GENOTYPING AN *EMILIANIA HUXLEYI* (PRYMNESIOPHYCEAE) BLOOM EVENT IN THE NORTH SEA REVEALS EVIDENCE OF ASEXUAL REPRODUCTION

Stacy A. Krueger-Hadfield[1],

Marine Biological Association, Citadel Hill Laboratory, Plymouth, PL1 2PB, UK

College of Charleston, Grice Marine Laboratory, 205 Fort Johnson Road, Charleston, SC 29412, USA[2]

Cecilia Balestreri[1],

Marine Biological Association, Citadel Hill Laboratory, Plymouth, PL1 2PB, UK

Department of Earth Sciences, Oxford University, South Parks Road, Oxford OX1 3AN, UK

Joanna Schroeder[1], *Andrea Highfield*,

Marine Biological Association, Citadel Hill Laboratory, Plymouth, PL1 2PB, UK

Pierre Helaouët

Sir Alister Hardy Foundation for Ocean Science, Citadel Hill Laboratory, Plymouth, PL1 2PB, UK

Jack Allum

School of Biological Sciences, Faculty of Science, University of Plymouth, Plymouth PL4 8AA, UK

Roy Moate

Plymouth Electron Microscope Centre, Faculty of Science, University of Plymouth, Plymouth PL4 8AA, UK

Kai T. Lohbeck,

Evolutionary Ecology of Marine Fishes, GEOMAR Helmholtz-Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany,

Biological Oceanography, GEOMAR Helmholtz-Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany

Peter I. Miller

Plymouth Marine Laboratory, Prospect Place, The Hoe, Plymouth, PL1 3DH

Ulf Riebesell,

Biological Oceanography, GEOMAR Helmholtz-Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany

Thorsten B. H. Reusch,

Evolutionary Ecology of Marine Fishes, GEOMAR Helmholtz-Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany

Ros E. M. Rickaby,

Department of Earth Sciences, Oxford University, South Parks Road, Oxford OX1 3AN, UK

Jeremy Young

Department of Earth Sciences, University College London, WC1E 6BT

Gustaaf Hallegraeff

School of Plant Science, University of Tasmania, Private Bag 55, Hobart, Tasmania 7001, Australia

Colin Brownlee and Declan C. Schroeder[2]

Marine Biological Association, Citadel Hill Laboratory, Plymouth, PL1 2PB

[1]shared first author

[2] author for correspondence

ABSTRACT

Due to the unprecedented rate at which our climate is changing, the ultimate consequence for many species is likely to be either extinction or migration to an alternate habitat. Certain species might, however, evolve at a rate that could make them resilient to the effects of a rapidly changing environment. This scenario is most likely to apply to species that have large population sizes and rapid generation times, such that the genetic variation required for adaptive evolution can be readily supplied. *Emiliania huxleyi* (Lohm.) Hay and Mohler (Prymnesiophyceae) is likely to be such a species as it is the most conspicuous extant calcareous phytoplankton species in our oceans with generation times of 1 day^{-1}. Here we report on a validated set of microsatellites, in conjunction with the coccolithophore morphology motif genetic marker, to genotype 93 clonal isolates collected from across the world. Of these, 52 came from a single bloom event in the North Sea collected on the D366 United Kingdom Ocean Acidification cruise in June-July 2011. There were 26 multilocus genotypes (MLGs) encountered only once in the North Sea bloom and 8 MLGs encountered twice or up to six times. Each of these repeated MLGs exhibited P_{sex} values of less than 0.05 indicating each repeated MLG was the product of asexual reproduction and not separate meiotic events. In addition, we show that the two most polymorphic microsatellite loci, EHMS37 and P01E05, are reporting on regions likely undergoing rapid genetic drift during asexual reproduction. Despite the small sample size, there were many more repeated genotypes than previously reported for other bloom-forming phytoplankton species, including a previously genotyped *E. huxleyi* bloom event. This study challenges the current assumption that sexual reproduction predominants during bloom events. Whilst genetic diversity is high amongst extant populations of *E. huxleyi*, the root cause for this diversity and ultimate fate of these populations still requires further examination. Nonetheless, we show that certain CMM genotypes are found everywhere; while others appear to have a regional bias.

1. INTRODUCTION

The coccolithophore, *Emiliania huxleyi* (Lohm.) Hay and Mohler (Prymnesiophyceae), is thought to be the main calcite producer on Earth (Westbroek et al. 1993) and is present in all but extreme polar oceans. It regularly forms extensive "white water" blooms in high latitude coastal and shelf ecosystems which extend over thousands of square kilometres and may persist for many months. In the later stages these blooms become visible to satellites such as the Moderate Resolution Imaging Spectroradiometer (MODIS) due to the mass shedding of highly scattering calcium carbonate coccoliths following large scale cell death (Holligan et al. 1993). During and post-bloom events, coccoliths sink towards the bottom of the water column taking large amounts of organic carbon with them (i.e., ballast effect), where a significant proportion become lost to the carbon cycle for millennia (Coxall et al. 2005, Riebesell et al., 2009). While the process of calcification results in decreased alkalinity of surface waters, potentially reducing the drawdown of CO_2 from the atmosphere, coccolithophores are also thought to contribute to reductions in atmospheric CO_2 by creating a net export of carbon to the seabed (Robertson et al. 1994, Riebesell & Tortell 2011).

Current estimates are that as much as 27% of the anthropogenic CO_2 produced from burning of fossil fuels released between 1959 -2011 has been absorbed by the oceans (Le Quéré et al. 2013). As CO_2 reacts with seawater, it generates dramatic changes in carbonate chemistry, including decreases carbonate ions and pH (ocean acidification) and an increase in bicarbonate ions. The consequences of this overall process are commonly referred to as ocean acidification. Moreover, ongoing atmospheric warming is expected to cause significant changes to the ocean climate by the end of this century (the average temperature of the upper layers of the ocean having increased by 0.6°C over the past 100 years, IPCC, 2007). The oceans are, therefore, experiencing unprecedented levels of change, raising

concerns about the impacts on key biological species such as *E. huxleyi*. The nature of such impacts will have important biological, ecological, biogeochemical and societal implications (Turley et al. 2010). Langer et al. (2009) found that different clonal *E. huxleyi* isolates vary in their phenotypic traits such as growth and calcification rate, suggesting a potential role for selection on standing genetic variation in shaping future populations. This mechanism was demonstrated by Lohbeck et al. (2012) who identified pH-driven selection on 6 clonal isolates from an *E. huxleyi* bloom near Bergen, Norway. Functional diversity within this set of clones allowed selective sorting over only 500 generations of exponential growth. These findings raise questions about the pace and relevance of such clonal sorting under natural conditions. Unfortunately, very little is known about the population biology of this key phytoplankton species and hence, forecasting how future populations will respond is difficult.

Future *E. huxleyi* populations could have a very different set of phenotypes when compared with present-day populations. This shift in phenotypic traits would have profound implications on ecosystem function and biogeochemical cycles. However, before we can address the effects of a rapidly changing climate on *E. huxleyi*, we must understand the very basic properties of its genetic diversity and ecological interactions. Martínez et al. (2007 & 2012) described a genetically rich, but stable *E. huxleyi* population using the coccolithophore morphology motif (CMM) in the North Atlantic. The CMM lies within the 3' untranslated mRNA region of the coccolith polysaccaride associated protein GPA, which is implicated in controlling coccolith structure (Schroeder et al., 2005). In addition, Iglesias-Rodríguez et al. (2006) and Hinz (2010) found high levels of intraspecific microsatellite genetic diversity in different *E. huxleyi* bloom events. In contrast to the CMM, microsatellites appear to be highly polymorphic markers that can resolve neutral genetic diversity within populations. The authors concluded that this is most likely driven by high rates of sexual reproduction. However, for species with large population sizes and rapid generation times, sex is not the

sole driver for high genetic diversity. Indeed, in species exhibiting large dispersal potential and geographic ranges, very high levels of genetic diversity are expected (i.e., molecular hyperdiversity, Cutter et al. 2013). In the natural environment *Saccharomyces* yeasts only reproduce sexually one in every 1000 to 3000 effective generations (Tsai et al. 2008). The mycorrhizal fungi (phylum Glomeromycota) are among the oldest and most successful symbionts of land plants and show no evidence of sexual reproduction (Van Kuren et al. 2012). Indeed, a combination of intra-individual polymorphism and effective population sizes in the Glomeromycota contribute to its evolutionary longevity.

The 10 polymorphic microsatellite markers used in Iglesias-Rodríguez et al. (2006) and Hinz (2010) were developed without the benefit of genome sequence information for this species (Read et al. 2013). In this study, we revisited 10 polymorphic microsatellite markers developed by Iglesias-Rodríguez et al. (2002, 2006), thoroughly tested and critically evaluated them in order to begin characterizing genetic diversity in an *Emiliania huxleyi* bloom event sampled during the D366 Sea Surface Consortium UK Ocean Acidification cruise (http://www.surfaceoa.org.uk/). The estimated genetic diversity, as defined by both the CMM and microsatellite markers, was used to critically revise the predominant mode of reproduction during an *E. huxleyi* bloom. Moreover, clonal diversity in the North Sea bloom event is compared to a biogeographic phytoplankton data set and the adaptive potential of future *E. huxleyi* populations facing a changing ocean is discussed.

2. MATERIALS AND METHODS

2.1 Validation of microsatellite primers

(i) Ten polymorphic microsatellite sequence primer pairs (AJ487304-17; AJ494737-42, Table 1) were blasted (blastn) against the CCMP1516 genome (Read et al. 2013) in order to verify the amplification of a single site within the genome.

(ii) PCR conditions used are as those described in Iglesias-Rodríguez et al. (2002, 2006), using the following modified PCR mix: 20 µL final volume, 2 µL of at least 10 ng DNA template, 1x reaction buffer, 1.5 mmol L^{-1} $MgCl_2$, 0.25 mmol L^{-1} deoxyribonucleotide triphosphate, 250 mmol L^{-1} each of unlabeled forward and reverse primers and 1 U of taq polymerase (GoTaq Flexi, Promega). In addition, the loci which produced repeatable PCR results and for which single-locus genetic determinism was verified were tested with an annealing temperature of 54 ^{o}C in order to facilitate the multiplexing of loci in the future. Initial PCR amplification trials were visualized using 1.8% agarose gels with a 50 bp ladder (New England Biolabs, MA, USA). Each reliable locus produced the same results as when tested with the original annealing temperature. Therefore, all subsequent reactions were run at 54 ^{o}C, though for the purposes of this study, all reactions were done in simplex.

(iii) In order to investigate the stability of alleles at each locus, strain no. 62 used in Lohbeck et al. (2012, 2013) was genotyped at the start of the experiment and after 1300 generations of exponential growth under a set of different CO_2 conditions (i.e. mapping any changes between June 2010 to November 2012). A second strain, CCMP1516 (Read et al. 2013), was also used spanning multiple generations, varying culture conditions under alternating exponential and stationary growth conditions that resulted in loss of coccolith production.

2.2 Microsatellite amplification

For optimization purposes, all successful PCR products were transferred to an ABI 3130 xL genetic analyzer (Applied Biosystems, Foster City, CA, USA) equipped with a 36 cm capillary array. The PCR mix was updated to include a fluorescently labeled forward primer: 150 mmol L^{-1} of the labeled forward primer, 100 mmol L^{-1} of the unlabeled forward primer and 250 mmol L^{-1} of the unlabeled reverse primer, where all other mix components

remained unchanged. Two µL of each PCR product was added to 10 µL of loading buffer containing 0.3 µL of size standard (GeneScan – 500 Liz, Applied Biosystems, Foster City, CA, USA) plus 9.7 µL of Hi-Di formamide (Applied Biosystems, Foster City, CA, USA). The loading mix was denatured at 92°C for 3 minutes. A positive and negative control was electrophoresed with each set of samples run on the sequencer.

After optimization, a subset of known genotypes was transferred to SourceBioScience Nottingham for fragment analysis on a 3730xL DNA analyser run on a 50 cm capillary array. For all clonal isolates, 7 µL of each PCR product was sent to SourceBioScience, including positive and negative controls for each sequencer run. All genotypes were scored manually using GENEMAPPER ver. 4 (Applied Biosystems, Foster City, CA, USA).

2.3 UK Ocean Acidification Research Cruise

The RV Discovery, cruise number 366, circumnavigated the British Isles in June/July 2011 as part of the UK Ocean Acidification research programme (http://www.surfaceoa.org.uk/). Samples used in this study were collected mainly in the North Sea (5 stations, Figure 1) and also in the Western coast of Scotland, Bay of Biscay and Western English Channel (Table 2).

2.4 Satellite Imagery

Ocean colour data from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor on the Aqua satellite were acquired from NASA OceanColor Website and processed to version R2013.0 using the PML Generic Earth Observation Processing System (GEOPS) (Shutler et al., 2005). Chlorophyll-a concentration was estimated using the OC3M algorithm, and a 7-day median composite calculated from the cloud-free pixels to gain a synoptic view. The enhanced colour view is obtained from 7-day median composites of

remote sensing reflectance at 547nm, 488nm and 443nm, combined as the red, green and blue channels respectively of an RGB image; hence this enhances the green-blue section of the visible spectrum. These images are useful for distinguishing different types of plankton or sediment: pure water looks blue; plankton blooms appear green or brown-red for more dense blooms; suspended sediment appears whitish/yellow; and *E. huxleyi* blooms appear brighter turquoise.

Sea-surface temperature (SST) data were generated from Advanced Very High Resolution Radiometer (AVHRR) data on NOAA satellites, acquired by NEODAAS-Dundee, and processed using the Panorama system (Miller et al., 1997). The NOAA non-linear SST (NLSST) algorithm was applied, and again the 7-day median composite used to reduce the effect of clouds.

2.5. E. huxleyi *clonal isolates*

Culture strains used in this study are listed in Table 2. The D366 samples were screened and sorted using a flow cytometer (FACSORT, BD Biosciences, San Jose, CA, USA) and cell counts were assessed using a flow cytometer (Accuri C6, BD Biosciences, San Jose, CA, USA) at the following thresholds: FSC 2000 and FL3 800. A dilution factor was calculated in order to obtain a starting concentration of approximately 1000 cells/mL. Each sample was subjected to a dilution-to-extinction regime in order to isolate individual cells and obtain clonal uni-algal cultures. All the cultures, including those additional geographically diverse strains resourced from various culture collection repositories (Table 2), were maintained in f/2 -Si medium (Guillard, 1975) in a constant temperature room at 15 °C and irradiated by a photon flux of 40-55 μmol $m^{-2}s^{-1}$ on a 16:8 hours LD cycle. The Qiagen DNeasy Blood and Tissue protocol (QIAGEN, Valencia, CA, USA) was used to extract DNA from each isolate.

2.6 Scanning Electron Microscopy

All of the samples were filtered using a 0.45 µm cellulose nitrate membrane filter, mounted onto metallic stubs using adhesive tape and coated in a thin layer of gold (Au) using an Au sputter coater. These were visualized using a JEOL 5600 Low Vacuum Scanning Electron Microscope. Scanning electron micrographs were captured at magnifications ranging between x8,000 - x20,000, and electron beam damage was minimized by operating the microscope at 15 kV. A total of 152 micrographs were captured, 62 from the environmental samples and 90 from the clonal isolates. All coccoliths were measured mainly at x20,000 magnification using ImageJ v1.38 software (http://rsb.info.nih.gov/ij/). Morphometrics included in analysis were distal shield length and width, central area length and width, average element length and width, and coccosphere diameter. To reduce bias and maintain a randomized sampling method during examination the surface area of the stubs was divided into nine squares. For each sample, six squares were randomly allocated using a random number generator, and examined for coccospheres with coccoliths lying flat on the substrate.

2.7 CMM amplification and sequencing

Amplification of the coccolith morphology motif (CMM, Schroeder et al. 2005) was achieved using a set of nested primers qCBP_F (5'-AGTCTCTCGACGCTGCCTC-3') and qCBP_R (5'-TGGCCTAGCACCAGTCTTTGG-3') corresponding to position 1203-1221 and 1283-1303, respectively, for the GPA mRNA of strain L (AF012542). The template DNA was added to 12.5 µL of QuantiTect Multiplex PCR NoROX kit master mix (Qiagen) and 1 µL for each probe (2 pmol), final volume of 25 µL for each reaction. PCR products were incubated with ExoSAP-IT (USB corporation) before being sequenced using the ABI

Big Dye terminator cycle sequencing ready reaction kit version 3.1 (Applied Biosystems) at Geneservice, Cambridge, UK.

