An investigation into the potential impacts of ocean acidification and ocean fertilisation on the genetic diversity of marine bacterial assemblages

by

John Woolven-Allen

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Biology

University of Warwick, Department of Biological Sciences

September 2008

Contents

List of Figures

List of Tables

Acknowledgements

I thank Dr Martin Mühling of Plymouth Marine Laboratory (PML), my principle supervisor, for four years of support and guidance. Martin was especially thorough in reading and constructively criticising written work and presentations I prepared. This was most appreciated for the several draft copies of this thesis that were produced. He also collaborated on a journal paper based on some of the results presented here, and provided initial training in the laboratory techniques I used throughout the study.

Dr Ian Joint (PML) and Prof J. Colin Murrell (Warwick) also provided academic guidance, and practical advice throughout the study period. I also thank Colin for checking the thesis script and making valuable comments.

Fellow students and staff at PML, where I undertook this work, were supportive and often an important source of academic, practical and administrative knowledge. Some of those particularly helpful in shedding light on matters to do with molecular biology, microbiology and bioinformatics include Mr Paul Pickerill, Dr Mike Allen, Mrs Karen Weinberg, Dr Victoria Goddard, Dr Ronnie Henderson, Dr Puyasloke Bhradury, Dr Paul Somerfield and Dr Susan Kimmance.

I made great use of the facilities at Warwick University Library, and also those at the National Marine Biological Library, where Linda Noble was very helpful.

Finally I am indebted to my family. My mother and father provided continual support which must not be taken for granted. Most important and fundamental was the support of my wife, Sarah, who was unwavering with support in all senses of the word, and provided a light for me at those few times during the project when things got dark. Also I must mention my son, Toby, who gave me phenomenal uplift and motivation in the last few months.

I acknowledge with thanks financial support from the Natural Environment Research Council (UK).

Declaration

I hereby declare that the work presented in this thesis was conducted by me under the direct supervision of Dr Martin Mühling, with the exception of those instances where the contribution of others has specifically been acknowledged.

None of the work presented herein has been previously submitted for any other degree.

Some of the thesis data, presented in Chapters 1, 2, 3 and 4 have been published:

Mühling, M., Woolven-Allen, J., Murrell, J.C., and Joint, I. (2007) Improved group-specific PCR primers for denaturing gradient gel electrophoresis analysis of the genetic diversity of complex microbial communities. *The ISME Journal*. **2**: 379-392.

John Woolven-Allen

24th September 2008

Abstract

Based on the increase of 16S rRNA gene sequences in databases it is possible to design improved oligonucleotide primers for this gene. Primers were designed *in silico* to specifically amplify fragments of the gene from the *Alpha*, *Beta* and *Gamma* subgroups of the *Proteobacteria*, as well as from *Bacteroidetes*, *Firmicutes*, *Cyanobacteria* and *Planctomycetes* and tested *in silico* and *in vitro*.

The aim was to investigate bacterioplankton diversity and reveal greater fingerprint diversity within these groups than is possible using primers specific for the entire domain *Bacteria*, and also to reduce clone library redundancy. It was then aimed to investigate the potential impacts of increased pCO_2 and ocean fertilisation with iron (Fe) and phosphorus (P), on bacterioplankton diversity. Group-specific clone libraries representing contrasting marine regions were analysed, and the usefulness and specificity of the primers validated. The clone libraries showed members of the oligotrophic marine group (OMG) to be present in an *in situ* coastal mesocosm supplemented with nutrients.

The newly-developed group-specific primers were used in combination with an improved method of denaturing gradient gel electrophoresis (DGGE) to profile in detail bacterial communities in mesocosms, which were maintained at 750 ppm of pCO_2, the level projected for the global surface ocean in the year 3000, and 380 ppm of CO_2, the present level. Increased pCO_2 correlated with a decrease in abundance of some members of the *Gammaproteobacteria.* Otherwise there was little impact on diversity due to raised pCO_2.

The same DGGE protocol was applied to samples from an ocean Fe and P fertilisation experiment. Diversity change due to Fe was not evident. However in seawater amended with P there was an explosive growth of some cells with 16S rRNA genes similar to those of the SAR86 clade, and others with similarity to *Gammaproteobacteria* with large genomes such as *Oceanospirillum* sp. and *Psychromonas* sp.

Abbreviations

16S rDNA	gene encoding prokaryotic small subunit rRNA
16S rRNA	prokaryotic small subunit rRNA
18S rRNA	eukaryotic small subunit rRNA
23S rRNA	prokaryotic large subunit rRNA
AMT	Atlantic meridional transect
APS	ammonium persulphate
ARISA	automated ribosomal intergenic spacer analysis
BLAST	basic local alignment search tool
BLASTn	BLAST nucleotide sequences
BSA	bovine serum albumin
C	carbon
CFB	*Cytophaga-Flexibacter-Bacteroides*
Chl *a*	chlorophyll *a*
CO_2	carbon dioxide
CO_3^{2-}	carbonate
crDNA	copy ribosomal DNA
CTD	conductivity, temperature, density
ddNTP	dideoxynucleotide triphosphate
DGGE	denaturing gradient gel electrophoresis
DIC	dissolved inorganic carbon
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DMS	dimethyl sulphide
DOM	dissolved organic matter
dsDNA	double stranded DNA
EDTA	ethylenediamine tetra-acetic acid
EisenEx	*in situ* iron addition experiment
EMBL	European Molecular Biology Laboratory
ENSO	el niño-southern oscillation
Fe	iron
FeeP	*in situ* iron and phosphate addition experiment
$FeOH_3$	iron hydroxide

FISH	fluorescence *in situ* hybridisation
GOM	gelatinous organic matter
H	hydrogen
H_2O^{MQ}	Milli-Q purified water
HCO_3^-	bicarbonate
HNLC	high-nutrient low-chlorophyll
IMS	industrial methylated spirit
IronEX I	*in situ* iron addition experiment I
IronEX II	*in situ* iron addition experiment II
KCl	potassium chloride
KEOPS	Kerguelen: ocean and plateau surface study
LB	Luria-Bertani
LH-PCR	length heterogeneity PCR
MDS	multidimensional scaling
$MgCl_2$	magnesium chloride
$MgSO_4$	magnesium sulphate
MV	Motor Vessel
N	nitrogen
NaCl	sodium chloride
NaOH	sodium hydroxide
NCBI	National Center [US] for Biotechnology Information
O	oxygen
OIF	ocean iron fertilisation
OMG	oligotrophic marine group
OMZ	oxygen minimum zone
OPF	ocean phosphorus fertilisation
OTU	operational taxonomic unit
P	phosphorus
pCO_2	partial pressure of CO_2
PCR	polymerase chain reaction
PeECE	pelagic ecosystem CO_2-enrichment study
PETM	Palaeocene-Eocene thermal maximum
pH	power of hydrogen
POM	particulate organic matter

psi	pounds per square inch
QC-PCR	quantitative competitive PCR
rDNA	ribosomal DNA
RDP	Ribosomal Database Project
RFLP	restriction fragment length polymorphism
RISA	ribosomal intergenic spacer analysis
RNA	ribonucleic acid
rRNA	ribosomal RNA
RRS	Royal Research Ship
S	sulphur
SARST	serial analysis of ribosomal sequence tags
SAZ	sub-Antarctic zone
SEEDS	sub-Arctic iron experiment for ecosystem dynamics study
SOC	super optimal broth with catabolite repression
SOFeX	southern ocean iron experiment
SOIREE	southern ocean iron release experiment
SSCP	single strand conformation polymorphism
TAE	tris HCl-acetic acid-EDTA
TEMED	N, N, N', N'-tetramethylethylenediamine
TRFLP	terminal fragment RFLP
Tris-HCl	tris (hydroxymethyl) aminomethane-hydrogen chloride
UV	ultraviolet
V1, V2, etc.	variable regions within a 16S rRNA gene sequence

Chapter 1

Introduction

1.1 Marine environment

The seawater of the world's oceans is a fluid environment which exhibits variable properties. Excluding extreme conditions, such as those produced by volcanic eruptions, temporal change occurs on daily, seasonal, annual and cosmic cycles, whilst spatial change occurs over micrometre to kilometre scales, and with increasing latitude and depth.

Changing physical parameters include temperature (-2 to 28 ^{o}C), and pressure (1 to >1100 atm) (Proujan, 1979). The intensity of visible light incident at the surface depends on season and atmospheric conditions. Depending on the trophic state of the water, up to 10 % incident light remains below 50 m, and this is largely restricted to blue and green wavelengths (Thurman, 1990). Salinity in open oceans ranges from 25 to 40 ‰ (Thurman, 1990). Currents are responsible for large scale horizontal movement of water bodies with boundary mixing of their planktonic and abiotic components. Upwelling and waves generate turbulence which mixes water in the vertical plane. As well as salts, many other dissolved and particulate organic and inorganic compounds occur in seawater. Living phytoplankton and particulate organic matter (POM) in seawater provide a habitat for particular communities of bacteria associated with them, and which in turn influence the growth characteristics of the phytoplankton and the related production of POM and dissolved organic matter (DOM) (Grossart and Simon, 2007). Sub-micrometre particles, largely resulting from protistan grazing of bacteria, occupy the boundary between particulate and dissolved matter, creating a size continuum of organic material in the ocean (Isao *et al.*, 1990). Dissolved inorganic carbon (DIC) content controls the acidity of seawater.

The majority of free-living, non-filamentous marine bacteria range in size from the ubiquitous SAR11 candidatus *Pelagibacter ubique* (0.01 μm^3) (Giovannoni *et al.*, 2005) and the cyanobacterium *Prochlorococcus* sp. (0.1 μm^3) to heterotrophic *Proteobacteria* such as *Vibrio* sp. (2.0 μm^3). The microlitre environment around the cell can be considered the bacrerium's micro-environment. The properties of this small volume of water are to what it has to adapt. Biopolymeric substances, derived from DOM, form networked gels on a scale of a few to several hundreds of micrometres in the water making the bacterium's environment more structured and less fluid than that perceived at a larger scale (Verdugo *et al.*, 2004). This gel organic matter (GOM) amounts to 10 % of organic carbon in DOM, and exceeds global marine biomass by a factor of 50. GOM links patches of DOM and POM in a micro-

structural continuum, the density of which varies with the turbidity state of the sea (Verdugo *et al.*, 2004). Free-living bacteria must cope with a changing external environment with both temporal cycles, and transport of their microenvironment both relative to latitude and depth. Some bacteria, such as certain *Vibrio* sp., are capable of limited motion by flagella (Perry *et al.*, 2002). This may aid association with, or evasion of, other biotic components in the microenvironment, including viruses, protists, algae and other bacteria. Beyond this microenvironment bacteria may be considered immobile relative to the fluid medium in which they live.

1.2 Marine bacterial communities

1.2.1 The origin of marine bacteriology

Antony van Leeuwenhoek first reported observing bacteria in 1673. It was not until 1838 that Otto Müller listed bacteria and classified them into *Monera* and *Vibrio*, and divided the *Vibrio* into four genera based on cell morphology. By 1853 Ferdinand Cohn had advanced the morphological division of bacteria into groups and genera, although he placed them within the kingdom *Plantae*. Cohn was confident that a robust classification of bacteria was possible, "I have become convinced that the bacteria can be separated into just as good and distinct species as other lower plants and animals, and it is only their extraordinary smallness and the variability of the species which makes it impossible for us with our present day methods to differentiate the various species which are living together in mixed array." But forseeing problems that have persisted to this day he warned, "In general it will be difficult to determine the species of individual bacteria with certainty" (Cohn, 1875).

Although marine bacteria were isolated in the early twentieth century, their role was considered confined to decomposition of organic matter. They were not seen to play a significant part in global nutrient cycling, some part of which had been attributed to terrestrial bacteria. As late as 1956 Sir Alister Hardy omitted them from his synthesis of global plankton because "very little is known about their occurrence in the plankton" (Hardy, 1970).

The abundance, diversity and importance of marine bacteria began to be realised in the 1970s with a general upswing in understanding of, and scientific interest in, the oceans. Lawrence Pomeroy expedited the changing views of marine microbial ecologists with his theory of the microbial loop (Pomeroy, 1974). This stipulated bacteria-mediated recycling of DOM at many levels, back into the food web, thus enhancing overall productivity (Pomeroy, 1974). Sieburth (1976) argued that dissolved organic carbon (DOC), estimated at up to 2 mg L^{-1} of carbon (C), must sustain a far greater heterotrophic bacterial population. By the 1990s that population was put at 5 x 10^5 cells mL^{-1} (global total 360 x 10^{26} cells) in the < 200 m photic zone, and 0.5 x 10^5 cells ml L^{-1} (global total 650 x 10^{26} cells) in the water column below 200 m (Whitman *et al.*, 1998). Photic zone bacteria have an average turnover of 6.25 days, while those inhabiting the sub-photic zone have an average turnover of 290 days. These pelagic bacteria account for ~ 2.4 % of the total bacteria in the biosphere, most of which are subterranean (Whitman *et al.*, 1998). In the 1970s however,

bacterial function beyond phototrophic primary production and heterotrophic consumption (secondary production) remained obscure. Little diversity was described beyond broad physiological, and microscopic morphological, groupings.

1.2.2 Use and limitations of cultured isolates

Isolated strains of bacteria were first cultured in 1876 by Robert Koch, with the growth and characterisation of *Bacillus anthracis* in a serum medium (Koch, 1876). Culture collections have since provided a wealth of information on their representative bacterial species. This information has mainly been metabolic, morphological, physiological and biochemical in nature, for example Gram staining characteristics, preferred carbon source, and cell shape, and has informed artificial nomenclature, classification and phylogeny. However, it has long been appreciated that there is a large discrepancy between the observed bacterial cell number in a sample, and the number of colonies culturable from the same sample. This "great plate count anomaly" (Staley and Konopka, 1985), is strongest in the oceans where nutritional needs are often unknown, and especially in oligotrophic waters where bacteria have evolved in an environment of very dilute nutrients. The portion of microorganisms in marine environments capable of forming colonies on solidified agar media ranges from 0.001 % to 0.1 % (Amann *et al.*, 1995).

However, technical advances in the last two decades have seen the cultivation of marine representatives previously labelled "unculturable" (Jannasch and Jones, 1959). Dilution culturing of mixed natural communities enabled isolates to be characterised from cultures of a few mixed types (Li and Dickie, 1985; Schut *et al.*, 1993). Button and colleagues (1993) extended this method by diluting samples in natural seawater as low as 1-10 cells per tube, and applying these samples to flow cytometry. Extinction culturing, as it was called, led to the description of the oligotrophic alphaproteobacterium *Sphingomonas alaskensis* (Button *et al.*, 1998; Vancanneyt *et al.*, 2001). Increased sample throughput with samples as small as 200 μL, media diluted to *in situ* nutrient levels, detection of cultures as dilute as 10^3 cells mL^{-1}, and cell encapsulation methodology led to further progress. Oligotrophic *Alphaproteobacteria* SAR11 members, *Betaproteobacteria* OM43 members, and *Gammaproteobacteria* SAR92, OM60 and OM241 members, all abundant in the ocean, were isolated and characterised (Connon and Giovannoni, 2002).

Kaeberlein *et al.* (2002) suggested microorganisms require "signals originating from their neighbours indicating the presence of a familiar environment." The absence of such signals would explain the inability to culture many bacterial species. A growth chamber, housing a simulated natural environment, and separated from an untreated marine sediment by a 0.03 μm pore-size membrane to allow diffusion of pheromones, permitted the growth of colonies of several candidate *Bacteroidetes* species. This method, developed using sediment microorganisms, has the potential to be adapted for their pelagic counterparts. Simu and Hagström (2004) reported that several alpha- and betaproteobacterial isolates did not form colonies due to inherent growth behaviour involving immediate dispersal of progeny cells away from parental cells. This strategy, while inhibiting colony formation on solid media *in vitro*, may maximise the interception of new substrate *in vivo* in the pelagic environment. Biochemical, metabolic and physiological analyses of cultured isolates remains the fundamental source of knowledge of the functional properties of bacteria.

The production of single strain axenic cultures provides important genetic/genomic data, informing taxonomy and phylogeny, as well as providing metabolic and physiological data. However bacterial types *in situ* live in complex and interacting and interdependent communities. These dependencies, together with dispersal behaviour, may be what ultimately prevent the majority of them being grown and studied in isolation.

1.2.3 Culture-independent approaches

The ability to identify the majority of bacterial types in the sea, and to form these into a meaningful phylogeny, cannot be achieved through culture technology. The veil over the identity of the unknown majority was lifted with the advent of direct identification of community members from the environment by sequencing the semantic macromolecules deoxyribonucleic acid (DNA), ribonucleic acid (RNA) and proteins. Zuckerkandl and Pauling (1965) envisioned a molecular phylogeny, ideally based on genomic DNA. Such a phylogeny for prokaryotes was realised two decades later based, by consensus, on genes encoding ribosomal ribonucleic acid (rRNA) (Woese, 1987; Pace, 1997). Twelve bacterial phyla were represented, all based on DNA sequences from cultured bacteria. By 2003 the collection of DNA sequences was divided into 53 bacterial phyla, the majority having marine representatives, 26 of

which had no cultured representatives (Giovannoni *et al.*, 1990; Rappé and Giovannoni, 2003).

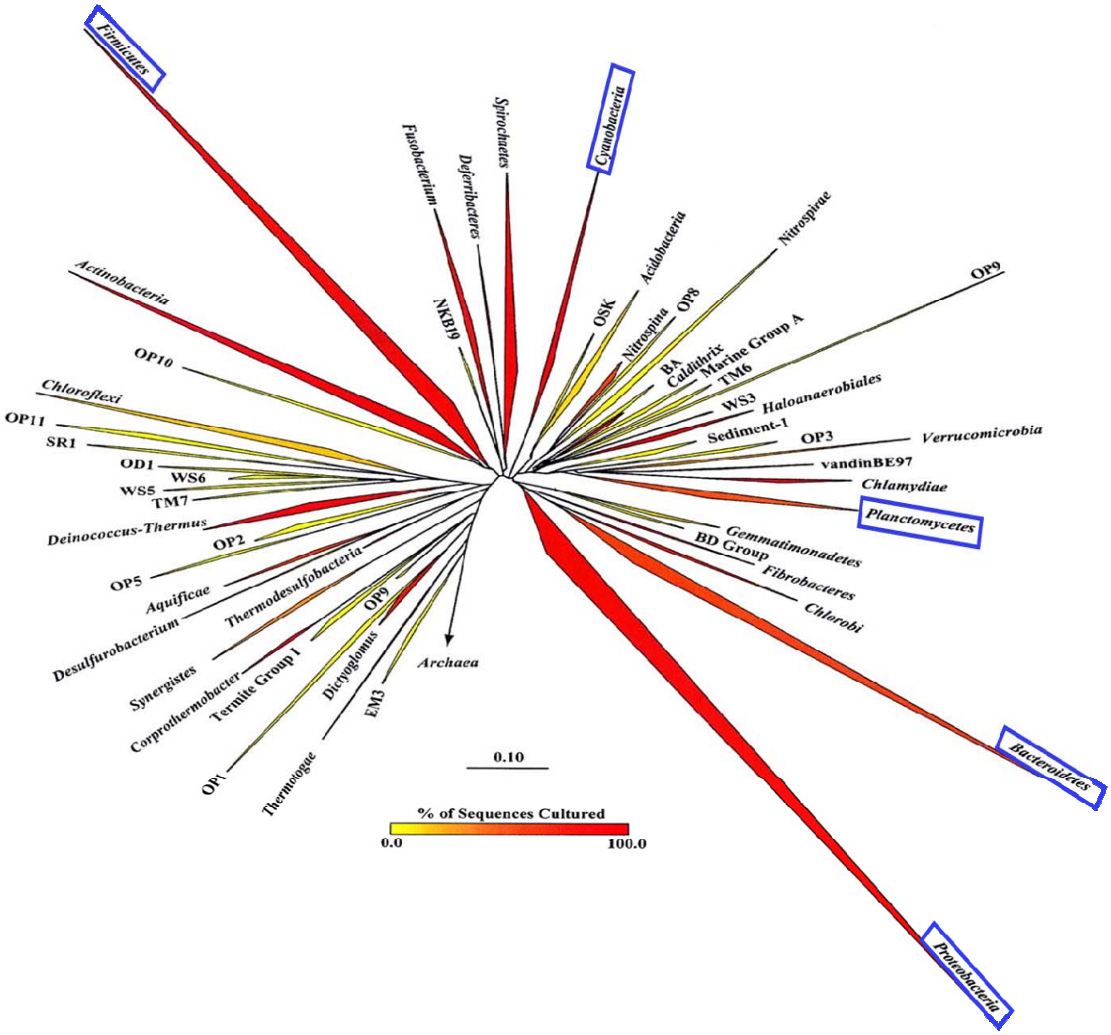


Figure 1.1. Phylogenetic tree showing the domain *Bacteria* divided into 53 phyla. Established phyla with cultured representatives have italicised names. Non-italicised names are of candidate phyla with no cultured representatives. The tree was constructed using ARB (Ludwig *et al.*, 2004) with 16,964 >1000bp 16S rRNA gene sequences. The distal angle of each wedge is proportional to the abundance of sequences from each phyla. The length of each wedge indicates the extent of evolutionary branching in the phylum. Degree of redness correlates with the percentage of sequences from each phylum, used to construct the tree, which originated from cultured isolates. Phyla studied here using specific PCR primers for DGGE analysis of marine pelagic communities are enclosed in blue boxes. Adapted from Handelsman (2004).

Not only did the new molecular methods reveal a greater variety of bacteria at higher taxonomic levels, as shown in the phylogenetic tree in Figure 1.1, than that

elucidated by a culture-dependent approach, but the proportions of groups common to both approaches differed as shown in Figure 1.2, (Suzuki *et al.*, 1997; Hagström *et al.*, 2002).

