH145-EKD10
OLI 11001
OLI 11038
DH148-EKD22
AT8-27
RA0012 19.40
OLI 1175.69
OLI 11011
OLI 11029
OLI 11033
BL010320.28
AT4-47
DH147-EKD18
Perkinsus marinus
Perkinsus sp.
Parvilucifera infectans
Cryptosporidium baileyi
Cryptosporidium parvum
Cryptosporidium serpentis
Ophryocystis elektroscirrha
Mattesia geminata
Monocystis agilis
Selenidium terebellae
Marine parasite from Tridacna crocea
Sarcocystis muris
Toxoplasma gondii
Cyclospora colobii
Theileria youngi
Babesia equi
Hepatozoon sp.
Colpodella pontica
BOLA914
Colpodella edax
Halteria grandinella
Stylonychia pustulata
Strombidium purpureum
Oxytrichia granulifera
Uronychia transfuga
Diophrys appendiculata
Metopus contortus
Epistylis galea
Paramecium tetraurelia
Furgasonia blochmanni
Ephelota sp.
Achlya bisexualis
Bolidomonas mediterranea
Labyrinthuloides minuta
0.1
Group II
(Syndiniales?)
Dinoflagellates
Group I
Perkinsids
Apicomplexans +
Gregarines +
Cryptosporidium +
Colpodella
Ciliates

Fig. 4. Bayesian phylogeny of alveolates based on the analysis of 108 nearly complete SSU rRNA sequences. Three sequences of stramenopiles were treated as outgroup. The Gblock software retained 1473 positions for phylogenetic analyses from the 2133 initial positions. Neighbor-joining and Maximum parsimony bootstrap values (1000 replicates in each case, values >70% shown) are indicated at the nodes of major branches. The scale bar corresponds to 0.1% sequence divergence. Sequences in dark belong to known parasites from the Syndiniales. The clade, supported by high bootstrap values, formed by *Amoebophrya* and environmental sequences and regrouping clusters 1 to 5 is outlined in grey. This tree topology is identical to that obtained by ML analysis, using a GTR+G+I model (selected by Modeltest 3.06), using the following parameters: Lset Base = (0.2757 0.1787 0.2494), Nst = 6, Rmat = (1.0746 2.9600 1.2514 1.0820 4.7079), Rates = gamma, Shape = 0.7091, Pinvar = 0.2330 (–lnL 39844.13337)

Our analysis, nevertheless, yields some general trends in the ecological distribution of these enigmatic alveolates. Although widespread in marine waters, these alveolate lineages have not been detected in freshwater ecosystems. This is consistent with the fact that all Syndiniales described so far are marine organisms. Groups I and II have been retrieved from very different marine habitats, but some clusters appear to be restricted to specific habitats. In fact, only 2 clusters from Group I (I-1 and I-4) have been retrieved from all the marine habitats explored, which range from marine surface waters (both oceanic and coastal) to anoxic habitats, including deep ocean waters and hydrothermal sediments. Therefore, these 2 clusters must contain anaerobic (or anoxy-tolerant organisms), as well as truly heterotrophic organisms. Within these 2 groups, Group I-1 contains the largest number of environmental clones (59). This is equivalent to some marine stramenopile clusters (MAST) that are widely distributed among different marine systems and that account for most clones in environmental genetic libraries, such as MAST-1, MAST-3, MAST-4 and MAST-7 (Massana et al. 2004b). In contrast, some clusters belonging to either Group I or II have only been detected up to now in 'extreme' environments (anoxic or hydrothermal, see I-2, I-3, II-9, and II-15). Group I-2 and Group I-3 contain very closely related environmental sequences that were retrieved independently from anoxic waters and hydrothermal sediments. In fact, deep-sea vents and anoxic organic rich ecosystems (whale skeletons for example) are known to contain invertebrates that are similar with respect to their phylogeny and ecology, as both ecosystems are highly dependent on symbiotic associations with prokaryotic organisms (Van-Dover et al. 2002). These 2 clusters are not directly related; consequently they probably constitute secondary divergent adaptations to extreme environments rather than having retained common ancestral characters. These clusters could be targeted to identify protists highly specialized to these habitats. Many Group II clusters were only detected in coastal ecosystems. Coastal waters represent 1/3 of the marine habitats that have been explored by environmental eukaryotic rDNA libraries. Consequently, it is probable that these clusters have been undersampled in other ecosystems. Still, Group II is the dominant eukaryotic lineage in the majority of coastal genetic libraries, except in coastal Mediterranean Sea waters (Blanes, Spain). Groups II-1 to II-5 form a monophyletic assemblage that contains several sequences obtained from different isolates belonging to the species complex of *Amoebophrya ceratii*, a parasite that infects marine dinoflagellates. Because these 5 clusters are monophyletic and supported by high bootstrap values, it is reasonable to hypothesize that all these sequences belong to *Amoebophrya ceratii*-like organisms, with a similar parasitic trophic mode. It is important to note that the '*Amoebophrya ceratii*-like' clusters do not contain sequences from sediments or anoxic ecosystems. In fact, all the environmental sequences from the *Amoebophrya ceratii*-clusters have been retrieved from surface waters with the exception of one (DH148-EKD27) obtained from deep waters collected in the Antarctic Ocean. Within the species complex *Amoebophrya ceratii*, strains that are genetically different have different host specificity (Coats & Park 2002, Gunderson et al. 2002). This specificity may drastically increase the genetic diversity of the whole group, each genotype being specific for a reduced number of hosts. However, in our analyses, the phylogeny of these parasites does not appear to be linked with the phylogeny of their hosts. As an example, the strain of *Amoebophrya* that parasites the dinoflagellate *Prorocentrum minimum* is more closely related to the strain of *Amoebophrya* that infect *Karlodinium micrum* than the strain able to infect *Prorocentrum micans* (Fig. 3). The evolutionary history of these parasites is therefore probably different from that of their host. Dinoflagellates contain both autotrophic and heterotrophic organisms and constitute an important component of the plankton able to colonize the whole water column. They are also successful bloomers in coastal waters. Other described species of *Amoebophrya* are able to infect very different protists such as ciliates, Acantharians, or even other parasites (Cachon & Cachon 1987). Considering that this group is widespread, parasitism could therefore be an important, but largely unexplored process in marine ecology, on the same level as viral lysis or flagellate predation.