2.8 CMM probe design and multiplex assay

Dual labeled probes (Table 3, Figure 2) were designed based on multiple sequence alignments from reference CMM sequences (Schroeder et al. 2005) and sequences generated from section 2.7. The probes were designed to be specific to a particular CMM group I to IV. Based on the sequence variation, two different probes were designed for CMM II and IV. The probes were divided into two multi-probe sets according to their fluorescent dyes and melting temperatures to allow for multiplexing (Table 3). The multiplex probe assay was carried out using a Corbette Rotor-Gene™ 6000 (QIAGEN, Valencia, CA, USA). The PCR proceeded with an initial denaturation at 95 °C for 15 minutes, followed by 40 cycles of a two step PCR: 94 °C for 60 seconds and 68 °C for 60 seconds for the first probe-set (probes I, II and III) and 94 °C for 60 seconds and 64 °C for 90 seconds for the second probe-set (probes IIb, IV and IVb). The fluorescence was acquired at the end of each annealing/extension step on the green, yellow and crimson channels.

2.9 Microsatellite multilocus genotype analyses

For each of the following analyses, the biogeographic (MLG-Geo) and North Sea (MLG) bloom clonal isolates (Table 2) were treated separately.

Prior to analyses, the number of repeated identical multilocus microsatellite genotypes (MLG) was computed using the Mutlilocus Matches option in GENALEX, ver. 6.5 (Peakall and Smouse, 2006, 2012). This option automates detection of repeated genotypes within a dataset. The genotypic richness (*R*) was calculated as:

$$R = \frac{G-1}{N-1}$$

where *G* is the number of distinct multilocus genotypes and *N* is the total number of studied

individuals (Dorken and Eckert, 2001). This modification of Ellstrand and Roose's (1987) index of clonal diversity was proposed by Dorken and Eckert (2001) such that the smallest possible value in a mono-clonal bloom is always 0, independently of sample size, and the maximum value is still 1, when all the different samples analyzed correspond to distinct clonal lineages.

Repeated MLGs may occur due to repeated sampling of the same genet which are produced through asexual reproduction (i.e., sampling many clones of the same genotype) or two distinct sexual events wherein the resulting cells share the exact same alleles at all loci. In order to estimate whether putative genets shared the same MLG, GENCLONE 2.0 was used (Arnaud-Haond and Belkhir, 2007). For each repeated MLG, P_{sex}, which is the probability for a given multilocus genotype to be observed in N samples as a consequence of two different sexual reproductive events, was calculated. For $P_{sex} > 0.05$, duplicated multilocus genotypes were considered as different genets having arisen from two independent sexual recombination events). If $P_{sex} < 0.05$, the duplicated multilocus genotypes were considered clones of the same genet (i.e., products of asexual reproduction).

2.10 Null alleles and linkage disequilibria

The frequency of null alleles was estimated using a maximum likelihood estimator in the software ML-NULLFREQ (Kalinowski and Taper, 2006). Linkage disequilibrium was tested for using GENEPOP, ver. 4.1 (Rousset, 2008). In addition to physical linkage on a chromosome, disequilibria may be due to a lack of recombination caused by clonal propagation or selfing (mating system) or to differences in allele frequencies among populations (spatial genetic structure). Significance testing was done using 1,000 permutations and Bonferroni correction (Sokal and Rohlf, 1995).

2.11 Sampling effort

Variation in allelic richness depends, essentially, on population size—large samples are expected to have more alleles, especially rare ones, than small samples. Rarefaction (in-silico) analyses involve subsampling each sample without replacement at a range of depths. By considering these subsamples taken from each sample, samples originally of different sizes can be compared and unbiased estimates of allelic richness computed (Kalinowski, 2005). Using rarefaction, as implemented in the program HP-RARE, ver. 1.0 (Kalinowski, 2005), the mean number of alleles (i.e., the number of alleles averaged over the total number of loci used) expected with a sample size of 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 and 75 were computed. In addition, the accumulation of different genotypes sampled in the North Sea bloom was calculated for CMM and the microsatellites separately using the FASTGROUPII web-based calculator (Yu et al., 2006).

2.12 Genetic distance

Bruvo et al.'s (2004) approach was used to calculate a genetic distance matrix from the alleles observed at the five microsatellite markers. The genetic distance between two `individuals' at a single microsatellite marker reflects the probability that the alleles of one individual mutated to the other. Probabilities are calculated using a model which assumes that slipped-strand mis-pairing is the main cause of changes in microsatellite length, resulting in single-step mutations. Notably, the Bruvo et al. (2004) calculation is independent of the microsatellite mutation rate, which in this study, and the majority of other studies, is unknown. A genetic distance matrix (comparing all samples) was computed for each microsatellite marker and the average of these matrices used in the analyses described. The Polysat package (Clark and Jasienuik, 2011) was used with R version 3.0.0 to perform the computations.

The genetic distance matrix was then analyzed using a permutational multivariate analysis of variance implemented in the R community ecology package `Vegan' (ver. 2.0-7, Oksanen et al., 2012). Termed ADONIS in the software package, the function partitions the variation observed in the distance matrix into sums of square distance matrices, characterising variation attributable to specified sources. This method is a robust alternative to parametric MANOVA (multivariate analysis of variance) and to ordination methods for describing how variation is attributed to different uncontrolled covariates. ADONIS is also an alternative to AMOVA (nested analysis of variance, Excoffier et al., 1992) for genetic data when there are some samples with limited numbers of individuals. Significance is assessed using *F*-statistics on sequential sums of squares from permutations of the raw data. In this study, permutational multivariate analysis of variance (ADONIS) was used to partition distance matrices among the following sources of variation in Sea Surface Temperature (SST), Northern vs. Southern Hemisphere and Locality. These tests were considered across all samples (i.e. the full genetic distance matrix) and within samples of specific CMM genotypes (i.e. submatrices of samples extracted from the full genetic distance matrix according to CMM genotype).

2.13 Global SSTs determination

Gridded (1° x 1°) Sea Surface Temperature (SST) data originated from the Hadley Centre (http://www.metoffice.gov.uk/hadobs/hadisst/). For those samples that fell outside the Hadley Centre SST coverage, i.e. the extreme coastal, their nearest SST values in a latitudinal direction were used instead. Similarly, *in situ* SST data were used for the Oslo Fjord strains. The matrices have been calculated by averaging SST values for the sampling effort (from January 2006 to December 2011). The samples were then clustered using a hierarchical clustering algorithm (termed hclust) implemented in R (version 3.0.0). The algorithm starts

with each sample as a cluster in itself and merges clusters together sequentially using Ward's minimum variance criterion (Ward, 1963). The sequential merging was continued until all samples were contained in a single cluster and the subsequent tree describing how the clusters merged was `cut' to yield three clusters. These clusters formed the low, medium and high SST groups.

3. RESULTS

3.1 Genetic inheritance, polymorphism and stability of the microsatellite markers

Loci P01F08 and P02A08 did not produce any PCR products after repeated attempts and alteration of PCR conditions (Table 1). These two markers were, therefore, the first to be eliminated from the suite of loci. In addition, there were no hits against the CCMP1516 genome for either of these two primer pairs (Table 1). Of the remaining eight markers that produced products, P02E11, P02E10 and EHMS15 resulted in multi-allelic (i.e., more than two, the maximum number of alleles possible for a diploid) profiles. There were at least three distinct peaks corresponding to at least three different alleles (Supplement, Fig. S1). Altering PCR conditions resulted in different allelic peaks rendering these loci unrepeatable. Moreover, P02E10 and EHMS15 primer pairs were found five and two times, respectively, in the CCMP1516 genome (Table 1). The multiple hits suggested these primer pairs may have amplified more than one region in the genome which corresponded to the multi-peaked profiles observed. As they were not repeatable and did not follow single-locus genetic determinism, they were rejected from further analyses.

EHMS37, P01E05, P02F11, P02E09 and P02B12 produced consistent results at their original annealing temperatures as well as the modified PCR program with an annealing temperature of 54 °C. For each of these polymorphic markers, single-locus Mendelian inheritance was assumed as only one (i.e., homozygous) or two peaks (i.e., heterozygous)

were observed for each of the clonal isolates tested. For the 15 samples (5 replicates, 3 different CO_2 conditions) from Lohbeck et al. (2012) extracted at the start of the CO_2 selection experiment in 2010, there were no differences between replicates and treatments. Further, in the same replicate selection lines extracted after 1300 generations of exponential growth, there was no change in the alleles present at each locus (Table 4). However, CCMP1516 showed variation in allele number and size for both EHMS37 and P01E05; the two most polymorphic loci (section 3.5). When comparing the genome sequence (Read et al. 2013) and previously characterized microsatellite data for this strain (Mackinder et al. 2011b) to our PCR amplicons, variation extended to the locus P02E09. The loss of the 137 PO1E05 allele in strain CCMP1516 genotyped in this study coincided with the loss of calcification, i.e. failure to produce a coccolithosphere. Unfortunately, Mackinder et al. (2011b) did not look at this allele (Table 4). Moreover, CCMP1516 can no longer produce haploid flagellate life-forms (P. von Dassow, *personal communication*), therefore these genetic modifications were not due to sexual recombination.

3.2 D366 E. huxleyi *cultures*

The techniques used to isolate clonal uni-algal *E. huxleyi* strains from the D366 cruise, selected only for calcified (diploid) forms that were cultured. We successfully produced 104 isolates from single cells, 88 (85%) remained viable (data not shown). Of these, 65 D366 isolates were successfully genotyped (Table 2), 52 of which originated from the North Sea bloom event (Figure 1). *E. huxleyi* morphotype A was the only morphotype to be identified (Figure 3). The mean coccosphere diameter was 5.4 μm (range 3.9-7.5 μm). Coccolith dimensions (Figure 4) were consistent with the classic morphotype A phenotype. The mean coccolith distal shield length was 3.2 μm ranging between 2.1-4.4 μm (Figure 4a), and the mean distal shield width was 2.6 μm ranging from 1.5 to 4 μm (Figure 4b). The

mean central area length was 1.6 μm (range 1.2-2.5 μm), and the mean central area width was 1.1 μm (range 0.7-1.7 μm). The mean average element length was 0.63 μm (range 0.25-0.95 μm), and the mean average element width was 0.12 μm (range 0.09-0.16 μm). All consistent with the classic morphotype A phenotype (Young et al. 2003).

3.3 Biogeographic E. huxleyi *cultures*

A select group of 26 *E. huxleyi* strains were chosen based mainly on origin and date of isolation. Our aim was to include strains from diverse geographic locations, from both northern and southern hemispheres and disparate climatic environments. In addition, we wanted to restrict the age of the cultures to lessen the influence of genetic drift from the point of isolation. Our final data set comprised strains not more than 5 years older than D366 strains, with the only exception being strain CH25/90 (Table 2) as the most recent and only one of two reference strains for morphotype B (CMM II) still in culture (Schroeder et al. 2005). The majority (84%) of all the biogeographic samples, including the D366 cultures, were isolated in 2011. Twenty isolates originate from the Southern hemisphere, while 6 isolates were isolated from the Mediterranean Sea, Oslo fjord, Irish Sea and Tsushima Strait, Japan (Table 2). The SST experienced by these strains ranged from 4.1 to 21.2 ºC (Figure 5). All strains could be clustered into three SST groups, namely low, <5 ºC, medium >5 & <14.3 ºC, and high >14.3 ºC (Table 2). The North Sea SSTs as observed by AVHRR (Figure 1C) are consistent with the SST clustering ranges that were based on Hadley Centre temperatures (Table 2).

3.4. CMM genotyping

Isolates in our reduced D366 dataset could be divided into three main CMM groups, namely homozygous for CMM I, homozygous for CMM IV and heterozygous for CMM I/IV

& III/IV (Table 2). It is, however, important to note that two of the 13 isolates that did not make the final reduced D366 dataset, produced complex MLGs and CMM profiles; all indicative of the presence of multiple genotypes in the same sample (data not shown). For technical reasons, these and the remainder 11 strains were not included in later analyses.

The CMM identity was mainly determined by applying the multiplex CMM probe assays (Supplement, Figs. S2 and S3), with sequencing of CMM amplicons from a few isolates to validate the probe assay results (Table 2). Note that multiple CMM probes were designed to account for the additional sequence variation outside the designated CMM region (Figure 2). When this was taken into account for two of the main affected CMMs, namely CMM II and IV, both sets of probes improved the sensitivity of the assay.

Of the North Sea D366 clonal isolates, 38 were homozygous for CMM I, 3 were homozygous for the CMM IV and 11 were heterozygous for CMM I/IV (Table 2). Therefore, CMM I was the most numerically abundant genotype. "CMM I in a homozygous state was also found in other geographic strains, seven were of Chilean and two of Norwegian origins (Table 1).". Similarly, CMM IVs were distributed widely geographically, while CMM I/IVs where restricted to the Northern hemisphere.

No CMM IIs were detected in our D366 dataset. The five B/C and C morphotypes from the Southern Ocean and Tsushima Strait, respectively, were however shown only to have the CMM II genotype (Table 2). There are 91 samples in this data set and of these 6 are homozygous CMM II (including the homozygous CMM II morphotype B Ch25/90 reference strain - Schroeder et al. 2005). Furthermore, exactly these 6 samples are characterized by a morphotype other than type A (morphotype R being a Southern Ocean over-calcified variant of A). The probability that these non-morphotype A samples are the only CMM II genotypes by chance is $1/({}^{91}C_6) = 1.5e\text{-}09$. The number ${}^{91}C_6 = 666563898$ is the total number of ways 6

samples can be selected from 91, it suggests the observed result is highly unlikely to have occurred by random chance.

3.5 Microsatellite genotyping

There were significantly greater amplification rates in this study (Table 2) compared to Iglesias-Rodríguez et al. (2006; t = 5.18, df = 5, p = 0.004), but no difference between this study and Hinz (2010; t = 0.75, df = 4, p = 0.493). However, the amplification rate at locus P02B12 in Hinz (2010) was only 66%, whereas in this study it was 100%.

One hundred and eight alleles were characterized across the five microsatellite loci. The number of alleles ranged from two to ten in the North Sea bloom, whereas there were five to 17 alleles encountered on a global scale (Table 2). Each of the loci corresponded to a stepwise mutation model. EHMS37 was the most polymorphic locus whereas P02B12 was the least polymorphic locus. Allele frequencies are available upon request.

Of the 52 clonal isolates genotyped in the North Sea bloom, 26 MLGs were only encountered once, five MLGs were encountered twice, two MLGs were encountered three times, one MLG was encountered five times and, finally, one MLG was encountered six times. The genotypic richness, *R*, in the North Sea was 0.667, the smallest value reported during a phytoplankton bloom. Moreover, each duplicated MLG was characterized by P_{sex} values much smaller than 0.05 (Table 2). In other words, it was extremely unlikely that they were the product of two independent meiotic events. All repeated microsatellite MLGs also shared the same CMM allele. Consequently, all repeated MLGs were considered descendants of the same genotype. In addition, there was also a repeated microsatellite MLG encountered three times in a bloom sampled off the coast of Chile in 2011 (Table 2). This repeated MLG exhibited P_{sex} values much smaller than 0.05 (Table 2) and, as above, was considered descendants of the same genotype.

There was no evidence of linkage disequilibrium in the North Sea bloom (i.e., all p-values were > 0.05 before Bonferroni correction). There was evidence of null alleles at each locus except P02F11 in the North Sea bloom. The null allele frequencies varied from 0.194 at EHMS37 to 0.258 at P01E05. However, as demonstrated by Krueger-Hadfield et al. (2011 & 2013), null allele frequencies calculated in diploid stages of haploid-diploid life cycles could be biased due to violation of some of the assumptions underlying maximum likelihood estimators. Therefore, null alleles may be present in our diploid strains (i.e., a diploid strain may have been scored as homozygous at locus EHMS37, but was in fact a heterozygote for the allele amplified and for an allele that was not amplified due to, for example, a possible mutation in the primer binding site). However, the frequency estimates are likely upwardly biased and the actual numerical value should be treated with caution as we are unsure of certain parameters of the *E. huxleyi* life cycle (i.e., mating system), which could bias the maximum likelihood estimator.

3.6 Sampling effort

There was a difference between CMM and the microsatellites in that the rarefaction curve for CMM genotypes reached a plateau whereas the microsatellites did not (Supplement, Fig. S4). Although the microsatellite rarefaction curve did not plateau, at the point at which sampling was ceased, the gradient of curve was not as steep as that observed in other studies (e.g., Hinz, 2010). That said, a slight increase did occur between 50 and 75 genes sampled (Supplement, Fig. S4).

3.7 Population genetic structure at different spatial scales

Using the ADONIS method to attribute variation in microsatellite Bruvo genetic distances (Figure 6) to variation in SST, Northern vs Southern hemispheres and locality

yielded weak correlations: between 8 and 31% of the variation in the distance matrix was explained by these variables (Table 5). In addition, the morphotypes did not cluster together on the basis of microsatellite genetic distance, notably the four B/C morphotypes from the cooler Australian waters were dispersed between other morphotypes (Figure 6). Within CMM genotypes, locality explained the most variation out of the three covariates.