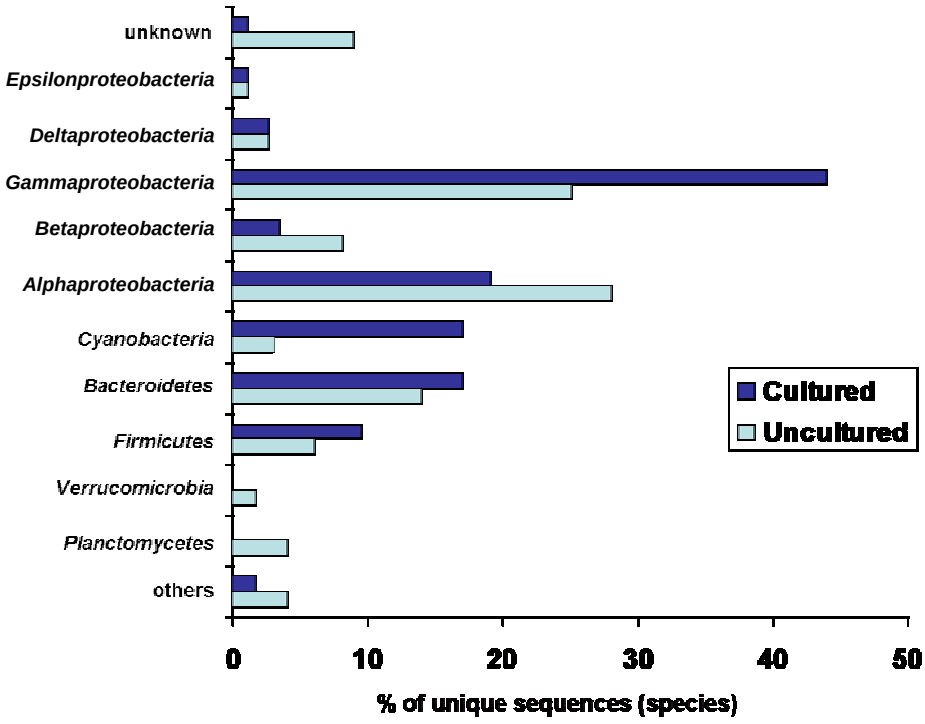


Figure 1.2. Taxonomic distribution of catalogued 16S rRNA bacterial sequences obtained from cultured (n=508) and uncultured (n=609) organisms. Numbers of sequences have grown for all groups , but the proportions of the total sequences, uncultured or cultured in origin, ascribed to each bacterial group remain similar to those shown here. Ref: Munn, 2004, from Hagström *et al*., 2002.

Molecular methods for identification and classification of bacteria are applied to DNA from cultured isolates, as well as that from environmental samples. This is important as these fully identified and characterised representatives provide phylogenetic anchors, and establish linkage between phylogeny and function. DNA sequences from cultured isolates allow also classification of sequences to the genus or species level. Thus they provide reliable reference points in growing phylogenetic trees, and this is one reason continued efforts to culture novel bacteria are of fundamental importance (Hagström *et al.*, 2000).

There are several molecular techniques for measuring location, abundance and diversity of bacteria in the marine environment based on comparison of homologous genetic sequences. These techniques are made possible by, and incorporate, directed amplification, by polymerase chain reaction (PCR), of specific sections of genomic

DNA (Mullis and Faloona, 1987). PCR-generated sequences can be separately ligated into plasmid vectors, which are in turn amplified in transformed host cells, commonly *E. coli*. These clone libraries of individual sequences give a proportional (partiality of sampling and molecular analysis notwithstanding) representation of the diversity of such sequences present in the environment. If the amplified sequence is part of a universal evolutionary marker such as the 16S (small subunit) rRNA gene, then the library should be representative of the universal (prokaryotic) diversity in the sampled environment. The method becomes finance- and labour-intensive if sufficient clones are to be generated and sequenced to reveal the rarer members of the community. These rarer members of the community may present a flat rank abundance when sampled. Their minimum diversity can still be constrained. The diversity of marine bacteria rare in clone libraries, and masked by the abundant distribution of common species, has been estimated to be very large (Lunn *et al*. 2004). Clone libraries however, have been instrumental in pioneering studies of marine bacterial communities including those of the epipelagic Mediterranean (Moeseneder *et al.*, 2005; Zaballos *et al.*, 2006), nutrient-rich Beagle Channel (Prabagaran *et al.*, 2007), oligotrophic Sargasso Sea (Giovannoni *et al.*, 1990; Fuhrman *et al.*, 1993; Field *et al.*, 1997), subtropical north east Pacific open ocean (Fuhrman *et al.*, 1993), subtropical north east Pacific coastal ocean (Field *et al.*, 1997), a detailed phylogenetic study of the temperate coastal north west Atlantic Ocean (Acinas *et al.*, 2004), and deep-ocean communities, such as that associated with a volcanic seamount in the north east Pacific (Huber *et al.*, 2007). Cloned 16S rRNA gene sequences were used to resolve the detailed phylogeny of specific communities such as the oligotrophic *Cyanobacteria* (Urbach *et al.*, 1998) and *Vibrionaceae* (Kita-Tsukamoto *et al.*, 1993). Such nucleotide sequences from PCR amplified 16S rRNA genes, whether whole or partial, environmental or cultured are only representative of their host organisms, and there are limits to the taxonomic and phylogenetic information that can be inferred (Cohan, 2002). These sequences are often called operational taxonomic units (OTUs), “phylotypes” or “ribotypes”. They are treated as unique genetic entities, irrespective of multiple occurrence of different types within a single genome (organism), or common occurrence of the same type in different genomes (organisms). Ribotypes, meaning partial but specific 16S rRNA gene amplicons, rather than the more general total RNA complements of cells, is the term that will be used in this study.

There have been many advances in molecular techniques aimed at revealing the structure of *in situ* marine microbial communities. Denaturing gradient gel electrophoresis (DGGE), described in detail in Section 1.2.4, is a fingerprinting technique that can quickly and repeatably generate community profiles directly from environmental DNA without the need for cloning (Muyzer *et al.*, 1993; Schäfer and Muyzer, 2001). DGGE can also be used to screen large clone libraries and remove the otherwise high level of sequence redundancy (Gonzalez *et al.*, 2003). A related profiling method is single strand conformation polymorphism (SSCP) in which single stranded PCR-amplified environmental DNA is resolved on account of each molecule assuming a secondary structure according to its sequence, thus uniquely influencing its electrophoretic mobility (Hayashi, 1991). SSCP is yet to be used to directly assess marine bacterial communities, but has been used to describe bacterial diversity in soil (Stach *et al.*, 2001).

Another clone-free method, fluorescence *in situ* hybridisation (FISH), uses oligonucleotide probes to reveal *in situ* complementary genomic DNA binding sites within cells. The method can measure organism abundance as well as spatial relationships, but, like the use of oligonucleotide primers in a PCR, it is restricted to the investigation of microbial taxa possessing target DNA of known sequence (Amann and Ludwig, 2000; Pernthaler *et al.*, 2001). FISH has been used to determine the spatial distribution and proportional abundances of major community members of free-living picoplankton using enrichment cultures (Uphoff *et al.*, 2001), of members of a phytoplankton-associated community using confocal microscopy (Biegala *et al.*, 2002), and of the *Planctomycetes* phylum from several contrasting environments (Neef *et al.*, 1998). Cotterell and Kirchman (2000) contrasted analysis of bacterioplankton using FISH, which involves considerable sample manipulation, with analysis using clone libraries, which are subject to the bias of the PCR, in which environmental sequence amplification is directed by primers designed *in silico* using sequence databases not fully comprehensive of the environment. They found the *Alphaproteobacteria* were under-represented by FISH and the *Bacteroidetes* over-represented, amongst other discrepancies. Castle and Kirchman (2004) reported a discrepancy in abundances measured by FISH as compared to DGGE, with FISH showing *Betaproteobacteria* as the dominant group in a coastal seawater community, but DGGE showing *Betaproteobacteria* absent from the same community.

A potentially more accurate measure of abundance of component community members is by real-time PCR (Walker, 2002). Again restricted by knowledge of taxon-specific sequence information, and additionally by gene copy number per genome, real-time PCR is nevertheless potentially capable of much higher throughput. It has been employed to measure the abundance of pelagic denitrifying bacteria (Grüntzig *et al.*, 2001; Labrenz *et al.*, 2004), and the changing proportions of the picocyanobacteria *Prochlorococcus* and *Synechococcus* (Suzuki *et al.*, 2000).

Rapid community fingerprints can be gained using length heterogeneity (LH) PCR, which distinguishes organisms by length polymorphisms in specific genomic locations, such as the internal transcribed spacer (ITS) between the 16S (small subunit) and 23S (large subunit) rRNA genes, when it is known as rRNA intergenic spacer analysis (RISA) (Suzuki *et al.*, 1998; Acinas *et al.*, 1999). Profiling is coarse but requires no sequence knowledge, beyond that of priming sites, and enables rapid tracking of changes in community structure. Carlson and co-workers used LH-PCR to monitor the mesopelagic bacterial community in the Sargasso Sea following nutrient enrichment of the sub-photic population due to convective overturn (Carlson *et al.*, 2004). LH-PCR can also be used to screen clone libraries for clones with desired inserts prior to sequencing, making library analysis more efficient (Suzuki *et al.*, 2004).

In automated RISA (ARISA) high throughput is enabled by automated electrophoresis (Fisher and Triplett, 1999). Rapid, repeatable profiles are generated which can be used to inform decisions about further analysis with more accurate, but intensive, techniques such as cloning and sequencing (Brown *et al.*, 2005). ARISA has been used successfully to compare spatial and temporal developments in bacterial communities in a freshwater lake (Fisher and Triplett, 1999), an offshore river water plume (Hewson *et al.*, 2006), and a coastal ocean community sampled through seasonal changes (Brown *et al.*, 2005).

Library clones are profiled using restriction fragment length polymorphism (RFLP) in which DNA is digested into fragments according to the positions of specific restriction sites (Moyer *et al.*, 1994). Screening clones this way is fast, and does not necessarily require sequencing, without which identification is restricted to an electrophoresis pattern database. Sequences recovered from RFLP gels lack the universality and therefore phylogenetic usefulness of 16S rRNA gene sequences recovered from DGGE gels. In this way the technique is less informative than DGGE,

and indeed it is mostly used as a screening tool to inform choices about further analysis such as sequencing. Bacterial diversity has been screened, by RFLP, in clone libraries constructed from environmental DNA originating along an estuarine transect (Henriques *et al.*, 2004), in contrasting open ocean and coastal ocean locations (Rappé *et al.*, 2000), in mobile deltaic sediments (Todorov *et al.*, 2000), in a community of generic phototrophs and their co-occurring viruses (Mühling *et al.*, 2005; Mühling *et al.*, 2006), and in fine resolution within a single bacterial class in the oligotrophic open ocean (Jameson *et al.*, 2007). RFLP of library clones was used in conjunction with FISH and sequencing to identify a novel bacterial phylum associated with marine sponges from the Sargasso and Mediterranean Seas (Fieseler *et al.*, 2004).

Whole community profiles can be obtained using only the terminal restriction fragments which are fluorescently labelled. Terminal RFLP (TRFLP), as it is known, gives an accurate community profile based on the lengths of terminal fragments, and is comparable in resolution to DGGE (Liu *et al.*, 1997; Moeseneder *et al.*, 1999; Díez *et al.*, 2001). Like DGGE it requires much research into choice of primers used for PCR and the specific combination of restriction enzymes used. Bacterial sequences in databases need to be analysed before they can be matched to specific terminal length fragments, thus being assigned a "ribotype", and there is currently a three-fold redundancy between sequences (catalogued but not analysed) and ribotypes (sequences assigned a taxonomic identity). Ultimately identification of organisms relies on alignment with known and catalogued genomic DNA sequences, such as those of 16S rRNA genes originating from either cultured bacteria, or from whole-genome sequences in Genbank.

Serial analysis of ribosomal sequence tags (SARST) is a high throughput, cloning-free procedure (Neufeld *et al.*, 2004). Restriction enzymes digest environmental 16S rRNA gene PCR products to generate a pool of homologous rRNA hypervariable region gene sequences. These sequences or "tags" are concatemerised by ligation, and can then be sequenced. Originally targeting the V1 region of the 16S rRNA gene, a variation on the method targets the V6 region and produces community profiles of slightly different proportions (Kysela *et al.*, 2005).

1.2.4 Denaturing gradient gel electrophoresis (DGGE)

The collection of double stranded (ds) DNA molecules that result from one or more rounds of PCR using environmental DNA and oligonucleotide primers binding to sites within the 16S rRNA gene are all roughly of the same length. These homologous sequences may exhibit a limited length polymorphism, but the majority cannot be separated on the basis of length heterogeneity by conventional agarose gel electrophoresis. Such a mixture of equal length DNA fragments can be resolved using DGGE; first shown in medical tissue typing to result in the separation of over 50 % of fragments between 100 and 1000 base pairs, which differed by a single base substitution (Sheffield *et al.*, 1989). Electrophoresis is through an acrylamide gel containing a gradient of increasing concentration of the denaturants urea and formamide. The DNA fragments begin to denature as they encounter higher denaturant concentration. Denaturation proceeds according to the distribution, between and within sequences, of melting domains. These stretches of base sequence denature at certain positions in the gel according to the abundance and pattern of A:T base pairs, which bond with two hydrogen bonds, compared to the three of the G:C base pair. The final spread of DNA bands through

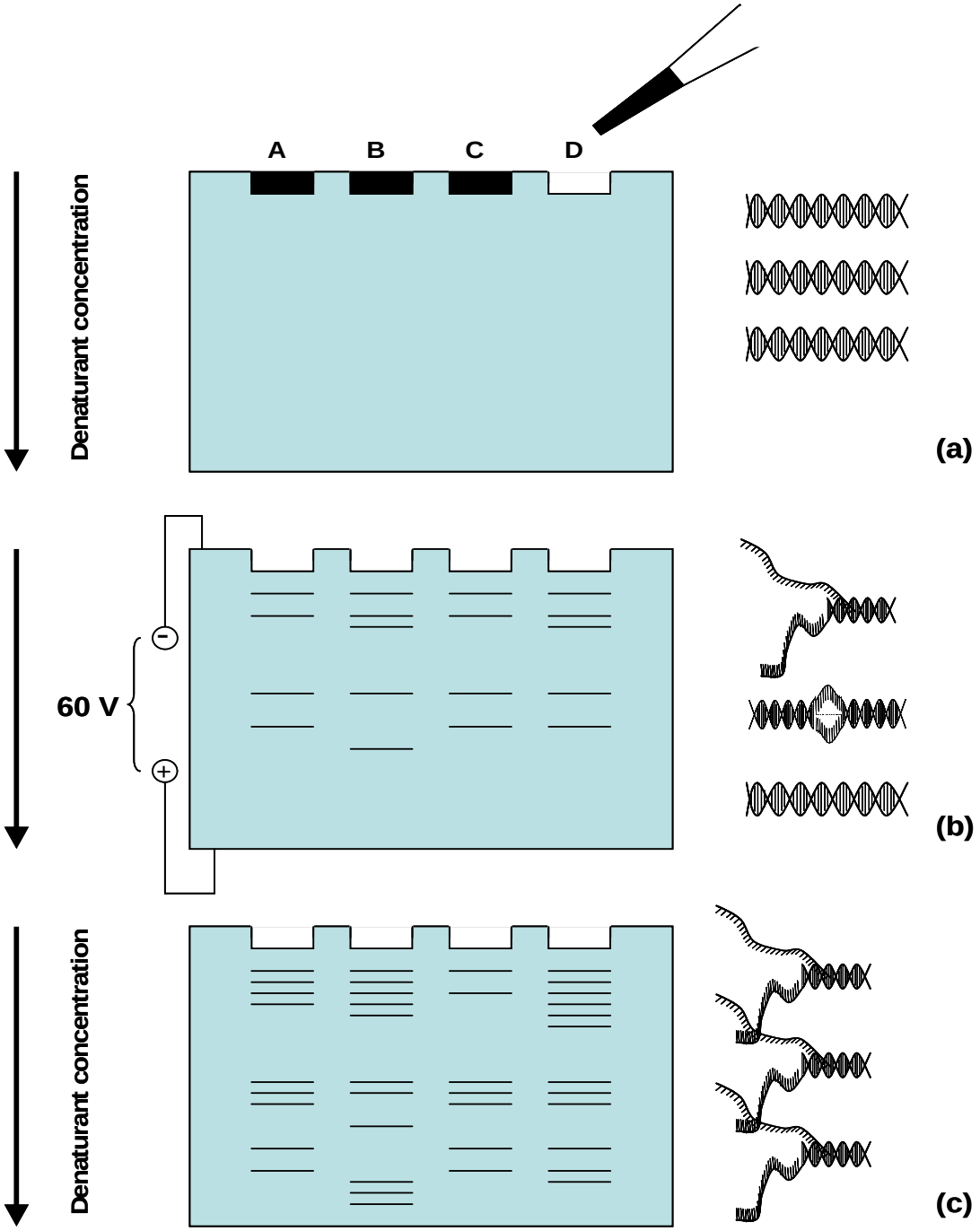


Figure 1.3. Theory of denaturing gradient gel electrophoresis. **(a)** Mixed 16S rRNA gene PCR products generated with the same primers, but with different environmental template DNA (A, B, C, D) are loaded into wells in a polyacrylamide gel containing an increasing concentration gradient of denaturants (formamide and urea). Double-stranded DNA fragments incorporate GC-clamps. **(b)** A potential of 60 V is applied across the gel. After two hours there has been differential migration of dsDNA fragments possessing differing melting domains. Melting properties of each fragment are governed by the fragment's base sequence and the contrasting strengths of the hydrogen bonding between A:T and G:C base pairs. Partially melted fragments have reduced mobility, while fully melted fragments, held together only by their GC-clamps, are arrested in the gel. **(c)** Electrophoresis is complete after 17 hours and all fragments have fully melted and separated according to the smallest differences. Banding patterns are visualised under UV illumination after staining of DNA with ethidium bromide, an example of which can be seen in Figure 2.5. Occurrence of bands (consisting of ribotypes – 16S rRNA gene sequences) in profiles ranges from unique (occurring in single profiles in isolation) to universal (common to all profiles).

the gel reflects the melting properties of each unique sequence. Fragments possessing predominantly GC-rich melting domains travel further through the gel than those containing predominantly AT-rich melting domains. The mechanism responsible for

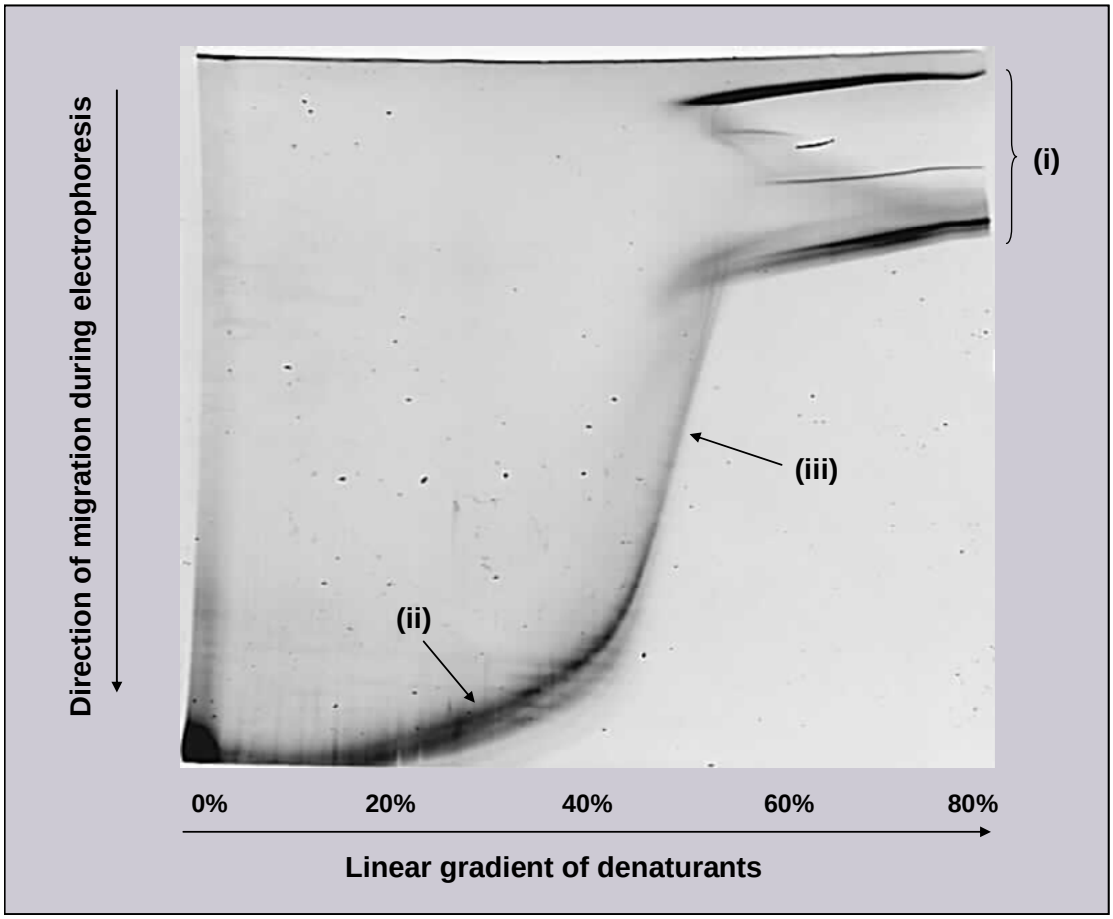


Figure 1.4. Estimating denaturing gradient range using perpendicular DGGE. Acrylamide gel is cast such that the gradient of denaturants runs horizontally. 400 µL PCR product was loaded into a continuous well at the top of the gel. Sample was obtained from mixed genomic DNA from cultured *Bacillus* sp. and, *Paenibacillus* sp. with nested PCR using *Firmicutes*-specific primers listed in Table 2.1. (i) On the high denaturant side of the gel dsDNA fragments are stopped high in the gel when they partially melt, and there is some separation according to variation in melting domain properties, a result of sequence variation. This is analogous to the separation of mixed fragments in a parallel denaturing gel. (ii) On the low denaturant side of the gel dsDNA fragments travel quickly through the gel without denaturing. There is little separation due to sequence-induced variation in melting behaviour. The DNA sample will run through the gel completely if electrophoresis continues. (iii) The portion of the DNA sample melted and thus stopped between the high and low denaturant regions of the gel indicates the range of denaturant concentration which will separate the majority of the sample on a parallel denaturant gradient gel. In this case that range is roughly between 40% and 55%. Electrophoresis was for 15 hours at 60 V. [Cultured isolates from M. Mühling's collection]

stopping migration is mechanical; the uncoupled single stranded portions of molecules cannot easily travel through the matrix of the gel. A GC-rich sequence (GC-clamp) ligated at one end of each fragment, and originating from an especially elongated PCR primer (see "Primers used for re-PCR for DGGE" in Table 3.2), prevents complete denaturation and excessive electrophoretic mobility and the loss of single stranded DNA fragments from the gel. A standardised mixture of known fragments may be run in one or more lanes of a gel to allow inter-gel comparison. Banding profiles, or fingerprints, generated from the same DNA template using the same PCR protocol are repeatable.

