Genetic characterisation of *Emiliania huxleyi* (Haptophyta)

L.K. Medlin [a,*], G.L.A. Barker [b], L. Campbell [c], J.C. Green [d], P.K. Hayes [b], D. Marie [e], S. Wrieden [a], D. Vaulot [e]

[a] *Alfred-Wegener-Institute for Polar and Marine Research, Am Handelshafen 12, D-27570, Bremerhaven, Germany*
[b] *University of Bristol, School of Biological Sciences, Bristol, BS8 1UG, UK*
[c] *University of Hawaii, Department of Oceanography, Honolulu, HI 96822, USA*
[d] *Plymouth Marine Laboratory, Citadel Hill, Plymouth PL1 2PB, UK*
[e] *Station Biologique, CNRS et Université Pierre et Marie Curie, BP 74, 29682 Cédex, Roscoff, France*

Abstract

Amongst the coccolithophorids, *Emiliania huxleyi* is the most successful and can form large scale blooms under a variety of environmental conditions. This implies extensive genetic variation within this taxon. Physiological, morphological and antigenic differences between clonal isolates support this suggestion. Our investigations into the level of genetic variation within the morphological species concept of *E. huxleyi* indicate that it is such a young taxon that sequence comparisons of both coding and non-coding regions cannot resolve the issue of how many separate taxonomic entities are involved. However, PCR-based genetic fingerprinting techniques do reveal extensive genetic diversity, both on a global scale and within major bloom populations in both space and time. Cell DNA content can also separate cells with morphotype A coccoliths from those with morphotype B coccoliths. Taken together with physiological and morphological evidence, these data suggest that the morphotypes of *E. huxleyi* should be separated at the variety level. We have used both nuclear and plastid rRNA sequence comparisons to confirm the place of *E. huxleyi* within the Haptophyta.

1. Introduction

Many important members of the marine phytoplankton are haptophyte microalgae. Of these, the coccolithophorids, which precipitate calcium carbonate in the form of coccoliths, play a vital role in the global carbon cycle. Amongst the coccolithophorids, *Emiliania huxleyi* (Lohmann) Hay & Mohler is the most successful, forming large scale blooms in both oceanic and neritic waters from sub-polar to tropical latitudes (Brown and Yoder, 1994). The ability to bloom under such a multitude of environmental conditions implies extensive genetic variation within this taxon. This hypothesis is supported by physiological, morphological, and biochemical evidence that suggests that *E. huxleyi* may include a number of distinct entities, possibly separated at the species level.

At present, three morphological coccolith variants, designated A, B and C, are recognised and can be sorted using standard microscopical techniques (Young and Westbroek, 1991; Van Bleijswijk et al., 1991). There are significant ultrastructural differ-

* Corresponding author. E-mail: lmedlin@awi-bremerhaven.de.

ences between the coccoliths of the three morpho-types (Young and Westbroek, 1991). Cells bearing the A type coccoliths appear to be the most common. They are slightly smaller than cells bearing B type coccoliths, B morphotype cells tending to have more coccoliths per cell. The C morphotype has been referred to as the "cold-water form" (McIntyre and Bé, 1967), although this is no longer recommended (Young and Westbroek, 1991). They resemble the B type coccoliths in structure but are smaller. Cells with the C type of coccolith have never been obtained in culture. Antisera raised against coccolith-associated polysaccharides can differentiate cells carrying the A, B or C coccoliths (Van Bleijswijk et al., 1991; Young and Westbroek, 1991; Van der Wal, pers. commun.); however, antisera raised against naked cells cannot (Campbell et al., 1989).

Growth rates of cells with the A and B morpho-types differ significantly (Van Bleijswijk et al., 1994), as do acclimatised growth rates for individual clones of cells with a single morphotype from the same water sample (Brand, 1982). Different key biomarker compounds (Conte et al., 1995) and different pigment compositions (Mantoura, unpubl.) can discriminate oceanic from neritic strains of *E. huxleyi* but not A from B morphotypes.

On the basis of the above evidence it seems likely that at least two taxa may exist within the general morphological species concept of *E. huxleyi*, each exhibiting considerable genetic variability. However, it is not clear how this genetic variability is partitioned on a spatial or temporal scale. If genetically distinct populations exist within such a complex (albeit different populations of the same taxon, different life history stages of the same taxon, or even different sympatric taxa) and have varying tolerances to different environmental conditions, then this might provide a mechanism through which *E. huxleyi* can respond rapidly to ecological changes to form and maintain blooms on a large geographic scale.

Our research into the genetic variation within *E. huxleyi* is designed to determine if *E. huxleyi* represents a multi-species complex, to place it within a phylogenetic reconstruction of the Haptophyta, to investigate the structure of bloom populations (natural large scale blooms and blooms induced within artificial mesocosms) and to determine the feasibility of making species-specific probes.

2. Methods

2.1. Culture / collection methods

A major collection of global isolates of *E. huxleyi* is available at the Plymouth Marine Laboratory. These isolates originate from most ocean basins as well as from time series taken through two blooms: a natural bloom in the North Atlantic and a second induced bloom in the mesocosm experiments established at the University of Bergen Marine Biological Field Station in Korsfjorden, Norway. The North Atlantic strains were collected during 1989–1991 UK-JGOFS Biogeochemical Ocean Flux Study (BOFS) cruises. On these cruises, a 1 ml aliquot was withdrawn from 500 ml of near-surface water, that was obtained either from water bottles attached to a rosette sampler or from the non-toxic seawater supply (4 m depth) and had previously been concentrated by gravity filtration through a 3 μm Nucleopore filter. This aliquot was inoculated into 30 ml of enriched seawater culture medium (Keller et al., 1987) and incubated at 10–15°C with a 12:12 h LD cycle until they were returned to Plymouth Marine Laboratory. Single cell isolation methods were then undertaken to establish clonal cultures of *E. huxleyi*. For the Norwegian mesocosm experiments (23 April–29 May, 1992), single cell isolates and non-clonal, back-up cultures were established in the same medium as above, from concentrated seawater taken from mesocosm bags 3, 5 and 6 placed in the Raunefjord outside the University of Bergen field station at Espegrend, Norway, and from fjords outside the experimental area. Mesocosm bags 3, 5 and 6 represent natural waters with added N:P ratios of 16:1, 16:0.5, and 16:5, respectively (Egge and Heimdal, 1994). Those strains selected for genetic analysis are listed in Table 1.

Clonal isolates were inoculated into one litre of "K" enriched seawater media, monitored daily and harvested for DNA analysis and for key biomarker compounds (Conte et al., 1995) during log phase either by differential centrifugation at low speeds, or by filtering through 3 μm Nuclepore filters to reduce bacterial contamination.

2.2. DNA methods

DNA was extracted using a 3% (w/v) CTAB (Doyle and Doyle, 1990) or phenol/chloroform/iso-

amyl alcohol extraction and purified twice by CsCl/ethidium bromide density gradient centrifugation for the RAPD analysis. In some cases the DNA was purified with Glass-max spin columns (Gibco, BRL).

Nuclear and plastid-encoded small subunit rRNAs (ssu-rDNA) were amplified using the polymerase chain reaction (PCR; Saiki et al., 1988) with eukaryote-specific or plastid-specific primers, respectively, containing multiple restriction sites for cloning (Medlin et al., 1988, 1994, 1995; Barker, 1995). Alternatively, ssu-rDNAs were amplified with biotin-labelled primers, and single-stranded templates were purified for sequence determination (Dynabeads M-280 Streptavidin, DYNAL A.S., Oslo, Norway). Both coding and non-coding strands were completely sequenced (Sanger et al., 1977; Elwood et al., 1985; Medlin et al., 1995). The products of no fewer than 6 PCR reactions in each orientation were pooled for the direct sequencing reactions; for cloned ssu-rDNAs a total of at least five independent clones were pooled to provide the sequencing templates. The non-coding region separating the genes encoding the two subunits of ribulose 1,5 bisphosphate carboxylase (*rbcL-rbcS* spacer) was also amplified using biotin-labelled primers and directly sequenced (Barker, 1995).

2.3. Sequence analysis

The nuclear ssu-rDNA sequences of *Emiliania huxleyi*, *Gephyrocapsa oceanica* Kamptner, and *Coccolithus pelagicus* (Wallich) J. Schiller were aligned with over 300 eukaryote rRNA sequences (Neefs et al., 1991; Andersen et al., 1993; Medlin et al., 1991, 1993, 1994) using maximum primary and secondary structural similarity. Positional homology was assumed for 1493 positions, of which 671 were informative for parsimony analyses of the nuclear-encoded gene. The final data set contained 42 taxa; all of these sequences are available from GenBank. *Dictyostelium discoideum* Raper, a myxomycete was used as outgroup. The *rbcL-rbcS* spacer regions from two B morphotypes and five A morphotypes of *E. huxleyi* were aligned with sequences from two other haptophytes, one heterokont alga, four red algae, and one glaucocystophyte alga (Barker, 1995).

Maximum parsimony analyses of the nuclear ssu-

rDNAs were implemented with the PAUP computer program (Swofford, 1993). Introduced gaps were treated as missing data; informative characters were treated as multistate unordered. Unweighted maximum parsimony trees were obtained using the tree-bisection-reconnection (TBR) branch swapping option in a heuristic search with random taxon addition. The most parsimonious trees (MPT) and the data matrix were loaded into the MacClade computer program (Maddison and Maddison, 1992) to infer a weighted data set. Positions were weighted inversely porportional to the number of base substitions at that position in the data set (Scale 1-100) as described in the MacClade program. The type of substitution at each position was also weighted as the cost of going from one state to another (Scale 1-100) also as described in the MacClade program. These weighting schemes greatly enhance the ability of the maximum parsimony analyses to recover the correct tree when multiple substitutions have occurred over sites (Hillis et al., 1994). Stability of monophyletic groups in weighted maximum parsimony trees was estimated with a bootstrap analysis (Felsenstein, 1985). The information content of this data set was measured with the G_1 statistic, which tests for the distribution skewness of 10,000 random trees (Hillis and Huelsenbeck, 1992). The more negative the value of the G_1, the greater phylogenetic signal contained within the data set.