4. DISCUSSION

The use of a validated set of microsatellites and the CMM functional genetic marker demonstrated clear evidence of asexual reproduction prevailing during a single *E. huxleyi* bloom event in the North Sea in 2011. Eight genotypes were encountered between two to six times across the sampling dates and locations of the bloom event. Despite the small sample size, there were many more repeated genotypes than previously reported for other bloom-forming phytoplankton species, including a previously genotyped *E. huxleyi* bloom event. This study challenges the assumption that sex drives genetic diversity within and between *E. huxleyi* populations. Whilst genetic diversity is high amongst extant populations of *E. huxleyi*, the root cause for this diversity still requires further examination in order to be able to predict the impacts of unprecedented levels of climate change are having on key biological species such as *E. huxleyi*.

4.1. Asexual dominance in the D366 North Sea Bloom

For population genetics, the key benefit of microsatellites is the high inter-individual variation, which makes it possible to study both intra- and inter-population genetic diversity. The evolutionary dynamics, biological function, genomic distribution and practicality of microsatellites have been summarized in a wide variety of reviews (see Schlötterer 1998, Selkoe and Toonen 2006). As a down-side, mutation rates may be so high that appreciable

genotypic changes may occur during an observational period (e.g., Tesson et al., 2013). However, whether these are real mutations or mis-scoring (discussed again below) would need more careful analysis. Microsatellite mutation rates vary, but the typical range is thought to be 10^{-2} to 10^{-6} mutations per locus per generation (Li et al., 2002). Hinz (2010) estimated the number of mutations per microsatellite locus per generation in *E. huxleyi* to be between $7x10^{-3}$ to 142 over a 15 year culture period. Assuming this calculation is meaningful for certain strains, 1 mutation per 1000 generations is expected statistically within each lineage. As each of these mutations would be selectively neutral, the probability of fixation would be negligible and would be dependent upon the size of the asexual population. In other words, even if occasional mutations occurred in uni-algal cultures, it would not be possible to detect - as seen for the Lohbeck et al. (2012) strain that did not show any changes based on microsatellite genotyping during 1300 asexual generations. However, we investigated a second strain (CCMP1516) that originates from the warmer tropical Pacific environment and has been in culture since 1991 (Schroeder et al. 2005). In contrast, the strain used in Lohbeck et al. (2012) originates from Bergen (relative cooler environment) and was maintained in culture for a lot less time (i.e., since 2009) and under continuous exponential growth. Our data suggests the change in selective pressure incurred due to culturing in artificial laboratory conditions over a 20 year time period has had a compounding effect on fitness. While adaptation to high pCO2 conditions had little effect on Lohbeck strains ability to calcify, (i.e. cells never lost their ability to produce coccoliths),, we predict that the same would not be true for CCMP1516. We predict that it would have behaved very differently as it often loses its ability to calcify under current pCO_2 scenarios. Replicate cultures of CCMP1516 have to be kept to ensure that the calcified form of CCMP1516 is not lost for good.

Mis-scoring of alleles was certainly a problem for CCMP1516 (Table 4). The variations observed in the EHMS37 and P02F11 are likely as a result of noise, user interpretation and between sequencer shifts associated with the stutter peaks surrounding the "dominant" microsatellite peak (expanded upon again later). By contrast, the variations observed in P01E05 and PO2E09 are more intriguing. What is the source of this variation? Could the P01E05 loci be informative about the state of calcification? We know that the allele size 137 for PO1E05 was likely present in the genome sequence dataset (Read et al. 2013) but was omitted from the final genome due to the complexities of assembly, i.e. assembly of genomes of diploid organisms eliminates subtle variation and reports mainly on a single consensus chromosomal copy. However, the disappearance of this allele in the 2012 non-calcifying strain (Table 4) raises important questions regarding the role of this genomic region in the calcification process. What is certain however is that some genomic regions within *E. huxleyi* are subject to greater genetic drift or rearrangements within an asexually maintained state. Until we determine the source and the nature of these variations and understand the effect and extent of the changes on the fitness of a diversity of strains, estimation of microsatellite mutation rates per locus for *E. huxleyi* would be futile. This in turn raises questions of the usefulness of these particular microsatellites in *E. huxleyi* population genetics

Microsatellites have previously been used to explore genetic diversity and population structure in several bloom-forming phytoplankton (e.g., diatoms: Rynearson and Armbrust, 2000, 2004, 2005, Evans et al., 2005; dinoflagellates: Alperman et al., 2009, Erdner et al., 2011, Casabianca et al., 2011; coccolithophores: Iglesias-Rodríguez et al., 2006). High levels of intraspecific genetic variability have been reported in all phytoplankton groups, but often these results are discussed as somewhat of a paradox. A bloom event should be dominated by asexual reproduction, as asexual reproduction is likely the only mode by which such large

biomass can be generated over short time periods. Yet, the paradigm of sexual reproduction being the source of exceptional genetic diversity during bloom periods has pervaded the microbial literature. For *E. huxleyi*, we have seen that sexual recombination was not the cause of the microsatellite variation observed in CCMP1516. This has been documented in other asexually reproducing organisms, such as fungi. Sexual recombination was thought to only occur between two fungal strains of opposite mating types; however, Lin et al. (2005) demonstrated recombination in isogenic mating types. We have no evidence that recombination between diploid *E. huxleyi* cells are the source for the genetic variation observed, but this merely highlights the many possibilities that could explain high levels of genetic variation within species. Due to the high levels of genetic diversity and linkage equilibrium observed in our study, genetic drift had occurred, but was unlikely to have contributed to genetic diversity directly during the D366 North Sea bloom. Indeed, rare recombination events can erase any signatures of clonality, such as heterozygote excess and linkage disequilibrium (Halkett et al., 2005). Yet, the fact that many genotypes were re-sampled indicates that asexual reproduction was driving the bloom formation.

This is one of the only studies which calculated P_{sex} values in order to demonstrate the origin of the repeated MLGs (sexual or asexual events). In contrast, Iglesias-Rodríguez et al. (2006) and Hinz (2010) reported few, if any, repeated MLGs in two previous studies on *E. huxleyi* blooms, but this is likely due to several features of these studies which do not arise directly from the biology of this coccolithophore. First, the sample size used to calculate genetic diversity from a sampling location or time point (Iglesias-Rodríguez et al. 2006) or a particular mesocosm or time point (Hinz 2010), was small and therefore, repeated genotypes may not be detected due to chance or isolation techniques. Second, Iglesias-Rodríguez et al. (2006) included several loci which have been shown in this study to be multi-allelic and are, therefore, not suitable for genotypic diversity estimates. Further, only seven out of the 85

isolates tested amplified at all ten loci. It is unclear from Iglesias-Rodríguez et al. (2006) what the genotypes were for the validated five loci used in this study and whether these genotypes were in fact different. Third, in Iglesias-Rodríguez et al. (2006), the authors used two microsatellites, P01E05 (potentially mutating after long periods of time in culture) and EHMS15 (multi-allelic), in isolation to describe the geographic distribution of genotypes and potential reductions in gene flow. However, if one uses restricted data sets to perform these calculations, such as between Northern and Southern hemisphere strains, spurious results will be encountered. For example, we demonstrated that SST, Northern vs. Southern hemisphere and Locality does not explain the overall clustering of the strains based on CMM or microsatellite profiling.

Iglesias-Rodriguez et al. (2006) also estimated the number of genotypes in the environment to be, at the minimum, 2.4×10^{20}. Yet, the computational method of calculating this value depends on locus independence. There were no calculations of linkage disequilibrium, but if one assumes the loci are independent and in linkage equilibrium based on the results of the current study, this would not be a major violation. However, the method likely overestimates the number of different genotypes. If there were four alleles at a locus, then in Iglesias-Rodríguez et al.'s (2006) method, there would be six different heterozygous combinations plus the four possible homozygous states. This would then be multiplied by the next figure at the next locus and so on. The computational method used does not take into account the manner in which certain alleles are encountered or that some combinations are never found. Capture-recapture statistics is a preferred method to estimate the number of lineages within a bloom in a conservative manner.

One issue with studies, such as this in coccolithophores (also see Cook et al., 2013) or in diatoms, as in Rynearson and Armbrust (2005), is the sample size of clonal isolates from a given "site." For macroalgae, it is necessary to sample at least 30 diploids and haploids (for

those which have haploid-diploid life cycles) from a population (Krueger-Hadfield 2011). However, due to the difficulty of single cell extractions in some phytoplankton and the large scale of their distribution and bloom events, more than 30 samples of at least the diploid phase are likely to be necessary. For example, the daily sample size of clonal isolates from Rynearson and Armbrust (2005) varied from 20 to 76 with values of *D* ranging from 0.87 to 1.0. Plotting the *N* versus *R* resulted in a significant negative slope (r^2= 0.456, b = -0.001, p < 0.023), indicating that increasing the sample size of clonal isolates increases the chances of re-encountering a MLG.

Yet, even values in Rynearson and Armbrust (2005) with apparently sufficient sample size to detect repeated MLGs, there were still more unique MLGs encountered than in the North Sea *E. huxleyi* bloom studied here. This might be expected due to the nature of diatom blooms. Diatoms continue dividing until they reach a critical size when sexual reproduction is triggered (Chepurnov et al., 2005). However, Rynearson and Armbrust (2005) did not find any sexual stages during the sampling of a *Ditylum brightwellii* bloom event in Puget Sound. Therefore, the high genotypic diversity in the diatom bloom may have been due to past sexual events, but also resting stages of *D. brightwellii*. Resting stages can act as inocula for blooms and provide an additional diversifying effect.

4.2 A place for CMM

Ascribing a genetic basis to a particular coccolithophore morphotype has been attempted in several studies which were able to show some genetic differentiation among the strains tested (gpa/CMM: Schroeder et al., 2005, tufA: Cook et al., 2011, cox1b and atp4 Hagino et al., 2011). There are four main morphotypes: Type A [E. huxleyi var huxleyi] has varying levels of calcification, global distribution and is the most prevalent in bloom events (Hagino et al. 2011, Cook et al., 2011 & 2013). The other three, namely C [*E. huxleyi* var

kleijniae Young & Westbroek ex Medlin & Green] (Young et al. 2003), B [*E. huxleyi* var *pujosae* (Verbeck) Young & Westbroek ex Medlin & Green] and B/C [*Emiliania huxleyi* var *aurorae* Cook & Hallegraeff] are found in the most northern and southern latitudes (van Bleijswijk et al. 1991, Young et al. 2003, Cook et al. 2013). Two other morphotypes, R (Young et al. 2003, Cook et al. 2011) & O (Hagino et al. 2011) have been reported in the southern and northern latitudes, respectively. Schroeder et al. (2005) used the CMM to reinforce the partitioning of the A & B morphotypes. In addition, morphotype A has a combination of CMM I, CMM III or CMM IV alleles, while morphotype B was only found associated with CMM II. The present study has expanded on this finding by showing that the morphotype R is likely an over-calcified form of A, and more surprisingly linking morphotypes C and B/Cs to B. While the latter share a similar biogeography, their cell sizes span the smallest (C – 2.5 µm) to the largest (B – 7 µm) for this species.

CMM I was the numerically dominant allele in the form of homozygous CMM I and heterozygous CMM I/IV. However, CMM IV was the second most abundant genotype and the most widely distributed. This was partially supported by the ADONIS variation test (i.e. locality being the greatest influence on the genetic variation for homozygous CMM IV), but also by the discovery of a CMM IV repeated MLG in the North Sea and the Western English Channel (see Table 2, MLG 34).

CMM II, on the other hand, was not detected in the North Sea locality. One of the original B morphotype strains, CH25/90, originated from the North Sea (van Bleijswijk et al. 1994) at a location not too dissimilar from the D366 North Sea sampling sites. In addition, Martinez et al. (2012) reported the presence of CMM II in the North Sea in 1999. The absence of morphotype B or CMM II in our D366 culture collection raises important questions as to whether the well-documented increase in SSTs over the past decade could have negatively affected the natural habitat for this morphotype. We know that CMM IIs,

including B/C and C's, predominantly or even exclusively occupy the more northern and southern latitudes. It is conceivable to predict that in the case of the North Atlantic the morphotype B's could have moved further north to cooler environments. Helaouët et al. (2011) showed a similar northward movement for the copepod, *Calanus*, over the past decade. Higher spatial and temporal resolution is required before we can conclude that climate change could also have attributed to the range restriction of morphotype B. Taken together, morphotype A appears to be more resilient and thus dominates at a regional and global scale while morphotype B is more sensitive and thus likely to be more specific to the niche it occupies.

The true biological function of the calcium binding protein, GPA, which CMM is thought to influence (Schroeder et al., 2005), remains to be resolved. Recent studies have shown that GPA is most likely not directly involved in the production of coccoliths in *E. huxleyi* (Mackinder et al. 2011b, Rokitta et al. 2011) but there is evidence to suggest GPA binds Ca^{2+} (Corstjens et al. 1998). The link between CMM and morphotypes observed in this study is clear (i.e., one in one and a half billion chance of all six CMM II's being randomly associated with morphotypes other than the dominant A morphotype). Interestingly, the plastid gene *tufA* (Cook et al. 2011) supports the division of *E. huxleyi* into two main subgroups or varieties (Cook et al. 2013), while the mitochondrial (mtDNA) *cox1b-ATP4* genes (Hagino et al. 2011) found that no genetic distinction could be made. The most parsimonious explanation for this apparent discrepancy is that the chromosomal (CMM) and plastid (*tufA*) alleles are under different selection pressure, possibly as a function of their individual attributes to fitness, while the mtDNA genes provide an insight into the ancestral history of this species through their maternal line. Such discrepancies between mtDNA and chromosomal phylogenies are well documented in animal systems. For example, apparent discrepancies exist between the distributions of the lineages of mtDNA and of the two major

Y-chromosome lineages in mice (Boissinot and Boursot 1997). Some subspecies share the same mtDNA lineage but have different chromosome lineages or vice versa (Boissinot and Boursot 1997). Partitioning *E. huxleyi* into different CMM subgroups certainly has its place in population genetics as it appears to be more informative than when using microsatellites in isolation.

4.3 Implications for future research in microalgal population genetics

The bloom population in *E. huxleyi* appears to be relatively stable over consecutive blooms in a similar location, as also documented in *Ditylum brightwellii* (Rynearson and Armbrust 2005). Martínez et al. (2007) demonstrated a stable inter-annual population using CMM genotypes using environmental DNA. However, this has been a limiting step as microsatellites necessitate clonal cultures or individuals. Preliminary data suggest certain allelic combinations are found in different years in the North Atlantic (*unpublished data*). Yet, this raises a critical point. As microalgae inhabit such a stochastic environment that changes rapidly, how should genotypes be scored? As gradations of allele frequencies or distinctive diagnostic genotypes? Schuller et al. (2012) demonstrated genetic difference in *Saccharomyces cerevisiae* were due to fine-scale allelic changes rather than diagnostic genotypes (i.e., very different allele sizes). The authors cautioned that though microsatellites are useful for population-level analyses, sub-strain level discrimination may occur due to their relatively high mutation rates. In this study, there was noise around the dominant allele of several base pairs, suggesting these alleles were recent mutations from the dominant (i.e., 100 and 104 alleles surrounding the 102 allele in P02F11, Table 4). Therefore, it might be necessary to treat microalgae in a similar manner to yeast. Does this represent something biological or is it simply noise? Are other bloom events in other basins dominated by the same or different alleles? Applying the techniques used in this study will enable us to

respond to these questions and in so doing begin to describe the genetic structure of *E. huxleyi* in more detail. This is a critical step for further exploring host-viral dynamics (e.g., Martinez et al. 2007), the occurrence of meta-population dynamics (Rynearson et al. 2009), associated levels of genetic diversity (Walser and Haag, 2012) and understanding how this species will respond to climatic change or ocean acidification. High standing genetic variation and the fact that bloom events do not appear to cause a genetic bottleneck indicate that phytoplankton populations have the potential to adapt fast enough to keep pace with ongoing climate change. *E. huxleyi* is a relatively new species, having only appeared less than 300,000 years ago (Raffi et al. 2006). Therefore, it will be interesting to explore the population genetics of this species in more detail in order to determine how this species has and is evolving.

ACKNOWLEDGEMENTS

Special thanks go to Sue Cook, Bente Edvardsen, Ian Probert and Kyoko Hagino for either supplying us with DNA or live cultures for the biogeographic comparison. Thanks also go to Stephen Cotterell, Matt Hall and Gideon Mordecai for the technical advice and assistance; Mairi Knight for use of the capillary sequencer at Plymouth University. This project has been supported by Interreg IV Marinexus project (Ref: 1956/4073) and the UK Ocean Acidification programme.

REFERENCES

Alpermann, T. J., Beszteri, B., John, U., Tillmann U., and Cembella, A. D.: Implications of life-history transitions on the population genetic structure of the toxigenic marine dinoflagellate *Alexandrium tamarense*, Mol. Ecol., 18, 2122-2133, 2009.