In summary, the analysis provided in this paper serves 3 purposes. First it provides the first detailed phylogenetic analyses of 2 important novel alveolate

lineages. Second, it suggests strongly that some clusters within each of the 2 lineages correspond to highly specialized habitats or life modes (parasitism). Third, it forms the base to define novel oligonucleotide probes that will help to visualize these organisms in natural samples.

Acknowledgements. This work was supported by the EU project PICODIV (EVK3-CT1999-00021), as well as project funded by the 'Centre de Ressources Biologiques' and the GIS 'Génomique Marine' from France.

LITERATURE CITED

Amaral-Zettler LA, Gomez F, Zettler E, Keenan BG, Amils R, Sogin ML (2002) Microbiology: eukaryotic diversity in Spain's river of fire. Nature 417:137

Berney C, Fahrni J, Pawlowski J (2004) How many novel eukaryotic 'kingdoms'? Pitfalls and limitations of environmental DNA surveys. BMC Biol 2:13

Cachon J, Cachon M (1987) Parasitic dinoflagellates. In: Taylor FJR (ed) The biology of dinoflagellates, Vol 21. Blackwell Scientific Publications, Oxford, p 571–611

Cavalier-Smith T (1993) Kingdom Protozoa and its 18 phyla. Microbiol Rev 57:953–994

Coats DW, Park MG (2002) Parasitism of photosynthetic dinoflagellates by three strains of *Amoebophrya* (Dinophyta): parasite survival, infectivity, generation time, and host specificity. J Phycol 38:520–528

Corliss JO (1984) The kingdom Protista and its 45 Phyla. Biosystems 17:87–126

Dawson SC, Pace NR (2002) Novel kingdom-level eukaryotic diversity in anoxic environments. Proc Natl Acad Sci USA 99:8324–8329

Delong EF (1992) Archaea in coastal marine environments. Proc Natl Acad Sci USA 89:5685–5689

Díez B, Pedrós-Alió C, Massana R (2001) Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. Appl Environ Microbiol 67:2932–2941

Edgcomb VP, Kysela DT, Teske A, de Vera Gomez A, Sogin ML (2002) Benthic eukaryotic diversity in the Guaymas basin hydrothermal vent environment. Proc Natl Acad Sci USA 99:7658–7662

Fuhrman AJ, McCallum K, Davis AA (1992) Novel major archae-bacterial group from marine plankton. Nature 356: 148–149

Giovannoni SJ, Britschgi TB, Moyer CL, Field KG (1990) Genetic diversity in Sargasso Sea bacterioplankton. Nature 345:60–63

Gruebl T, Frischer ME, Sheppard M, Neumann M, Maurer AN, Lee RF (2002) Development of an 18S rRNA gene-targeted PCR-based diagnostic for the blue crab parasite *Hematodinium* sp. Dis Aquat Org 49:61–70