Distance analysis was performed on the nuclear ssu-rDNAs using the PHYLIP computer program (Felsenstein, 1993). Dissimilarity values (Fitch and Margoliash, 1967), based on pairwise comparisons of sequences, were transformed into distances using the Kimura-two-parameter-model (Kimura, 1980). Distance matrices were converted into trees using the neighbor-joining method (Felsenstein, 1993); a bootstrap analysis (Felsenstein, 1985) was also performed for these sequence comparisons. Maximum likelihood (ML) analyses were done on the data set using the fastDNAml program (v. 1.0; Larsen et al., 1993).

Phylogenetic relationships inferred with the above methods may reflect shared base compositional bias rather than a monophyletic origin. Therefore, an analysis using the LogDet evolutionary model, which attempts to correct for base compositional bias (Lockhart et al., 1994), was done to determine if this had any affect on the phylogenetic analysis.

Table 1

Strains of *Emiliania huxleyi* and other prymnesiophytes and the regions of the genome or the primers used in the genetic analysis. (B) = B morphotype. + = coding or non-coding region sequence completed. Numbers below the RAPD-PCR columns indicate the number of bands produced from each primer (see Table 2)

Strain	Location	Date	28S	18S	16S	Rub-isco	PucF	ITSF	1400F	OE2	OE18	OE20	OG6	OG11	OG13
Emiliania huxleyi															
NE Atlantic Iceland Basin [1]															
G1779a	60°N, 20°W	14.6.89	+	+				10+	10+						
G1779c	60°N, 20°W	14.6.89					8	10+	10+	6					
DWN60/3/8 [2]	60°54N, 22°54W	23.6.91								7	10+	4	6	2	0
DWN60/6/1	61°N, 15°W	24.6.91								9	10+	3	3	1	0
DWN60/6/2	61°N, 15°W	24.6.91								10+	10+	4	3	3	2
DWN60/10/1	61°14N, 15°W	24.6.91								10+	7	2	0	3	2
DWN61/3/2	58°41N, 15°54W, 10 m	11.7.91								10+	7		1	0	1
DWN61/3/7	58°41N, 15°54W, 10 m	11.7.91								0	0				
DWN61/4/17	58°41N, 15°54W, 10 m	11.7.91								10+	10+	5	4	5	0
DWN61/6/2	58°41N, 15°54W, 30 m	11.7.91								10+	7		9	7	1
DWN61/6/10	58°41N, 15°54W, 30 m	11.7.91								10+	8		3	6	0
DWN61/7/3	58°41N, 15°54W, 40 m	11.7.91								10+	9	3	4	1	2
DWN61/7/14	58°41N, 15°54W, 40 m	11.7.91								10+	10+		5	5	3
DWN61/12/2	61°05N, 20°W, 10 m	12.7.91								10+	10+		9	4	2
DWN61/12/6	61°05N, 20°W, 10 m	12.7.91								10+	5		10+	4	1
DWN61/67/2	60°42N, 20°17W, 8 m	23.7.91						10+	10+	10+	7	4	5	5	2
DWN61/67/5	60°42N, 20°17W, 8 m	23.7.91								10+	10+		0	5	0
DWN61/81/5	60°33N, 19°24W, 4 m	12.7.91								10+	10+	2	0	3	1
DWN61/87/17	60°33N, 19°24W, 40 m	12.7.91						10+		10+	9	2	3	3	1
NE Atlantic, subtropical															
DWN53/74/6	24°03N, 20°02W, 4 m	08.10.90	+			+	6	10+	10+	0					
Sargasso Sea															
MCH-1 (B)	off Bermuda	30.1.67					10+	10+		2					
English Channel/North Sea															
CH25/90 (B)	57°26N, 6°13E	7.90				+	10+	10+	10+						
92D (B)	50°02N, 4°22W	6.75	+	+	+	+	10+	10+	8	4					
Norwegian Fjords and Mesocosms 1992 [3]															
B011	Meso3, prebloom	30.4.92					10+	10+	10						
B013	Meso3, prebloom	30.4.92						10+	5		10+	4	1	2	9
B021	Trengereid Fj.	30.4.92					5	10+							
B027	Meso3, prebloom	01.5.92					8	4	7						
B028	Meso3, prebloom	01.5.92						4	10						
B043	Meso3, prebloom	02.5.92					2								
B049	Meso3, bloom	04.5.92					10+	10+							
B079	Meso3, bloom	07.5.92						3	8						
B086	Meso3, bloom	12.5.92					10+	10+	8		4	8	0	1	2
B131	Meso3, bloom	14.5.92					10+	10+	8–10+		10+	10+	1	10	10+
B170	Meso3, bloom	16.5.92						4	8						
B208	Meso3, postbloom	20.5.92						5	7–10+		10	10+	1	6	9

Table 1 (continued)

Strain	Location	Date	Sequence data				RADP-PCR Fingerprinting								
			28S	18S	16S	Rub-isco	PucF	ITSF	1400F	OE2	OE18	OE20	OG6	OG11	OG13
B317	Meso3, postbloom	24.5.92					4	10+	8						
B065	Meso6, prebloom	03.5.92						10+	3	7	8	2	3	4	
B126	Meso6, bloom	14.5.92						10+	6	7	6	1	3	9	
B174	Meso6, bloom	16.5.92						10+	10+	10	10+	0	7	10+	
B202	Meso6, postbloom	20.5.92						10+	10+	10+	10+	3		10+	6
'L' Strain	60°N, 11°E, Oslo Fjord	1959		+	+	10+	7	2	4						
Indian Ocean, subtropical															
S. Africa	Durban Pier	March 83				+	5		4						
NE Pacific, temperate															
Van55	50°11N, 145°W	?				+	9		3						
Van556	49°05N, 144°40W	?					10+								
SW Pacific, subtropical															
EH2	Great Barrier Reef	Dec 90													
Additional Species															
Gephyrocapsa oceanica															
SW Pacific, subtropical															
EH1	Great Barrier Reef	Dec 90				+									
NW Pacific, temperate															
GO1	41°N, 141°E, Japan	Oct 91		+	+	+	8		7						
SW Pacific, temperate															
AB1	36°16S, 174°48E New Zealand	Oct 92				+									
C. pelagicus															
182G	English Channel	Nov 92					10+		5						

[1] Where locations differ from those in Conte et al. (1995), then these are correct.
[2] Sampling code = cruise number/bottle cast/isolate.
[3] See Egge and Heimdal (1994).

2.4. Antibody cross-reactions

Two A and two B morphotypes of *E. huxleyi* and two strains of *Gephyrocapsa oceanica* were tested for their cross-reactivity to the antibody raised against *E. huxleyi* isolate BT6 (CCMP 373, Campbell et al., 1989) from the Sargasso Sea using the method described in Shapiro et al. (1989). These data were compared with those obtained from a survey of cross-reactions of various coccolithophorids with the antibodies raised against the coccolith-associated polysaccharides from the *E. huxleyi* A and B morphotypes (Van der Wal, pers. commun.).

2.5. Relative genome size

Relative amounts of DNA per cell from four A and two B morphotypes were estimated using flow cytometry. Cells were fixed in methanol (Olson et al., 1986) and stained with chromomycin A3, Hoechst 33342, or propidium iodide for analysis (Vaulot et al., 1994). Strains BT6 and CH25/90 (Texel B) were added as internal standards.

Table 2

Primers used for the RADP analysis of *E. huxleyi* clones. Primers ITSF and 1400F were designed for use as the nuclear rRNA ITS forward PCR primer and ssu rRNA forward sequencing primer, respectively. The Operon primers were manufactured specifically for RAPD analysis. The GC% is given for each primer together with the approximate melting temperature (*Tm*) based on the formula Tm (°C) $= 2 (A+T) + 4 (G+C)$ where A, T, G and C are the number of these nucleotides in the primer

Primer	Length	Sequence	GC%	$Tm(°C)$
ITSF	17-mer	AGCTGGTTCTCCAGCCG	64.7	56
1400F	17-mer	TGYACACACCGCCCGTC	64.7	56
OPE-02	10-MER	GCTGCGGGAA	70	34
OPE-18	10-MER	GGACTGCAGA	60	32
OPE-20	10-MER	AACGCTGACC	60	32
OPG-06	10-MER	GTGCCTAACC	60	32
OPG-11	10-MER	TGCCCGTCGT	70	34
OPG-13	10-MER	CTCTCCGCCA	70	34

2.6. RAPD analysis

RAPD-PCR reaction procedures for 17-mer and 10-mer (Operon Technologies, Almeda, CA) oligonucleotide primers and gel running conditions have been described previously (Barker et al., 1994). The RAPD data were analysed as three separate sets: (1) global isolates, (2) 1991 North Atlantic bloom study and (3) 1992 Norwegian mesocosm bloom study. The sequences of the two 17-mer and six 10-mer oligonucleotide primers used in the RAPD analysis are listed in Table 2.

Prior to analysis, RAPD-banding patterns for the global isolates, the North Atlantic Bloom isolates and isolates representing a time series through Norwegian mesocosm Bag 3 were captured using a CCD-camera, converted to bitmap format image files, and then to negative images (PAINTSHOP PRO). These images were used as input to the GEL COMPARE software Package (Applied Maths, Kortijk, Belgium). A series of size standards on each gel was fitted to a standard profile so that the "normalised" banding patterns could be compared from one gel image to one another. Pearson similarity coefficient matrices for each possible primer/template combination tested were produced automatically from the normalised gel images. Cluster analysis of the similarity matrix data was performed by the "neighbor-joining method" program option (Saitou and Nei, 1987), with output in the form of unrooted dendrograms.

Where possible, similarity matrices and accompanying dendrograms were created with data from as many as 5 separate oligonucleotide primers using the "combined gel" option of Gel Compare (Tables 1 and 2).

For the Norwegian mesocosm experiments, RAPD banding patterns from a reduced time series of isolates from Bag 3 and Bag 6 were converted to binary data and analysed with the split-decomposition program (Bandelt and Dress, 1992, 1993). A similar analysis was performed on the global isolates amplified with OPE-02 primer and for a data set compiled from five primers for the North Atlantic bloom study.