Arnaud-Haond, S., and Belkhir K..: GenClone 1.0 a new program to analyse genetics data on clonal organisms, Mol. Ecol. Notes, 7, 15–17, 2007.

Bruvo, R., Michiels N. K., D'Souza, T. G., and Schulenburg, H.: A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level, Mol. Ecol., 13, 2101–2106, 2004.

Boissinot, S, and Boursot, P.: Discordant phylogeographic patterns between the Y chromosome and mitochondrial DNA in the house mouse: selection on the Y chromosome?, Genetics, 146, 1019–1034, 1997.

Casabianca, S., Penna, A., Pecchioli, E., Jordi, A., Basterretxea, G., and Vernesi, C.: Population genetic structure and connectivity of the harmful dinoflagellate *Alexandrium minutum* in the Mediterranean Sea, Proc. R. Soc. B, 279, 129-138, 2010.

Chepurnov, V. A., Mann, D. G., Sabbe, K., Vannerum, K., Casteleyn, G., Verleyen E., Peperzak, L., and Vyverman, W.: Sexual reproduction, mating system, chloroplast dynamics and abrupt cell size reduction in *Pseudo-nitzschia pungens* from the North Sea (Bacillariophyta), Eur. J. Phycol., 40, 379-395, 2005.

Clark, L.V. and Jasieniuk, M.: Polysat an R package for polyploid microsatellite analysis, Mol. Ecol. Res., 11, 562—566, 2011.

Cook, S.S., Whittock, L., Wright, S.W., and Hallegraeff, G.M.: Photosynthetic pigment and genetic differences between two Southern Ocean morphotypes of *Emiliania huxleyi* (Haptophyta), J. Phycol., 47, 615-626, 2011.

Cook, S.S., Jones, R.C., Vaillancourt, R. E., and Hallegraeff, G.M.: Genetic differentiation among Australian and Southern Ocean populations of the eubiquitous coccolithophore *Emiliania huxleyi* (Haptophyta), Phycologia, 52, 368-374. 2013

Corstjens, P., van der Kooij, A., Linschooten, C., Brouwers, G. J., Westbroek, P. and deVrind-de Jong, E.W.: GPA, a calcium binding protein in the coccolithophorid *Emiliania huxleyi* (Prymnesiophyceae), J. Phycol., 34, 622-630, 1998.

Coxall, H. K., Wilson, P. A., Pëlike, H., Lear, C. H. and Backman, J.: Rapid stepwise onset of Antarctic glaciation and deeper calcite compensation in the Pacific Ocean, Nature, 433, 53-57, 2005.

Cutter, A. D., Jovelin, R. and Dey, A.: Molecular hyperdiversity and evolution in very large populations, Mol. Ecol. 22, 2074-2095, 2013.

Dorken, M. E. and Eckert, C.G.: Severely reduced sexual reproduction in northern populations of a clonal plant, J. Ecol., 89, 339-350, 2001.

Ellstrand, N. C. and Roose, M. L.: Patterns of genotypic diversity in clonal plant species, Am. J. Bot., 74, 123-131, 1987.

Erdner, D. L., Richlen, M., McCauley, L. A. R., and Anderson, D. M.: Diversity and dynamics of a widespread bloom of the toxic dinoflagellate *Alexandrium fundyense*. PLoS ONE, 6, e22965, 2011.

Evans, K. M., Kühn, S. F.,and Hayes, P. K.: High levels of genetic diversity and low levels of genetic differentiation in North Sea *Pseudo-nitzchia pungens* (Bacillariophyceae) populations, J. Phycol., 41, 506-514, 2005.

Excoffier, L., Smouse, P., and Quattro, J.: Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data, Genetics, 131, 479–491, 1992.

Guillard, R.R.L.: Culture of phytoplankton for feeding marine invertebrates. W.L. Smith, M.H. Chanley (Eds.), Culture of Marine Invertebrates Animals, Plenum, New York, pp. 296–360, 1975.

Hagino, K., Bendif, E. M., Young, J. R., Kogame, K., Probert, I., Takano, Y., Horiguchi, T., De Vargas, C. and Okada, H.: New evidence for morphological and genetic variation in the cosmopolitan coccolithophore *Emiliania huxleyi* (Prymnesiophyceae) from the *cox1b-ATP4* genes, J. Phycol., 47, 1164-1176, 2011.

Halkett, F., Simon, J. C., and Balloux, F.: Tackling the population genetics of clonal and partially clonal organisms. Trends Ecol. Evol., 20, 194-201, 2005.

Helaouët, P., Beaugrand, G. and Reid, P. C.: Macrophysiology of *Calanus finmarchicus* in the North Atlantic Ocean, Prog. Oceanogr., 91, 217-228, 201.

Hinz, D. J.: *Emiliania huxleyi* and climate change: a genetic and biogeographic investigation of bloom dynamics for a key phytoplankton species in the global carbon cycle, University of Southampton, PhD Thesis, 169 pp, 2010.

Holigan, P. M., Fernandez, E., Aiken, J., Balch, W. M., Boyd, P., Burkill, P. H., Finch, M., Groom, S. B., Malin, G., Muller, K., Purdie, D. A., Robinson, C., Trees, C. C., Turner, S. M. and van der Wal, P. A biogeochemical study of the coccolithophore, *Emiliania huxleyi*, in the North Atlantic, Glob. Biogeochem. Cycle, 7, 879-900, 1993.

Iglesias-Rodriguez, M. D., Garcia Saez, A., Groben, R., Edwards, K. J., Batley, J., Medlin, L. K. and Hayes, P. K.: Polymorphic microsatellite loci in global populations of the marine coccolithophorid *Emiliania huxleyi*, Mol. Ecol. Notes, 2, 495-497, 2002.

Iglesias-Rodriguez, M. D., Schofield, O. M., Batley, J., Medlin, L. K. and Hayes, P. K.: Intraspecific genetic diversity in the marine coccolithophore *Emiliania huxleyi* (Prymnesiophyceae): the use of microsatellite analysis in marine phytoplankton population studies, J. Phycol., 42, 526-536, 2006.

IPCC: Climate Change 2007: Synthesis report. Contribution of working groups I.

Kalinowski, S. T.: HP-RARE 1.0 a computer program for performing rarefaction on measures of allelic richness, Mol. Ecol. Notes, 5, 187-189, 2005.

Kalinowski, S. T. and Taper, M. L..: Maximum likelihood estimation of the frequency of null alleles at microsatellite loci, Conserv. Genetics, 7, 991-995, 2006.

Krueger-Hadfield, S. A.: Structure des populations chez l'algue rouge haploid-diploïde *Chondrus crispus*: système de reproduction, differérenciation génétique et épidémiologie. Université de Pierre et Marie Curie, Paris, Pontificia Universidad Católica de Chile, Santiago, PhD Thesis, 2011.

Krueger-Hadfield, S. A., Collén, J., Daguin, C., and Valero, M.: Distinguishing among gents and genetic population structure in the haploid-diploid seaweed *Chondrus crispus* (Rhodophyta), J. Phycol., 47, 440-450, 2011.

Krueger-Hadfield, S. A., Roze, D., Mauger, S., and Valero, M.: Intergametophytic selfing and microgeographic genetic structure shape populations of the intertidal red seaweed *Chondrus crispus* (Rhodophyta), Mol. Ecol., 22, 3242-3260, 2013.

Langer, G., Nehrke, G., Probert, I., Ly, J., and Ziveri. P.: Strain-specific responses of *Emiliania huxleyi* to changing seawater carbonate chemistry, Biogeosciences, 6, 2637-2646, 2009.

Le Quéré, C., Andres, R. J., Boden, T., Conway, T., Houghton, R. A., House, J. I., Marland, G., Peters, G. P., van der Werf, G. R., Ahlström, A., Andrew, R. M., Bopp, L., Canadell, J. G., Ciais, P., Doney,S. C., Enright, C., Friedlingstein, P., Huntingford, C., Jain, A. K., Jourdain, C., Kato, E., Keeling, R. F., Klein Goldewijk, K., Levis, S., Levy, P., Lomas, M., Poulter, B., Raupach, M. R., Schwinger, J., Sitch, S., Stocker,B. D., Viovy, N., Zaehle, S., and Zeng, N.: The global carbon budget:1959-2011, Earth System Science Data, 5, 165-185, 2013.

Li, Y-C., Korol, A. B., Fahima, T., Beiles, A., and Nevo, E.: Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review, Mol. Ecol., 11, 2453-2465, 2002.

Lin, X., Hull, C. M., and Heitman, J.: Sexual reproduction between partners of the same mating type in *Cryptococcus neoformans*, Nature, 434, 1017-1021, 2005.

Lohbeck, K. T., Riebesell, U. and Reusch, T. B. H.: Adaptive evolution of a key phytoplankton species to ocean acidification, Nature Geoscience, 5, 346-351, 2012.

Lohbeck, K. T., Riebesell, U., Collins, S., and Reusch. T. B. H.: Functional Genetic Divergence in high CO_2 adapted *Emiliania huxleyi* populations, Evolution, 67, 1892-1900, 2013.

Mackinder, L., Bach, L, Schulz, K., Wheeler, G., Schroeder, D., Riebesell, U. and Brownlee, C.: The molecular basis of inorganic carbon uptake mechanisms in the coccolithophore *Emiliania huxleyi*, Eur. J. Phycol., 46, 142-143, 2011a.

Mackinder, L., Wheeler, G., Schroeder, D., von Dassow, P., Riebesell, U. and Brownlee, C.: Expression of biomineralization-related ion transport genes in *Emiliania huxleyi*, Environ. Microbiol., 13, 3250- 3265, 2011b.

Martínez-Martínez, J., Schroeder, D. C., Larsen, A., Bratbak, G. and Wilson, W. H.: Molecular dynamics of *Emiliania huxleyi* and co-occurring viruses during two separate mesocosm studies, Appl. Environ. Microbiol., 73, 554-562, 2007.

Martinez Martinez, J., Schroeder, D., and Wilson, W. H.: Dynamics and genotypic composition of *Emiliania huxleyi* and their co-occurring viruses during a coccolithophore bloom in the North Sea, FEMS Microbiol. Ecol., 81, 315-323, 2012.

Miller, P., Groom, S., McManus, A., Selley, J. and Mironnet, N.: Panorama: a semi-automated AVHRR and CZCS system for observation of coastal and ocean processes.

RSS97: Observations and Interactions, P. Rem. Sens. Soc., Reading, pp. 539-544, 1997.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R.,. O'Hara, R. B , Simpson, G. L., Solymos, P., Stevens, M. H. H. and Wagner, H.: vegan Community ecology package, R package version 2.0-5., http://CRAN.R-project.org/package=vegan, 2012.

Peakall, R., and Smouse, P. E,: GENALEX 6.2: genetic analysis in Excel. Population genetic software for teaching and research, Mol. Ecol. Notes, 6, 288–295, 2006.

Peakall, R. and Smouse, P. E.: GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update, Bioinformatics, 28, 2537-2539, 2012.

Raffi, I., Backman, J., Fornaciari, E., Pälike, H., Rio, D., Lourens, L., and Hilgen, F.: A review of calcareous nannofossil astrobiochronology encompassing the past 25 million years. Quaternary Sci. Rev., 25, 3113-3137, 2006.

Read B. A., Kegel J., Klute M. J., Kuo A., Lefebvre S. C., Maumus F., Mayer C., Miller J., Monier A., Salamov A., Aguilar M., Claverie J-M., Frickenhaus S., Gonzalez K., Herman E.K., Lin Y-C., Napier J., Ogata H.i, Sarno A. F., Shmutz J., Schroeder D., de Vargas C., Verret F., von Dassow P., Valentin K., Van de Peer Y., Wheeler G., Emiliania huxleyi Annotation Consortium, Dacks J. B.,. Delwiche C. F, Dyhrman S. T., Glöckner G., John U., Richards T., Worden A. Z., Young J., Zhang X. and Grigoriev I. V. *Emiliania*'s pan genome drives the phytoplankton's global distribution, Nature, 499, 209–213, 2013.

Riebesell, U., Körtzinger, A., and Oschlies, A.: Sensitivities of marine carbon fluxes to ocean change, Proc. Natl Acad. Sci. USA, 106, 20802-20809, 2009.

Riebesell, U., Tortell, P.D.: Effects of ocean acidification on pelagic organisms and ecosystems. In: *Ocean Acidification*, Gattuso, J.-P., Hansson, L. (eds.) Oxford University Press, pp. 99-121, 2011.

Robertson, J. E., Robinson, C., Turner, D. R., Holligan, P., Watson, A. J., Boyd, P., Fernandez, E., and Finch M.: The impact of a coccolithophore bloom on oceanic carbon uptake in the northeast Atlantic during summer 1991, Deep Sea Res. Part I: Oceanogr. Res. Papers, 41, 297-314, 1994.

Rokitta, S. D., de Nooijer, L. J., Trimborn, S., de Vargas, C., Rost, B., and John, U.: Transcriptome analyses reveal differential gene expression patterns between the life-cycle stages of *Emiliania huxleyi* (Haptophyta) and reflect specialization to different ecological niches, J. Phycol., 47, 829-838, 2011.

Rousset, F.: Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux, Mol. Ecol. Res., 8, 103-106, 2008.

Rynearson, T. A., and Armbrust, E. V.: DNA fingerprinting reveals extensive genetic diversity in a field population of the centric diatom *Ditylum brightwellii*, Limnol. Oceanogr., 45, 1329-1340, 2000.

Rynearson, T. A., and Armbrust. E. V.: Genetic differentiation among populations of the planktonic marine diatom *Ditylum brightwellii* (bacillariophyceae), J Phycology, 40, 34-43, 2004.

Rynearson, T.A, and Armbrust. E. V.: Maintenance of clonal diversity during a spring bloom of the centric diatom *Ditylum brightwellii*, Mol. Ecol., 14, 1631-1640, 2005.

Rynearson, T. A., Lin, E. O., Horner, R. A., and Armbrust. E. V.: Gene flow and metapopulation structure in the planktonic diatom *Ditylum brightwellii*, Protist, 160, 111-121, 2009.

Schlötterer, C. Are microsatellites really simple sequences? Current Biology 8: 132-134, 1998.

Schroeder, D. C., Biggi, G. F., Hall, M., Davy, J., Martinez Martinez, J., Richardson, A., Malin G. and Wilson, W. H.: A genetic marker to separate *Emiliania huxleyi* (Prymnesiophyceae) morphotypes, J. Phycol., 41, 874-879, 2005.

Schuller, D., Cardoso, F., Sousa, S., Gomes, P., Gomes, A. C. , Santos, M. A. S. and Casal, M.: Genetic diversity and population structure of *Saccharomyces cerevisiae* strains isolated from different grape varieties and winemaking regions, PLoS ONE, 7, e32507, 2012.

Shutler, J. D., Smyth, T. J., Land, P. E. and Groom, S. B.: A near-real time automatic MODIS data processing system, I. J. Rem. Sens., 26, 1049-1055, 2005.

Seber, G. A. F.: Multivariate Observations. Hoboken, NJ: John Wiley & Sons, Inc., 1984.

Selkoe, K. A. and Toonen, R. J.: Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers, Ecol. Lett., 9, 615-629, 2006.

Sokal, R. R. and Rohlf, F. J.: Biometry: the principles and practice of statistics in biological research. 3rd ed. W. H. Freeman and Co., New York, 1995.

Spath, H.: Cluster Dissection and Analysis: Theory, FORTRAN Programs, Examples. Translated by J. Goldschmidt. New York: Halsted Press, 1985

Tesson, S. V. M., Legrand, C., van Oosterhout, C., Montresor, M., Kooistra, W., and Procaccini, G.: Mendelian inheritance pattern and high mutation rates of microsatellite alleles in the diatom *Pseudo-nitzschia multistriata*, Protist, 164, 89-100, 2013.

Turley, C., Brownlee, C., Findlay, H.S., Mangi, S., Ridgwell, A., Schmidt, D.N. and Schroeder, D.C. Ocean Acidification in MCCIP Annual Report Card 2010-11, MCCIP Science Review, 27pp. www.mccip.org.uk/, 2010.

Tsai, I. J., Bensasson, D., Burt, A. and Koufopanou, V.: Population genomics of the wild yeast *Saccharomyces paradoxus* quantifying the life cycle, Proc. Natl Acad. Sci. USA, 105, 4957–4962, 2008.

van Bleijswijk, J. D. L., van der Wal, P., Kempers, E. S., Veldhuis, M.J.W, Young, J. R., Muyzer, G., de Vrind-De Jong, E., and Westbroek, P.: Distribution of two types of *Emiliania huxleyi* (Prymnesiophyceae) in the Northeast Atlantic region as determined by immunofluorescence and coccolithmorphology, J. Phycol., 27:566-570, 1991.

van Bleijswijk, J. D. L., Kempers, R., Veldhuis, M. J., and Westbroek, P.: Cell and growth characteristics of types A and B of *Erniliania huxleyi* (Prymnesiophyceae) as determined by flow cytometry and chemical analysis, J. Phycol., 30:230-240, 1994.