DGGE was first used to separate DNA fragments of equal length in the early 1980s by Fischer and Lerman who were looking to separate phage λ sequences differing by single nucleotide mutations (Fischer and Lerman, 1983). Greater utility was given to the technique, which had become popular in medical point mutation diagnostics, with the introduction of GC-clamped primers (Sheffield *et al.*, 1989). Gerhard Muyzer and coworkers (1993) demonstrated the potential of the technique for resolving mixed samples of ribosomal gene fragments amplified directly from the environment, thus adapting the technique to microbial ecology.

Since then many researchers have employed DGGE to elucidate the separate component members of microbial communities in a wide range of natural (for example see Grossart *et al.*, 2006b) and artificial (for example see Lopez *et al.*, 2003) Habitats.

In the marine environment there has also been a wide range of studies using DGGE to describe variation within and between different communities, and also changes in such communities over time (for examples see Ferrari and Hollibaugh, 1999; Schäfer *et al.*, 2002; and Rink *et al.*, 2007).

Although several housekeeping genes encoding proteins and enzymes, and some intragenic stretches of genomes, have been used, the genomic DNA of choice for this community profiling is the 16S rRNA gene. Many PCR primers have been designed, and primer combinations used, for DGGE, amplifying varying portions of the 16S rRNA gene, in a variety of taxonomic groups of prokaryotes.

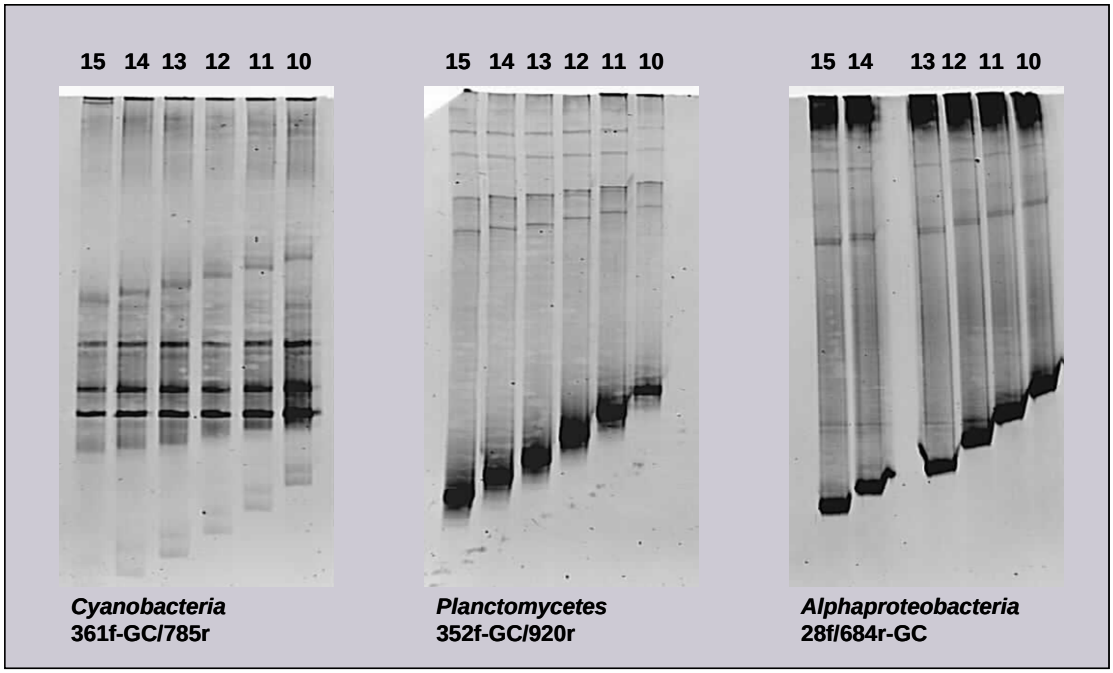


Figure 1.5. Optimising electrophoresis duration in a DGGE time series experiment. Similar PCR products, generated using the primers indicated and environmental DNA from an ambient CO_2 mesocosm (Section 4.2.1), are run in adjacent lanes at 60 V for varying lengths of time (hours). Over shorter time spans different bands may fail to separate, or fail to migrate significantly from the well. Over longer time spans separated bands may come together again, and some fragments may be lost through the gel completely. The range of concentration of the denaturants formamide and urea in the gels is from 30 % to 60 %.

1.2.5 The choice of taxonomic and phylogenetic marker

That a semantic molecule might act as an evolutionary chronometer was first suggested by Zuckerkandl and Pauling, who realised also that genomic DNA had the greatest potential (Zuckerkandl and Pauling, 1965). The 16S rRNA gene has become the "gold standard" marker in the study of bacterial evolution and ecology (Case *et al.*, 2007). The three domain phylogeny of life, pioneered by Woese (1987), is the accepted framework for bacterial phylogenetics, and is currently founded on over 30,000 near full-length 16S rRNA gene sequences (Ludwig and Schleifer, 2005). The latest release of Bergey's Manual of Systematic Bacteriology, the consensus reference for classification of species, is also based on 16S rRNA gene sequences (Garrity *et al.*, 2001).

Any phylogenetic marker needs to satisfy four requirements. It must (i) be of universal occurrence, at least within the group being studied, (ii) exhibit functional

constancy indicating an authologous (relating to a single common ancestor) descent, (iii) possess sufficient sequence conservation to span the evolutionary time required and (iv) possess sufficient sequence variation to distinguish between closely-matching sequences such as those of generic species.

16S rRNA genes satisfy the requirements for universal occurrence and sequence conservation and variability very well. However their existence in genomic operons as non-identical paralogues, raises the possibility of functional duplication within lineages, causing confusion in resulting phylogenies, such as the overestimation of species number (Ueda *et al.*, 1999). Phylogenetic precision can be restored to some degree by referring to "ribotypes" or "phylotypes," in which trees constructed on 16S rRNA gene sequences are labelled to show the relationships only between the 16S rRNA genes, and not necessarily the bacteria possessing them (Jaspers and Overmann, 2004). Recombination in the 16S rRNA gene due to genetic crossing-over between different genomes is rare in bacteria but has been demonstrated, and this is possibly a more intractable problem for phylogenetic trees (Sneath, 1993).

Recently researchers have begun to voice support for other markers of universal phylogeny, notably protein-encoding genes. The genes encoding the beta and gamma subunits of RNA polymerase, *rpoB* and *rpoC* respectively, exist only in single copy, and have been shown to overcome the difficulties encountered in the multiple intracellular ribotypes of 16S rRNA genes (Dahllöf *et al.*, 2000; Mühling *et al.*, 2006; Case *et al.*, 2007). Some markers, such as the elongation factors *Tu* and *1alpha* in *Bacteria* and *Archaea* respectively , and *rpoB* and *rpoC*, as well as those for heat shock proteins, and some genes for aminoacyl-tRNA sythetases produce phylum and domain level phylogenies supporting the three-domain tree described by 16S rRNA genes (Ludwig and Schleifer, 2005). Other markers do not support the basic three-domain tree of life. These include ATPase genes, DNA gyrase subunit genes, *RecA*, and several aminoacyl-tRNA synthetase genes (Ludwig and Schleifer, 2005).

Some markers may satisfy all the criteria for phylogenetic application except universality. However such markers can still be used effectively to define phylogenies within the evolutionary groups that they occur. For example *psbA*, a gene encoding a protein subunit of Photosystem II, was used to elucidate evolutionary relationships among the oceanic *Cyanobacteria* (Zeidner and Béjà, 2004).

Other markers can distinguish members of even more specific communities. A population of the coccolithophore *Emiliania huxleyi* and its co-occurring viruses was fractionated into subtypes, with specific infecting viruses, using the respective markers *GPA*, a gene encoding a calcium-binding protein, and one encoding the major capsid protein of the virus (Martínez Martínez *et al.*, 2007). The related 18S rRNA gene of eukaryotes was used to generate a phylogeny of marine nematodes (Cook *et al.*, 2005), and of pelagic phytoplankton (Moon-van der Staay *et al.*, 2001).

Plesiomorphy is problematic in any sequence based phylogeny. Especially in variable regions, any base in bacterial genomic DNA may be the result of several mutations resulting in a return to the original base. Thus trees may be based on false homogeneity (Gupta and Griffiths, 2002). This is difficult to detect, but its likelihood decreases in regions of conservation, and its effect is diluted in longer sequences.

Marine bacterial community studies employing DGGE with markers other than the 16S rRNA gene include those using *rpoB* to assess bacteria associated with a macroalgae (Dahllöf *et al.*, 2000), and a coral (Bourne and Munn, 2005). *PsbA*, which reconfirmed the 16S rRNA gene phylogeny, was used to assess a marine cyanobacterial population (Zeidner *et al.*, 2003; Zeidner and Béjà, 2004).

Whole genome studies, or at least studies based on several phylogenetic and functional genetic markers, look set to succeed those based on the 16S rRNA gene. Databases housing whole genome sequences of bacteria are expanding as those for the 16S rRNA gene once did. However there is a risk of whole genome datasets providing increased genetic information of unknown function, and therefore without a clear context. For the moment the gene is still highly useful, and was therefore the choice of marker for an investigation of bacterial communities, using DGGE, in this study.

1.2.6 DGGE using the 16S rRNA gene as a taxonomic and phylogenetic marker

DGGE has been used, in conjunction with PCR of environmental 16S rRNA genes, to study a wide variety of bacterial communities. Non-marine habitats explored include freshwater environments (Casamayor *et al.*, 2002; Goddard *et al.*, 2005), soil (Nakatsu *et al.*, 2000; McCaig *et al.*, 2001; Girvan *et al.*, 2005), an arctic saline spring (Perreault *et al.*, 2007), sewerage (Boon *et al.*, 2002; Dar *et al.*, 2005), household waste (Mayrhofer *et al.*, 2006), wine (Lopez *et al.*, 2003), and the human body (Li *et al.*, 2006a).

Within the marine environment DGGE-based studies have helped to elucidate community diversity and structure in specific habitats, and to track temporal and spatial changes in communities. Castle and Kirchman (2004) used DGGE and FISH to assess community composition and component abundance along a transect through the surface waters of an estuary. In another study seasonal changes in the composition of a pelagic bacterioplankton community in offshore Mediterranean water were recorded. Small changes were reported, although primers used were universal for the *Bacteria* domain, and results were described mostly at the class level (Alonso-Sáez *et al.*, 2007). Eiler and Bertilsson (2006) used *Vibrio*-specific primers for QC-PCR together with DGGE to ascertain the variety of *Vibrio* strains in seawater samples containing fewer than 200 target cells.

Other bacterial communities resolved using DGGE include those associated with sponges (Li *et al.*, 2006b; Wichels *et al.*, 2006), with seasonal phytoplankton blooms (Rink *et al.*, 2007), with pelagic diatoms (Schäfer *et al.*, 2002; Grossart *et al.*, 2005), with mesopelagic marine snow particles (Grossart *et al.*, 2006b), and with free-living pelagic communities throughout the global ocean (e.g.: Ferrari and Hollibaugh, 1999; Riemann *et al.*, 1999; West and Scanlan, 1999). Depth-specific spatial changes have also been reported for various bacterial taxa (Blümel *et al.*, 2007).

1.2.7 Some important bacterial phyla; *Bacteroidetes, Cyanobacteria, Firmicutes, Planctomycetes, Proteobacteria*

Bacteroidetes. This phylum of aerobic or facultatively anaerobic chemoheterotrophs was recently described to include the monophyletic majority of the previous grouping *Cytophaga-Flavobacterium-Bacteroides* (CFB) which was polyphyletic (Giovannoni and Rappé, 2000). Members of the genera *Cytophaga/Cellulophaga* and *Flavobacterium* possess unusual flexirubin and carotenoid pigments, and there is still taxonomic confusion within the *Bacteroidetes* lineage (an example of work to resolve this can be found in Johansen *et al.*, 1999). Gliding motility and production of extracellular degradative enzymes are characteristic properties of the group, the latter being important in decomposing POM, some being able to degrade the agar of culture plates. Some genera are psychrophillic. Blümel and co-workers (2007) demonstrated shifts in composition of the *Bacteroidetes* component of the bacterioplankton with depth. These changes, they

suggest, probably reflect changing properties of seawater, with the quantity and quality of DOM and POM being important, as well as temperature and salinity. The phylum as a whole is important in the heterotrophic cycling of organic carbon, and members of the orders *Flavobacteriales* and *Sphingobacteriales* were reported to comprise over 50 % of temperate surface water bacterial communities in spring, with 20 % of the phylum's representatives coming from a single *Cytophaga* species (Eilers *et al.*, 2001). The genus *Bacteroides* is normally found in mammalian intestines, but can persist for a long time in the sea, and may be considered an indicator of faecal pollution (Munn, 2004).

Cyanobacteria. The oxygenic phototrophic *Cyanobacteria* include the most important marine fixers of carbon (*Synechococcus* and *Prochlorococcus*) and nitrogen (*Trichodesmium*), and are therefore key players in the global cycling of these two elements (Zubkov *et al.*, 2003). The *Cyanobacteria* contain the green pigment chlorophyll *a* and red or blue phycobilins, with *Prochlorococcus* also containing chlorophyll *b* (Chisholm *et al.*, 1992). Their oxygenic autotrophy was responsible for turning the marine world from one dominated by heterotophic anaerobes to one dominated by photoautotrophs and heterotrophic aerobes around three billion years ago (Kasting and Seifert, 2002).

Their original classification as simple algae, and division into 150 genera and over 1000 species, has been found to be polyphyletic and unreliable using molecular methods. Many common genera are filamentous such as *Nostoc,* the nitrogen-fixing *Trichodesmium*, and the aggregate-forming *Cyanocystis*.

The most important photosynthetic organisms in the pelagic environment belong to the order *Chroococcales* and the genera *Prochlorococcus* (Chisholm *et al*., 1988) and *Synechococcus* (Waterbury *et al.*, 1979). After their discovery in the late 1980s and 1970s respectively, studies on oceanic primary production ascribed a greater role, within the phytoplankton, to cyanobacteria (Johnson and Sieburth, 1979; Krempin and Sullivan, 1981), and now the *Prochlorococcus* lineage alone is believed responsible for 60 % to 80 % of carbon fixation in oligotrophic regions of the oceans (Platt *et al.*, 1983). Both genera are ubiquitous in tropical and temperate waters where their density distribution correlates positively with temperature and negatively with nitrogen level (Jiao *et al.*, 2005). *Prochlorococcus* alone occurs at 10^5 to 10^6 cells mL^{-1} between 30^o N and 30^o S, and down to 200 m. It fixes between 15 % and 40 % of all

carbon entering the marine food chain (Partensky *et al.*, 1999). There are genetically distinct *Prochlorococcus* and *Synechococcus* ecotypes inhabiting vertically defined niches differing in levels of irradiance and nutrient concentration (Moore *et al.*, 1998; West and Scanlan, 1999; Rocap *et al.*, 2002). Although less than 2% divergent in 16S rRNA gene sequence, these ecotypes might be assigned as different species. This case may represent an example of the limitations of 16S rRNA-based phylogeny and taxonomy.

It is accepted dogma that chloroplasts of eukaryotic phytoplankton, as well as multicellular macroalgae and plants, originated from a symbiosis between a cyanobacterium and an ancestral eukaryote, possibly on multiple occasions (Archibald and Keeling, 2005). However the symbiont was not a direct ancestor of the modern *Prochlorococcus*, which lacks phycobilins, but probably was a common predecessor of both *Prochlorococcus* and the chloroplast's ancestral free-living cyanobacterium (Yoon *et al.*, 2002; Munn, 2004).

Firmicutes. Also known as the low GC-content Gram-positive bacteria, the phylum *Firmicutes* is best known from marine sediments. Pelagic species will likely be confined to aerobic genera such as *Bacillus*, and rarer representatives of *Staphylococcus*, *Lactobacillus* and *Listeria* (Munn, 2004). A diagnostic feature of the group is the production of reproductive endospores which allow for wide spatial and temporal distribution, potentially remaining viable for thousands of years (Munn, 2004). The existence of dormant spores also impacts on the distinction between the total cells present in a body of water, sampled in 16S rRNA gene-based studies, and the metabolically active community, made apparent in 16S rRNA-based studies.

Planctomycetes. There are only a few described genera within the phylum *Planctomycetes* (Ward *et al.*, 1995). These cells attach to particulate matter using a stalked proteinaceous appendage, and lack peptidoglycan, their cell walls consisting only of a more basic S-layer of geometrically arranged glycoproteins. Cells have internal membrane-bound compartments including, in some, a membrane-bound nucleus, previously thought to be definitive of eukaryotes (Munn, 2004). These features make the phylum one of few within the *Bacteria* identifiable by morphological taxonomic methods. Described marine genera include *Pirelulla* and *Planctomyces* (Ward *et al.*, 1995; Vergin *et al.*, 1998), as well as more recently

Brocadia (Strous *et al.* 2006), a stalked Planctomycete and the first organism identified as being responsible for the anammox reaction (the combined reduction of nitrate (NO^{2-}) and oxidation of ammonium (NH_4^+) to produce N_2 gas). They attach to marine snow particles and degrade organic compounds, reproducing by budding new cells from the non-attached pole (Munn, 2004).

Proteobacteria. An ancient origin and wide dispersal makes the *Proteobacteria* the most diverse and abundant bacterial phylum. It includes the well known *Escherichia* genus. The phylum is currently divided into five subgroups, *Alpha*, *Beta*, *Gamma*, *Delta* and *Epsilon* based on 16S rRNA gene sequences. The *Alpha*, *Beta* and *Gamma* subgroups each contain more species than most other bacterial phyla (Gupta and Griffiths, 2002). Since the growth of 16S rRNA gene sequence databases it has become apparent that *Gammaproteobacteria* were over-represented in culture collections, while *Alphaproteobacteria* were under-represented (Eilers *et al.*, 2001; Rappé and Giovannoni, 2003; Cho and Giovannoni, 2004).

Pelagic marine representatives include the anoxygenic photosynthesisers *Roseobacter* (*Alpha*) which may operate aerobically or anaerobically. The *Roseobacter* lineage may not be a phylogentically coherent group within the *Alphaproteobacteria* (Buchan *et al*., 2005), and its physiological diversity means it does not appear as a distinct family or order in Bergey's Manual of Systematic Bacteriology's 2003 Taxonomic Outline of the Prokaryotes (Garrity *et al*., 2003; Buchan *et al.*, 2005). The group, significant in all pelagic niches studied, contains cultured isolates for most sub-clusters which match closely with environmental clones. Members of this lineage often comprise 25 % or more of bacterial communities in coastal and polar oceans (Wagner-Döbler and Biebl, 2006).

The purple non-sulphur bacterium *Rhodospirillum* (*Alpha*) can utilise organic or inorganic carbon sources. Numerous free-living heterotrophs are typified by the aerobic organic carbon consuming *Pseudomonas* (*Gamma*) and *Halomonas* (*Alpha*), while nitrification is carried out by *Nitrospira* (*Gamma*). Light-driven membrane proton pumps, called proteorhodopsins, help power cellular processes within the SAR86 (*Gamma*) and other groups. However the sequences of genes for these proteins describe phylogenetic lineages in contrast to those constructed using 16S rRNA genes (Sabehi *et al.*, 2004). *Vibrio* (*Gamma*) species tend to be associated with pelagic animals, particularly zooplankton (Heidelberg *et al.*, 2002).

Deltaproteobacteria include the aphotic Marine Group B and the intracellular parasite *Bdellovibrio* sp. (Perry *et al.*, 2002; Munn, 2004).

1.2.8 The need for deeper, more reliable bacterial identification

There have been many studies investigating the phylogenetic relationships within and between the phyla of the *Bacteria*, and investigating the geographical distribution, biogeochemical roles, and food web roles of their representatives in the marine environment. However those involving PCR of environmental DNA, or *in situ* oligonucleotide hybridisation, have largely relied on 16S rRNA gene primers or probes loosely targeted at the entire *Bacteria* domain. Such universal oligonucleotides may provide only limited resolution or skewed coverage of an actual community. For example Castle and Kirchman (2004) could not detect *betaproteobacteria* in community samples using DGGE with universal *Bacteria* primers, which proved to be present in the same samples when analysed by other approaches. Attempts to estimate the total bacterial community diversity have used partial analysis of the total community (that is clone library screening) combined with theoretical models. This approach, however, does not reveal the identity of the less abundant components of the assemblage. Increasing the number of clones per clone library has been successful in detecting novel bacterial clades, (for example Chouari *et al.*, 2005), or comparing different marine environments in terms of bacterial community composition (for example Rappé *et al.*, 2000; Zaballos *et al.*, 2006). However, despite the decreasing costs for nucleotide sequencing, the success of this approach is still limited because of the huge scale of bacterial diversity — perhaps as many as 2×10^6 different species in the oceans (Curtis *et al.*, 2002). The recent arrival of pyrosequencing of 16S rRNA tags (Sogin *et al.*, 2006) may represent an alternative because of the much lower cost per sequence; but pyrosequencing still does not allow analyses and comparison of the bacterial diversity in different environments on a routine basis.