Gunderson JH, John SA, Boman WC, Coats DW (2002) Multiple strains of the parasitic dinoflagellate *Amoebophrya* exist in Chesapeake Bay. J Eukaryot Microbiol 49:469–474

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucl Acids Symp Ser 41:95–98

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755

Koeppen N (1903) *Hyalosaccus ceratii* nov. gen. et sp.' parasit Dinoflagellat'. Zapiski Kiev Obshch 16:89–135

Kuvardina ON, Leander BS, Aleshin VV, Mylnikov AP, Keeling PJ, Simdyanov TG (2002) The phylogeny of colpodellids (Alveolata) using small subunit rRNA gene sequences suggests they are the free-living sister group to apicomplexans. J Eukaryot Microbiol 49:498–504

Leander BS, Clopton RE, Keeling PJ (2003a) Phylogeny of gregarines (Apicomplexa) as inferred from small-subunit rDNA and b-tubulin. Int J Syst Evol Microbiol 53:345–354

Leander BS, Keeling PJ (2004) Early evolutionary history of dinoflagellates and apicomplexans (Alveolata) as inferred from HSP90 and actin phylogenies. J Phycol 40: 341–350

Leander BS, Kuvardina ON, Aleshin VV, Mylnikov AP, Keeling PJ (2003b) Molecular phylogeny and surface morphology of *Colpodella edax* (Alveolata): insights into the phagotrophic ancestry of Apicomplexans. J Eukaryot Microbiol 50:334–340

Lefranc M, Thénot A, Lepère C, Debroas D (2005) Genetic diversity of small eukaryotes in lakes differing by their trophic status. Appl Environ Microbiol 71:5935–5942

López-García P, Philippe H, Gail F, Moreira D (2003) Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the mid-Atlantic ridge. Proc Natl Acad Sci USA 100:697–702

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

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

Massana R, Castresana J, Balagué V, Guillou L, Romari K, Groisillier A, Valentin K, Pedrós-Alió C (2004b) Phylogenetic and ecological analysis of novel marine stramenopiles. Appl Environ Microbiol 70:3528–3534

Medlin LK, Metfies K, Wiltshire K, Mehl H, Valentin K (2006) Picoplankton diversity at the Helgoland time serie site as assessed by three molecular methods. Microb Ecol (in press)

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

Moreira D, López-García P (2002) The molecular ecology of microbial eukaryotes unveils a hidden world. Trends Microbiol 10:31–38

Norén F, Moestrup Ø, Rehnstam-Holm A-S (1999) *Parvilucifera infectans* Norén et Moestrup gen. et sp. nov. (Perkinsozoa phylum nov.): a parasitic flagellate capable of killing toxic microalgae. Eur J Protist 35:233–254

Patterson DJ (1999) The diversity of eukaryotes. Am Nat 154: 96–124

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818

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

Saldarriaga JF, McEwan ML, Fast NM, Taylor FJR, Keeling PJ (2003) Multiple protein phylogenies show that *Oxyrrhis marina* and *Perkinsus marinus* are early branches of the dinoflagellate lineage. Int J Syst Evol Microbiol 53: 355–365

Saldarriaga JF, Taylor FJR, Cavalier-Smith T, Menden-Deuer S, Keeling PJ (2004) Molecular data and the evolutionary history of dinoflagellates. Eur J Protist 40:85–111

Slapeta J, Moreira D, López-García P (2005) The extent of protist diversity: insights from molecular ecology of freshwater eukaryotes. Proc R Soc Ser B 3195:1–9

Stoeck T, Epstein S (2003) Novel eukaryotic lineages inferred from small-subunit rRNA analyses of oxygen-depleted marine environments. Appl Environ Microbiol 69: 2657–2663

Stoeck T, Taylor GT, Epstein SS (2003) Novel eukaryotes from the permanently anoxic Cariaco basin (Carribean Sea). Appl Environ Microbiol 69:5656–5663

Swofford DL (2002) PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sinauer associates, Sunderland, MA

Taylor FJR (2004) Illumination or confusion? Dinoflagellate molecular phylogenetic data viewed from a primarily morphological standpoint. Phycol Res 52:308–324

Van-Dover CL, German CR, Speer KG, Parson LM, Vrijenhoek RC (2002) Evolution and biogeography of deep-sea vent and seep invertebrates. Science 295:1253–1257

Editorial responsibility: John Dolan, Villefranche-sur-Mer, France

Submitted: August 31, 2005; Accepted: December 7, 2005
Proofs received from author(s): March 13, 2006