Van Kuren, N. W., den Bakker, H. C., Morton, J. B., and Pawlowska, T. E.: Ribosomal RNA gene diversity, effective population size, and evolutionary longevity in asexual Glomeromycota, Evolution, 67, 207-224, 2012.

Walser, B. and Haag, C. R.: Strong intraspecific variation in genetic diversity and genetic differentiation in *Daphnia magna* - the effects of population turnover and population size, Mol. Ecol., 21, 851–861, 2012.

Westbroek, P., Brown C. W., van Bleijswijk J., Brownlee C., Brummer G.J., Conte M., Egge J., Fernández E., Jordan R., Knappertsbusch M., Stefels J., Veldhuis M., van der Wal P., and Young, J.: A model system approach to biological climate forcing. The example of *Emiliania huxleyi*, Glob. Planet. Change, 8, 27-46, 1993.

Young, J., Geisen, M., Cros, L., Kleijne, A., Probert, I., Sprengel, C., and Ostergaard, J.B. A guide to extant coccolithophore taxonomy, J. Nannoplankton. Res., 1, 1–124 Special Issue, 2003.

Yu, Y. N., Breitbart, M., McNairnie, P., Rohwer, F: FastGroupII: a web-based bioinformatics platform for analyses of large 16S rDNA libraries, BMC Bioinformatics, 7, 57, 2006.

FIGURELEGENDS

Figure 1. Earth observation 7-day composite data showing *Emiliania huxleyi* bloom development before, during and after cruise: (a) Enhanced ocean colour from Aqua-MODIS, showing coccoliths as bright patches and persistent cloud in black. (b) Chlorophyll-a concentration from Aqua-MODIS, with cloud in light grey. (c) Sea-surface temperature from AVHRR, where numbered circles indicate cruise stations listed in Table 2.

Figure 2 Alignment of CMM sequences produced in this study to reference CMMs (Schroeder et al. 2005). The CMM region is boxed. The dash line indicates the split between two subgroups of CMMs based on variation outside the CMM genotype. The bases shaded in grey show the positions of the probes (Table 3).

Figure 3. Scanning electron micrograph of a mixed *Emiliania huxleyi* culture prior to single cell isolation originating from D366 station 5 in the North Sea. Bar = 5µm.

Figure 4. Frequency distribution histograms of all the measurements taken for distal shield length (a) and width (b): 95% t-confidence for mean is shown.

Figure 5. Average Sea Surface Temperature (SST) values for the sampling effort from January 2006 to December 2011 for the world's oceans. The four regions that include Europe, Japan, Chile and Australia that represent all our dataset are shown in greater detail. Key: temperature colour index from blue to red, 0ºC to 25ºC, respectively.

Figure 6. Multi-dimensional scaling plots constructed using Bruvo et al.'s (2004) genetic distance creating a 2-dimensional representation of the dissimilarity matrix used for the

permutational multivariate analysis of variance (ADONIS {VEGAN} community ecology package in R) for all the samples.

Figure 7 Multi-dimensional scaling plots constructed using Bruvo et al.'s (2004) genetic distance creating a 2-dimensional representation of the dissimilarity matrix used for the permutational multivariate analysis of variance (ADONIS {VEGAN} community ecology package in R) in the biogeographic group: a) CMM I, b) CMM IV & c) CMM II

Table 1. Characteristics of the 10 microsatellite markers isolated in *Emiliania huxleyi* by Iglesias-Rodríguez et al. (2002, 2006). N_{Bio}, total number of distinct alleles observed over the biogeographic data set and N_{NS}, total number of distinct alleles observed over the North Sea Bloom data set).

Locus	**Acc. No.**	**Fluorescent Dye**	**Profile***	**BLAST**	**Amplification proportion**	**A-range (bp)**	**N_{Bio}**	**N_{NS}**
EHMS37	AJ494737 AJ494738	PET	one	1	0.93	194-340	37	12
P01E05	AJ494739 AJ494740	6-FAM	one	1	0.96	106-190	28	10
P02F11	AJ487316 AJ487317	NED	one	0	0.98	98-192	21	8
P02E09	AJ494741 AJ494742	PET	one	1	0.99	82-172	10	7
P02B12	AJ487310 AJ487310	NED	one	0	1.00	204-224	11	4
P02E11	AJ487312 AJ487313	VIC	multiple	1	-	-	-	-
P02E10	AJ487314 AJ487315	6-FAM	multiple	5	-	-	-	-
EHMS15	AJ487304 AJ487305	VIC	multiple	2	-	-	-	-
P01F08	AJ487306 AJ487307	-	none.	0	-	-	-	-
P02A08	AJ487308 AJ487309	-	none	0	-	-	-	-

*: number of loci amplified

Table 2. *Emiliania huxleyi* isolates used in this study

Strain number	Strain ID	MLG	P_{sex} value	Station	Year	Location	Latitude	Longitude	Morphotype	CMM sequencing	CMM probe assay		EHMS37		P01E05		P02F11		P02E09		P02B12		SST	SST cluster	Culture Collection
1	D366 106-2	MLG 1	3.01×10^{-16}	1	2011	North Sea	57.2	3.48	A	I (■)	I	I	210	210	142	142	98	102	102	102	208	208	10.544	medium	Plymouth
2	D366 17-1	-	-	5	2011	North Sea	56.5	3.65	A	I (■)	I	I	210	210	142	142	98	102	102	102	208	208	10.544	medium	Plymouth
3	D366 26-1	-	-	5	2011	North Sea	56.5	3.65	A	-	I	I	210	210	142	142	98	102	102	102	208	208	10.544	medium	Plymouth
4	D366 124-1	MLG 2	7.54×10^{-05}	4	2011	North Sea	57.45	5.53	A	I (■)	I	I	210	210	142	142	102	102	102	102	208	208	10.639	medium	Plymouth
5	D366 40-1	-	-	5	2011	North Sea	56.5	3.65	A	I (■)	I	I	210	210	142	142	102	102	102	102	208	208	10.544	medium	Plymouth
6	D366 32-5	-	-	5	2011	North Sea	56.5	3.65	A	I (□)	I	I	210	210	142	142	102	102	102	102	208	208	10.544	medium	Plymouth
7	D366 124-2	-	-	4	2011	North Sea	57.45	5.53	A	-	I	I	210	210	142	142	102	102	102	102	208	208	10.639	medium	Plymouth
8	D366 124-5	-	-	4	2011	North Sea	57.45	5.53	A	-	I	I	210	210	142	142	102	102	102	102	208	208	10.639	medium	Plymouth
9	D366 35-1	-	-	5	2011	North Sea	56.5	3.65	A	-	I	I	210	210	142	142	102	102	102	102	208	208	10.544	medium	Plymouth
10	D366 25-3	MLG3	-	5	2011	North Sea	56.5	3.65	A	I (■)	I	I	210	210	142	142	102	102	102	102	212	212	10.544	medium	Plymouth
11	D366 112-1	MLG4	2.05×10^{-07}	2	2011	North Sea	57.56	4.2	A	I (■)	I	I	214	214	142	142	102	102	98	102	208	208	10.571	medium	Plymouth
12	D366 30-1	-	-	5	2011	North Sea	56.5	3.65	A	I (■)	I	I	214	214	142	142	102	102	98	102	208	208	10.544	medium	Plymouth
13	D366 40-5	-	-	5	2011	North Sea	56.5	3.65	A	-	I	I	214	214	142	142	102	102	98	102	208	208	10.544	medium	Plymouth
14	D366 33-3	-	-	5	2011	North Sea	56.5	3.65	A	-	I	I	214	214	142	142	102	102	98	102	208	208	10.544	medium	Plymouth
15	D366 112-3	-	-	2	2011	North Sea	57.56	4.2	A	-	I	I	214	214	142	142	102	102	98	102	208	208	10.571	medium	Plymouth
16	D366 112-2	MLG5	-	2	2011	North Sea	57.56	4.2	A	-	I	I	214	214	132	142	102	106	98	102	208	208	10.571	medium	Plymouth
17	D366 21-5	MLG6	-	5	2011	North Sea	56.5	3.65	A	-	I	I	206	218	132	142	102	106	102	102	208	208	10.544	medium	Plymouth
18	D366 22-4	MLG7	-	5	2011	North Sea	56.5	3.65	A	-	I	I	196	206	132	142	102	106	102	102	208	208	10.544	medium	Plymouth
19	D366 124-3	MLG8	2.74×10^{-05}	4	2011	North Sea	57.45	5.53	A	I (☒)	I	I	202	210	142	142	102	102	102	102	208	208	10.639	medium	Plymouth
20	D366 124-4	-	-	4	2011	North Sea	57.45	5.53	A	-	I	I	202	210	142	142	102	102	102	102	208	208	10.639	medium	Plymouth
21	D366 26-3	MLG9	0.035	5	2011	North Sea	56.5	3.65	A	I (☒)	I	I	206	206	142	142	102	102	102	102	208	208	10.544	medium	Plymouth
22	D366 26-4	-	-	5	2011	North Sea	56.5	3.65	A	-	I	I	206	206	142	142	102	102	102	102	208	208	10.544	medium	Plymouth
23	D366 26-5	MLG10	-	5	2011	North Sea	56.5	3.65	A	I (☒)	I	I	206	206	142	142	102	130	102	102	208	208	10.544	medium	Plymouth
24	D366 33-1	MLG11	5.16×10^{-06}	5	2011	North Sea	56.5	3.65	A	I (■)	I	I	206	206	152	152	118	134	102	102	208	208	10.544	medium	Plymouth
25	D366 33-2	-	-	5	2011	North Sea	56.5	3.65	A	-	I	I	206	206	152	152	118	134	102	102	208	208	10.544	medium	Plymouth
26	D366 17-3	MLG12	-	5	2011	North Sea	56.5	3.65	A	-	I	I	206	206	152	152	102	102	102	102	208	212	10.544	medium	Plymouth
27	D366 35-2	MLG13	4.15×10^{-08}	5	2011	North Sea	56.5	3.65	A	I (■)	I	I	206	206	132	132	102	102	102	102	208	212	10.544	medium	Plymouth
28	D366 35-3	-	-	5	2011	North Sea	56.5	3.65	A	-	I	I	206	206	132	132	102	102	102	102	208	212	10.544	medium	Plymouth
29	D366 36-4	-	-	5	2011	North Sea	56.5	3.65	A	-	I	I	206	206	132	132	102	102	102	102	208	212	10.544	medium	Plymouth
30	D366 20-4	MLG14	-	5	2011	North Sea	56.5	3.65	A	I (■)	I	I	206	206	132	132	102	102	102	102	208	208	10.544	medium	Plymouth
31	D366 21-3	MLG15	-	5	2011	North Sea	56.5	3.65	A	I (■)	I	I	206	214	148	148	102	102	102	102	208	208	10.544	medium	Plymouth
32	D366 89-5	MLG16	-	5	2011	North Sea	56.5	3.65	A	-	I	I	210	254	132	132	102	102	102	102	208	208	10.544	medium	Plymouth
33	D366 26-2	MLG17	1.61×10^{-05}	5	2011	North Sea	56.5	3.65	A	I (■)	I	I	210	218	126	142	102	102	102	102	208	208	10.544	medium	Plymouth
34	D366 30-2	-	-	5	2011	North Sea	56.5	3.65	A	-	I	I	210	218	126	142	102	102	102	102	208	208	10.544	medium	Plymouth
35	D366 24-1	MLG18	-	5	2011	North Sea	56.5	3.65	A	-	I	I	202	210	130	130	102	102	106	106	208	208	10.544	medium	Plymouth
36	D366 37-4	MLG19	-	5	2011	North Sea	56.5	3.65	A	-	I	I	200	206	132	142	102	102	106	106	208	212	10.544	medium	Plymouth
37	D366 120-1	MLG20	-	3	2011	North Sea	57.91	4.85	A	-	I	I	206	210	132	152	102	102	102	102	208	208	10.571	medium	Plymouth
38	D366 40-3	MLG21	-	5	2011	North Sea	56.5	3.65	A	I (■)	I	I	206	218	126	142	102	102	102	102	208	208	10.544	medium	Plymouth
39	PVDCH1	MLG1-Geo	-	-	2011	Pacific Ocean, Chile	-30.25	-71.7	R	-	I	I	276	290	146	146	102	102	102	102	204	208	15.645	high	Roscoff
40	PVDCH8	MLG2-Geo	-	-	2011	Pacific Ocean, Chile	-30.25	-71.7	R	-	I	I	282	294	150	150	102	102	102	102	204	208	15.645	high	Roscoff
41	PVDCH6	MLG3-Geo	3.36×10^{-08}	-	2011	Pacific Ocean, Chile	-30.25	-71.7	R	-	I	I	282	282	156	156	102	102	102	102	208	208	15.645	high	Roscoff
42	PVDCH112	-	-	-	2011	Pacific Ocean, Chile	-30.16	-71.56	R	-	I	I	282	282	156	156	102	102	102	102	208	208	15.645	high	Roscoff
43	PVDCH47	-	-	-	2011	Pacific Ocean, Chile	-30.16	-71.56	R	-	I	I	282	282	156	156	102	102	102	102	208	208	15.645	high	Roscoff
44	PVDCH140	MLG4-Geo	-	-	2011	Pacific Ocean, Chile	-30.16	-71.56	R	-	I	I	276	284	160	160	102	102	88	102	208	208	15.645	high	Roscoff
45	PVDCH148	MLG5-Geo	-	-	2011	Pacific Ocean, Chile	-30.16	-71.56	R	-	I	I	268	276	146	146	102	130	98	154	208	208	15.645	high	Roscoff
46	UIO262	MLG6-Geo	-	-	2010	Oslo fjord	59.25	10.71	A	-	I	I	206	214	150	156	102	130	102	102	208	208	12	medium	University of Oslo
47	UIO269	MLG7-Geo	-	-	2010	Oslo fjord	59.25	10.71	A	-	I	I	206	214	146	156	102	130	102	102	208	208	12	medium	University of Oslo
48	D366 106-4	MLG22	-	1	2011	North Sea	57.2	3.48	A	-	I	IV	208	218	126	142	102	102	102	102	208	208	10.544	medium	Plymouth
49	D366 106-5	MLG23	-	1	2011	North Sea	57.2	3.48	A	-	I	IV	210	212	132	132	102	102	102	116	208	208	10.544	medium	Plymouth
50	D366 120-2	MLG24	-	3	2011	North Sea	57.91	4.85	A	-	I	IVb	206	210	150	150	102	102	102	102	208	208	10.571	medium	Plymouth
51	D366 31-3	MLG25	-	5	2011	North Sea	56.5	3.65	A	-	I	IVb	206	210	132	142	102	102	102	102	208	208	10.544	medium	Plymouth