RAPD banding patterns produced using primer OPE-02 for isolates from the North Atlantic bloom and from the extended time series through mesocosm bag 3 were used to calculate population genetic parameters including estimates of allele frequencies and gene diversity (Hartl, 1980). For all species studied, most RAPD markers have been shown to be inherited in a Mendelian fashion and can be scored as phenotypes (Hadrys et al., 1992). For example, only 4 out of 133 RAPD markers were not inherited as dominant Mendelian markers (Levitan and Grosberg, 1993). For our purposes, each discrete band size class on the gels was treated as a genetic locus, where the presence of a band indicated the presence of at least 1 dominant allele (genotype = AA or Aa), and absence indicated a null homozygote (aa). All population genetic parameters were calculated with the methods described in Lynch and Milligan (1994), under the assumption that genotype frequencies are in Hardy-Weinberg equilibrium. Rare null alleles were omitted from the analysis, because these give rise to biased parameter estimates (Lynch and Milligan, 1994). An estimate of gene diversity, Hj, for each locus in each population, a mean over all loci, $Hj(L)$, and their variance in the populations were also calculated (Lynch and Milligan, 1994).

3. Results and discussion

3.1. Comparisons of coding and non-coding regions

The nuclear and plastid-encoded ssu-rRNA genes from the selected A and B morphotypes of *E. huxleyi* as well as those from the GO1 and AB1 strains

of *Gephyrocapsa oceanica* (Table 1) are identical. The spacer regions separating the plastid-borne genes encoding the large and small subunits of ribulose 1,5 bisphosphate carboxylase (*rbcL-rbcS*) in the two *E. huxleyi* morphotypes are also identical; however, we were unable to obtain clear sequence data from this region from *G. oceanica* for comparison. Fujiwara et al. (1994) also found that the sequence of *rbcL*, encoding the large subunit of ribulose 1,5 bisphosphate carboxylase, was conserved at the nucleotide level between *Emiliania* and *Gephyrocapsa*. The rate of divergence in the nuclear-encoded rRNA gene in higher plants, green algae and the diatoms averages 1% in 20–25 m.y. (see references in Kooistra and Medlin, 1996), whereas non-coding regions, such as the ITS (internal transcribed spacer) and the RUBISCO spacer evolve 8–10 times faster (see ref-

erences in Barker, 1995) and have been routinely used in algal studies to differentiate at the population level (Maggs et al., 1992; Bakker et al., 1992). The lack of variation in these coding and non-coding regions strongly suggests that the genetic separation of the A and B morphotypes of *E. huxleyi* from each other or from their last common ancestor shared with *G. oceanica* is too recent to be detected by base substitutions in the regions we have selected. It is believed that these genera separated from each other approximately 268,000 yr ago (Verbeek, 1990). Other faster evolving regions of the genome, such as hypervariable regions of the mitochondria, may be able to resolve the *Emiliania/Gephyrocapsa* divergence.

Despite the similarity of the nuclear-encoded ssu coding regions of *Emiliania* and *Gephyrocapsa*, we were able to determine one region specific for these
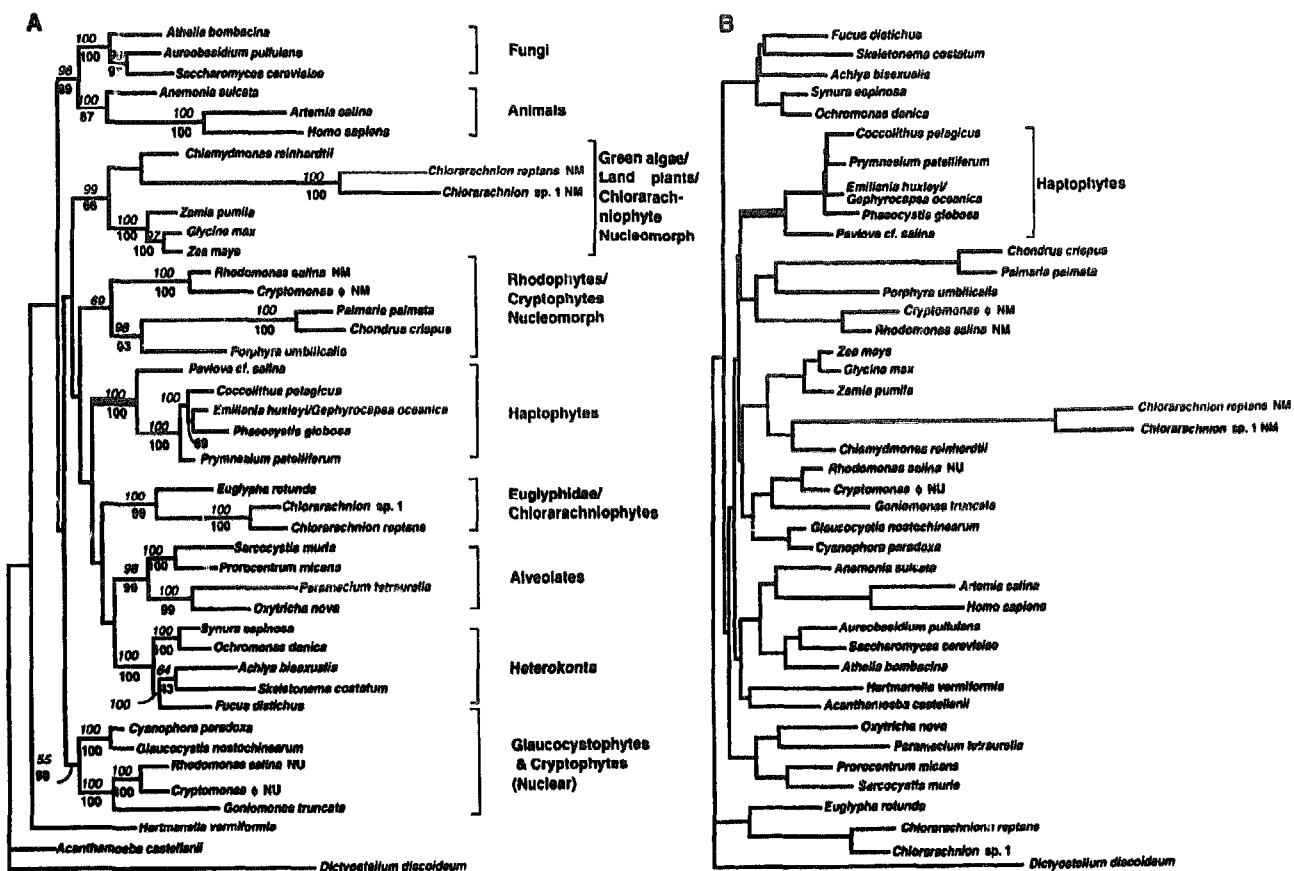
Fig. 1. A. Phylogeny of the Haptophyta inferred from ssu rRNA sequence comparisons of 42 taxa with *Dictyostelium discoideum* as outgroup with the ML method (fastDNAml). Italic figures above the internal nodes are bootstrap values based on a weighted maximum parsimony analysis; bold figures below the internal nodes are bootstrap values based on neighbor-joining analysis. Bootstrap values < 50% are not shown. The consistency index improved from 0.380 to 0.655 in the weighted analysis. B. Phylogeny of the Haptophyta inferred with the neighbor-joining method using a LogDet evolutionary model.

two genera. We have begun testing the specificity of a fluorescein-labelled oligonucleotide probe to this region using flow cytometric methods (data not shown). Our preliminary screening results indicate that we need to increase the stringency of our hybridisation conditions and use competitor probes to increase the ratio between probe and no probe labelling to partition *Emiliania* and *Gephyrocapsa* from other haptophyte algae.

We have used sequence comparisons of the nuclear ssu-rDNAs to determine the position of *Emiliania / Gephyrocapsa* within the Haptophyta and the position of the Haptophyta relative to other eukaryotic organisms. The position of the plastid-encoded ssu-rRNA has been determined elsewhere (Medlin et al., 1995). Results of the maximum likelihood, neighbor-joining and weighted maximum-parsimony analyses are summarised in Fig. 1a. The primary differences between the three analyses lie in the branching order of the major eukaryotic lineages, which are not supported in any of the bootstrap analyses. This rapid radiation is documented by the short branch lengths separating these lineages. Bootstrap values (> 50%) from weighted maximum parsimony and neighbor-joining analyses are placed at the internal nodes that were shared by these and the maximum likelihood analyses and strongly support the monophyly of the major lineages but not their relationships to one another. A G_1 of -0.87932 ($P < 0.1$) indicates that the data set contains a strong phylogenetic signal.

Two most parsimonious trees were generated in the unweighted and weighted maximum parsimony analyses (trees not shown). They differed in the position of the clade containing *Acanthamoeba* and *Hartmanella*, which was positioned either inside or outside a lineage containing green algae/ cryptophytes/ glaucocystophytes. The bootstrap analyses using a weighted data matrix based on both of these trees were nearly identical and differed only in recovering bootstrap support for the red algal/cryptophyte endosymbiont nucleomorph (NM) clade and for a clade containing the chlororachniophytes/animals/fungi.

In all analyses, the haptophytes are a monophyletic lineage with strong bootstrap support. In none of the analyses were the haptophytes sister taxon (i.e. branching at the same point in the tree) to

the heterokont algae. However, in the ML tree they are the sister group to a lineage containing the euglyptophytes/ chlororachniophytes/ alveolates/ heterokont algae. There is no bootstrap support for the branching order within this radiation. Earlier comparisons of ssu rDNAs with limited sampling of the haptophytes have also been unable to establish a clear sister group for the haptophytes (Bhattacharya et al., 1992; Leipe et al., 1994; Medlin et al., 1994). Within the haptophyte lineage, *Pavlova* is the first deep divergence. The remaining haptophytes diverge into two groups. Species of *Emiliania* and *Gephyrocapsa* are most closely related to *Phaeocystis*.

The phylogenetic relationships inferred from the Log-Det evolutionary model are presented in Fig. 1b. The branching order of the major clades are very different from those summarised in Fig. 1a. In this analysis, the haptophytes are still a distinct monophyletic lineage very distant to the heterokonts. After the initial divergence of *Pavlova*, a very rapid radiation appears to have occurred amongst the other haptophytes.