Alternative approaches, such as DGGE, are routinely used to determine diversity because they avoid large-scale sequencing efforts. However, these are likely also to detect only a small fraction of the total diversity. The use of universal bacterial PCR primers is likely to miss minor fractions of the microbial community because most of the PCR product will be composed of the more abundant species. Faint DNA bands on DGGE gels are unlikely to be detected or their identity determined. To overcome this limitation and to detect less abundant sequence clones, (Holben *et al.*,

2004) fractionated 16S rRNA gene sequences from a microbial community according to their G+C content before DGGE analysis. However, with the possible exception of the high G+C-containing *Actinobacteria*, this method has limited application and does not separate bacteria by phylogeny. Combining bromodeoxyuridine immunocapture and DGGE has been proposed to separate the DNA of the actively growing bacteria from the rest of the environmental DNA (Hamasaki *et al.*, 2007). However, this results in the analysis of subgroups of bacteria that are not defined on phylogenetic criteria, and does not allow screening of the whole range of genetic bacterial diversity.

An alternative approach is the application of group-specific oligonucleotide probes. This has mainly been used in combination with FISH. This has been a very effective technique but, in contrast to the analysis of clone libraries produced from PCR fragments, it cannot reveal unknown bacterial groups. It only detects and quantifies those bacteria that the probes were designed to detect.

Group-specific primers used in DGGE will facilitate description of community structure at improved resolution, and with reduced overall PCR bias. They will improve detection of sequences rare in the environment, and will thus be more sensitive to spatial and temporal changes in community structure due to environmental perturbation. Phylogenetic trees based on genomic DNA sequences such as the 16S rRNA gene are becoming increasing annotated with the metabolic functions of the bacteria those sequences represent. Precisely targeted oligonucleotide-based experiments will thus be able to assess temporal and spatial changes in bacterioplankton populations. This will yield results not only describing changes to genotype, such as changing 16S rRNA gene sequences, but also describing functional metabolic shifts in bacterial communities. The use of 16S rRNA genes for this purpose will be furthered by the linking of function to ribosomal RNA genes specific for narrower taxonomic groups. The eight phylum- and class-specific primer sets used here will help further this goal.

1.3 Ocean acidification

1.3.1 The nature of ocean acidification

The proportion of the Earth's atmosphere made up of carbon dioxide (CO_2), a significant greenhouse gas, has been steadily rising over the last quarter millennium. Mankind's burning of fossil fuels and destruction of forest ecosystems are the primary reasons for this increase, from an atmospheric CO_2 component of 280 ppm in 1750 to 380 ppm in 2000 (Figure 1.9, Caldeira and Wickett, 2003). However, the current level of CO_2 in the atmosphere does not account for all the anthropogenic production of the gas over the last 250 years. Separate from the terrestrial biosphere, the global ocean has absorbed, from the atmosphere, approximately half of all anthropogenic CO_2 ever produced (Sabine *et al.*, 2004). The Atlantic Ocean is responsible for ~ 60 % of this CO_2 sink, with the remainder accounted for by the Indian and Southern Oceans, with the Pacific Ocean having a neutral net flux (Takahashi *et al.*, 1997). Thus an equilibrium, pertaining to the gas, is maintained between the sea and the air.

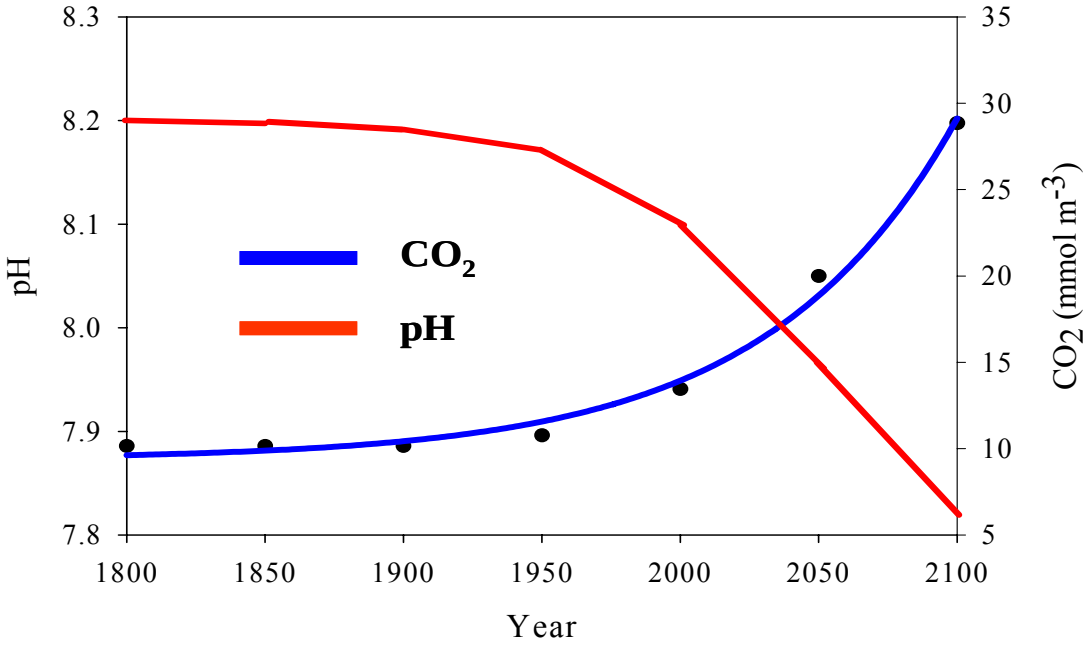


Figure 1.6. Past and projected future surface seawater pH, and pCO_2, for the global ocean (redrawn from Wolf-Gladrow *et al.*, 1999).

Once dissolved in seawater, CO_2 exists as a weak acid, carbonic acid. It is otherwise known as dissolved inorganic carbon (DIC) and exists in three forms; CO_2, carbonate (CO_3^{2-}) and bicarbonate (HCO_3^-) which vary in proportion depending on

temperature and acidity (pH). Therefore a rising partial pressure of CO_2 (pCO_2) in seawater has potential consequences for marine life due to a lowered pH, or ocean acidification, and an alteration in concentration of the three constituents of DIC (Raven *et al.*, 2005). The relationship between pH and DIC constitution is illustrated in Figure 1.7. Pelagic microorganisms, including bacteria, will both respond to, and be an influence on, seawater pCO_2. Generally photosynthesis and sedimentation of organic matter consume CO_2, while respiration releases it. Oceanic physico-chemical parameters influence these microbial populations, and events such as El Niño – Southern Oscillation (ENSO) drive decadal changes in surface water pCO_2 of Δ5 ppm yr^{-1} (Takahashi *et al.*, 2003). Oscillations in pCO_2 driven by ENSO include a reduction in the strength of the CO_2 sink in drought-induced increasingly saline waters. While clearly climate-related, such temporal shifts in oceanic carbon cycling need to be separated from long-term, and anthropogenically caused, ocean acidification (Dore *et al.*, 2003).

A distinct effect of lowering the pH of seawater is an alteration of the ratio of the DIC species present, as shown in Figure 1.7. Significantly CO_3^{2-} declines with lower pH, and the saturation point of this compound rises in the water column. Below this saturation point, the compensation depth, which tends to be higher for aragonite than calcite, organisms that rely on mineralisation of $CaCO_3$ for protective shells, such as corals and pteropods (aragonite) and coccolithophores and foraminiferans (calcite), will find it increasingly difficult to build, and maintain the integrity of, shells by mineralising dissolved $CaCO_3$ (Feely *et al.*, 2004). Calcifying planktonic organisms act as a significant buffer in the ocean carbonate system and were the primary players in the development of a deep-sea carbonate sink (Ridgwell *et al.*, 2003). The compensation depth for calcite was shown to rise by over 2 km globally in the space of 10,000 years around 55 x 10^6 years ago during the Palaeocene-Eocene thermal maximum (PETM) (Zachos *et al.*, 2005). While not directly influencing bacteria, this occurrence may have affected bacteria associated with calcifying organisms such as the abundant coccolithophores and formainiferans. Shelled molluscs and arthropods have their own associated bacterial communities, and they are locally important producers of DOM. However some widespread and abundant phytoplankton species, such as the coccolithophore *E. huxleyi*, form shells of minute $CaCO_3$ liths. Riebesell and co-workers (2000) noted reduced calcification in coccolithophorid species at pH 8, brought about by pCO_2 of 800 ppm, and a

conservative estimate of ocean acidity for the end of this century (Caldeira and Wickett, 2003). Carbon fixation and organic matter release by eukaryotes, such as *E. huxleyi*, impact significantly on the heterotrophic bacterial community throughout the water column, which is dependent on DOC to supply both energy generating and anabolic biochemical pathways.

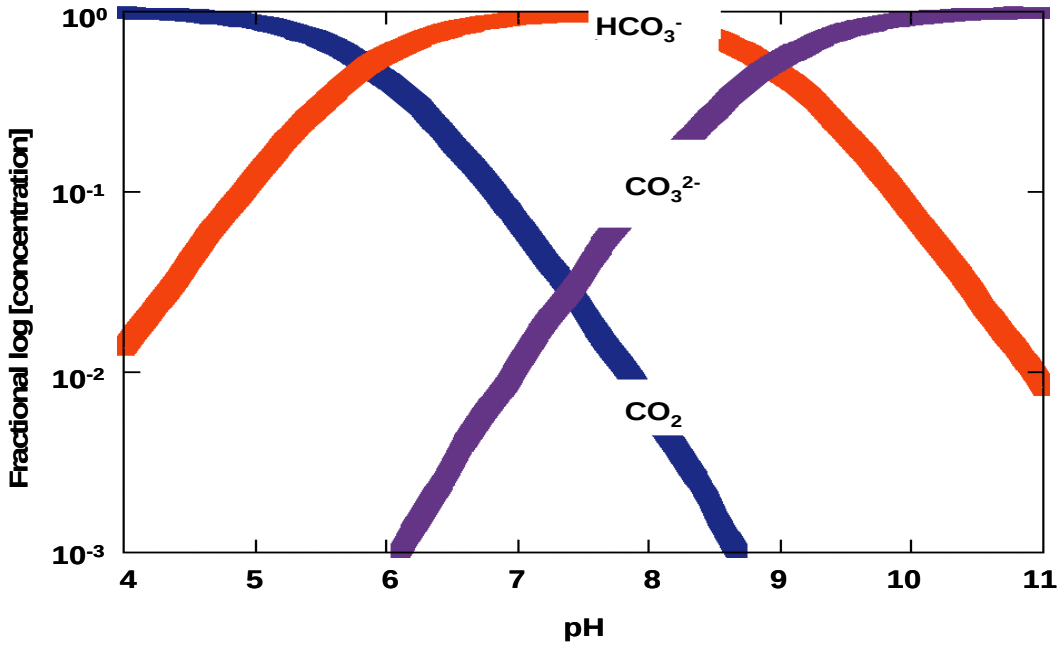


Figure 1.7. Relationship in seawater between the components of DIC and pH (Raven *et al.*, 2005).

Essentially, atmospheric CO_2, increasing in concentration at an unprecedented rate, due mainly to anthropogenic emissions, is absorbed by, and reacts with, seawater to become carbonic acid (H_2CO_3) and the associated DIC compounds bicarbonate (HCO_3^-) and carbonate (CO_3^{2-}), increasing the concentration of free protons. As well as an overall fall in pH, associated changes to the property of seawater, as shown in Figure 1.7 include rising pCO_2 and falling pCO_3^{2-} (Raven *et al.*, 2005). This decreased availability of CO_3^{2-} may adversely affect eukaryotic organisms, such as the unicellular *Coccolithophoridae* and multicellular *Mollusca* and *Echinodermata*, dependent on the compound for the manufacture of $CaCO_3$ shells (Raven *et al.*, 2005). As the concentration of pCO_3^{2-} declines the dissolving depth for $CaCO_3$ moves nearer the surface.

Bacteria associated with these organisms will be indirectly affected. Many bacterial taxa may be affected by falling pH and changing seawater chemistry

directly. These environmental changes may drive changes in the composition of bacterial communities by acting on individual cells. Such potential changes can occur via two mechanisms. These are altered gene expression, i.e. phenotypic acclimation, and selection for naturally occurring mutant (homologous) genes in the community, i.e. natural selection.

1.3.2 Past episodes of ocean acidification

During the course of its possible 3.8 x 10^9 years of existence (Conway Morris, 1998), from the end of the Hadean period, the chemical and physical state of the world ocean has changed continuously. The early atmosphere exhibited a greenhouse effect due to CO_2 over 2 x 10^9 years ago, which maintained the nascent oceans in a liquid state (Ohmoto *et al.*, 2004). Initially anoxic, the rise of oxygenic *Cyanobacteria* created the current high levels of oxygen in the atmosphere and seas. During this oxic period of Earth history the pH level of the ocean has risen and fallen, due to geological events such as massive release, and subsequent oxidation, of methane during the PETM leading to pH lowering (Zachos *et al.*, 2005). This ocean acidity variation has been mainly due to variation in atmospheric, and therefore oceanic, CO_2 concentration, the source of which has been mainly volcanism, and the sink of which

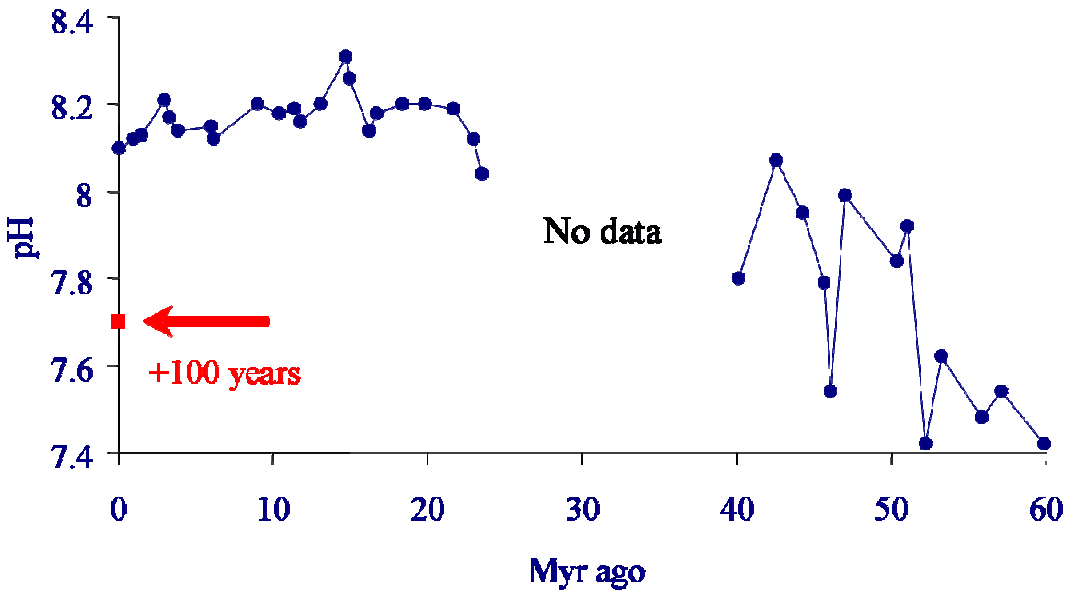


Figure 1.8. Past levels of surface ocean pH obtained by proxy analysis of boron isotopes in fossilised foraminiferans. Red arrow indicates projected global average pH level for surface ocean for year 2100 (Pearson and Palmer, 2000).

has been mainly organic life (Raven *et al.*, 2005 review of ocean acidification throughout Earth's history).

Along with variations in other physico-chemical environmental parameters such as temperature and light intensity, cellular marine life has had to cope with this varying pH. Past fluctuations in ocean acidity have ranged higher and lower than that forecast for the Anthropocene era due to industry-driven rises in CO_2 emissions, largely from the burning of fossil fuels.

However, past extinction episodes probably hit hardest multicellular organisms such as animals and vascular plants. Although evidence for or against the occurrence of ancient bacterial mass extinction events is scant, it is intuitive and reasonable to suppose that prokaryotic populations will be more resilient to future climate-led changes, while metazoans, such as *Homo* sp., will be more sensitive.

1.3.3 Future predictions

The main difference between the current episode of ocean acidification and past episodes is the much higher rate of change. A global surface average pH of 8.3 in 1950 is set to drop to 7.8 by 2100 (Caldeira and Wickett, 2003). As shown in Figure 1.8, not only will this be the lowest level for over 25 million years, but this rate of change (0.5 pH units per 150 years) is markedly higher than a previous comparable drop in global ocean pH when, according to boron isotope measurements in fossilised *Foraminifera*, pH fell from 8.0 to 7.5 between 47 and 46 million years ago, although finer timescale fluctuations within this period cannot be accurately measured (Figure 1.8; Pearson and Palmer, 2000).

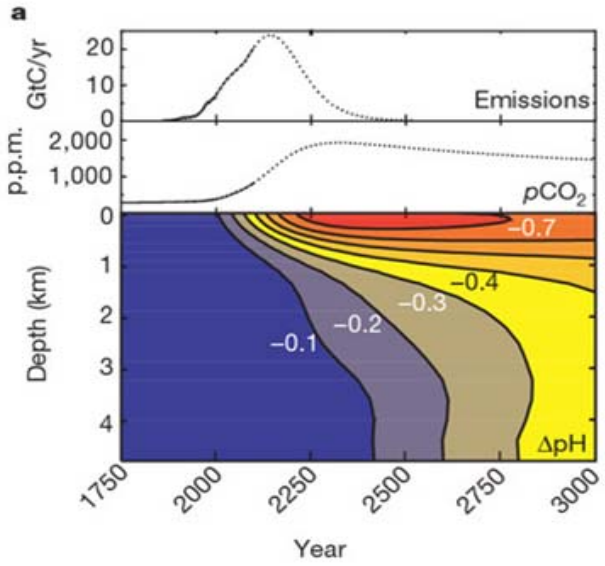


Figure 1.9. Anthropogenic carbon emission-driven changes to global atmospheric pCO_2 and oceanic pH since the industrial revolution in Europe and projected to the year 3000, assuming continued exploitation of fossil fuel reserves (Caldeira and Wickett, 2003).

Caldeira and Wickett (2003) document a widely held consensus of the anthropogenic effects on global ocean acidity and temperature. Their model predicts that organisms in the euphotic zone (< 250 m) will have to cope with a drop in seawater pH of 0.7 units over the next 200 years, and that this acidity peak will last for a further 500 years. Although water at greater depths will be affected to a lesser extent, it is in the euphotic zone that nearly all the global marine primary production occurs (Munn, 2004). That portion of primary production based on geothermal energy is hardly significant in its contribution to global carbon fixation (Munn, 2004).

Much research has been conducted into the effects of predicted ocean acidification on other seawater properties and on phytoplankton species and communities. These include the coccolithophorid *E. huxleyi* (Iglesias-Rodríguez *et al.*, 2002; Sciandra *et al.*, 2003; Engel *et al.*, 2004), inorganic/organic carbon ratios in seawater (Zondervan *et al.*, 2002), ratios of C, N and P, known as Redfield ratios, in different species (Burkhardt *et al.*, 1999), rates of primary production (Burkhardt *et al.*, 2001; Rost *et al.*, 2003; Schippers *et al.*, 2004), combined effects with nitrogen-limitation (Sciandra *et al.*, 2003), and evolutionary effects on phytoplankton phenotypes (Collins and Bell, 2004). Rather less research has been conducted into the possible effects on bacteria. Stretton and colleagues (1996) showed how CO_2 might

induce changes in gene expression thus altering physiological phenotypes *in situ*, while Grossart and co-workers' study was the first to report dynamic changes to heterotrophic bacterial populations responding to altered pCO_2 (Grossart *et al.*, 2006a).

This study shows predominant stability with a few significant changes occurring in the composition of bacterioplankton communities at the phylum and class level following a relatively short duration of elevated pCO_2.

1.4 Mitigation of ocean acidification

1.4.1 Natural carbon fixation

The carbon cycle in the sea is instrumental in buffering increases in atmospheric CO_2, and therefore has a potential damping effect on climate change. The biological carbon pump engaged in long-term removal of carbon from shorter-term cycling, however, may not be able to keep pace with increased levels of DIC in the upper ocean, which result from a flux maintaining an atmosphere/ocean equilibrium for pCO_2 (Pahlow and Riebesell, 2000; Hopkinson and Vallino, 2005). Such a model predicts there will be a build up of CO_2 in the upper ocean, and pH will fall (Caldeira and Wickett, 2003).

Further, over millennia seawater acidification may be limited by negative feedback mechanisms of carbon remineralisation. Hypothetically, increased CO_2 abundance in seawater should drive increased primary production, and dead organic matter should sink to the seafloor and be buried in sediments. However, with other nutrients such as iron and phosphorus limiting cellular growth, this process may not happen fast enough to buffer the anticipated changes in pH in the upper ocean.

1.4.2 Deep sea and sub-seafloor sequestration

Civil engineers have proposed to mitigate the build up of CO_2 in the atmosphere and ocean by capturing the gas at locations such as power station exhausts and storing it in subterranean, including sub-seafloor, chambers such as those created by the removal of fossil fuels. Long term security, i.e.: prevention of steady or catastrophic leakage, must also be factored into any solution. Other proposals include injection of captured CO_2 into the deep ocean above the sea floor (Herzog *et al.*, 2000; Ormerod *et al.*, 2002). Thistle and co-workers (2005) demonstrated abyssal sediment metazoan infauna, which comprises the bulk of deep sea metazoan diversity and abundance, to be highly sensitive to CO_2 sequestered into deep ocean water.

1.4.3 Ocean nutrient fertilisation

As a result of atmosphere/ocean flux, driven primarily by photosynthetic carbon fixation, CO_2 is taken out of the atmosphere, and is therefore not available to act as a greenhouse gas. Artificially increasing the total global primary production accounted for by marine phytoplankton, particularly in high nutrient-low chlorophyll (HNLC) regions, has been proposed (Coale *et al.*, 1996). Since nutrients other than

carbon and nitrogen (phosphorus and iron especially) usually limit the growth of phytoplankton and cyanobacteria in the euphotic zone, addition of these elements should increase marine carbon fixation. Several large scale *in situ* iron-fertilisation experiments have been carried out, prominent examples of which include IronEX-I (Martin *et al.*, 1994), IronEX-II (Coale *et al.*, 1996), EisenEX (Gervais *et al.*, 2002), SOFEX-S (Buesseler *et al.*, 2004) and EIFEX (Hoffmann *et al.*, 2005). To be taken out of the medium-term carbon cycle this extra production of organic matter needs to sink via the biological carbon pump to the sea bed. There the carbon will be sequestered in sediments for geological ages. Whether this carbon sink is actually enlarged, and to what extent, by increased euphotic zone primary production is unproven (Gnanadesikan *et al.*, 2000). Other concerns about this carbon mitigation strategy surround the impacts of nutrient fertilisation on the various marine ecosystems affected (ecosystem responses to perturbation are reviewed in Gunderson, 2000).