52	D366 34-1	MLG26	-	5	2011	North Sea	56.5	3.65	A	-	I	IV	210	210	132	142	102	102	82	102	208	208	10.544	medium	Plymouth
53	D366 36-5	MLG27	-	5	2011	North Sea	56.5	3.65	A	-	I	IVb	210	210	132	132	102	102	102	102	208	208	10.544	medium	Plymouth
54	D366 34-3	MLG28	-	5	2011	North Sea	56.5	3.65	A	-	I	IV	210	210	132	132	102	106	102	102	208	208	10.544	medium	Plymouth
55	D366 19-2	MLG29	-	5	2011	North Sea	56.5	3.65	A	-	I	IVb	196	210	142	142	102	106	102	102	208	208	10.544	medium	Plymouth
56	D366 36-2	MLG30	-	5	2011	North Sea	56.5	3.65	A	-	I	IVb	210	210	132	156	102	106	102	102	208	208	10.544	medium	Plymouth
57	D366 21-4	MLG31	-	5	2011	North Sea	56.5	3.65	A	-	I	IVb	202	206	132	142	102	106	102	102	208	208	10.544	medium	Plymouth
58	D366 31-2	MLG32	-	5	2011	North Sea	56.5	3.65	A	-	I	IV	210	210	160	160	102	102	102	102	208	208	10.544	medium	Plymouth
59	D366 48-3	MLG8-Geo	-	-	2011	Western coast of Scotland	56.78	-7.4	A	-	I	IVb	196	206	142	142	102	126	102	102	208	208	11.384	medium	Plymouth
60	D366 80-1	MLG9-Geo	-	-	2011	Bay of Biscay	45.7	-7.16	A	-	I	IV	214	214	132	146	102	102	102	102	208	212	15.523	high	Plymouth
61	D366 97-5	MLG10-Geo	-	-	2011	Western English Channel	50.08	-4.61	A	-	I	IV	206	226	132	142	102	102	102	102	208	212	13.007	medium	Plymouth
62	D366 48-5	MLG11-Geo	-	-	2011	Western coast of Scotland	45.7	-7.16	A	-	III	IVb	196	202	150	150	102	102	102	102	208	208	11.384	medium	Plymouth
63	D366 126-2	MLG33	-	4	2011	North Sea	57.45	5.53	A	IV (▲)	IV	IV	202	202	132	142	102	102	102	102	212	212	10.639	medium	Plymouth
64	D366 30-4	MLG34	-	5	2011	North Sea	56.5	3.65	A	IV (▶)	IV	IV	196	206	132	132	102	134	82	102	208	208	10.544	medium	Plymouth
65	D366 91-2	MLG35	-	5	2011	North Sea	56.5	3.65	A	-	IV	IVb	196	210	132	142	102	106	102	102	208	208	10.544	medium	Plymouth
66	D366 48-2	MLG12-Geo	-	-	2011	Western coast of Scotland	56.78	-7.4	A	IV (◀)	IV	IV	196	264	138	138	98	134	102	102	208	208	11.384	medium	Plymouth
67	D366 80-3	MLG13-Geo	-	-	2011	Bay of Biscay	45.7	-7.16	A	-	IV	IVb	276	276	146	146	102	102	102	102	208	208	15.523	high	Plymouth
68	D366 80-4	MLG14-Geo	3.24×10^{-07}	-	2011	Bay of Biscay	45.7	-7.16	A	-	IV	IVb	276	276	146	160	102	102	102	102	208	208	15.523	high	Plymouth
69	D366 80-5	-	-	-	2011	Bay of Biscay	45.7	-7.16	A	-	IV	IVb	276	276	146	160	102	102	102	102	208	208	15.523	high	Plymouth
70	D366 71-1	MLG15-Geo	4.42×10^{-05}	-	2011	Bay of Biscay	46.2	-7.21	A	IV (▼)	IV	IV	202	202	146	146	102	118	102	102	212	212	15.149	high	Plymouth
71	D366 71-4	-	-	-	2011	Bay of Biscay	46.2	-7.21	A	-	IV	IV	202	202	146	146	102	118	102	102	212	212	15.149	high	Plymouth
72	D366 98-1	MLG34	-	-	2011	Western English Channel	50.08	-4.61	A	IV (▲)	IV	IV	196	206	132	132	102	134	82	102	208	208	13.007	medium	Plymouth
73	D366 J31	MLG16-Geo	-	-	2011	Irish Sea	52.46	-5.9	A	-	IV	IV	196	230	150	150	98	102	82	102	208	208	12.268	medium	Roscoff
74	D366 J7	MLG17-Geo	-	-	2011	Irish Sea	52.46	-5.9	A	-	IV	IV	238	290	132	150	102	134	98	102	208	208	12.268	medium	Roscoff
75	PVDCH250	MLG18-Geo	-	-	2011	Pacific Ocean, Chile	-34.1	-79	R	-	IV	IV	206	214	130	130	102	102	82	82	208	220	16.429	high	Roscoff
76	PVDCH280	MLG19-Geo	-	-	2011	Pacific Ocean, Chile	-34.1	-79	R	-	IV	IV	202	260	142	160	102	102	102	102	208	208	16.429	high	Roscoff
77	PVDCH288	MLG20-Geo	-	-	2011	Pacific Ocean, Chile	-34.1	-79	R	-	IV	IV	202	260	150	150	102	102	98	102	208	208	16.429	high	Roscoff
78	BOUM6	MLG21-Geo	-	-	2008	Mediterranean Sea, Spain	39.1	5.35	A	-	IV	IV	202	260	150	156	102	102	102	102	208	208	19.395	high	Roscoff
79	BG10-6	MLG22-Geo	-	-	2007	Irish Sea	49.5	-10.5	A	IV (▲)	IV	IV	206	214	106	118	98	98	102	102	208	208	13.417	medium	Roscoff
80	EHSO_50.28	MLG23-Geo	-	-	2007	Southern Ocean	-49.58	149.25	A	IV (△)	IVb	IVb	210	210	142	142	102	120	102	154	208	208	9.133	medium	UTAS
81	EHSO_5.25Q	MLG24-Geo	-	-	2006	Southern Ocean	-49.58	149.25	A	-	IV	IV	210	210	132	156	102	102	102	154	208	208	9.133	medium	UTAS
82	EHSO_50.14	MLG25-Geo	-	-	2006	Southern Ocean	-49.58	149.25	A	IV (▼)	IV	IV	210	210	132	156	102	102	102	154	212	214	9.133	medium	UTAS
83	EHSO_50.25	MLG26-Geo	-	-	2006	Southern Ocean	-49.58	149.25	A	IV (▼)	IV	IV	202	206	132	156	102	102	102	154	212	212	9.133	medium	UTAS
84	EHSO_50.3	MLG27-Geo	-	-	2006	Southern Ocean	-49.58	149.25	A	-	IV	IV	210	210	132	170	102	120	102	154	208	212	9.133	medium	UTAS
85	EHBi_21	MLG28-Geo	-	-	2006	Bicheno, EastTasmania	-41.11	148.16	A	-	IVb	IVb	272	284	132	156	102	102	82	102	208	212	15.523	high	UTAS
86	CH25_90	MLG29-Geo	-	-	1990	North Sea	57.43	1.22	B	II	II	II	282	282	132	138	102	134	102	102	204	204	-	-	Plymouth
87	NG26	MLG30-Geo	-	-	2011	Tsushima Strait, Japan	32.42	128.67	C	II	IIb	IIb	206	210	142	142	102	130	98	102	208	208	21.257	high	Roscoff
88	EHSO_65.06	MLG31-Geo	-	-	2007	Southern Ocean	-54.11	146	B/C	II (●)	IIb	IIb	202	206	132	142	102	102	102	102	208	208	4.177	low	UTAS
89	EHSO_8.15	MLG32-Geo	-	-	2007	Southern Ocean	-53.55	145.55	B/C	II (●)	IIb	IIb	206	206	132	142	102	102	102	102	204	208	4.960	low	UTAS
90	EHSO_8.15Q	MLG33-Geo	-	-	2007	Southern Ocean	-53.55	145.55	B/C	-	IIb	IIb	206	206	132	132	102	102	102	170	204	208	4.960	low	UTAS
91	EHSO_65.17	MLG34-Geo	-	-	2007	Southern Ocean	-54.11	146.16	B/C	II (○)	IIb	IIb	206	214	132	132	102	102	102	170	212	212	4.177	low	UTAS

Table 3. *Emiliania huxleyi* dual labeled probes for the CMM probe assay.

Multiplex	Probe	CMM	Sequence (5' -> 3')	Tm (°C)	Dye (5')	Quencher (3')	Channel	Excitation/Detection
	Probe I	I	CCTGACGGGTGGTGGGCGGCG		6-FAM	BHQ1	Green	470 nm/ 510 nm
1	Probe II	II	CGGCGATTTTTATGCGCCCACCA	68	ATTO680	BBQ650	Crimson	680 nm/ 712 nm
	Probe III	III	GATCGAGAGGCCTGACGGGTGG		CY5	BBQ650	Red	625 nm/ 660 nm
	Probe IIb	II	CGGCGATTTTATGCGCCCACCA		HEX	BHQ1	Yellow	530 nm/ 555 nm
2	Probe IV	IV	GGCGGCGATTTTTATGCCCGCCCCA	64	ATTO680	BBQ650	Crimson	680 nm/ 712 nm
	Probe IVb	IV	GGGGCGGCAATTTTATGCCCGCCCCA		6-FAM	BHQ1	Green	470 nm/ 510 nm

Table 4. Microsatellite stability over multiple generations.

Sample	Year	Generations	EHMS37		P01E05		P02F11		P02E09		P02B12		Source
Lohbeck*	2010	0	208	214	124	148	102	104	102	104	208	208	this study
	2012	1300	208	214	124	148	102	104	102	104	208	208	this study
CCMP1516	*2007*	-	*341*		*158*		*no hit*		*100*		*no hit*		Read et al (2013)^
	2010	-	339	339	ND		119	193	96	102	212	216	Mackinder et al. (2011a)
	2010"	-	339	339	ND		119	193	96	102	212	216	Mackinder et al. (2011a)
	2010	ND	338	340	137	153	120	192	100	106	212	216	this study
	2011	ND	340	340	137	153	120	192	100	106	212	216	this study
	2012"	ND	338	340	153	153	120	192	100	106	212	216	this study

*: Lohbeck et al. (2013)

": independent loss of coccolithsphere production

^: from the genome

ND: not determined

Table 5. ADONIS output with three different clustering variables: SST, Northern vs. Southern hemispheres (North vs. South) and Locality. Each model is fitted to all samples, CMM type I/I samples only, CMM type II/II samples only, CMM type I/IV samples only, and CMM type IV/IV samples only.

Clustering variables	**Samples (N)**	**R^2 ”**	**DF^**
SST	All (71)	12.9	2
	I-I (28)	19.8	1
	II-II (5*)	39.0	1
	I-IV (15)	9.0	1
	IV-IV (23)	15.8	1
North vs. South	All (71)	8.8	1
	I-I (28)	19.8	1
	II-II(5*)	39.0	1
	I-IV (15)	NA	NA
	IV-IV (23)	12.8	1
Locality	All (71)	31.1	9
	I-I (28)	25.6	2
	II-II(5*)	39.0	1
	I-IV (15)	33.6	3
	IV-IV (23)	51.3	7

N: sample size
*: small sample size
“: R^2 indicates the proportion (%) of variability accounted for by the clustering variable
^: DF is the number of free parameters in the model

FIGURELEGENDS

Figure 1. Earth observation 7-day composite data showing *Emiliania huxleyi* bloom development before, during and after cruise: (a) Enhanced ocean colour from Aqua-MODIS, showing coccoliths as bright patches and persistent cloud in black. (b) Chlorophyll-a concentration from Aqua-MODIS, with cloud in light grey. (c) Sea-surface temperature from AVHRR, where numbered circles indicate cruise stations listed in Table 2.

Figure 2 Alignment of CMM sequences produced in this study to reference CMMs (Schroeder et al. 2005). The CMM region is boxed. The dash line indicates the split between two subgroups of CMMs based on variation outside the CMM genotype. The bases shaded in grey show the positions of the probes (Table 3).

Figure 3. Scanning electron micrograph of a mixed *Emiliania huxleyi* culture prior to single cell isolation originating from D366 station 5 in the North Sea.

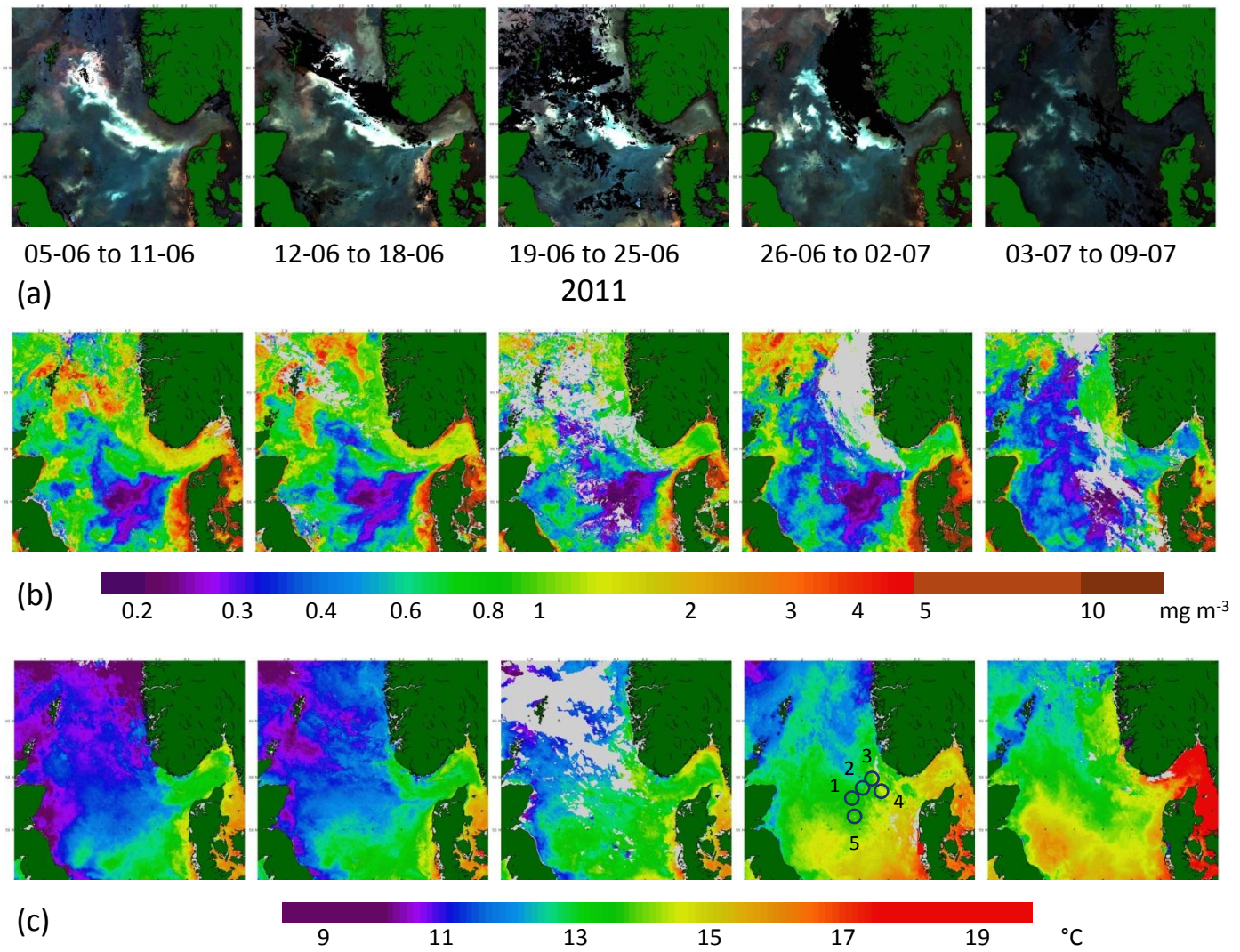
Figure 4. Frequency distribution histograms of all the measurements taken for distal shield length (a) and width (b): 95% t-confidence for mean is shown.

Figure 5. Average Sea Surface Temperature (SST) values for the sampling effort from January 2006 to December 2011 for the world's oceans. The four regions that include Europe, Japan, Chile and Australia that represent all our dataset are shown in greater detail. Key: temperature colour index from blue to red, 0ºC to 25ºC, respectively.

Figure 6. Multi-dimensional scaling plots constructed using Bruvo et al.'s (2004) genetic distance creating a 2-dimensional representation of the dissimilarity matrix used for the

permutational multivariate analysis of variance (ADONIS {VEGAN} community ecology package in R) for all the samples.