3.2. Comparisons of antibody cross-reactions

Antibodies raised against the naked cells of *E. huxleyi*, clone BT6 from the Sargasso Sea (Shapiro et al., 1989) cross-react with both morphotypes of *E. huxleyi* as well as with *Gephyrocapsa oceanica* (Table 3). In an earlier study, the specificity of this antibody was tested against the coccolithophorid,

Table 3
Immunofluorescence results from cross reactions within the A and B morphotypes of *Emiliania huxleyi* and *Gephyrocapsa oceanica* with antiserum directed against *Emiliania hyxleyi* (BT6). Reaction = rating for visual quality of staining: 4+ = bright stain; 3+ = less bright; 2+ = low intensity, but obvious compared with control; 1+ = low intensity but separable from control; – = no reaction. *Synechococcus* was used as the negative control

Strain	Morphotype	Stage	Reaction
CH25/90 = Texel B	B	naked, coccoid	4+
92D	B	naked, coccoid	2–3+
BT6/CCMP373	A	Naked, coccoid	4+
Ch24/90 = Texel A	A	coccoliths	2–3+
B011	A	coccoliths	2–3+
GO1	*Gephyrocapsa*	naked, coccoid	3+
AB1	*Gephyrocapsa*	coccoliths	2–3+

Coccolithus pelagicus, and other organic-scaled members of the Haptophyta (Shapiro et al., 1989). This antibody did not cross-react with any of the haptophyte algae tested, although it should be noted that only one coccolithophorid species was included in the survey. In contrast, antisera raised against the coccolith-associated polysaccharides from the A and B morphotypes (Van Bleijswijk et al., 1991) can discriminate the two morphotypes. However, a recent survey of the reaction of these antibodies against other coccolithophorid taxa indicates that the A morphotype antibody will cross-react with the type C morphotype of *E. huxleyi*, as well as with species of *Alisphaera, Acanthoica*, and *Umbellosphaera*; the B morphotype antibody cross-reacts with *Calcidiscus leptoporus* and with species of *Gephyrocapsa* and *Umbilicosphaera* (Van der Wal, pers. commun.). Although the specificity of these two antibodies is less than previously believed, they can still distinguish a variety of coccolithophorid taxa in a predictable fashion. These results indicate that all of the existing antibodies raised against *E. huxleyi* are not genus-or species-specific and should be tested against a wider spectrum of coccolithophorid taxa before being used for routine phytoplankton analysis.

3.3. Comparisons of genome size

The genome size of four *E. huxleyi* clones with coccoliths of the A morphotype and two clones with coccoliths of the B morphotype were compared using flow cytometry. Clones BT6 (A morphotype) and CH25/90 (B morphotype) were used as internal standards (Table 4). All three fluorochromes used in this study gave similar results; only the data obtained using chromomycin A3 are presented.

At least 2 or more different genome sizes are present within the *E. huxleyi* complex and these appear to segregate the A and B morphotypes. This has also been secondarily confirmed with a selection of strains analysed at NIOZ (Van Bleijswijk, pers. commun.). Of the two strains with B coccoliths, the CH25/90 isolate contains almost exactly half the DNA content of the 92D isolate; these isolates have been designated haploid and diploid, respectively. Amongst those with A coccoliths, BT6 has approximately half the DNA of the others and so has also been designated a haploid genotype. Strain B011 from Norway contains two DNA levels, which may represent both haploid and diploid components. Haploid cells represented only a small fraction of the cells in the culture (approximately 10%), and could correspond to the flagellate monad stage. The presence of this stage was not verified microscopically in the culture analysed but was present in small numbers in the stock culture. A comparison of a mixture of strains B011 (morphotype A) and CH25/90 (= Texel B, morphotype B) is shown in Fig. 2 to illustrate the relative difference in the DNA content of the two morphotypes.

The two "Darwin" strains, when compared with CH25/90 (= Texel B), show a similar level of DNA to that of the "diploid" component of B011 (Table 4). We have no expanation as to why the two *Darwin* isolates do not appear to contain the same amount of DNA as the "diploid" component of B011 when compared using the BT6 strain as the internal standard.

It is now accepted that many, if not all, coccolithophorids have a life cycle involving either scale-bearing flagellates and heterococcolith-bearing cells (sometimes with a benthic plase) or an alternation of

Table 4

Flow cytometric determination of relative amounts of DNA from A and B morphotypes of *Emiliania hyxleyi* using chromomycin A3 on methanol fixed cells. CH25/90 (= Texel B, morphotype B) and BT5 (morphotype A) used as internal standards

| Strain	Morphotype	Stage	DNA/Texel B	DNA/BT6	Preliminary interpretation group
Texel B	B	naked, coccoid	1.00	1.31	haploid group 1
92D	B	naked, coccoid	1.97	2.57	diploid group 1
BT6/CCMP373	A	naked, coccoid	0.73	1.00	haploid group 2
B011	A	flagellates & coccoliths	0.73 + 1.34	1.0 + 2.0	haploid + diploid group 2
DWN/61/81/5	A	naked, coccoid	1.30	1.69	diploid group 2 or ?3
DWN/61/87/17	A	naked, coccoid	1.30	1.71	diploid group 2 or ?3

Relative DNA content
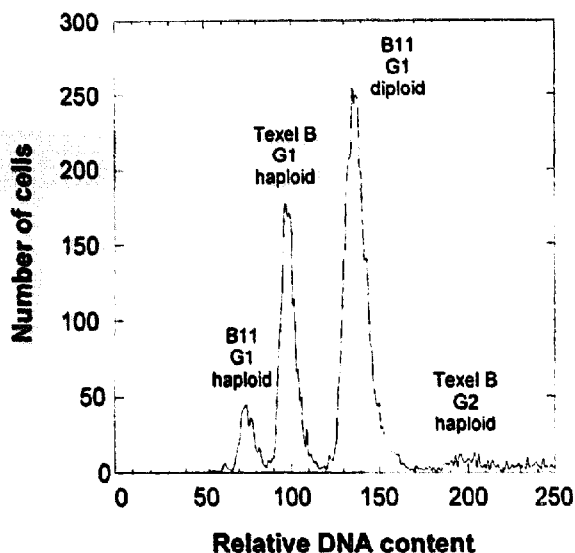
Fig. 2. Mixture of strains B11 and CH25/90 (= Texel B) stained with the DNA specific dye Chromomycin A3 and analysed by flow cytometry. The first and third peaks correspond respectively to the haploid and diploid B11 cells (morphotype A) in the G_1 phase of the cell cycle. Few G_2 haploid cells are also probably included in the third peak. The second peak corresponds to haploid G_1 CH25/90 cells (= Texel B, morphotype B). The fourth minor peak is likely to be constituted of haploid G_2 CH25/90 cells, although the presence of diploid G_1 CH25/90 cells cannot be ruled out.

generation life cycle with diploid heterococcolith stages cycling with haploid holococcolith stages (Billard, 1994; Thomsen et al., 1991). Based on the early life cycle work by Klaveness (1972), Billard (1994) has proposed that a diploid coccolith-bearing stage of *E. huxleyi* alternates with a haploid scaly monad stage. The presence of haploid scaly monad flagellates in cultures of diploid coccolith-bearing stages of *E. huxleyi* has recently been confirmed with flow cytometry (Course et al., 1994; Green et al., 1996-this issue). However, the flagellated monads have never produced any coccolith-bearing stage in culture.

Our preliminary results from the measurement of genome size with flow cytometry clearly indicate that there is a two to four-fold variation in the DNA content amongst the cells with the two types of coccoliths. There are several possible interpretations of the data. First, the life cycle proposed by Billard (1994) may be correct and we may have many taxa within *E. huxleyi*, each being diploid and capable of producing haploid flagellate monads to complete the

life cycle. Alternatively, the life cycle proposed by Billard (1994) may be incomplete. If, between the isolates bearing the B coccoliths, 92D is diploid relative to CH25/90 (Texel B), and amongst the isolates bearing the A coccoliths, all are diploid relative to BT6, then the life cycle of *E. huxleyi* may involve an alternation of both diploid and haploid coccolith-bearing generations as well as scale-bearing flagellates. However, it should be noted that, although all strains analysed were producing coccoliths at the time of isolation and identification, they were not (except for B011) at the time of analysis. The absence of organic body scales on the hetero-coccolith cells prevents determination of ploidy states using the morphology of the two faces of the scales as in other coccolithophorids (Billard, 1994). If these results are confirmed, then they might explain why the life cycle of *E. huxleyi* has proved so difficult to elucidate. Given the wealth of isolates of *E. huxleyi* presently in culture, further careful comparison of coccolith-bearing stages and re-isolated flagellate and other naked cells should be done with flow cytometry to determine the ploidy states of *E. huxleyi* and the types of cells involved in the life cycle.

If our hypotheses on ploidy levels are correct, then haploid genome sizes appear to be similar in clones of the same morphotype, but different between the two morphotypes. The B morphotype may be a polyploid of the A morohotype. Such differences in genome size have been recently established for strains of *Phaeocystis* originating from distinct oceanic regions (Vaulot et al., 1994) and shown to correlate with differences in phenotypic characters, such as pigment composition (Vaulot et al., 1994), and in genotypic characters, such as nuclear ssu-rDNA sequences (Medlin et al., 1994).

It should be noted that Van Bleijswijk and Veldhuis (pers. commun. in Green et al., 1996-this issue) have interpreted the GC ratios from a number of clones as evidence for several sub-taxa within the *E. huxleyi* A-morphotype.

3.4. Comparisons of randomly amplified polymorphic DNA (RAPD) fingerprints

A total of nine 17-mer and forty 10-mer oligo-nucleotides were screened for use as RAPD primers (Barker, 1995); of these, only two 17-mer and six

10-mer primers gave banding patterns suitable for analysis (Tables 1 and 2). The reproducibility of these banding patterns was tested first by running duplicate reactions on many occasions, and second by analysing three separate DNA preparations of isolate G1779a with four different 10-mer oligonucleotides. RAPD banding patterns were found to be relatively stable, with nearly all bands being reproduced perfectly: the most commonly encountered differences appearing as variations in band intensity rather than presence vs. absence (Barker, 1995). Some differences were found amongst the different DNA preparations of isolate G1779a. Examples of reproducibility in *E. huxleyi* RADP banding patterns have been published previously (Medlin et al., 1994).