The natural buffering capacity of the oceans to maintain pH operates on a timescale of 10^3 to 10^5 years (Boyle, 1997). For humanity's much shorter term interests these natural processes are not applicable. Potential strategies for engineering the mitigation of ocean acidification include direct sequestration of CO_2 into sub-seafloor geological recesses (Plasinskl *et al.*, 2007). Another strategy to reduce atmospheric pCO_2 involves injecting CO_2 directly into the deep ocean along with an alkali substance, such as magnesium hydroxide ($Mg(OH)_2$) or CO_3^{2-} from terrestrial limestone, to counter its effect on acidity (Kheshgi, 1995; Rau and Caldeira, 1999, 2002). CO_2 finding its way into the oceans either naturally, or indirectly or directly as a result of human activity, will drive down pH. Efforts to counter this fall in pH, rather than to prevent it, may involve either treatment with alkali, or augmenting primary production by the existing photosynthetic community to fix more inorganic carbon. That primary production in nutrient-rich waters was limited by the bioavailablility of iron was first suggested by Martin and Fitzwater (1988).

There are proposals for commercial ocean iron fertilisation (OIF) projects based on the newly-emerging carbon unit trading system. The California-based Planktos company plans straightforward kilotonne fertilisation of HNLC zones with haematite, while the New South Wales-based Ocean Nourishment Corporation plans kilotonne fertilisation of the Philippine Sea with urea (Young, 2007). Such plans are controversial in view of a dearth of evidence for their efficacy and side effects. In

March 2008 Planktos suspended its plans indefinitely in light of these concerns (News 21 Feb. 2008, *Nature* **451** page 879).

Mitigation strategies which themselves radically alter the environment have the potential to impact on the structure of bacterial communities in similar ways to ocean acidification; i.e. by those genetic mechanisms outlined above in Section 1.3.1. The chemical changes to the pelagic environment brought about by fertilisation, such as with Fe or P, may be locally more intense, and cause greater ecological impact, than the changes due to anthropogenic acidification. It is therefore important not only to ascertain the efficacy of nutrient fertilisation experiments for amplifying the ocean carbon sink, but also to try to ascertain the ecological impact of the experiments themselves. This study therefore aimed to show how naturally occurring pelagic bacterial communities would change in structure, resolved to the class and phylum level, as a result of adaptation to increased levels of Fe and P during a large-scale Fe and P addition experiment in the subtropical northeast Atlantic Ocean.

Iron is an essential element in all bacterial cells. It has a functional role in enzymes such as nitrogenase, involved in the fixation of N_2, and cytochromes and iron-sulphur proteins comprising the energy-generating electron transport systems of phototrophs. Genes encoding iron-containing proteins are highly conserved and deeply rooted in bacterial lineages. They likely originate in an era when the oceans were not oxic, and reduced iron was abundant in the sea (Munn, 2004).

The rise of oxygenic cyanobacteria gave birth to the oxidising marine environment which persists still. Iron is oxidised rapidly to an insoluble state, inaccessible to microorgnisms. Bioavailable, reduced iron exists in the surface ocean in picomolar concentrations, 99 % in the form of iron hydroxide ($Fe(OH)_3$) and tightly associated with organic matter. As with other solutes, iron concentration generally declines away from the continents, and rises with depth. C:Fe intracellular ratios of < 10,000:1 (Kirchman, 2000) mean there is competition for iron amongst microorganisms, and the element may be seen as scarce and limiting for growth. Fe-scavenging mechanisms have evolved. These include siderophores; extracellular reduced iron-chelating protein mesh structures, which are subsequently internalised. Some cells can internalise iron-rich siderophores produced by unrelated organisms. Photo-breakdown of siderophores increases the bioavailablility of iron, and this may be a major source of the element in tropical and subtropical seas. Viral lysis and protistan grazing of bacteria releases iron contained in porphyrins, which may be

internalised by phytoplankton and be their major source of the element during coastal blooms (Kirchman, 2000; Munn, 2004). A generalised oceanic iron cycle is represented in Figure 1.10.

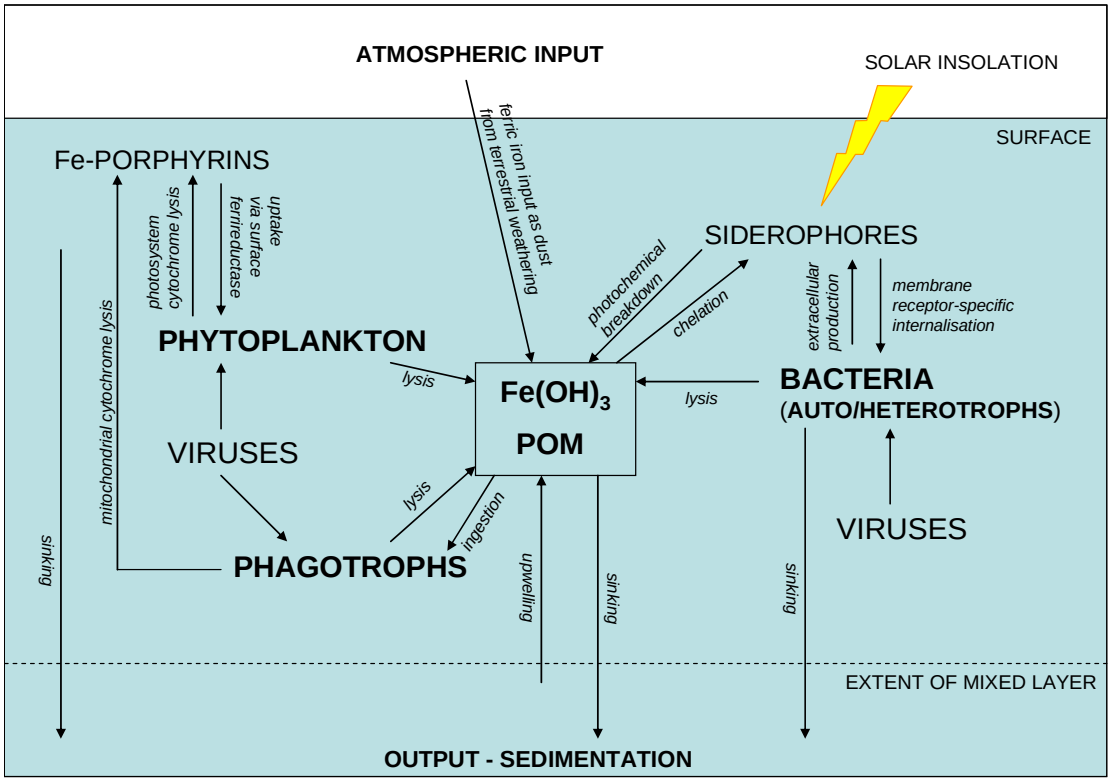


Figure 1.10. Generalised representation of the cycling of iron in the ocean (inspired by Munn, 2004).

In low-Fe waters competition for iron between microorganisms may be a basic driver of the community structure (Hutchins *et al.*, 2001b). Competition between cyanobacteria, heterotrophs and phytoplankton for available iron in its different chemical states may lead to niche partitioning (Thingstad, 2000). Some diatoms use domoic acid to bind and store iron when supply outweighs cellular demand. Some bacteria use ferritin storage proteins in a similar way, although this has not been demonstrated in marine species. Forty percent of reduced iron available in the surface ocean is used by heterotrophic bacteria (Tortell *et al.*, 1996).

The six most abundant elements in microorganisms, are H, C, O, N, P and S (Kirchman, 2000). O, H and S are not normally in short supply in the ocean. Inorganic carbon is readily available to photoautotrophs, whereas organic carbon has been shown to limit heterotrophic growth (Thingstad, 2000). N, P and Fe are each potentially a factor limiting bacterial growth in the photic zone. Experimental

evidence showing which element limits bacterioplankton growth, has yielded different results at different times, and in different ocean locations. Growth limitation due to a lack of organic carbon has been demonstrated in the oligotrophic tropical open Pacific Ocean (Kirchman and Rich, 1997), and in the Caribbean Sea (Rivkin and Anderson, 1997). A combined shortfall of C and N was demonstrated for the sub-Arctic Ocean. N and P have each been shown to be the limiting element for the growth of marine microbial biomass (Elser *et al.*, 1995; Pomeroy *et al.*, 1995; Cotner *et al.*, 1997; Thingstad *et al.*, 1998). Organic carbon added to oligotrophic environments usually stimulates greater microbial growth. However it does not necessarily follow that following eutrophication this will cease to be the case. Any of several elements may still be limiting in apparently nutrient-replete conditions. It should also be borne in mind, as noted by Thingstad (2000), that any element may be deficient in the environment, and yet not be the growth-limiting factor.

Iron has been shown to be the nutrient limiting primary production, and therefore indirectly, if not directly, bacterial growth, at least in the Southern Ocean (Pakulski *et al.*, 1996).

Phosphorus is the fifth most abundant element in bacteria, crucial for metabolism, and also as a structural component of nucleic acids and membrane phospholipids. Thingstad (2000) lists P as the most common bacterial growth-limiting nutrient after C and N. The C:P content of pelagic heterotrophic bacteria averages around 50, contrasting with the figure of 108 for phytoplankton (Kirchman, 2000). The higher concentration of P in bacteria is mostly inorganic orthophosphate. This ratio in bacteria is variable however, and reflects that of the aqueous environment. It also is a reflection of small cell volume and an increased membrane: cytoplasm ratio.

In the ocean phosphorus is rapidly cycled between inorganic orthophosphate, available to phototrophs, and organic phosphates, which are degraded by heterotrophs, but also by extracellular autolytic enzymes. Unlike C and N, this process occurs largely within the euphotic zone so there is little loss through sinking. For this reason phosphorus traditionally was considered to be in plentiful supply for primary and bacterial production. However P is a candidate limiting nutrient in HNLC waters. Zweifel and colleagues (1993) reported P-limitation of bacterial growth in predator-free microcosms in the Mediterranean Sea, while Pomeroy and co-workers (Pomeroy *et al.*, 1995) demonstrated P to be low enough in the Gulf of Mexico to limit bacterial productivity. One hypothesis states that growth of microbial blooms

stimulated by iron fertilisation in oligotrophic waters, with phosphate concentrations commonly < 10 nM, will be checked by phosphorus supply once the iron supply has been used up (Mills *et al.*, 2004).

Pelagic bacterial communities are fundamentally important both in biogeochemical cycling and in ecology. *Cyanobacteria* may be responsible for over half of primary production in some ocean regions. Heterotrophic bacteria cycle and recycle organic matter via multiple pathways to maintain energy flow through food webs. The "microbial loop" is the cumulative name for these pathways (Pomeroy *et al.*, 2007). Bacterial cells themselves fuel oceanic energy transfer chains via side-on (viruses) and top-down (phagotrophs) predation. Therefore changes to bacterial community structure potentially have knock-on effects throughout entire food webs. If the theory is correct, that competition for growth-limiting nutrient elements underlies this community structure, then it follows that fertilisation with these same elements will profoundly alter microbial community composition. However there has been little investigation into these potential effects. A few studies have analysed bacterial community changes during artificial iron-stimulated phytoplankton blooms at the broadest taxonomic resolution. No such report into the effects of phosphorus fertilisation on community structure could be located.

Arrieta and colleagues (2004) working on the EisenEx experiment found Fe stimulated, and its absence therefore limited, bacterial abundance and production. However, using T-RFLP they reported "no major changes in the phylogenetic composition of the bacterioplankton community." This suggests the existing community phenotypically aclimatised, i.e. the community metagenome (the combined genomic variation within all the organisms comprising the community) possessed enough Fe-related genetic variation, evolved for other reasons, to cope with the environmental perturbation through a change in gene expression. This would not leave any change in the metagenome recordable within the 16S rRNA gene composition of the community. While alluding to changes in metagenomic composition, it is important to remember that all the microbial functions alluded to, including those affected by nutrient limitation, are performed, not by an amorphous mass but, by specific organisms with limited functions.

Hutchins *et al*. (2001a) used DGGE (universal primers 338f-GC/517r) to assess bacterial community change over seven days in three different shipboard incubation experiments. Although as expected the phytoplankton bloomed and was

dominated by large diatoms, and the resulting bacterial biomass rose, they found scant change in DGGE profiles. A slight increase in the dominance of *gammaproteobacteria* was noted in HNLC Fe-treated water from the eastern tropical Pacific Ocean and the sub-Arctic Ocean, and no change at all in a community from polar Southern Ocean water. Some limitations to these findings stem from the analysis of variation in DGGE band intensity which may result from PCR amplification bias, and therefore not be as reliable an indicator for abundance of cells, as straightforward band presence is for cell presence. Additionally the use of a single set of universal *Bacteria* primers, and the experiments' short duration, allowing little time for change in community structure to become apparent, especially when analysing DNA, rather than RNA.

The study presented here set out to investigate effects on the phylogenetic composition of the bacterial community during a joint OIF/OPF experiment in the subtropical northeast Atlantic Ocean, and to do so at a taxonomic resolution not previously achieved.

1.5 Aims of this study

DGGE performed with universal *Bacteria* primers for the 16S rRNA gene has been shown to suffer from two main limitations. Firstly the length of each gel limits the number of bands which may be resolved and subsequently excised for further analysis. Secondly universal primers show amplification bias. Primers used to amplify environmental DNA showed communities constructed almost entirely of *alphaproteobacteria* and *Bacteroidetes*, as shown in Figure 4.9d and Table 4.4. To address these limitations of universal *Bacteria* primers eight putative group-specific primer sets were designed *in silico.* These aimed to cover specifically the alpha, beta, gamma and delta subgroups of the *Proteobacteria*, the *Cyanobacteria*, the *Bacteroidetes*, the *Firmicutes* and the *Planctomycetes*. Clone libraries were made to establish the group-specificity and group-coverage of these eight group-specific primer sets.

It was then aimed to use the group-specific primers in PCR-DGGE experiments on environmental DNA, to test the following hypotheses:

Hypothesis I: Marine bacterial diversity varies as a consequence of geographical isolation

Hypothesis II: Marine bacterial diversity varies in relation to the level of seawater acidity

Hypothesis III: Marine bacterial diversity varies as a consequence of iron availability in seawater

Hypothesis IV: Marine bacterial diversity varies as a consequence of phosphate availability in seawater

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Oligonucleotides

All DNA oligonucleotides were purchased from MWG-Biotech (Ebersberg, Germany) at 0.2 μmol scale of synthesis. Each oligonucleotide was resuspended in water to a final concentration of either 500 μM or 250 μM depending on the total mass of DNA supplied. These stocks were diluted 50-fold or 25-fold respectively in water to give PCR-ready oligonucleotide solutions of 10 μM. All oligonucleotides were stored at -20^{o}C.

Table 2.1 Details of oligonucleotides used to prime PCRs in this study

Name	***Specificity***	***Sequence (5'-3')***	***Reference***
28f	*alphaproteobacteria*	ARCGAACGCTGGCGGCA	(Ashelford *et al.*, 2002)
684r		TACGAATTTYACCTCTACA	This study
359f	*betaproteobacteria*	GGGGAATTTTGGACAATGGG	(Ashelford *et al.*, 2002)
682r		ACGCATTTCACTGCTACACG	(Ashelford *et al.*, 2002)
395f	*gammaproteobacteria*	CMATGCCGCGTGTGTGAA	This study
871r		ACTCCCCAGGCGGTCDACTTA	This study
96f	*deltaproteobacteria*	AGTARAGYGGCGCACGGGTG	This study
495r		TTAGCCGGYGCTTCCT	This study
555f	*Bacteroidetes*	CCGGAWTYATTGGGTTTAAAGG	This study
968r		GGTAAGGTTCCTCGCGTA	This study
361f	*Cyanobacteria*	GGAATTTTCCGCAATGGG	This study
785(cya)r		GACTACWGGGGTATCTAATCC	This study
350f	*Firmicutes*	GGCAGCAGTRGGGAATCTTC	This study
814r		ACACYTAGYACTCATCGTTT	This study
352f	*Planctomycetes*	GGCTGCAGTCGAGRATCT	This study
920r		TGTGTGAGCCCCCGTCAA	This study
9bfm	*Bacteria*	GAGTTTGATYHTGGCTCAG	This study
1512uR	all prokaryotes	ACGGHTACCTTGTTACGACTT	(Weisburg *et al.*, 1991)
M13F	M13 region of vector	TGTAAAACGACGGCCAGT	
M13R		GGAAACAGCTATGACCATG	
GC-clamp	n/a	CGCCCGCCGCGCGCGGCGGGCGG GGCGGGGGCACGGGGGG	(Muyzer *et al.*, 1993)

2.1.2 Chemicals, reagents and laboratory consumables

General laboratory chemicals were obtained from Sigma-Aldrich (St. Louis, MO, USA) or VWR (Lutterworth, UK) and were of analytical grade or higher. Plastic ware and general laboratory consumables were obtained from Fisher-Scientific (Leicester, UK), Bibby-Sterilin (Stone, UK) or Sarstedt (Nümbrecht, Germany). All other reagents and consumables suppliers are listed in the text where appropriate. Ultrapure 18 MOhm double-filtered water (H_2O^{MQ}) was obtained from a Synergy 185

and Elix®3 water purifier (Millipore, Molsheim, France), and autoclaved at 121°C and 15 psi for 20 minutes prior to use.

GoTaq® DNA polymerase and its buffer, 25 mM $MgCl_2$, and bovine serum albumin (BSA) were supplied by Promega (Madison, WI, USA), while 100 mM dNTPs were supplied by Invitrogen (Paisley, UK), and mixed and diluted with H_2O^{MQ} to 2 mM for use in the PCR. PCR consumables such as pipette tips, 0.2 mL, 0.5 mL and 1.5 mL centrifuge tubes, as well as 96 x 0.2 mL-well plates, were supplied by Sarstedt.

2.1.3 Commonly used solutions

SET Lysis Buffer

0.75 M sucrose, 40 mM EDTA, 50 mM Tris-HCl (pH 9), store at room temperature

0.5 M ethylenediamine tetraacetic acid (EDTA)

186.1g EDTA.$2H_2O$ (disodium salt) dissolved in 1L H_2O^{MQ}, adjusted to pH 8 using NaOH

10% SDS

10 g sodium dodecyl sulphate (SDS) dissolved in 100 mL H_2O^{MQ}

50 x TAE

242 g Tris, 57.1 mL acetic acid (glacial), 100 mL 0.5 M EDTA, dissolved in 1 L H_2O^{MQ}

2.1.4 Growth media

Media for the growth of bacteria were prepared as follows:

Luria-Bertani (LB) Agar (solid)

15 g Bactoagar®, 10 g tryptone, 5 g yeast extract and 10 g NaCl were dissolved in 1 L H_2O^{MQ} and autoclaved. Upon cooling to 45°C, 300 µL X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside [Promega]), and 2.5 mL 20 mg/mL ampicilin (Invitrogen) were added. Aseptic technique was employed throughout.

SOC (liquid)

20 g Bacto-tryptone®, 5 g yeast extract, 0.5 g NaCl, 0.186 g KCl, 5 mL 2 M $MgCl_2$, 20 mL 1 M glucose and 10 mL 1 M $MgSO_4$ were dissolved in 1 L H_2O^{MQ}.

2.1.5 Materials specific for clone library construction

Clone libraries were made by transforming *E. coli* DH5α Library Efficiency® competent cells (Invitrogen) stored at -80°C, with the plasmid pGEM®-T Easy Vector System I (Promega) and the supplied ligation buffer and T4 DNA ligase, stored at -20°C.

2.1.6 Materials specific for DGGE

0% Denaturant Acrylamide Gel Solution

20 mL 40% bis acrylamide and 2 mL 50 x TAE, were dissolved in H_2O^{MQ} to a total volume of 100 mL.

80% Denaturant Acrylamide Gel Solution

20 mL 40% bis acrylamide, 2 mL 50 x TAE, 33.6 g urea, and 32 mL deionised formamide (Fluka, Buchs, Switzerland), were dissolved in H_2O^{MQ} to a total volume of 100 mL. Denaturing acrylamide solutions containing varying concentrations of urea and formamide were obtained by combining the above 0 % and 80 % solutions according to the different ratios shown in Table 2.2.

10 x DGGE Gradient Tracking Dye

0.05 g bromophenol blue, and 0.05 g xylene cyanole, were dissolved in 1 x TAE to a total volume of 10 mL.

6 x DGGE Loading Dye

0.025 g bromophenol blue, 0.025 g xylene cyanole, and 5 mL glycerol, were dissolved in H_2O^{MQ} to a total volume of 10 mL.

2.1.7 Materials specific for DNA sequencing

BigDye® 3.1 Ready Reaction Mix and its buffer were purchased from ABI Applied Biosystems (Warrington, UK).