Figure 7 Multi-dimensional scaling plots constructed using Bruvo et al.'s (2004) genetic distance creating a 2-dimensional representation of the dissimilarity matrix used for the permutational multivariate analysis of variance (ADONIS {VEGAN} community ecology package in R) in the biogeographic group: a) CMM I, b) CMM IV & c) CMM II


05-06 to 11-06
12-06 to 18-06
19-06 to 25-06
26-06 to 02-07
03-07 to 09-07
(a)
2011
(b)
0.2
0.3
0.4
0.6
0.8
1
2
3
4
5
10
mg m-3
(c)
1
2
3
4
5
9
11
13
15
17
19
°C

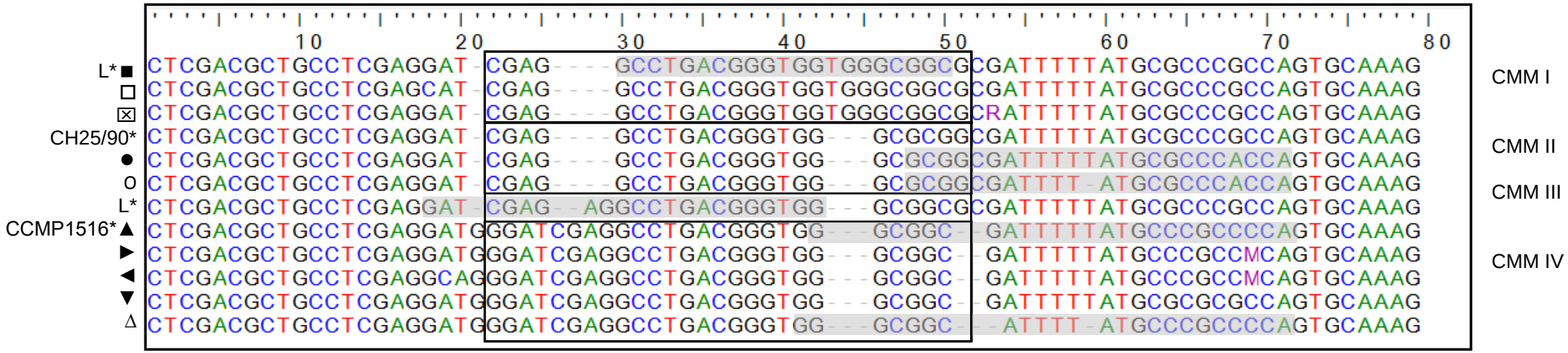

10 20 30 40 50 60 70 80
L* ■ CTCGACGCTGCCTCGAGGAT-CGAG----GCCTGACGGGTGGTGGGCGGCGCGATTTTTATGCGCCCGCCAGTGCAAAG
□ CTCGACGCTGCCTCGAGCAT-CGAG----GCCTGACGGGTGGTGGGCGGCGCGATTTTTATGCGCCCGCCAGTGCAAAG
⊠ CTCGACGCTGCCTCGAGGAT-CGAG----GCCTGACGGGTGGTGGGCGGCGCRATTTTTATGCGCCCGCCAGTGCAAAG
CH25/90* CTCGACGCTGCCTCGAGGAT-CGAG----GCCTGACGGGTGG---GCGCGGCGATTTTTATGCGCCCGCCAGTGCAAAG
● CTCGACGCTGCCTCGAGGAT-CGAG----GCCTGACGGGTGG---GCGCGGCGATTTTTATGCGCCCACCAGTGCAAAG
o CTCGACGCTGCCTCGAGGAT-CGAG----GCCTGACGGGTGG---GCGCGGCGATTTT-ATGCGCCCACCAGTGCAAAG
L* CTCGACGCTGCCTCGAGGAT-CGAG--AGGCCTGACGGGTGG---GCGGCGCGATTTTTATGCGCCCGCCAGTGCAAAG
CCMP1516* ▲ CTCGACGCTGCCTCGAGGATGGGATCGAGGCCTGACGGGTGG---GCGGC--GATTTTTATGCCCGCCCCAGTGCAAAG
▶ CTCGACGCTGCCTCGAGGATGGGATCGAGGCCTGACGGGTGG---GCGGC--GATTTTTATGCCCGCCMCAGTGCAAAG
◀ CTCGACGCTGCCTCGAGGCAGGGATCGAGGCCTGACGGGTGG---GCGGC--GATTTTTATGCCCGCCMCAGTGCAAAG
▼ CTCGACGCTGCCTCGAGGATGGGATCGAGGCCTGACGGGTGG---GCGGC--GATTTTTATGCGCGCGCCAGTGCAAAG
Δ CTCGACGCTGCCTCGAGGATGGGATCGAGGCCTGACGGGTGG---GCGGC--ATTTT-ATGCCCGCCCCAGTGCAAAG
CMM I
CMM II
CMM III
CMM IV

Figure 2

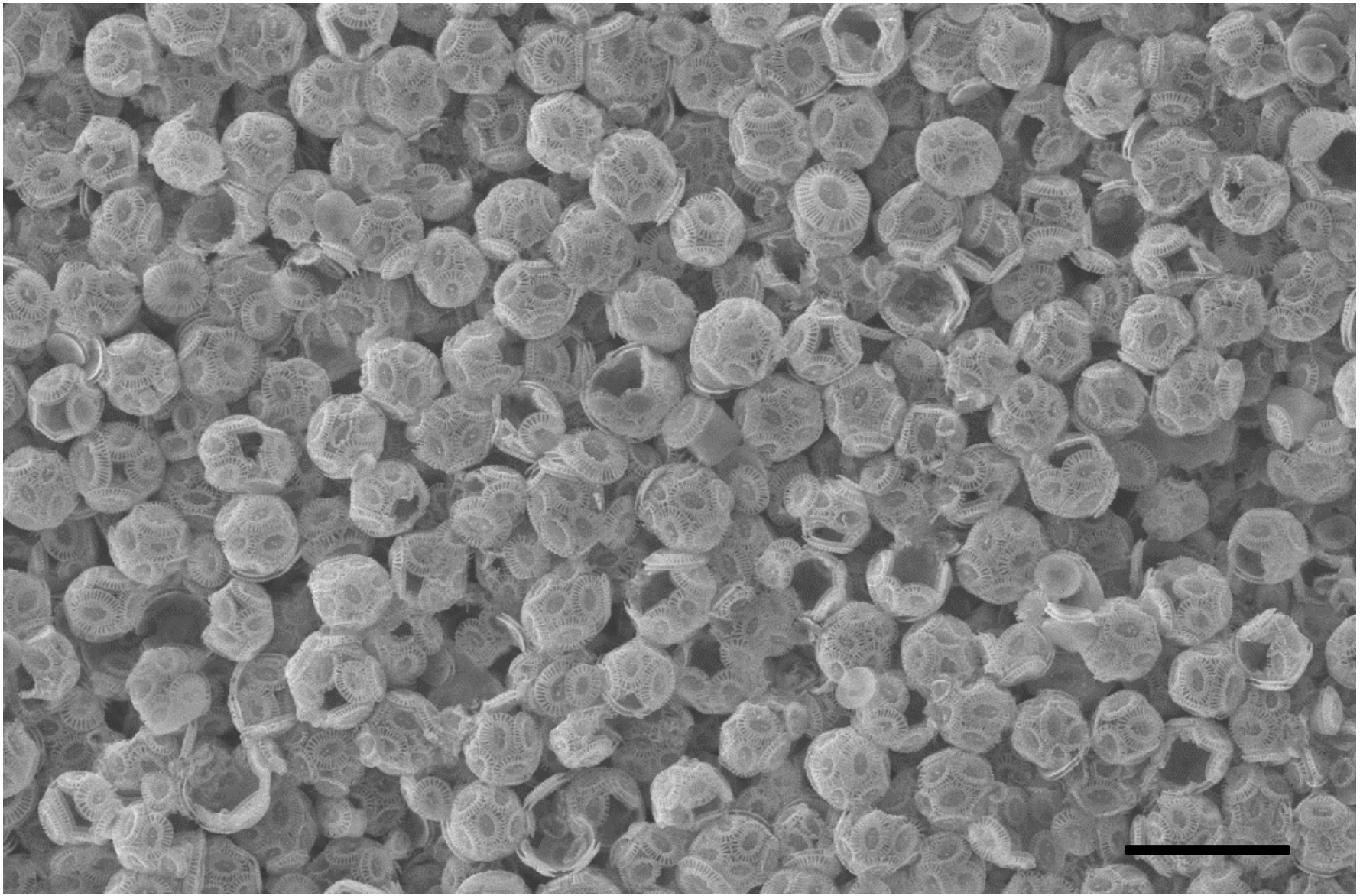
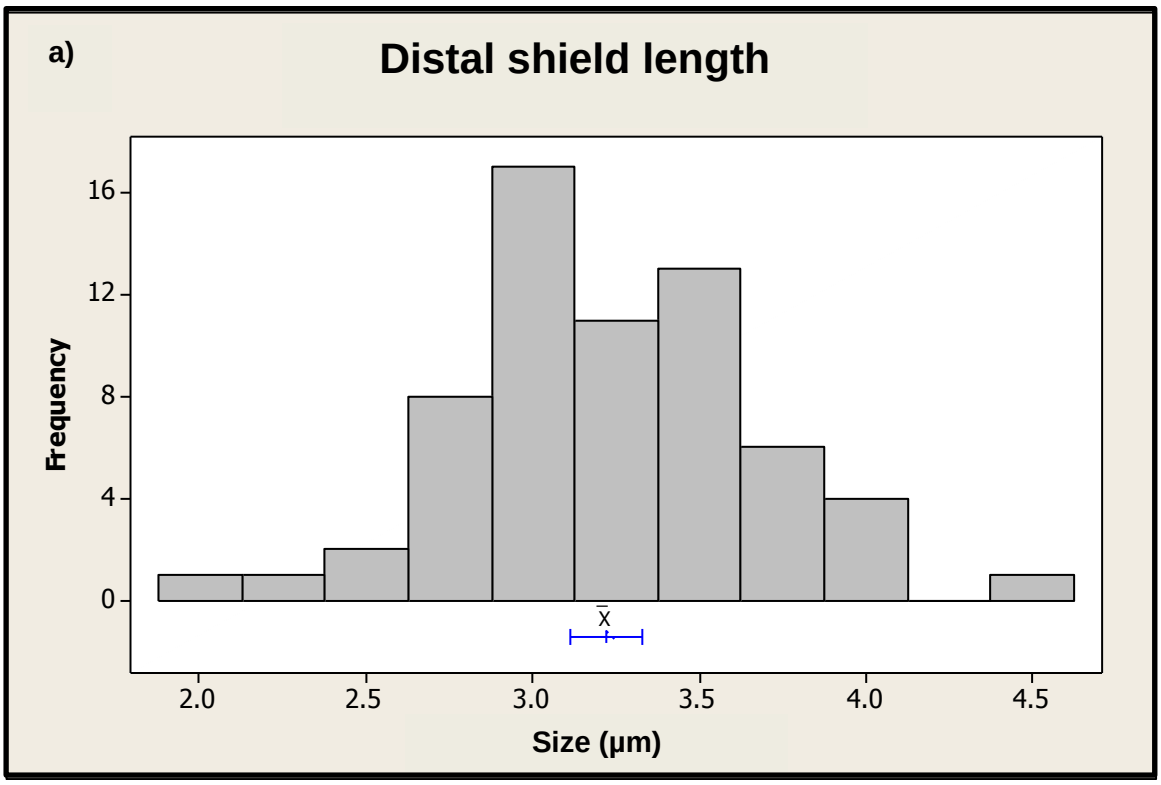


Figure 3


a)
Distal shield length
Frequency
16
12
8
4
0
x̄
2.0
2.5
3.0
3.5
4.0
4.5
Size (µm)

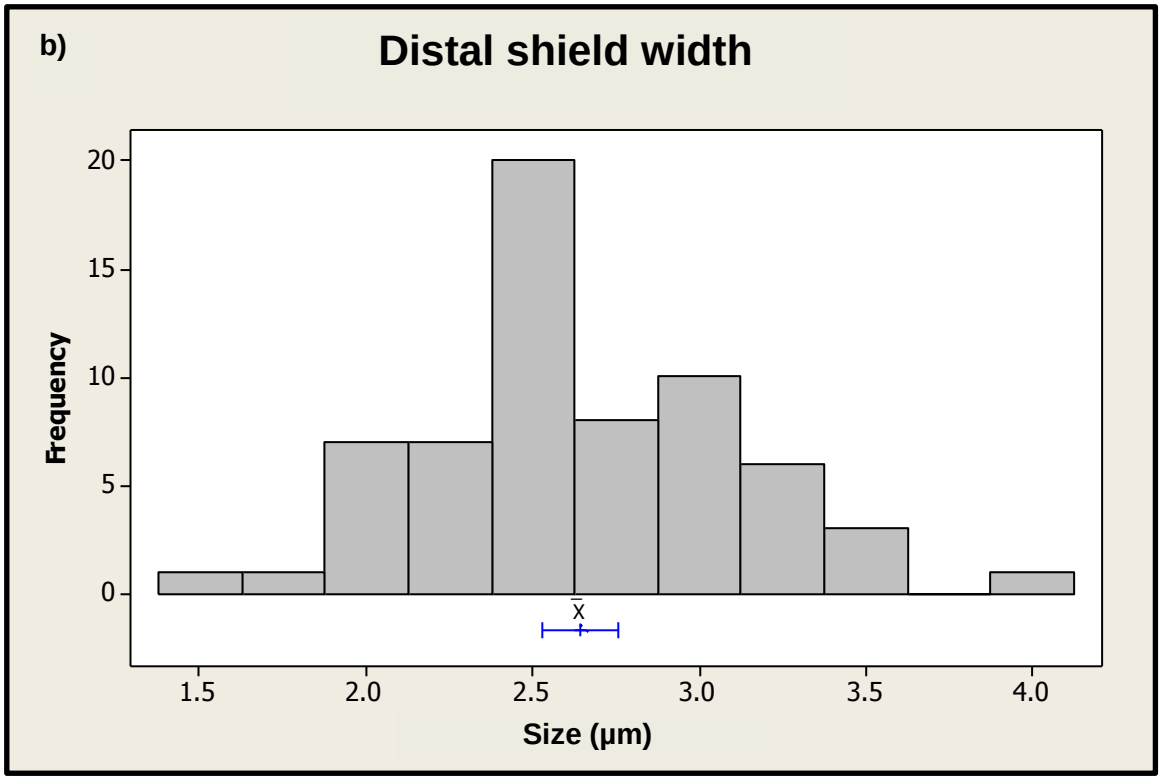

b)
Distal shield width
Frequency
20
15
10
5
0
x̄
1.5
2.0
2.5
3.0
3.5
4.0
Size (µm)