In most studies using RAPD markers, it is acknowledged that banding patterns are not fully reproducible, and that weaker bands in particular tend to be variable (e.g. Williams et al., 1990; Okamura et al., 1993; Black, 1993). However, few studies using RAPD analysis have included replicate DNA preparations of an isolate, so it is not clear how common this phenomena might be. It is possible that aberrant bands may be of bacterial origin, despite the precaution taken to avoid bacterial contamination. However, when prokaryotic genomes are mixed with eukaryotic genomes, RADP amplifications yield only eukaryotic-specific fragments, even when a 460-fold molar excess of prokaryotic DNA is present (Williams et al., 1993). Furthermore, it is likely that biological factors play a role in differential reproduction of RAPD bands. Patwary et al. (1993), for example, found that when haploid gametophytes of *Gelidium vagum* were self fertilised, the resulting sporophyte plants did not yield banding patterns identical to those of the original gametophyte, and Black (1993) also noted differences between aphid daughters produced parthenogenically from a single mother. Van Oppen et al. (1996) reported that RAPD banding patterns recovered from biological and synthetic offspring and their parents were not always as expected and suggested that the differences resulted from competition for primer binding sites related to ploidy level.

The scoring of weak bands is acknowledged to be an insurmountable problem in the analysis of RAPD banding patterns. Many weak bands are not scored

because of their potentially variable nature. Their absence should remove some "noise" from the data, but that is probably replaced by bias because a subjective decision regarding which band intensities to include in the data set will have to be made. In our study, any rejection of bands would have biased the gene diversity estimates, because it would, in some cases have meant scoring an allele as null, even when a band was clearly (albeit faintly) visible. For this reason, no markers were excluded from this study on the basis of their intensity.

RAPD banding patterns for global isolates and bloom isolates from the 1992 Norwegian mesocosm experiment and the 1991 BOFS North Atlantic Darwin cruises, screened with the primer OPE-02, are shown in Fig. 3. The gels show that no two isolates share identical banding patterns, which implies that neither bloom is clonal and that our global isolates, including A and B morphotypes, are also distinct.

3.4.1. Analysis of global isolates

"Gel compare trees" were constructed for all global isolates amplified with both the OPE-02 and ITSF primers or the OPE-02 primer alone (Fig. 4a,b, respectively). In the smaller data set from the combined primers, B morphotypes form a distinct clade (CH25/90 (= Texel B), 92D, and MCH-1) but one A morphotype (B131) from Norway also falls into this clade. The 1991 North Atlantic bloom isolates form a weak association, which also includes one Norwegian mesocosm isolate (B011). The Pacific isolate (Van 55) is separated by some distance from two other North Atlantic strains, G1779e and DWN53.74.06 isolated in 1989 and 1990, respectively. In the larger data set generated from the OPE-02 primer alone, similar relationships were recovered (Fig. 4b). The B morphotypes again cluster more or less together, although 92D is pulled slightly closer to other A morphotypes. The North Atlantic bloom isolates form a tighter cluster, although B011 from Norway and the South African clone are also included. The Norwegian mesocosm bloom isolates from Bag 3 form two clusters. The first contains pre- and post-bloom isolates (B028, B027, B317); the second consists of isolates B079 and B170, taken at the height of the bloom, and B208, a post-bloom isolate collected on the same day as B317. Only B131, a peak-bloom isolate, falls outside these two

clusters; it groups with the two outgroup taxa, *Coccolithus pelagicus* and *Gephyrocapsa oceanica*. This anomaly probably reflects poor PCR amplification or poor scoring of the bands in B131. Relationships amongst other North Atlantic isolates and the Pacific isolate are as in the smaller data set.
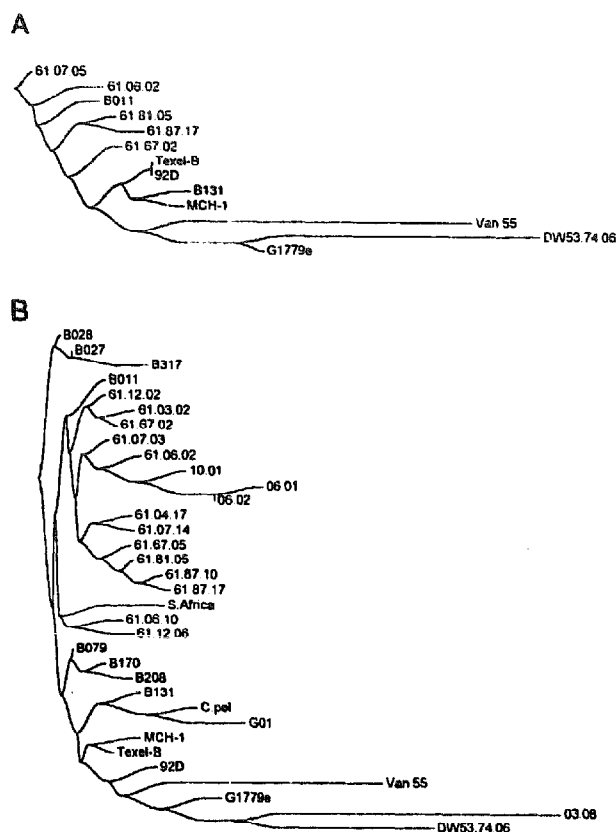
A

B
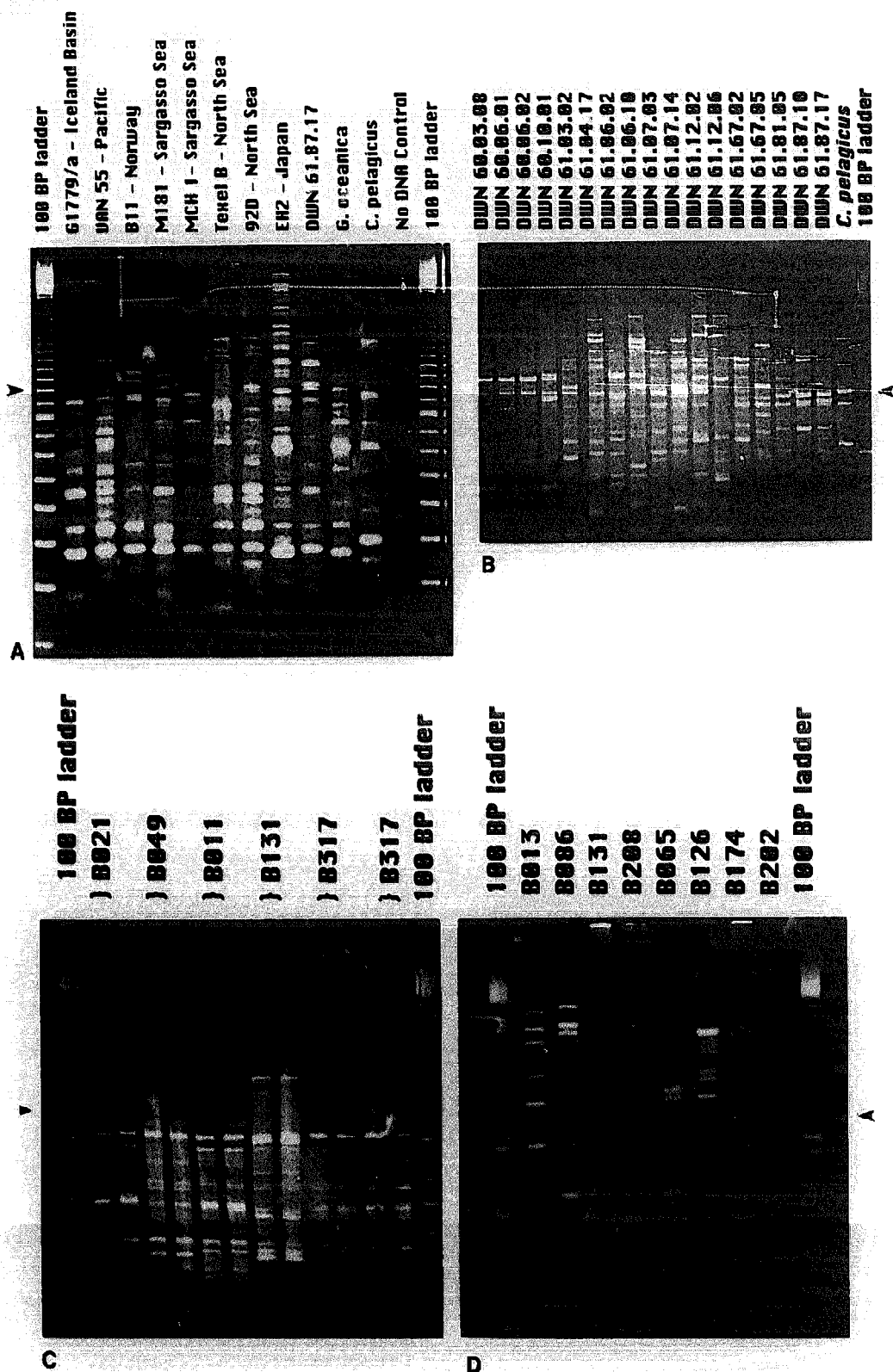
Fig. 4. Representation of the genetic distance amongst global strains of *Emiliania huxleyi* using the neighbor-joining method in the Gel Compare Program with RAPD data from (A) all global isolates compared with the OPE-02 and the ITSF primer, and (B) all global isolates compared only with the OPE-02 primer.