2.2 General molecular biology methods

2.2.1 Nucleic acid extraction

Cells filtered from *in situ* seawater mesocosms (Figure 4.1) onto Sterivex® cartridges (pore size 0.22 μm) were resuspended in 1.6 mL SET lysis buffer (0.75 M sucrose, 40 mM EDTA, 50 mM Tris-HCl). The suspension was incubated for 30 min at 37°C in a rotary oven (Hybaid) with lysozyme (1 mg/mL), followed by further incubation for 2 hrs at 55°C with 200 μL SDS (10 % w/v) and 55 μL proteinase K (20 mg/mL). The lysate was withdrawn from the Sterivex® cartridges and the cartridges rinsed of lysate residue with 1 mL SET buffer. The lysate was extracted twice with 2 mL phenol:chloroform:isoamyl alcohol (25:24:1) at pH 8, and once with phenol:chloroform (25:24). Aqueous phase DNA was precipitatied overnight with 1 mL 7.5 M ammonium acetate, 5 μL glycogen (Roche, Burgess Hill, UK) and 6 mL of absolute ethanol (100%) and pelleted by centrifugation at 14,000 x g

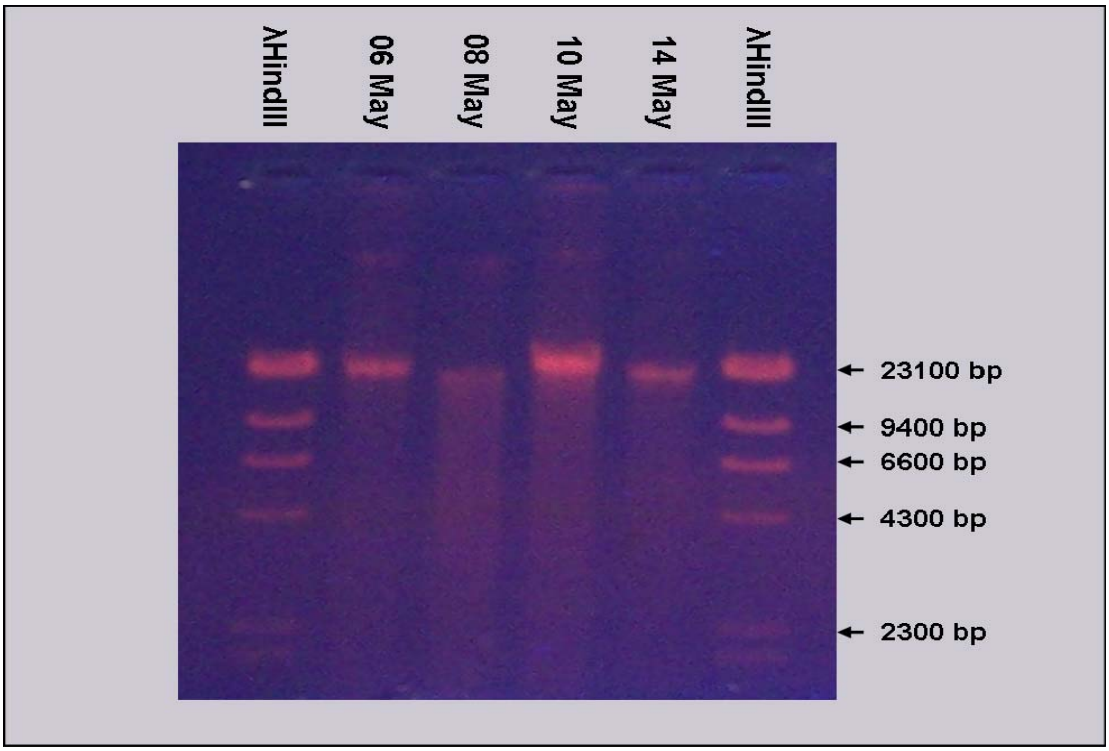


Figure 2.1. DNA extracted from 4 L seawater sampled from a mesocosm moored in a southern Norwegian fjord on four different days in 2006. Seawater was first filtered through a 0.22 μm Sterivex® filter, and DNA extracted using the protocol detailed in Section 2.2.1. 9 μL DNA per sample was run on a 1.0 % agarose gel for 1 hour at 100 volts.

for 30 minutes. DNA from the aqueous phase was washed twice in 2 mL of ethanol (80%) by centrifugation at 14,000 x g for 20 minutes. The DNA was resuspended in 200 µL sterilised H_2O^{MQ}. A sample of 9 µL of the DNA suspension was run on a 1.0 % agarose gel in 1 x TAE, as shown in Figure 2.1, to check for quantity and quality. The remainder was stored at -20 ^oC.

2.2.2 Polymerase chain reaction

The following basic 25 µL PCR reaction was performed throughout these experiments, except where otherwise stated:

5 x Green GoTaq® Flexi buffer		2.5 µL
$MgCl_2$ (25 mM)		2.0 µL
dNTPs (2mM)		2.0 µL
Forward primer (10 µM)		0.3 µL
Reverse primer (10 µM)		0.3 µL
GoTaq® DNA polymerase (5u/µL)		0.5 µL
BSA (2 mg/mL)		2.0 µL
Template DNA (~10 ng/µL)		4.0 µL
H_2O^{MQ}	(to total)	25.0 µL

General PCR cycling parameters:

An initial denaturation step of 96^oC for 2 minutes was followed by 35 cycles of denaturation at 96^oC for 1 minute, an annealing temperature (calculated for each primer pair, see Table 3.1) for 1 minute, and template-specific primer elongation at 74^oC for 1 minute per 1000 bp of amplified product length. A final elongation step consisted of 5 minutes at 74^oC. PCR products were stored at -20^oC until needed.

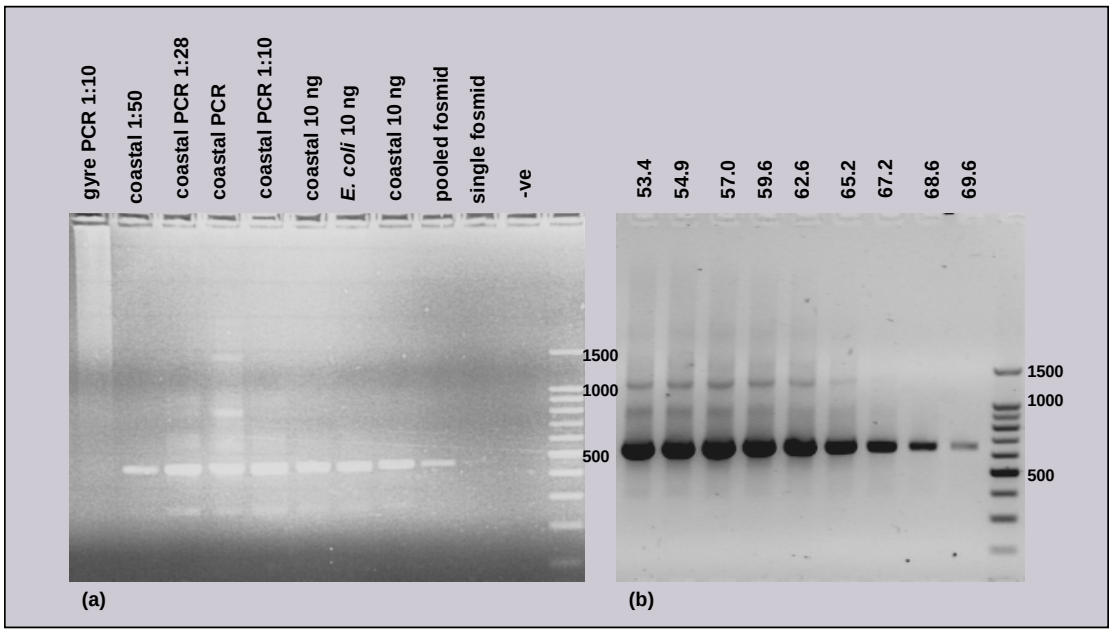


Figure 2.2. Effect on size and concentration of PCR product of varying template concentration and annealing temperature. **(a)** Template DNA concentration with other parameters constant. Samples are environmental DNA from a Norwegian fjord (coastal), DNA from an *E. coli* culture, DNA from a fosmid clone library, and various dilutions of PCR product using Norwegian fjord (coastal), and North Atlantic oligotrophic (gyre) DNA. **(b)** Annealing temperature (^{o}C) variation, with other PCR parameters, including template concentration, constant. Primers used in PCR were 395f-GC/871r (a) and 28f/684r-GC (b) detailed in Table 2.1.

2.2.3 Agarose gel electrophoresis

PCR products and the products of environmental DNA extraction were size-fractionated by electrophoresis on 1.8% and 1.0% agarose gels, respectively. For a 1.8% gel 1.8 g agarose was dissolved in 100 mL 1 x TAE by heating. After cooling ethidium bromide was added to a final concentration of 0.5 μg mL^{-1}, and the solution poured into a gel mold with comb. Once set the comb was removed and the gel placed in an electrophoresis tank, submerged in 1 x TAE running buffer. Either 5 μL PCR product, or 9 μL extracted environmental DNA, was added to each well, and run alongside 3.5 μL 100 bp ladder (markers at 100, 200…1000, and 1500 bp [Promega]), at 100 V (20 V per gel cm) for 30 minutes (for example see Figures 2.2 and 2.3). Lambda Hind III DNA digest (markers at 2.0, 2.3, 4.3, 6.6, 9.4 and 23.1 kbp) was run alongside extracted environmental DNA on a 1% gel (for example see Figure 2.1). Gels were visualised and photographed in a GeneGenius® UV gel documentation system (Synoptics, Cambridge, UK).

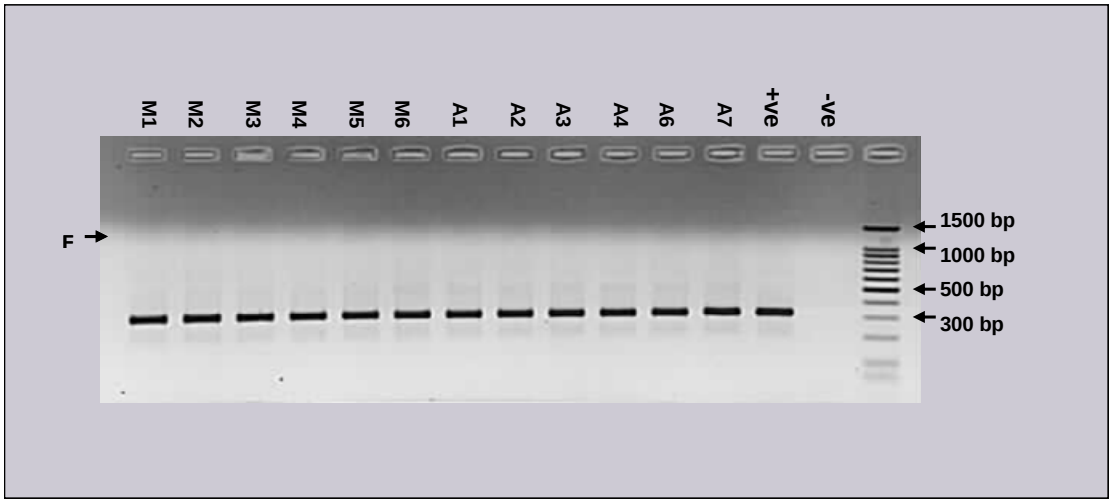


Figure 2.3. PCR products (5 µL) obtained with DGGE primers 518f-GC / 785r using PCR product, made using *gammaproteobacteria*-specific primers, as template, including positive (+ve) control using cultured *Alteromonas* sp. genomic DNA as template, and a negative (-ve) control containing no DNA template. Samples were run for 30 minutes on 1.0 % agarose gel at 10 volts.cm^{-1}, alongside 100 bp ladder (3.5 µL, Promega), and stained with ethidium bromide. F, front of ethidium bromide which, due to positive charge, migrates counter to DNA during electrophoresis. Environmental template DNA was extracted from a Norwegian fjord mesocosm (M) and the oligotrophic North Atlantic Ocean (A).

2.3 Methods specific for clone library construction and analysis

2.3.1 Vector ligation

PCR products (a mixture of 16S rRNA gene fragments representative of the diversity within the original environmental DNA sample) were ligated into plasmid vectors (pGEM®-T Easy, Promega) ready for transformation of competent *E. coli* cells according to the following protocol:

Between 1 and 3 µL of PCR product, chosen according to a visual estimate of band intensity, was immediately mixed with 1 µL T4 DNA ligase (3 Weiss units), 1 µL plasmid vector, 5 µL 2 x ligation buffer and sterilised H_2O^{MQ} to a total volume of 10 µL. Ligation of PCR products into plasmids proceeded at 4°C overnight, and the ligated plasmids were stored at -20°C.

2.3.2 Transformation of competent cells

Recombinant plasmids were introduced into Library-Efficiency® DH5α™ competent *E. coli* cells (Invitrogen) by the following transformation protocol:

2 µL ligation mix (0.5 ng DNA/µL) was added to 50 µL freshly thawed competent DH5α™ cells, and mixed on ice for 30 minutes. The mixture was given a 2 minute heat shock at 42°C and returned to ice for a further 2 minutes. 450 µL SOC medium (containing added 2M $MgCl_2$ at 5 mL/L and 1M glucose at 20 mL/L) was added and then the cells were incubated in a shaking incubator at 37°C for 90 minutes. 10% of the potentially transformed cells were spread on an LB agar (Section 2.1.4) plate using 50 µL of the cell solution. 90% of the potentially transformed cells were plated using a cell solution concentrated by centrifugation of the remaining cell solution at 10,000 x g for 1 minute, removal of 300 µL supernatant, and using 100 µL of the resuspended cell pellet. Plates were incubated at 37°C overnight.

2.3.3 Identification of recombinants

From the spread plates (Section 2.3.2) colonies of transformed cells, containing environmental cloned 16S rRNA gene fragment inserts, were selected. Non-transformed cells will not grow in the antibiotic-containing media. Transformed cells are distinguishable by the colour of the resultant colony, which indicates whether or not the vector/insert ligation was successful. The pGEM®-T Easy vector contains a gene encoding β–galactosidase which when expressed cleaves its pigment-ligated substrate (X-Gal) in the media resulting in a visible blue pigment, i.e.: colonies appear

blue. The pGEM®-T Easy insert-ligation site lies within this gene and successful insertion (ligation) of extraneous DNA (i.e. PCR product) prevents transcription of the β–galactosidase gene. The X-Gal substrate therefore remains unaltered and colonies appear white.

White colonies were picked from each "group-specific" spread plate onto a fresh Amp^+, $X\text{-}Gal^+$ LB plate in a grid formation. Colonies were again incubated overnight at 37°C after which resultant colonies were colour-checked again to confirm transformation. The plate was sealed with parafilm and stored at 4°C. Library plates were checked weekly for excess condensed evaporate from the LB media which can cause DNA from separate colonies to mix.

2.3.4 "Colony PCR"

Being derived from a single transformed cell, the cells of each colony (clone) should all contain identical 16S rDNA inserts. Clone libraries constructed from PCR products obtained with each putative group-specific primer pair were screened to ascertain: (i) the diversity of sequences, and abundance of each unique sequence, within the library and (ii) that all ribosomal DNA inserts cluster within the correct group. An aliquot of each colony was suspended in 30 µL H_2O^{MQ}, incubated at 98°C for 15 minutes, centrifuged at 10,000 g for 1 minute, and 0.5 µL supernatant taken for colony PCR template. Colony PCR with M13 primers was used to amplify vector inserts, the primers binding several bases either side of the site of the cloned insert. The size of each colony PCR product was assessed by electrophoresis in a 1.8% agarose gel (Figure 2.4).

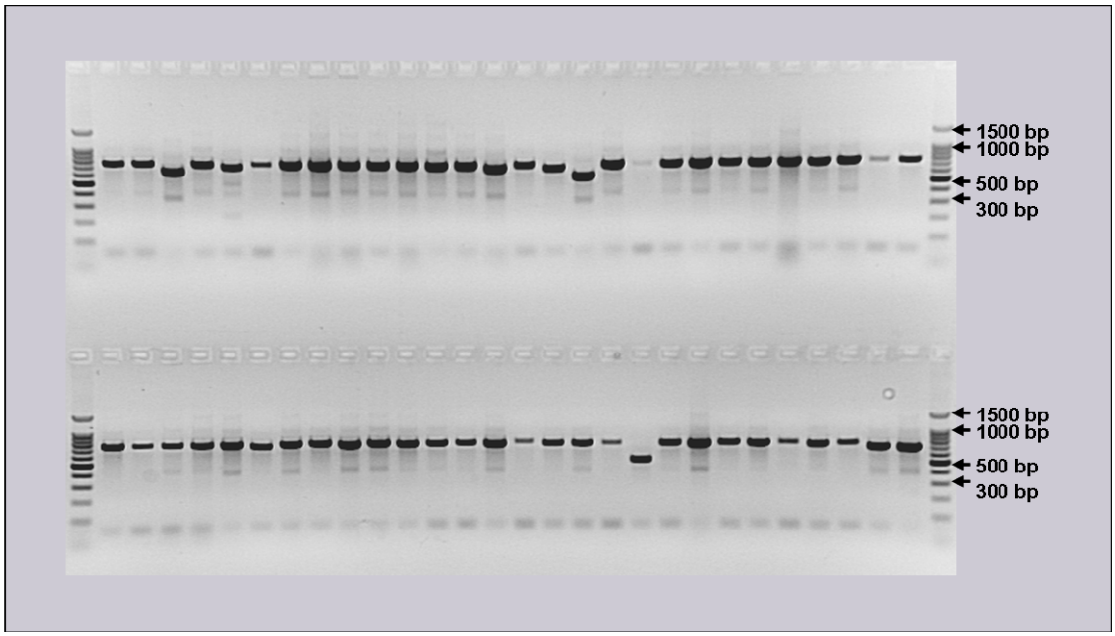


Figure 2.4. General example of "colony PCR" products run on a 1.8 % agarose gel for 30 minutes at 100 volts (20 V per gel cm). *Firmicutes*-specific 16S rRNA gene fragments were expected to be 524 bp in length, based on the the 16S rRNA gene reference sequence from *E. coli*. M13 primers amplify additional 233 bp of vector sequence. Successful *Firmicutes*-specific colony PCR products are ~ 757 bp. There has been some amplification resulting in colony PCR products shorter than 757 bp, possibly due to amplification of non-ligated vector sequence, or perhaps due to length polymorphism of the 16S rRNA gene within the *Firmicutes* sample population. The feint bands towards the bottom of the gel are unused primers and primer-dimers.

2.3.5 Sequencing cloned inserts

2.3.5.1 "Colony PCR" product clean-up

Each 2.5 µL colony PCR product (~50 ng DNA) was cleaned, prior to sequencing, with 1 µL ExoSAP-IT® (USB Corp., Cleveland, USA) by mixing and incubating at 37°C for 15 minutes. The ExoSAP-IT® enzymes were then deactivated by further incubating at 80°C for 15 minutes. ExoSAP-IT® consists of Exonuclease I which degrades single-stranded DNA such as residual primers, and shrimp alkaline phosphatase which hydrolyses unused deoxynucleotide triphosphates (dNTPs), both of which can interfere with subsequent PCR and sequencing reactions.

2.3.5.2 PCR sequencing using chain terminating ddNTPs

In order to generate the sequence of a fragment of DNA produced by PCR an array of sequences is produced varying in length by one base pair. Sequencing PCR employs a single sequence-specific primer (forward or reverse) and dideoxynucleotide triphosphates (ddNTPs) in addition to dNTPs which terminate DNA elongation when incorporated. A 1000 bp template thus yields an array of sequences from one to 1000 bp differing in length by one base pair. This array is the substrate for reading the sequence of bases using the Sanger-Coulson method (Sanger *et al.*, 1977).

The following PCR was employed to generate a sequence array for each colony PCR product, from which a sequence is produced on a sequence analyser:

Sequencing buffer	3.0 µL
Forward primer (M13F, 10 µM)	0.2 µL
Ready Reaction Mix (BigDye® 3.1)	1.0 µL
H_2O^{MQ}	12.3 µL
Template DNA (14 ng/µL)	3.5 µL
Total	*20.0 µL*

An initial denaturation step of 96°C for 1 minute was followed by 25 cycles of denaturation at 96°C for 10 seconds, 5 seconds annealing at 50°C, and primer extension at 60°C for 4 minutes. Reaction products were stored at 4°C until needed.

2.3.5.3 Ethanol clean up

Products of sequencing PCR were cleaned by ethanol precipitation prior to electrophoresis on a sequence analyser according to the following protocol:

To each 20 µL sample (PCR cycle sequencing product) was added 5 µL 125 mM EDTA and 60 µL 100% ethanol. Samples were mixed, incubated for 15 minutes at room temperature, and centrifuged at 3000 g for 30 minutes. The pellet was washed with 60 µL 70% ethanol and centrifuged at 1650 g for 15 minutes. DNA was air-dried and resuspended in 15 µL HiDi™ formamide (ABI), and stored at -20 °C prior to electrophoresis in a sequence analyser.

2.3.5.4 Sequencing

Sequencing of cloned inserts, using the Sanger-Coulson dideoxynucleotide incorporation method, was performed by Paul Pickerell of PML using a Prism®3100 Genetic Analyser (ABI) and ABI's Sequencing Analysis Software version 5.1 to check sequence quality and save sequences. The Chromas Pro software (http://www.technelysium.com.au/ChromasPro) was used to align and remove primer and vector sections of sequence, and also to identify and correct misread bases.

Some PCR products, such as excised DGGE bands, were treated with ExoSAP-IT® according to the manufacturer's instructions and used directly for sequence analysis. Generally only one strand of the DNA fragments was sequenced, this being initiated by either the forward or reverse primer. For the most part this proved to be sufficient for the taxonomic identification of the cloned 16S rRNA gene fragments. Both DNA strands (originating from forward and reverse primers) were sequenced in those cases where the base sequence of a single DNA molecule could not be read with confidence.

2.3.6 Sequence analysis

2.3.6.1 Sequence identification using the BLASTn algorithm and public sequence repositories

Edited partial 16S rRNA gene sequences were aligned in, and phylogenetically classified by, the Greengenes® 16S rRNA gene database (http://greengenes.lbl.gov). Alternatively sequences were imported into the NCBI's website and compared using the Basic Local Alignment Search Tool for nucleotide sequences (BLASTn) algorithm with all existing homologous sequences in the Genbank (http://www.ncbi.nlm.nih.gov/genbank), EMBL (http://www.ebi.ac.uk/embl/index) and RDP-II (http://rdp.cme.msu.edu) public repositories for ribosomal DNA and RNA sequences.

Taxonomic and phylogenetic information was recorded in Excel (Microsoft) spreadsheets, and then in the sequence information tables throughout the study.