Figure 4

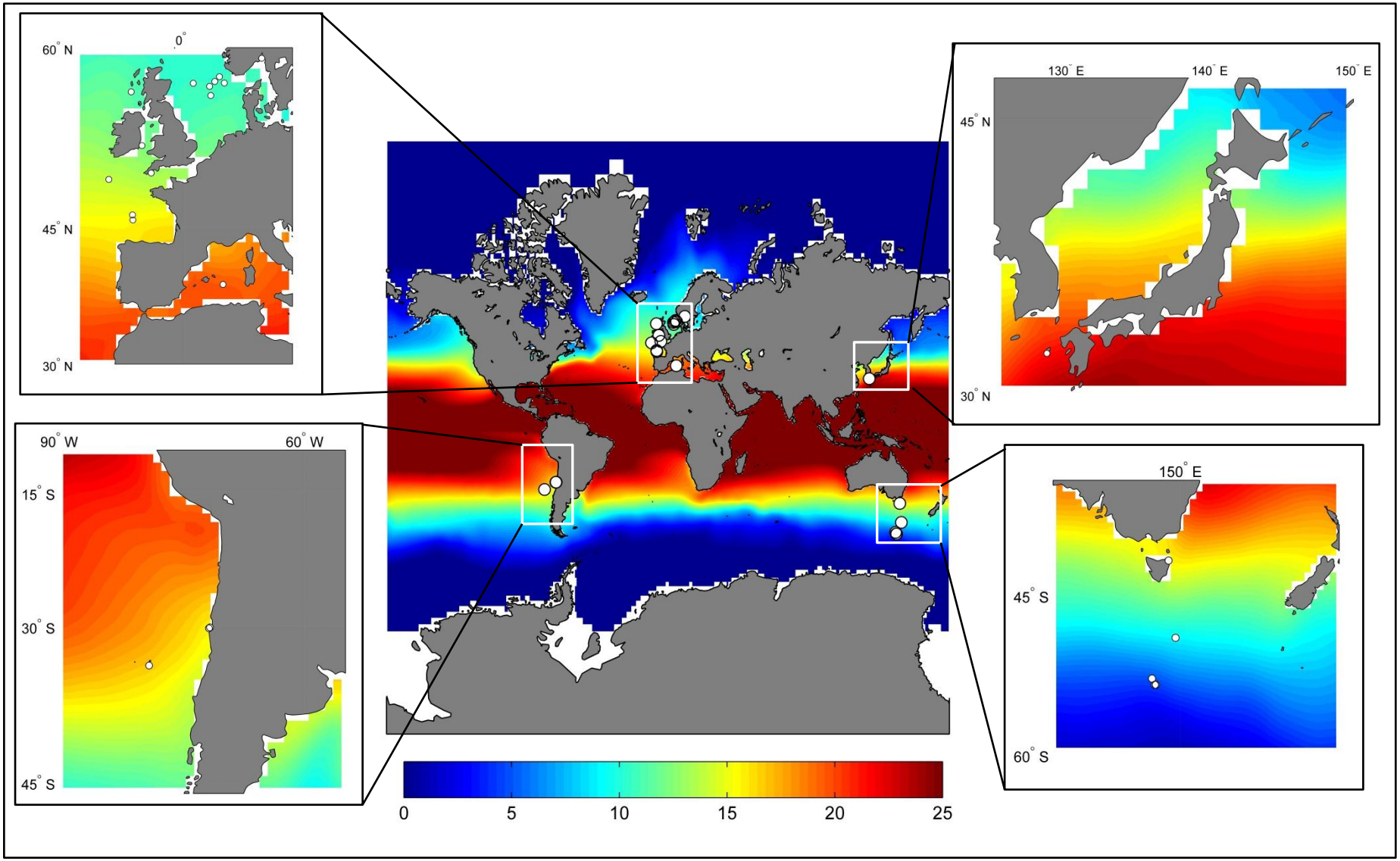

60° N
0°
45° N
30° N
130° E
140° E
150° E
45° N
30° N
90° W
60° W
15° S
30° S
45° S
150° E
45° S
60° S
0
5
10
15
20
25

Figure 5

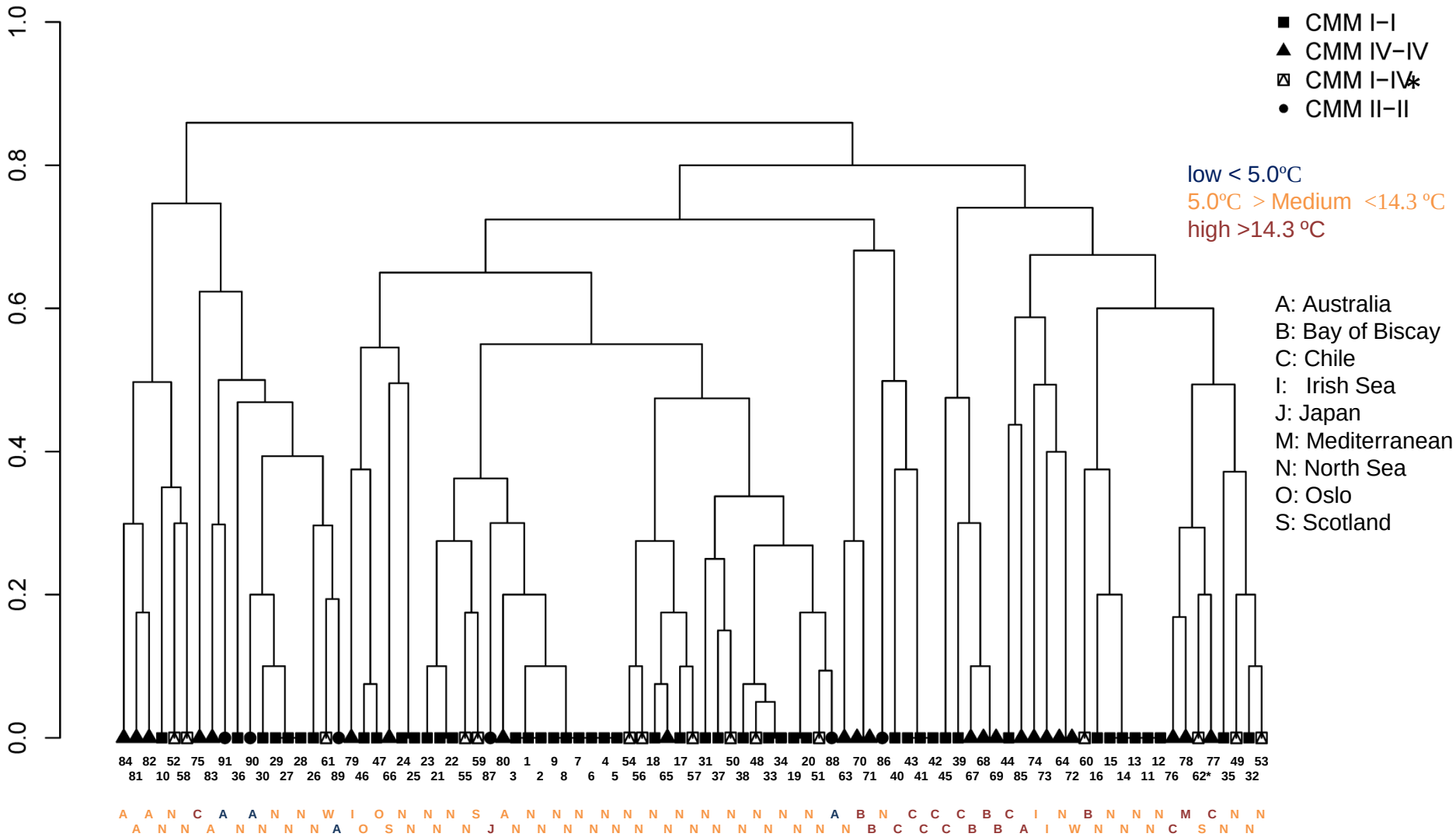

1.0
0.8
0.6
0.4
0.2
0.0
CMM I−I
CMM IV−IV
CMM I−IV*
CMM II−II
low < 5.0ºC
5.0ºC > Medium <14.3 ºC
high >14.3 ºC
A: Australia
B: Bay of Biscay
C: Chile
I: Irish Sea
J: Japan
M: Mediterranean
N: North Sea
O: Oslo
S: Scotland
84 82 52 75 91 90 29 28 61 79 47 24 23 22 59 80 1 9 7 4 54 18 17 31 50 48 34 20 88 70 86 43 42 39 68 44 74 64 60 15 13 12 78 77 49 53
81 10 58 83 36 30 27 26 89 46 66 25 21 55 87 3 2 8 6 5 56 65 57 37 38 33 19 51 63 71 40 41 45 67 69 85 73 72 16 14 11 76 62* 35 32
A A N C A A N N N W I O N N N S A N N N N N N N N N N N N A B N C C C B C I N B N N N M C N N
A N N A N N N N A O S N N N J N N N N N N N N N N N N N N N B C C C B B A I W N N N C S N N

Figure 6

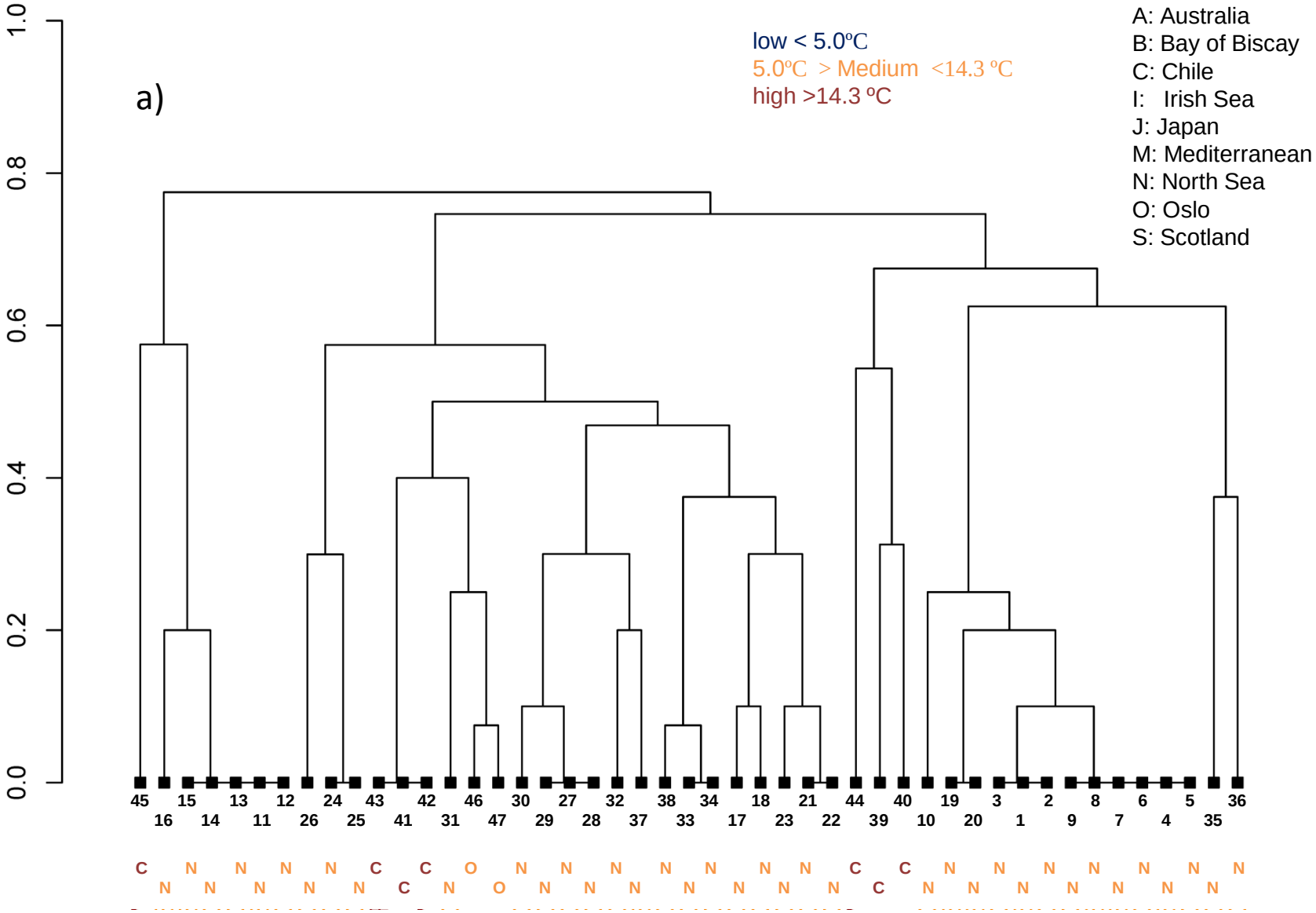

a)
low < 5.0ºC
5.0ºC > Medium <14.3 ºC
high >14.3 ºC
A: Australia
B: Bay of Biscay
C: Chile
I: Irish Sea
J: Japan
M: Mediterranean
N: North Sea
O: Oslo
S: Scotland
1.0
0.8
0.6
0.4
0.2
0.0
45 16 15 14 13 11 12 26 24 25 43 41 42 31 46 47 30 29 27 28 32 37 38 33 34 17 18 23 21 22 44 39 40 10 19 20 3 1 2 9 8 7 6 4 5 35 36
C N N N N C C O N N N N N N N N C C N N N N N N N
N N N N N C N O N N N N N N N C N N N N N N

Figure 7

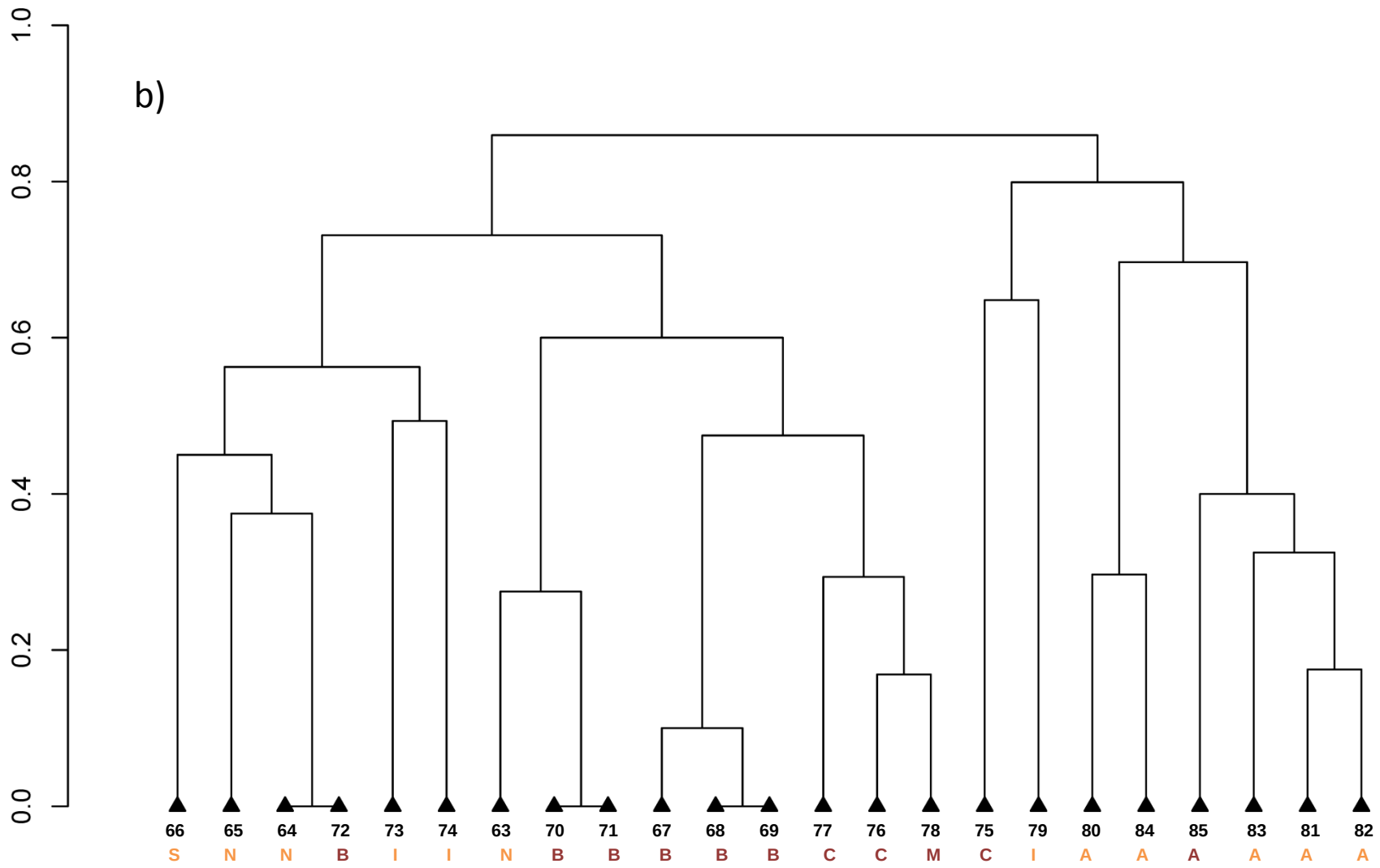

b)
1.0
0.8
0.6
0.4
0.2
0.0
66 65 64 72 73 74 63 70 71 67 68 69 77 76 78 75 79 80 84 85 83 81 82
S N N B I I N B B B B B C C M C I A A A A A A

Figure 7

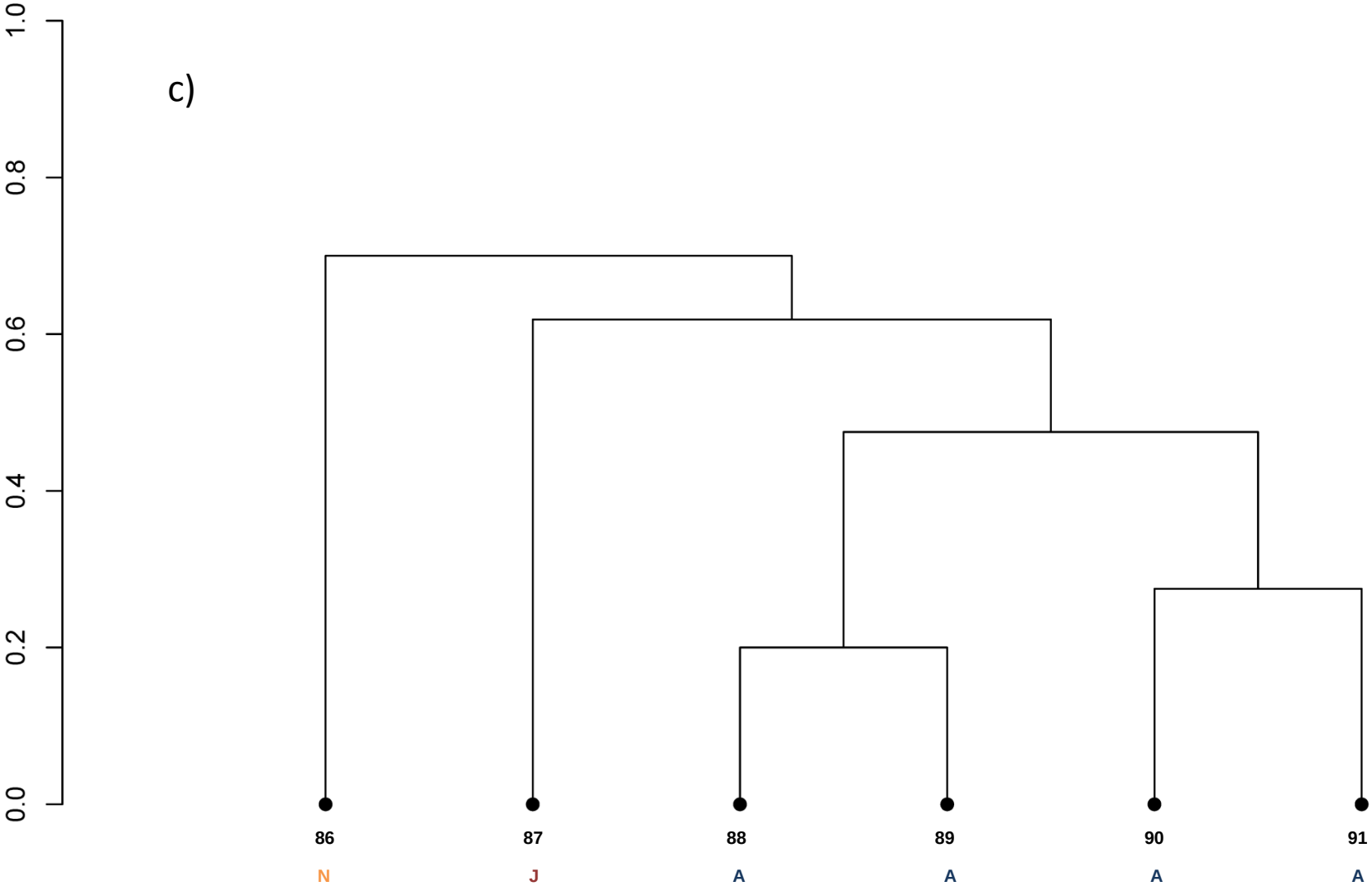


Figure 7