The RAPD bands of the global isolates amplified with primer OPE-02 (Fig. 3a) were converted to binary data and analyzed with the split decomposition program. This analysis partitions the phylogenetic content within a data set from the pattern obtained by reticulation or randomness. It decomposes a distance matrix into a number of components (splits) weighted by isolation indices, plus a residual component due to noise and homoplasy (convergence). A higher isolation index implies higher phy-

A

B
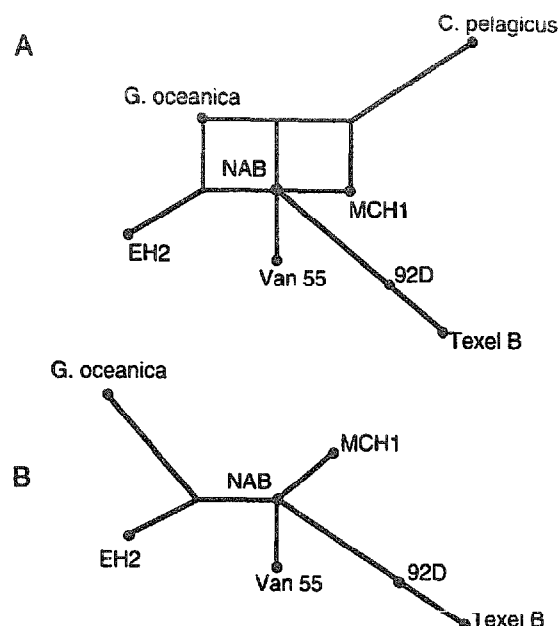
Fig. 5. Representation of the genetic distance amongst global strains of *Emiliania huxleyi* using the split-decomposition evolutionary network model with the outgroup *Coccolithus pelagicus* included (A) or excluded (B). Length of all lines proportional to the isolation index or the split between the taxa.

logenetic signal in the data sets; thus the length of an edge in the split diagram or the length of the line separating two taxa is proportional to the isolation index. The analysis was performed with and without the outgroup taxon, *Coccolithus pelagicus* (Fig. 5a,b). When *C. pelagicus* is included, homoplasy is recovered in the split-diagram (see box-like component to the tree, Fig. 5a). This is probably due to a comparison of non-homologous alleles/bands because *C. pelagicus* is too distantly related to *Gephyrocapsa* and *Emiliania* (see phylogenetic trees in Fig. 1). If omitted, then the homoplasy disappears (Fig. 5b), and the phylogenetic signal in the data set can be better interpreted. Both *G. oceanica* and the SW Pacific isolate of *E. huxleyi* (EH2) are distinct from the other *E. huxleyi* isolates. None of the North

Fig. 3. Agarose gel showing banding patterns produced with the RAPD-PCR technique using the 10-mer oligonucleotide primer OPE-02 from (A) globally distributed *E. huxleyi* isolates, (B) *E. huxleyi* isolates from both legs of the 1991 DWN cruise in the North Atlantic, (C) from a time series of *E. huxleyi* isolates from Bag 3 of the 1992 Norwegian Mesocosm Bloom Experiment, and (D) from a reduced time series through Bags 3 and 6 of the 1992 Norwegian Mesocosm Bloom Experiment. 100 bp ladder used as a size marker in all gels and the 1000 bp band is arrowed in each gel. See Table 1 for a list of all strains. Strain M181 was eliminated from all analyses because of a fungal contaminant.

Atlantic/mesocosm bloom isolates (G1779a, B011, DWN61.87.17) can be differentiated; however, the NE Pacific isolate (Van 55) and the B morphotype (MCH-1) from the Sargasso sea are distinct. In the distance analysis of the RAPD bands compiled from the two primers OPE-02 and ITSF, MCH1 also clustered with an A morphotype (Fig. 4a); both B morphotypes from North Sea/English Channel are quite distinct from the remaining A morphotypes despite being isolated into culture 15 years apart. Other clustering algorithms gave slightly different associations (Barker, 1995).

3.4.2. Analysis of 1991 North Atlantic bloom isolates

"Gel compare trees" were constructed from bands compiled from 5 primers (Table 1; Fig. 6). The distance analysis revealed a tendency for the isolates from the same bottle cast to be grouped together and for isolates from the two legs of the cruise to fall into different clusters. However, the split decomposition analysis showed that there was too much homoplasy/noise in the data set for any phylogenetic signal to be recovered below the terminal taxa. Branches leading to the terminal taxa ranged from up to 10 times the length of the edges of the split diagram seen in Fig. 5b. Nevertheless, 56% of the distance could be recovered in the split decomposition analysis. This suggests that although the bloom is not clonal, the resolution of the RAPD analysis has strayed from a phylogenetic analysis into the realm of tokogeny (i.e. evidence that lineages are interbreeding; Nixon and Wheeler, 1990). Van Oppen et al. (1995) also found that RAPD analysis failed to identify two cryptic species in populations of the red alga, *Phycodrys rubens*, sampled from the

Table 5

Population genetic parameters for North Atlantic and Norwegian mesocosm bloom populations

Population	Number of loci	$H_j(L)$	Variance	Standard error
BOFS 1991 Atlantic bloom	41 *	0.2429	2.2601×10^{-4}	0.0150
Mecocosm bag 3	24 **	0.2463	6.4618×10^{-4}	0.0254

* 2 loci omitted because the observed null allele frequency was less than $3/N$.

** 3 loci omitted because the observed null allele frequency was less than $3/N$.

North Sea to the Baltic, whereas isozyme and sequence analysis could.

3.4.3. Analysis of 1992 mesocosm bloom isolates

A "gel compare tree" was constructed from RAPD bands amplified using the OPE-02 primer and isolates taken from mesocosm Bag 3 (Table 1; Fig. 7a). This analysis suggests that the isolates taken at the pre- and post-stages of the bloom (B011, B027, B028, B208, B317) are more similar than those taken during the bloom peak (B079, B131, B170). The highly derived RAPD banding pattern scored in isolate B131 is clearly indicated in the gel (Fig. 3C). A limited combined analysis of mesocosm bag 3 and 6 isolates was undertaken because of the poor survival of the clonal isolates from mesocosm bag 6. Nevertheless, four isolates from each of the two bags taken during early, mid and late stages of the bloom were compared with six oligonucleotide primers. The split-decomposition analysis recovered 71% of the distance in the tree and suggested that the isolates from the two bags were separate from one another,
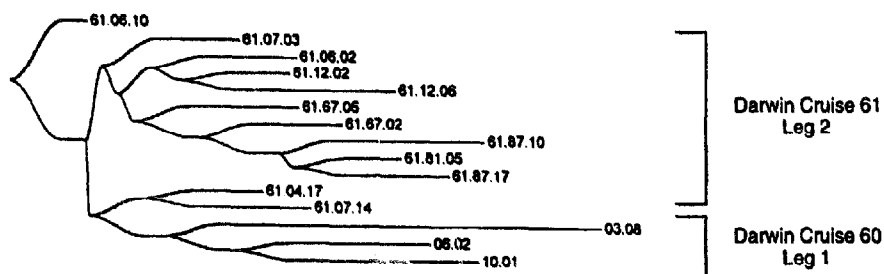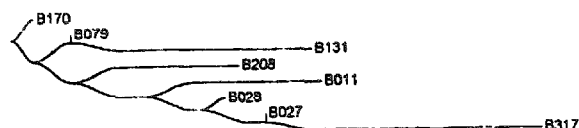

Fig. 6. Representation of the genetic distance amongst isolates of *Emiliania huxleyi* taken during the 1991 North Atlantic bloom using the neighbor-joining method in the Gel Compare Program with RAPD data compiled from five primers (41 bands). Legs 1 and 2 of the Darwin cruises are labelled.
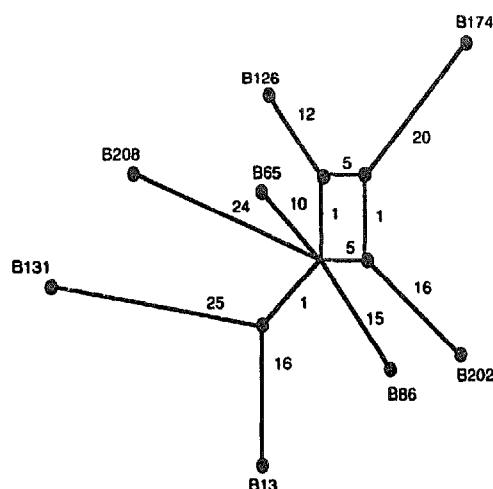
A

B

Fig. 7. Representation of the genetic distance amongst isolates of *Emiliania huxleyi* taken during the 1992 Norwegian Mesocosm Bloom Experiment (A) using the neighbor-joining method in the Gel Compare Program with RAPD data from a time series through Bag 3, and (B) using the split-decomposition evolutionary network model with RAPD data from a comparison of isolates from Bag 3 and Bag 6. Only the lengths corresponding to an isolation index of 1 are proportional. All other lengths leading to terminal taxa are not to scale and the length corresponding to an isolation index of 1 should be multiplied by the number above the line leading to each terminal taxon.

but very highly derived (see length of split leading to terminal taxa, Fig. 7b).

3.5. Comparisons of genetic diversity

Bands amplified with the OPE 02 primer from the North Atlantic bloom isolates and from the extended time series through Bag 3 were interpreted as alleles

at genetic loci. Estimates of gene diversity over all loci, $Hj(L)$ and its variance and standard error for each population are shown in Table 5. The gene diversity values for the two populations were found to be remarkably similar, the difference between the two populations being far less than the standard error of either value of $Hj(L)$.

Many studies have documented within population gene diversity in the past, but estimates have usually been based on allozyme data, which are likely to show less variation than RAPD data. At 0.2429 ± 0.0150, the diversity of the North Atlantic bloom is for example much greater than the average figure for the diatom *Skeletonema costatum* in Narragansett Bay (Gallagher, 1980) of 0.027 ± 0.029, or in human, mouse, *Drosophila*, horseshoe crab or *Lycopodium* populations (Hartl, 1980) calculated from allozyme variation. Despite the potential problems with estimating population genetic parameters from RAPD data discussed above, the standard errors for the estimates of gene diversity (H) calculated in this study are small in comparison with the values themselves, in contrast to Gallaghers' study where the standard error of H was greater than H itself. This shows an advantage of RAPD data. Although fewer samples were used in this study (17 from the North Atlantic bloom, compared with 457 *Skeletonema* isolates in Gallaghers' study), it was possible to score 41 loci from a single primer, as opposed to just 5 allozyme loci in the diatom study. It must be noted, however, that the standard error of H is the error due to sampling a finite number of individuals and loci (Lynch and Milligan, 1994) and does not account for systematic error, such as assigning non-related co-migrating bands to a locus.