2.4 Methods specific for DGGE

2.4.1 Gel casting and electrophoresis

The mixture of 16S rRNA gene sequences in the PCR products were separated by DGGE using a Universal D-Code® Mutation Detection System (Bio-Rad, Hercules, CA, USA). Eight percent [w/v] polyacrylamide solutions containing 0% and 80% denaturants were made (where 100% denaturant contains 7M urea and 40% formamide). The required range of denaturant concentration across each gel was optimised for each group-specific PCR product, by trial and error starting with a range of 30% to 70%, with the maximum and minimum denaturant concentration solutions made according to Table 2.2.

Table 2.2 Volumes (mL) needed of 80% and 0% stock denaturant acrylamide solutions to make varying gradients of denaturant acrylamide solutions ranging from 20% to 70%.

	Final concentration of denaturants in solution									
Stock	*20%*	*25%*	*30%*	*35%*	*40%*	*45%*	*50%*	*55%*	*60%*	*70%*
80%	3.75	4.7	5.6	6.5	7.5	8.5	9.4	10.3	11.25	13.0
80%	3.0	3.8	4.5	5.2	6.0	6.8	7.5	8.2	9.0	10.4
0%	11.25	10.3	9.4	8.5	7.5	6.5	5.6	4.7	3.75	2.0
0%	9.0	8.2	7.5	6.8	6.0	5.2	4.5	3.8	3.0	1.6

Unshaded figures are for a gel of dimensions 16 cm x 16 cm x 1.0 mm, shaded figures are for a gel of 16 cm x 16 cm x 0.75 mm

Polymerising agents (18 µL TEMED and 120 µL APS) were added to each solution, and 120 µL 10 x DGGE gradient tracking dye, to help visualise the denaturing gradient, was added to the solution containing the higher concentration of denaturants. The gel was cast using a Model 475 Gradient Delivery System (Bio-Rad) between glass plates separated by 1 mm plastic spacers, both having been cleaned thoroughly with 100% industrial methylated spirit (IMS). The number of wells was restricted to 16 to allow for greater sample volumes to be loaded in each, and to give a wider profile (track) front facilitating band resolution. Following polymerisation for 2 hours gels were immersed in 1 x TAE buffer at 60^{o}C and allowed at least an hour to equate to temperature. Wells were washed with 1 x TAE buffer to remove excess polymerised gel and denaturants. 15 µL 6 x DGGE loading dye was added to 45 µL final round PCR product (~500 ng DNA) to increase the sample's density and aid visualisation of electrophoresis, mixed, and a total of 55 µL added to each well. After returning the buffer to 60^{o}C, electrophoresis was run for 1000 volt hours, most commonly at 60V for 16 hours 40 minutes.

2.4.2 Staining and visualisation

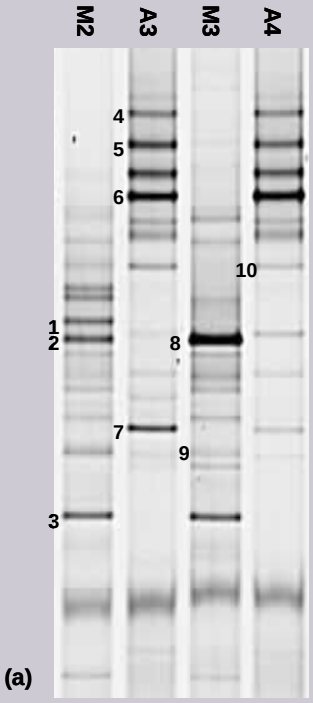
Following electrophoresis gels were stained in a shaking bath of ethidium bromide (1.0 μg/mL in H_2O^{MQ}) for 15 minutes, and then rinsed in H_2O^{MQ} for 30 minutes. Images were captured using a UV-transilluminator and camera gel documentation system (Syngene, Cambridge, UK) and saved as Windows bitmap (*.bmp*) format files.

2.4.3 Excision and sequence analysis of bands

Following gel image recording, the gel itself was placed on a benchtop UV-transilluminator and individual bands excised, using a sterile scalpel blade, under long wavelength UV light. Multiples of eight bands were excised and each placed in 30 μL H_2O^{MQ} at 4°C overnight, allowing diffusion of DNA out of the gel matrix into the aqueous solvent, facilitating its use in further molecular analysis.

The 16S rRNA gene fragments from the excised bands were reamplified by PCR (for protocol see Section 2.2.2) using the same primers used to generate the DGGE samples, but without a GC-clamp in order to facilitate subsequent sequencing. Template for this reamplification PCR was 5 μL of solution of resuspended DNA from excised bands.

Sequencing reamplified 16S rRNA gene fragments from excised bands from gels used the same protocol as that for sequencing fragments from a colony PCR (Sections 2.3.5.2 – 2.3.5.4). The same single forward primers were used in the sequencing PCR as for nested PCR in DGGE sample production; i.e.: 341f for *Bacteria*, and *Alphaproteobacteria*-specific gels, 518f for *Betaproteobacteria*-, *Gammaproteobacteria*-, *Cyanobacteria*-, *Firmicutes*-, and *Planctomycetes*-specific gels, and 555f for *Bacteroidetes*-specific gels. Unlike cloned sequences, re-amplified band sequences were occasionally undecipherable owing to the coincidental co-migration, and thus amplification, of additional but different DNA molecules in the gel. Such sequences were not usually recorded, but where the bands were more clearly resolved on DGGE gels of a different gradient range, the sequences were subsequently successfully recorded.



(a)

(b)

Band	Closest hit	Closest cultured isolate
1	Gammaproteobacteria16S rRNA clone ACE-1	Pseudomonas rhodesiae
2	Pseudomonas 16S rRNA clone F4C67	Azotobacter armeniacus
3	Gammaproteobacteria 16S rRNA cloneT31-155	Pseudomonas sp. 7049
4	Gammaproteobacteria 16S rRNA clone Surf 1.29	None
5	Gammaproteobacteria 16S rRNA clone JL-ESNP-I14	None
6	Gammaproteobacteria 16S rRNA clone SPOTSAUG01_5m25	None
7	Gammaproteobacteria 16S rRNA clone UST991130-003	Pseudomonas sp. F497
8	Pseudomonas 16S rRNA clone F4C67	Marinobacter sp. WMQ_A
9	Gammaproteobacteria 16S rRNA clone ARKDMS-9	Pseudomonas sp. 6-SLY
10	Alteromonadaceae 16S rRNA clone ESP450-K6V-55	Alteromonas sp. SDDC2

Figure 2.5. A DGGE gel, of *Gammaproteobacteria*-specific PCR product, amplified from seawater taken from the North Atlantic gyre (A) and a Norwegian fjord mesocosm (M), showing 16S rRNA gene bands and their corresponding taxonomic identity taken from their sequences. Sample information is in Table 2.3 (a). Typical information yielded from a BLASTn search of sequences obtained from DGGE bands, from which identity can be ascertained. *Azotobacter armeniacus* is a pelagic, aerobic, N_2-fixing *Gammaproteobacterium*. Although bands 2 and 8 yield identical closest hits they differ in the identity of their closest listed homologous isolates. Variation in the order of BLASTn results may stem from single nucleotide differences altering the phylogenetic positions of sequences relative to some of their nearest neighbours but not others. The process is similar for clone library-derived sequences (b).

2.4.4 Analysis of band profiles

2.4.4.1 Visual analysis

Bands in two separate gels cannot be compared directly without first either sequencing and classifying the ribotypes of those bands, or employing complex statistical analyses such as EquiBands (Huber and Peduzzi, 2004). However two DGGE gels showing comparable PCR products may be compared at least qualitatively in terms of gross patterns and profile changes apparent to the unaided eye, in a UV light photograph of the gels. Visual interpretation of DGGE gels may inform decisions about further analysis when resources, including time, are finite.

2.4.4.2 Statistical analysis

Binary matrix conversion. DGGE gel profile images were converted from Windows bitmap (*.bmp*) format to the higher resolution tagged image file format (*.tif*) and imported into the gel analysis software Phoretix 2D (Nonlinear Dynamics, Newcastle-upon-Tyne, UK). This software enabled the conversion of the banding pattern on the gel into a binary matrix. Matching lateral bands across adjacent tracks, especially at the edge of gels where a "smile" effect (see Figure 4.7d) can be caused by electrophoretic retardation due to the proximity of the gel frame, and deletion of bands erroneously identified due to dust particles and other visual noise, was done

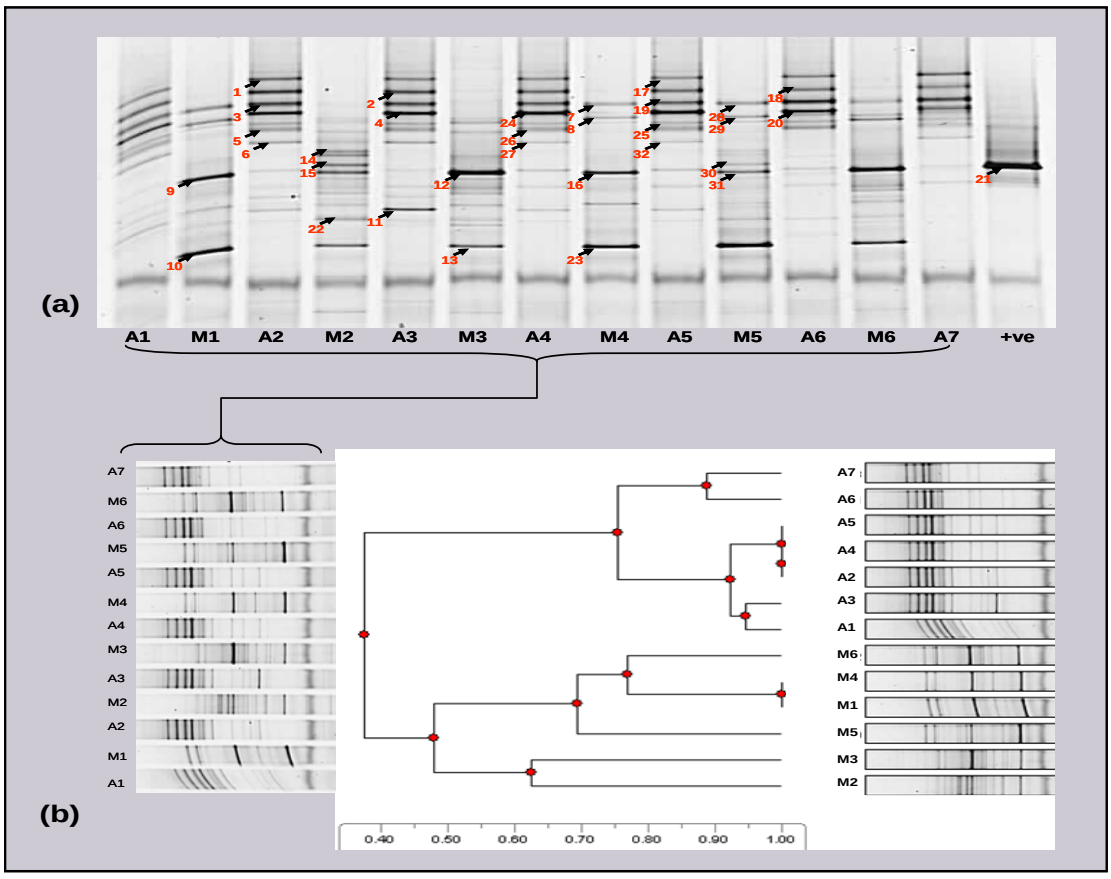


Figure 2.6. DGGE profiles of *Gammaproteobacteria*-specific 16S rRNA gene fragments amplified from environmental DNA extracted from two contrasting marine locations: the North Atlantic gyre (A) and a Norwegian fjord mesocosm (M). Sample details are listed in Table 2.3. Labelled bands were excised and sequenced. Taxonomic and ecological information from BLASTn alignment with sequences in Genbank are listed in Table 3.5 (a). Phoretix cluster analysis of profiles in (a) showing clear resolution into lineages according to geographical origin and experimental treatment. Scale bar indicates proportional similarity (b).

Table 2.3. Bacterioplankton samples referred to in Figures 2.5 and 2.6.

Sample	*Date*	*Location*	*Position*	*Environmental alterations*	*Depth*	*Depth to seafloor*	*Light level (blue/ green)*[a]	*Light level (orange / red)*[b]	*Experiment*
M1	11/05/03	Norwegian fjord mesocosm	61N 05E	n/a	2 m	80 m	98	85	PeECE CO_2 Enrichment
M2	17/05/03	Norwegian fjord mesocosm	61N 05E	n/a	2 m	80 m	98	85	PeECE CO_2 Enrichment
M3	23/05/03	Norwegian fjord mesocosm	61N 05E	n/a	2 m	80 m	98	85	PeECE CO_2 Enrichment
M4	11/05/03	Norwegian fjord mesocosm	61N 05E	900 ppm CO_2	2 m	80 m	98	85	PeECE CO_2 Enrichment
M5	17/05/03	Norwegian fjord mesocosm	61N 05E	900 ppm CO_2	2 m	80 m	98	85	PeECE CO_2 Enrichment
M6	23/05/03	Norwegian fjord mesocosm	61N 05E	900 ppm CO_2	2 m	80 m	98	85	PeECE CO_2 Enrichment
A1	03/05/04	Open ocean (Canary current)	28N 23W	n/a	1 m	2500 m	99	90	PML P & Fe Addition (FeeP)
A2	04/05/04	Open ocean (Canary current)	28N 23W	n/a	1 m	2500 m	99	90	PML P & Fe Addition (FeeP)
A3	06/05/04	Open ocean (Canary current)	28N 23W	P added	25 m	2500 m	65	5	PML P & Fe Addition (FeeP)
A4	17/05/04	Open ocean (Canary current)	27N 22W	Fe added	25 m	2500 m	65	5	PML P & Fe Addition (FeeP)
A5	19/05/04	Open ocean (Canary current)	27N 22W	Fe and P added	25 m	2500 m	65	5	PML P & Fe Addition (FeeP)
A6	25/09/04	N Atlantic subtropical gyre	28N 35W	n/a	50 m	5200 m	50	1	BAS AMT Cruise 14
A7	28/09/04	N Atlantic subtropical gyre	17N 40W	n/a	50 m	4800 m	50	1	BAS AMT Cruise 14

[a] Lower wavelength light of the visible spectrum used in rhodopsin-based photosynthesis, expressed as a percentage of that incident at the surface.
[b] Higher wavelength light of the visible spectrum used in chlorophyll-based photosynthesis, expressed as a percentage of that incident at the surface.

manually by eye. The binary matrix was saved in Excel and used for multivariate statistical analysis using PRIMER-E (Clarke and Gorley, 2006).

Distance matrix conversion. PRIMER-E analyses patterns within and between DGGE gels using a derived binary matrix as a starting point (for methods employing PRIMER-E see Clarke and Warwick, 2001). A binary matrix is a direct representation of the original gel profile pattern. Using the Jaccard similarity coefficient (Jaccard, 1908) this is converted into a distance matrix in which the profiles within a gel are compared, one with another, to assess their degree of similarity (expressed as a percentage where 100 % means two profiles are identical with respect to all bands either present or absent).

Jaccard similarity (for two DGGE profiles A and B) is calculated according to the formula:

$$S_{Jaccard} = N_{AB} / (N_A + N_B - N_{AB})$$

Where N_{AB} is the number of bands shared in common between profiles A and B, and N_A and N_B represent the total number of bands in profiles A or B respectively.

Unweighted pair-wise grouping with arithmetic mean (UPGMA). This is the simplest method for clustering DGGE profiles in a graphical form, or tree, which represents the distance matrix, calculated above using the Jaccard coefficient (Sections 4.4.1 and 4.4.2.7, Figures 2.6, 4.6, 4.7, 4.19 and 4.20).

UPGMA also is one of several methods that may be employed to cluster homologous DNA sequences, such as 16S rRNA gene fragments, into phylogenetic trees (Section 3.3.2.3).

Multidimensional scaling (MDS). This uses the distance matrix to reduce each DGGE profile to a single point in two-dimensional space. Comparison, or connection with lines, of these points allows visualisation of development of DGGE profiles that are either temporally or spatially related (Sections 4.4.2.3 and 5.3.2).

Ribotype accumulation. This measures temporal variance in the diversity of bacteria. The total number of ribotypes in a DGGE profile at the beginning of an

experiment is taken as the starting diversity. New unique ribotypes in subsequent profiles are added to the existing total. Community stability or variability translates into a flat curve or a steep curve, respectively (Section 5.3.3).

Chapter 3

Design and testing of group-specific PCR primers for DGGE analysis of the genetic diversity of complex microbial communities

3.1 Introduction

There have been many studies of the diversity of marine microbial communities in microbial ecology, with the 16S rRNA gene being the most frequently used phylogenetic marker (Amann *et al.*, 1995; Pernthaler and Amann, 2005). This marker gene has revealed the great genetic diversity of bacteria that is now assembled in the Ribosomal Database Project-II (RDP-II; http://rdp.cme.msu.edu/), which currently stores approximately 418,000 partial 16S rRNA gene sequences. It is however likely that there is some degree of redundancy and sequence anomalies within the database (Ashelford *et al.*, 2005). This extent of genetic diversity also makes it difficult to identify the whole range of diversity present in a single sample of a community.

The various methods used to assess 16S rRNA gene diversity within bacterial communities, including clone library analysis, SARST, SSCP and FISH, are outlined in Section 1.2.3. Several other methods, including RISA and RFLP, involving other genetic markers, such as the 16S/23S ITS and enzymes such as DNA-dependent RNA polymerase and tRNA synthetases, are described also in Section 1.2.3. In this chapter the focus is on PCR amplification of the 16S rRNA gene, the products of which are used to fingerprint the sampled community using DGGE (see Sections 1.2.4 and 2.4 for method and protocol).

To analyse the total bacterial community during changing environmental conditions and seasons, group-specific primers were developed for seven different taxonomic groups. Apart from expensive large-scale sequencing of clone libraries, DGGE remains one of the few profiling techniques currently available that allows analysis of the whole microbial community and identification, by sequence analysis of DGGE bands, of specific members of that community. Therefore, the approach of Dar *et al.* (2005) was used; that is re-amplification of the bacterial group-specific PCR products with a nested or semi-nested universal DGGE-PCR primer set.

3.1.1 Aims

As well as producing group-specific DGGE gels a specific aim was to show that each primer pair could produce distinct DGGE ribotype profiles for bacterial communities from two contrasting pelagic marine environments. These were near-shore waters in a nutrient-rich Norwegian fjord and the oligotrophic North Atlantic gyre.

3.2 Method

3.2.1 Primer design

PCR primers were designed using the ARB (Ludwig *et al.*, 2004) and PRIMROSE (Ashelford *et al.*, 2002) programs. The name given to each primer consists of a number, representing the position of the first base of the primer within the *E. coli* 16S rRNA gene sequence, and 'f' or 'r' indicating whether the primer is the forward or reverse primer, respectively. The sequences of the primers and their specificity are summarised in Table 3.1. The number of target and non-target matches of the primers was tested *in silico* using the PROBE MATCH function within the RDP-II database (http://rdp.cme.msu.edu) and the ROSE function of TOOLKIT (http://www.cf.ac.uk/biosi/research/biosoft) which identifies the number of sequences, among the 15,104 (October 2006) *Bacteria*-specific sequences in its database, to which the primer sequence is homologous.

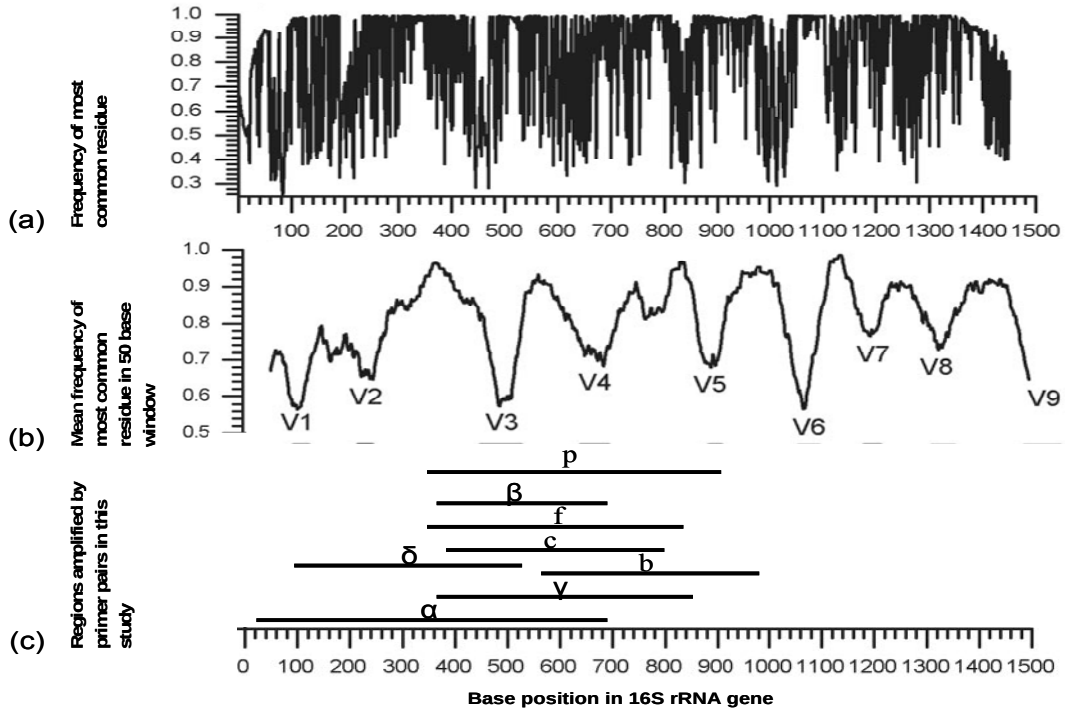


Figure 3.1. Variable and conserved regions within the 16S rRNA gene. (a) Frequency of the most common nucleotide at each base position, determined from 4383 type strains. (b) Mean frequency in a sliding 50 base window smoothes the plot line, revealing nine hypervariable regions, V1 to V9 (Comparative RNA website, http://www.rna.icmb.utexas.edu). (c) Regions amplified and sequenced in production of clone libraries using group-specific primers; (α) *alphaproteobacteria*, (β) *betaproteobacteria*, (γ) *gammaproteobacteria*, (δ) *deltaproteobacteria*, (b) *Bacteroidetes*, (c) *Cyanobacteria*, (f) *Firmicutes*, and (p) *Planctomycetes* (adapted from Ashelford *et al.*, 2005).