The gene diversity estimates for *E. huxleyi* populations were found to fall within the range of values calculated to date from RAPD data. Peever and Milligroom (1994), for example, found that gene diversity in populations of the Ascomycete fungus *Pyrenophora teres* ranged from 0.008 to 0.17, whereas Dawson et al. (1993) found the mean for H over 36 loci in 10 *Hordeum spontaneum* populations was 0.257. The distribution of *H. spontaneum* sampling sites is comparable in scale with that of the BOFS N. Atlantic *E. huxleyi* bloom, the sampling sites for *H. spontaneum* being spread over an area of approximately 300 × 50 km, whereas the samples

analysed from the BOFS cruise track cover an area ca. 350 × 110 km. The diversity found in the mesocosm bloom isolates (0.246 ± 0.0254) is also comparable with that found amongst 5 isolates of *H. spontaneum* from a single location (0.242). The level and spatial pattern of gene diversity found in *E. huxleyi* is therefore comparable with those in a sexual diploid land plant. However, variation in RAPD data has also been found in clonal populations reproducing asexually. In such cases, RAPD bands could be lost or gained by mutation over time, such that the degree of band sharing between clones depends on the mutation rate per generation, and the time since isolates shared a common ancestor (Brookfield, 1992). In addition, crossing over is known to occur during mitosis in yeast, *Drosophila* and mice (see Hastings, 1991 for review). RAPD analysis of the freshwater hydrozoan *Cristatella mucedo* (Okamura et al., 1993) showed that isolates from fresh water lakes in southern England were on average over 98% similar to one another, indicative of a clonal population. Gene diversity was not calculated for *Cristatella*, but examination of the RAPD banding pattern presented by Okamura et al. (1993) suggests that values would be very low. Without knowing the rate at which RAPD markers may be lost or gained by mutation in *E. huxleyi*, it is impossible to calculate whether there would be enough time for the level of variation seen in blooms to have accumulated by mutation alone in such a young species. Sexual reproduction however, even if occurring infrequently, could easily account for the observed diversity, as new combinations of alleles will periodically be brought together by crossing over at meiosis. Our results from the analysis of genome size and those of Course et al. (1994) and Green et al. (1996-this issue) indicate that an alternation of generation cycle, or at least a sexual cycle is probably present in *E. huxleyi*.

Both the presence of a sexual cycle compounded with the problem of ploidy differences could easily account for the large amount of genetic differences and homoplasy we have documented in both our global isolates and our two bloom studies. If we assume that *E. huxleyi* reproduces by asexual means most of the time, then closely related isolates within an area must share a more recent common ancestor than those that are widely separated. Mutation or

mitotic recombination since the divergence would then account for the differences observed between all isolates studied. In such a case the development of a bloom patch could stem from a single individual, or many closely related individuals. In the first instance, banding patterns would have to evolve very quickly to account for the differences between isolates, whereas in the second case, the genetic structure would have to be in place before the bloom, and remain there afterward to shape blooms in subsequent years. If we assume instead that *E. huxleyi* occasionally reproduces sexually, then the similarity between isolates in each area can be explained by interbreeding.

4. Conclusions

The separation of the A and B morphotypes of *Emiliania huxleyi* and *Gephyrocapsa oceanica* are too recent to be resolved using the coding and non-regions of the genome investigated here. However, the analysis of RAPD banding patterns and the measurement of genome size have reinforced the morphological differences observed between these taxa. Furthermore, physiological and biochemical differences between and amongst the A and B morphotypes of *E. huxleyi* probably reflect genetic variation within populations. Further investigations into the life cycle stages and ploidy levels in *Emiliania huxleyi* are warranted given the preliminary data shown here.

Young and Westbroek (1991) recommended that the A and B morphotypes of *E. huxleyi* should be recognised at the varietal level. Although they did not formalise their proposal, we feel that there is now sufficient evidence to complete this process. As well as the varieties based on morphoptypes A and B, it is necessary to validate the C morphotype as a variety. However, no genetic information is available from cells with C-coccoliths nor is there any information on the already existing *E. huxleyi* var. *coronata* (Okada et McIntyre) Jordan et Young (Jordan and Young, 1990).

4.1. Comparison of taxa

Emiliania huxleyi (Lohmann) Hay et Mohler in Hay et al., 1967, var. *huxleyi*

Coccoliths with moderately elevated distal shield, length (2.0) 2.5–3.5 (4.0) μm with elements of variable width, 0.05–0.12 μm; central area with a grill of curved rods.

Epitype: plate III, fig. 6 in Young and Westbroek, 1991, Mar. Micropaleontol., 18: 5–23.

Lectotype: Lohmann, 1902, Arch. Protistenkd., 1: table 4, figs. 1–9 and table 6, fig. 69. Note: Type A coccoliths of Young and Westbroek (1991) and Van Bleijswijk et al. (1991).

E. huxleyi var. *pujosae* (Verbeek) Young et Westbroek ex Medlin et Green var. nov.

Coccoliths large with elevated distal shields, length (3.4) 3.8–4.8 (5.0) μm and narrow shield elements, width 0.05–0.08 μm. Central area with lath-like elements.

Basionym: *Emiliania pujosae* Verbeek, 1990, Meded. Rijks Geol. Dienst, pp. 23–24, pl. 1, figs. 4–9.

Coccolithi magni, longitudine (3.4) 3.8–4.8 (5.0) μm, scutis elevatis distalibus, elementis scutorum angustis, latitudine 0.05–0.08 μm. Area centralis ex elementis tigillorum ad instar constans.

Note: Type B coccoliths of Young and Westbroek (1991) and Van Bleijswijk et al. (1991).

E. huxleyi var. *kleijniae* Young et Westbroek ex Medlin et Green var. nov.

Coccoliths small, distal shield 2.5–3.5 μm long, with well separated elements, 0.07–0.09 μm wide; central area elements lath-like, often forming a complete plate, sometimes absent.

Holotype: plate 6B in McIntyre and Bé, 1967, Deep-Sea Res., 14: 561–597.

Coccolithi parvi, scutis distalibus longitudine 2.5–3.5 μm, ex elementis distantibus, latitudine 0.07–0.09 μm constantibus; elementa areae centralis tigillorum ad instar, laminam totam saepe formantia.

Note: Type C coccoliths, illustrated by McIntyre and Bé as a cold-water ecophenotype, but possibly a form of var. *pujosae*.

Acknowledgements

References

Andersen, R.A., Saunders, G.W., Paskind, M.P. and Sexton, J.P., 1993. Ultrastructure and 18S rRNA gene sequence for *Pelagomonas calceolata* gen. et sp. nov. and the description of a new algal class, the Pelagophyceae classis nov. J. Phycol., 29: 701–715.

Bakker, F.T., Olsen, J.L., Stam, W.T. and Van den Hoek, C., 1992. Nuclear rDNA internal transcribed spacer regions (ITS1 and ITS2) define discrete biogeographic groups in *Cladophora albida* (Chlorophyta). J. Phycol., 28: 839–845.

Bandelt, H.J. and Dress, A.W.M., 1992. Split decomposition: A new and useful approach to phylogenetic analysis of distance data. Mol. Phylogenet., 1: 242–252.

Bandelt, H.J. and Dress, A.W.M., 1993. A relational approach to split decomposition. In: O. Opitz et al. (Editors), Information and Classification: Concepts, Methods and Applications. Springer, Heidelberg, pp. 123–131.

Barker, G.L.A., 1995. Genetic Diversity in *Emiliania huxleyi*. Thesis. Univ. Bristol.

Barker, G.L.A., Hayes, P.K. and Medlin, L.K., 1994. Preliminary results using the RAPD analysis to screen bloom populations of *Emiliania huxleyi* (Haptophyta). Sarsia, 79: 301–306.

Bhattacharya, D., Medlin, L., Wainwright, P.O., Ariztia, E.V., Bibeau, C., Stickel, S.K. and Sogin, M.L., 1992. Algae containing chlorophylls a + c are paraphyletic: molecular evolutionary analysis of the Chromophyta. Evolution, 46: 1808–1817.

Billard, C., 1994. Life cycles. In: J.C. Green and B.S.C. Leadbeater (Editors), The Haptophyte Algae. (Syst. Assoc. Spec. Vol., 51.) Clarendon Press, Oxford, pp. 167–186.

Black, W.C., 1993. PCR with arbitrary primers: approach with care. Insect Mol. Biol., 2: 1–6.

Brand, L.E., 1982. Genetic variability and spatial patterns of genetic differentiation in the reproductive rates of the marine coccolithophores *Emiliania huxleyi* and *Gephyrocapsa oceanica*. Limnol. Oceanogr., 27: 236–245.

Brookfield, J.F.Y., 1992. DNA fingerprinting in clonal organisms. Mol. Ecol., 1: 21–26.

Brown, C.W. and Yoder, J.A., 1994. Blooms of *Emiliania huxleyi* (Prymnesiophyceae) in surface waters of the Nova Scotian shelf and the Grand Bank. J. Plankton. Res., 15: 1429–1438.

Campbell, L., Shapiro, L.P., Haugen, E.M. and Morris, L., 1989. Immunochemical approaches to the identification of the ultraplankton: assets and limitations. In: E.M. Cosper et al. (Editors), Novel Phytoplankton Blooms. Springer, Berlin, pp. 39–56.

Conte, M.H., Thompson, A., Eglinton, G. and Green, J.C., 1995. Lipid biomarker diversity in the coccolithophorid *Emiliania huxleyi* (Prymnesiophyceae) and the related species *Gephyrocapsa oceanica*. J. Phycol., 31: 272–282.

Course, P.A., Tarran, G.A. and Green, J.C., 1994. The life-history of *Emiliania huxleyi*: Evidence for haploid and diploid states using flow cytometry. Phycologist, 37: 43.

Dawson, I.K., Chalmers, K.J., Waugh, R. and Powel, W., 1993. Detection and analysis of genetic variation in *Hordeum spontaneum* populations from Israel using RAPD markers. Mol. Ecol., 2: 151–159.

Doyle, J.J. and Doyle, J.L., 1990. Isolation of plant DNA from fresh tissue. Focus, 12: 13–15.

Egge, J.K. and Heimdal, B.R., 1994. Blooms of phytoplankton including *Emiliania huxleyi* (Haptophyta). Effects of nutrient supply in different N:P ratios. Sarsia, 79: 333–348.

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.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution, 39: 783–791.

Felsenstein, J., 1993. PHYLIP manual, Version 3.5. Dep. Genet., Univ. Washington, Seattle, WA.

Fitch, W.M. and Margoliash, E., 1967. Construction of phylogenetic trees: a method based on mutation distances as estimated from cytochrome *c* sequences is of general applicability. Science, 155: 279–284.

Fujiwara, S., Sawada, M., Someya, J., Minaka, N., Kawachi, M. and Inouye, I., 1994. Molecular phylogenetic analysis of the *rbc*L in the Prymnesiophyta. J. Phycol., 30: 863–871.

Gallagher, J.C., 1980. Population genetics of *Skeletonema costatum* (Bacillariophyceae) in Narragansett bay. J. Phycol., 16: 464–474.

Green, J.C., Course, P.A. and Tarran, G.A., 1996. The life cycle of *Emiliana huxleyi*: A brief review and a study of relative ploidy levels analyzed by flow cytometry. J. Mar. Syst., 9: 33–44.

Hadrys, H., Balick, M. and Schierwater, B., 1992. Applications of random amplified polymorphic DNA (RAPD) to molecular ecology. Mol. Ecol., 1: 55–63.

Hartl, D.L., 1980. Principles of Population Genetics. Sinauer, Sunderland, MA, 488 pp.

Hastings, I., 1991. Germline selection: population genetic aspects of the sexual/asexual life cycle. Genetics, 129: 1167–1176.

Hay, W.W., Mohler, H.P., Roth, P.H., Schmidt, R.R. and Boudreaux, J.E., 1967. Calcareous nannoplankton zonation of the Gulf Coast and Caribbean–Antillean area and transatlantic correlation. Trans. Gulf Coastal Geol. Soc., 17: 428–480.

Hillis, D.M. and Huelsenbeck, J.P., 1992. Signal, noise and reliability in molecular phylogenetic analyses. J. Hered., 83: 189–195.

Hillis, D.M., Huelsenbeck, J.P. and Cunningham, C.W., 1994. Application and accuracy of molecular phylogenies. Science, 264: 671–677.

Jordan, R.W. and Young, J.R., 1990. Proposed changes to the classification system of living coccolithophorids. Newsl. Int. Nannoplankton Assoc., 12: 15–18.

Keller, M.D., Selvin, R.C., Claus, M.H. and Guillard, R.R.L., 1987. Media for the culture of oceanic phytoplankton. J. Phycol., 23: 633–638.

Klaveness, D., 1972. *Coccolithus huxleyi* (Lohm.) Kampt. II. The flagellate cell, aberrant cell types, vegetative propagation and life cycles. Br. Phycol. J., 7: 309–318.

Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol., 16: 111–120.

Kooistra, W.F.C.H. and Medlin, L.K., 1996. Evolution of the diatoms (Bacillariophyta): IV. A reconstruction of their age from small subunit rRNA coding regions and the fossil record. Mol. Phyl. Evol., in press.

Larsen, L., Olsen, G.J., Maidak, B.L., McCaughey, M.J., Overbeek, R., Macke, R., Marsch, T.L. and Woese, C.R., 1993. The ribosomal database project. Nucl. Acids Res., 21(suppl.): 3021–3023.

Leipe, D.D., Wainwright, P.O., Gunderson, J.H., Porter, D., Patterson, D.J., Valois, F., Himmerich, S. and Sogin, M.L., 1994. The stramenopiles from a molecular perspective: 16S-like rRNA sequences from *Labyrinthuloides minuta* and *Cafeteria roenbergensis*. Phycologia, 33: 369–377.

Levitan, D.R. and Grosberg, R.K., 1993. The analysis of paternity and maternity in the marine hydrozoan *Hydractinia symbiolongicarpus* using randomly amplified polymorphic DNA (RAPD) markers. Mol. Ecol., 2: 315–326.

Lockhart, P.J., Steel, M.A., Hendy, M.D. and Penny, D., 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol., 11: 605–612.

Lohmann, H., 1902. Die Coccolithophoridae, eine Monographie der Coccolithen bildenden Flagellaten, zugleich ein Beitrag zur Kenntnis des Mittelmeerauftriebs. Arch. Protistenkd., 1: 89–165.

Lynch, M. and Milligan, B.G., 1994. Analysis of population genetic structure with RAPD markers. Mol. Ecol., 3: 91–99.

Maggs, C.A., Douglas, S.E., Fenety, J. and Bird, C.J., 1992. A molecular and morphological analysis of the *Gymnogongrus devoniensis* (Rhodophyta) complex in the North Atlantic. J. Phycol., 28: 214–232.

Maddison, W.P. and Maddison, D.R., 1992. MacClade. Analysis of Phylogeny and Character Evolution. Version 3. Sinauer, Sunderland.

McIntyre, A. and Bé, A.W.H., 1967. Modern coccolithophores of the Atlantic Ocean. I. Placoliths and Cyrtoliths. Deep-Sea Res., 14: 561–597.

Medlin, L.K., Elwood, H.J., Stickel, S. and Sogin, M.L., 1988. The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. Gene, 71: 491–499.

Medlin, L.K., Elwood, H.J., Stickel, S. and Sogin, M.L., 1991. Morphological and genetic variation within the diatom Skeletonema costatum (Bacillariophyta): evidence for a new species, Skeletonema pseudocostatum. J. Phycol., 27: 514–524.

Medlin, L.K., Williams, D.M. and Sims, P.A., 1993. The evolution of the diatoms (Bacillariophyta): I. Origin of the group and assessment of the monophyly of its major divisions. Eur. J. Phycol., 28: 261–275.

Medlin, L.K., Lange, M. and Baumann, M.E.M., 1994. Genetic differentiation amongst three colony-forming species of Phaeocystis: further evidence for the phylogeny of the Prymnesiophyta. Phycologia, 33: 199–212.

Medlin, L.K., Cooper, A., Hill, C., Wrieden-Prigge, S. and Wellbrock, U., 1995. Phylogenetic position of the Chromista plastids from 16S rDNA coding regions. Curr. Genet., 28: 560–565.

Neefs, J.-M., Van de Peer, Y., De Rijk, P., Goris, A. and De Wachter, R., 1991. Compilation of small ribosomal subunit RNA sequences. Nucl. Acids Res., 19 (suppl.): 1987–2015.

Nixon, K.C. and Wheeler, Q.D., 1990. An amplification of the phylogenetic species concept. Cladistics, 6: 211–223.

Okamura, B., Jones, C.S. and Noble, L.R., 1993. Randomly amplified polymorphic DNA analysis of clonal population structure and geographic variation in a freshwater bryozoan. Proc. R. Soc. Lond. B, 253: 147–154.

Olson, R.J., Vaulot, D. and Chisholm, S.W., 1986. Effect of environmental stresses on the cell cycle of two marine phytoplankton species. Plant Physiol., 80: 918–925.

Patwary, M.U., MacKay, R.M. and Van der Meer, J.P., 1993. Revealing genetic markers in Gelidium vagum (Rhodophyta) through the random amplified polymorphic DNA (RAPD) technique. J. Phycol., 29: 216–222.

Peever, T.L. and Milligroom, M.G., 1994. Genetic structure of Pyrenophora teres populations determined with random amplified polymorphic DNA markers. Can. J. Bot., 72: 915–923.

Saiki, R., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A., 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA-polymerase. Science, 239: 487–491.

Saitou, N. and Nei, M., 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol., 4: 406–425.

Sanger, F., Nicklen, S. and Coulsen, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA. 74: 5463–5467.

Shapiro, L.P., Campbell, L. and Haugen, E.M., 1989. Immunochemical recognition of phytoplankton species. Mar. Ecol. Progr. Ser., 37: 219–224.

Swofford, D.L., 1993. PAUP: Phylogenetic Analysis using Parsimony, Version 3.1.1. Ill. Nat. Hist. Surv., Champaign, Ill.

Thomsen, H.A., Østergaard, J.B. and Hansen, L.E., 1991. Heteromorphic life histories in arctic coccolithophorids (Prymnesiophyceae). J. Phycol., 27: 634–642.

Van Bleijswijk, J., Van der Wal, P., Kempers, R., Veldhuis, M. and Young, J.R., 1991. Distribution of two types of Emiliania huxleyi (Prymnesiophyceae) in the Northeast Atlantic region as determined by immunofluorescence and coccolith morphology. J. Phycol., 27: 566–570.

Van Bleijswijk, J.D.L., Kempers, R.S. and Veldhuis, M.J., 1994. Cell and growth characteristics of types A and B of Emiliania huxleyi (Prymnesiophyceae) as determined by flow cytometry and chemical analyses. J. Phycol., 30: 230–241.

Van Oppen, M.J.H., Olsen, J.L. and Stam, W.T., 1995. Genetic variation within and among North Atlantic and Baltic populations of the benthic alga Phycodrys rubens (Rhodophyta). Eur. J. Phycol., 30: 251–260.

Van Oppen, M.J.H., Klerk, H., De Graaf, M., Stam, W.T. and Olsen, J.L., 1996. Assessing the limits of random amplified polymorphic DNAs (RAPDs) in organisms with complex life histories: an example from the red algae. Mol. Ecol., submitted.

Vaulot, D., Birrien, J.L., Marie, D., Casotti, R., Veldhuis, M.J.W., Kraay, G.W. and Chrétiennot-Dinet, M.-J., 1994. Morphology, ploidy, pigment composition and genome size of cultured strains of Phaeocystis (Prymnesiophyceae). J. Phycol., 30: 1022–1035.

Verbeek, J.W., 1990. Late Quaternary calcareous nannoplankton biostratigraphy for the northern Atlantic Ocean. Meded. Rijks Geol. Dienst, 44: 13–33.

Williams, J.G.K., Kubelic, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl. Acids Res., 18: 6531–6535.

Williams, J.G.K., Hanafey, M.K., Rafalski, J. and Tingey, S., 1993. Genetic analysis using random amplified polymorphic DNA markers. Meth. Enzymol., 218: 704–740.

Young, J.R. and Westbroek, P., 1991. Genotypic variation in the coccolithophorid species Emiliania huxleyi. Mar. Micropaleontol., 18: 5–23.