As 16S rRNA gene sequence databases continue to grow, group-specific PCR primers must be continually re-evaluated for their specificity and range of sequence matches (Baker *et al.*, 2003). Furthermore, new and more sophisticated software packages, such as PRIMROSE and ARB, can help researchers to design 16S rRNA gene probes. These software packages use different sequence databases and each has its specific strengths. For example, as well as differences in the algorithm used for searching for priming sites, it is possible to use PRIMROSE to design degenerate probes, a function that is not available in ARB. ARB, however, allows the search for probes to be based on a greater number of specific parameters.

Table 3.1. Summary of group-specific 16S rRNA gene PCR primers, their specificity towards taxonomic groups as revealed by *in silico* analysis and the annealing temperatures used in the PCR reactions

Primer [a]	***Target group***	***Sequence (5' to 3')***	***Escherichia coli position***	***Includes variable regions***	***Identical matches within target group*** [b]	***Percentage matches within the target group (%)*** [c]	***Matches outside target group*** [b]	***AT group-specific PCR for clone libraries (°C)***	***16S rRNA gene fragment used for 'Probe Match'*** [d]	***Reference***
28f[e]	*Alphaproteobacteria*	ARCGAACGCTGGCGGCA	28–44	V1–V4	891	81.2 (83.6)	13	69	1–100	Ashelford *et al.* (2002) [f]
684r[e]	*Alphaproteobacteria*	TACGAATTTYACCTCTACA	684–702		1359	88.5 (89.3)	242		650–750	This study
359f	*Betaproteobacteria*	GGGGAATTTTGGACAATGGG	359–378	V3–V4	851	93 0 (88.8)	10	63	300–400	Ashelford *et al.* (2002) [f]
682r	*Betaproteobacteria*	ACGCATTTCACTGCTACACG	682–701		701	86.4 (82.4)	6		650–750	Ashelford *et al.* (2002) [f]
395f	*Gammaproteobacteria*	CMATGCCGCGTGTGTGAA	395–412	V3–V5	1412	52.8 (59.1)	203	54	350–450	This study
871r	*Gammaproteobacteria*	ACTCCCCAGGCGGTCDACTTA	871–891		1579	64.8 (62.1)	50		850–950	This study
555f	*Bacteroidetes*	CCGGAWTYATTGGGTTTAAAGGG	555–577	V4–V5	549	83.9 (85.0)	2	61	500–600	This study
968r	*Bacteroidetes*	GGTAAGGTTCCTCGCGTA	968–985		573	90.5 (92.2)	305		900–1000	This study
361f	*Cyanobacteria*, chloroplasts	GGAATTTTCCGCAATGGG	361–378	V3–V4	434	91.0 (92.2)	16	59	300–400	This study
785r	*Cyanobacteria*, chloroplasts	GACTACWGGGGTATCTAATCC	785–805		345	88.2 (87.8)	42		750–850	This study
352f[e]	*Planctomycetes*	GGCTGCAGTCGAGRATCT	350–367	V3–V5	209	84.0 (87.4)	140	68	300–400	This study
920r[e]	*Planctomycetes*	TGTGTGAGCCCCCGTCAA	920–937		103	98.1 (87.1)	6		900–1000	This study
350f[e]	*Firmicutes*	GGCAGCAGTRGGGAATCTTC	350–369	V3–V4	1140	24.9 (37.7)	7	57	300–400	This study
814r[e]	*Firmicutes*	ACACYTAGYACTCATCGTTT	814–833		1087	25.1 (30.2)	17		750–850	This study
9bfm	*Bacteria*	GAGTTTGATYHTGGCTCAG	9–27	V1–V9	3101	77.7 (86.4)	1	52	1–100	This study
1512uR	Universal (*Bacteria* and *Archaea*)	ACGGHTACCTTGTTACGACTT	1492–1512		3284	78.5 (80.0)		52	1450–1542	Weisburg *et al.* (1991)

Table 3.1.

Abbreviations: AT, annealing temperature; FISH, fluorescence *in situ* hybridization.

[a] The number in the primer name indicates the starting position of the primer sequence within the *E. coli* 16S rRNA gene sequence.
[b] Information based on analysis using the ROSE function within PRIMROSE; percentage of positive hits depends on number of sequences in the database with sufficient information; there is less information available for the termini compared to the centre of the 16S rRNA gene.
[c] Numbers in brackets indicate percentage matches within target group as obtained from using the online tool 'Probe Match' within the Ribosomal Database Project-II (RDP-II) database.
[d] Probe Match is a primer specificity test function in the RDP database (Section 3.2.1). Numbers indicate *E. coli* position; the longest 16S rRNA gene sequences in the RDP are 1542 bases.
[e] The template for the PCR was the PCR product obtained with primers 9bfm/1512uR.
[f] These primers were suggested for use as FISH probes by Ashelford *et al.* (2002), but have not yet been tested.

3.2.2 Nested PCR approach with DGGE primers

All of the group-specific primers described in this study were originally designed as DGGE primers, i.e. a GC clamp was attached to one of the primers of each pair. However, probably due to excessive PCR product length, the primers did not provide clear and reproducible DGGE patterns. The PCR-DGGE approach of Dar *et al.* (2005) was therefore adapted. This requires a nested PCR with *Bacteria* primers (one primer with a GC clamp attached to it) subsequent to the PCR with the group-specific primers, thus resulting in a two-step nested or, in the cases of the *Alphaproteobacteria*, the *Planctomycetes* and the *Firmicutes*, a three-step nested PCR-DGGE approach (Table 3.2). Tests of this approach showed that low yield of the PCRs that sometimes occurred with the specific primers was not a problem, since the nested re-amplification with the bacterial DGGE primer sets resulted in an amount of PCR product sufficient for subsequent DGGE analysis (Figure 3.2).
This nested PCR approach was successful with all primer pairs when applied to environmental DNA samples from both environments tested.

Table 3.2. Nested PCR approach with DGGE primers

Target group	*Primers used for group-specific PCR*[e]	*Primers used for re-PCR for DGGE* [a]	*AT (semi-) nested PCR (°C)*	*Denaturing gradient used for DGGE (%)* [b]
Alphaproteobacteria[c]	28f/684r[e]	341f-GC/518r	56	40–60
Betaproteobacteria	359f[e]/682r	518f-GC/682r[d]	60	40–55
Gammaproteobacteria	395f[e]/871r	518f-GC/785r	56	40–60
Bacteroidetes	555f[e]/968r	555f-GC/907r	64	40–60
Cyanobacteria	361f[e]/785r	518f-GC/785r[d]	56	40–55
Planctomycetes[c]	352f[e]/920r	518f-GC/907r	60	40–60
Firmicutes[c]	350f[e]/814r	518f-GC/785r	56	40–60
Bacteria	9bfm/1512uR[c]	341f-GC/518r	56	40–60

AT, annealing temperature

PCR products produced with the group-specific primers were re-amplified using a set of bacterial or universal DGGE primers.

[a] Nucleotide sequences of the primers, which have been published previously, and the GC clamp are: 341f, CCTACGGGAGGCAGCAG (Muyzer *et al.*, 1993); 518f, CCAGCAGCCGCGGTAAT (Muyzer *et al.*, 1993); 518r, ATTACCGCGGCTGCTGG (Muyzer *et al.*, 1993); 785r, CTACCAGGGTATCTAATCC (Lee *et al.*, 1993); 907r, CCGTCAATTCMTTTGAGTTT (Muyzer *et al.*, 1998); GC clamp, CGCCCGCCGCGCGCGGCGGGCGGGGCGGGGGCACGGGGGG (Muyzer *et al.*, 1993). Note the sequence of the bacterial primer 785r (*E. coli* position 785–803) matches 81.2% of all 16S rRNA gene sequences in the ROSE database, but only 1% of the *Cyanobacteria*/chloroplast 16S rRNA gene sequences.

[b] 100% denaturant contains 7 M urea and 40% (v/v) formamide.

[c] Group-specific PCR for *Alphaproteobacteria*, *Planctomycetes* and *Firmicutes* is nested within the *Bacteria* PCR, making PCR of these three specific groups three-step nested protocols.

[d] Semi-nested re-PCR; the reverse primers are identical to those used in the "Group-specific" step.

[e] Primer indicated originally incorporated the GC clamp, rendering the group-specific PCR product applicable to DGGE. Primers were used without GC clamps in the subsequent nested PCR approach

3.2.3 Sampling

Samples from two contrasting environments were used in this study. A coastal sample was collected in May 2003 from a 12,000 L mesocosm after 12 days of incubation during the 'Pelagic Ecosystem CO_2 Enrichment Study' (PeECE; http://peece.ifm-geomar.de/) at the EU Large-Scale-Facility, University of Bergen, Norway, located at Espeland in the Raunefjord (60°27'N 5°22'E), 10 km south of Bergen. During the experiment, nitrate, phosphate and silicate were added at typical winter concentrations to stimulate a phytoplankton bloom. An oligotrophic open ocean sample was obtained from 15 m depth in the North Atlantic gyre (35°N 20°W) collected during the Atlantic Meridional Transect (AMT)-15 cruise in September 2004 (Robinson *et al.*, 2006).

Four litre samples from the mesocosm and 7 L samples from the North Atlantic gyre were filtered through Sterivex® cartridges with 0.22 μm pore size filter membranes (Millipore, Watford, UK). The cartridges were stored at -70 °C until analysis. Total environmental DNA was isolated from the cartridges using the method described in Section 2.2.1.

3.2.4 Clone library analysis for primer validation

Clone libraries were prepared from PCR-amplified 16S rRNA gene fragments for each of the group-specific primer pairs. A nested PCR approach was required for the *Alphaproteobacteria*-, *Planctomycetes*- and *Firmicutes*-specific primer pairs to obtain sufficient PCR product for subsequent cloning. In these cases, aliquots of the PCR products obtained with the *Bacteria* primer pair 9bfm/1512uR were used as templates for a reamplification with a nested group-specific primer set. In all cases, the PCR (see Section 2.2.2) yielded only specific products, that is single bands as judged by electrophoresis of the PCR products on agarose gels (Section 2.2.3).

Aliquots of the products from group-specific PCRs were cloned into a TA vector using the pGEM®-T Easy Vector System I cloning kit (Promega, Section 2.3). Twenty or 50 clones were picked from each of the group-specific clone libraries prepared from the mesocosm and the North Atlantic gyre samples, respectively, and the 16S rRNA gene fragments were re-amplified using vector primers (M13, see "colony PCR," Section 2.3.4) . The PCR products were used for the sequencing reactions.

3.2.5 DGGE

DGGE was carried out on nested PCR products as described in Section 2.4. Improvements to gels, compared to the non-nested PCR-DGGE approach (Table 3.2), in terms of band resolution and abundance are illustrated in Figure 3.2.

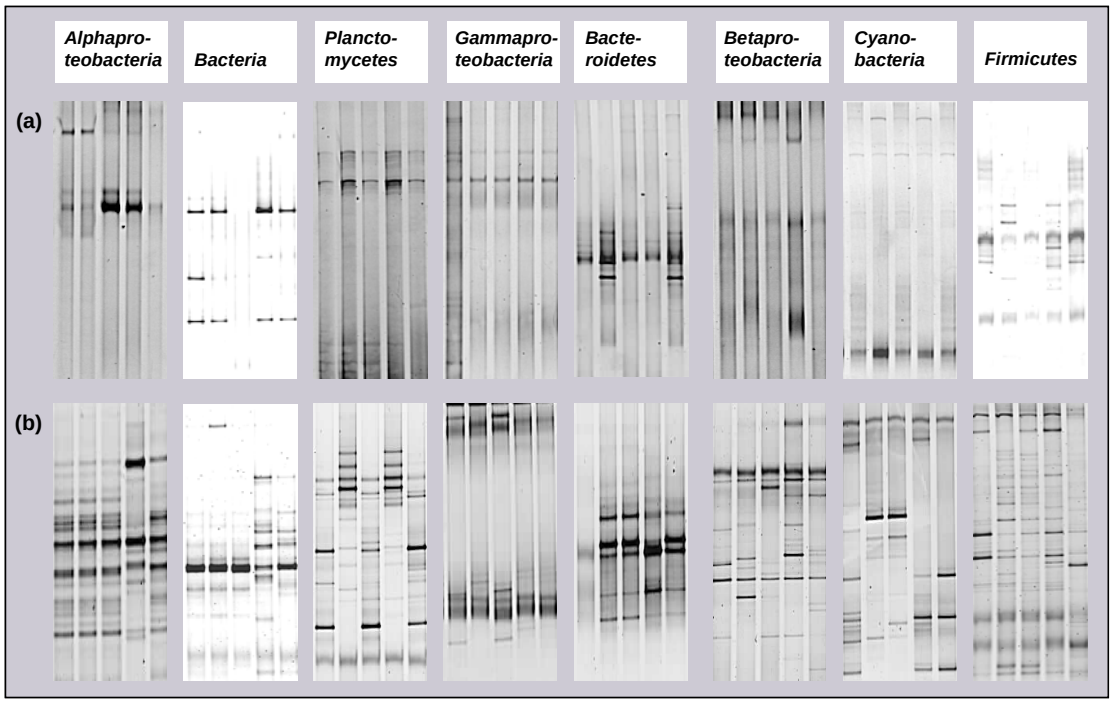


Figure 3.2. DGGE profiles obtained using original group-specific PCR primers with GC-clamps **(a)** outlined in Table 3.2, and improvements in band number and resolution with adoption of *Bacteria* DGGE primers nested within a group-specific PCR **(b)**, detailed in Table 3.2 under "re-PCR for DGGE."

3.2.6 Phylogenetic analysis

Sequences from the mesocosm and the North Atlantic gyre clone libraries prepared with the *Gammaproteobacteria*-specific primers were first compared to sequences stored in GenBank using the BLASTn algorithm. Subsequently, the sequences from these clone libraries, and sequences with high similarity as identified by the BLASTn searches, were imported into the ARB software program (http://www.mikro.biologie.tu-muenchen.de/pub/ARB) and aligned to other *Gammaproteobacteria* 16S rRNA gene sequences using the automated alignment tool within ARB (Ludwig *et al.*, 2004). Calculation of three phylogenetic trees was based on these sequence alignments using the neighbour-joining, maximum parsimony, and maximum likelihood methods each with Jukes–Cantor corrections.

Neighbour-joining method. All homologous sequences are compared and sorted into a distance similarity matrix, and then arranged as a star tree. The distance matrix is then modified so that each pair of nodes is calculated according to their average divergence from all other nodes in the star tree. The unidirectional tree is

begun by collapsing the least distant pair of nodes. As ancestral nodes are added, terminal nodes are removed. This pruning leaves multiple sequences occupying terminal nodes, as can be seen several times in Figure 3.3. The neighbour-joining tree keeps track of nodes as it expands or contracts, and input data does not need to be ultrametric, or conform to a constant rate of evolution (Saitou and Nei, 1987).

Maximum likelihood method. All possible trees are calculated and that with the highest probability of accounting for all the terminal (extant) lineages and nodes is the maximum likelihood tree. The method is robust to violations of evolutionary assumptions and sampling error. Different tree topologies are evaluated, making the final tree statistically well founded. However plesiomorphy is not accounted for, and tree generation requires much computer power.

Maximum parsimony method. All possible trees are again considered with the simplest (minimum total branch length) being chosen. Trees are inferred by minimizing the number of evolutionary steps required to gain the terminal (extant) set of lineages, or sequences. The maximum parsimony tree is that needing the minimum number of substitutions at each informative site. Informative sites are those where more than one nucleotide can exist, and each variant exists in more than one sequence. The method also requires much computer power.

For the phylogenetic analysis those branching points within a tree that were not supported by each of the three algorithms were collapsed within the neighbour-joining tree using a strict consensus rule until the branching was supported in all three analyses. The neighbour-joining tree (Section 3.3.2.3) depicts the phylogenetic relationship of the 16S rRNA genes of *Gammaproteobacteria* from environmental clones and cultured strains. Bootstrap values were calculated from 100 trees using the neighbour-joining method.

3.3 Results and discussion

3.3.1 Nucleotide sequence accession numbers

The sequence data of 16S rRNA gene fragments have been submitted to the EMBL database with accession numbers AM706671-AM707020 (North Atlantic gyre clone library), AM706537-AM706670 (mesocosm clone library), and AM747394-AM747468 (sequences of DGGE bands).

3.3.2 *In silico* and *in vitro* analysis of group-specific PCR primers

Blackwood *et al.* (2005) used ARB to develop five group-specific primers including primers for four of the groups of bacteria investigated in this study. Ashelford *et al.* (2002) compared in detail the PRIMROSE and ARB programs and concluded that in many cases it was possible to identify better oligonucleotide probes (judged by *in silico* analysis) using PRIMROSE rather than ARB. The development of our primers was based on the independent use of both software packages to ensure the best design. The theoretical specificities of all primers were tested with the ROSE program of the TOOLKIT software package, as well as the PROBE MATCH function within the RDP database (http://rdp.cme.msu.edu/).

PCR primers were developed for the amplification of 16S rRNA gene fragments that provide valuable additions to existing primers. *In silico* analyses indicate that they have generally a higher number of exact matches to the 16S rRNA gene sequences from members of the target group of bacteria for which they were designed, while their specificity is generally similar to that of published primers (Table 3.3).

However, some of the most suitable primers were identical to the FISH probes suggested by Ashelford *et al.* (2002): 28f (*alphaproteobacteria*), 359f and 682r (*betaproteobacteria*). Other probes suggested by Ashelford *et al.* (2002) or previously published primers (for example, Nübel *et al.*, 1997; Blackwood *et al.*, 2005) exploit similar priming sites to those used in this study but were, for example, of different length (for example, 684r, *alphaproteobacteria*, 555, *Bacteroidetes* and 785r, *Cyanobacteria*) and degeneracy (for example, 361f, *Cyanobacteria*). Blackwood *et al.* (2005) and Blümel *et al.* (2007) paired their group-specific primers with either the universal primer 1392r or the *Bacteria* primers Eub338 or 27f. In this study by contrast, two group-specific primers (that is a primer pair) were designed for each taxonomic group to increase specificity of the PCR.

Table 3.3. Comparison of 16S rRNA gene group-specific primers developed in this study with those used in previous studies by Blackwood *et al.* (2005) and Nübel *et al*. (1997)

Target group	*Primer pairs*	*Escherichia coli position*	*Percentage of matches within the target group*	*Matches outside target group*
Alphaproteobacteria	684r	684–702	88.5	242 [b]
	ADF681F	**682–698**	**87.3**	**47**
Betaproteobacteria	682r	682–701	86.4	6
	Beta680F	**680–694**	**85.7**	**5**
Planctomycetes	920r	920–937	47.5	6
	Plan930R	**931–947**	**43.1**	**3**
Firmicutes	350f	350–369	24.9	7
	BLS342F	**352–369**	**21.1**	**0**
Bacteroidetes	968r	968–985	90.5	16
	Cyt1020R	**978–995**	**8.5**	**0**
Cyanobacteria	361f	361–378	91.0	16
	CYA359F	**359–378**	**83.4**	**73**
	785r	785–805	88.2	42
	CYA781R [a]	**781–805**	**87.2**	**40**

The *in silico* analysis was carried out using the PRIMROSE program (Ashelford *et al.*, 2002). The primers used by Blackwood *et al.* (2005) and Nübel *et al.* (1997) are indicated in bold.

[a] Comparison is based on the mixture of primers CYA781R(a) and CYA781R(b) (Nübel *et al.*, 1997).

[b] The decrease in specificity of primer 684r did not compromise the specificity of the 28f/684r primer pair for the *Alphaproteobacteria* (Section 3.3.2.1).

The group-specific primer pairs were generally used to amplify 16S rRNA gene fragments directly from environmental DNA. However, primer pairs for the *Alphaproteobacteria*, the *Planctomycetes* and the *Firmicutes* resulted in only low yield when environmental DNA was used as template. Therefore, a nested PCR approach was designed to overcome this problem. Obviously, the forward primer used in the first PCR had to be upstream of the *Alphaproteobacteria*-specific forward primers (Alf28f), which were located close to the 5'-end of the 16S rRNA gene. Concerns have been raised for some time when the bacterial forward primers 8f and 27F are used because of limited amplification efficiency and potential mismatches with newly discovered strains or environmental 16S rRNA gene sequences (for example, Marchesi *et al.*, 1998). These primers were designed when the sequence databases consisted only of a few thousand clones. Testing *Bacteria* primers for specificity revealed that the sequences were homologous to only 56.5% (8f, Hicks *et al.*, 1992) or 72.9% (27F, Giovannoni *et al.*, 1996) of the over 15,000 sequences within the PRIMROSE database. The sequence of the universal (*Bacteria* and

Archaea) reverse primer 1512uR (Weisburg *et al*., 1991) was homologous to 78.5% of the sequences tested (Table 3.1).

The reverse primer 1512uR was used with a modified version of the forward primer 8f to account for sequence differences at this priming site. *In silico* analysis showed that the sequence of the new *Bacteria* primer 9bfm (Table 3.1) is homologous to 77.7% of the bacterial sequences in the PRIMROSE database but also to one archaea sequence (*Methanobrevibacter* sp. strain MB-9; accession AB017514). However, an archaeal 16S rRNA gene sequence was never detected, in either of the clone libraries constructed using primers 9bfm/1512uR and screened in this study.

However, despite the high percentage of exact matches to bacterial 16S rRNA gene sequences, primers may still be biased against certain groups of bacteria while matching the 16S rRNA gene sequence of most strains of other phylogenetic groups. Therefore, primers should always be tested with respect to the specific bacterial group of interest prior to use.

The newly developed primers were tested with environmental DNA samples from contrasting environments as described in Section 3.2.3.The following comparisons concerning specificity of the primers are based on the program ROSE of the TOOLKIT software package, though the specificity of the primers when compared using the PROBE MATCH function within the RDP database are also provided (Table 3.1). Clone libraries made using the group-specific primer pairs are shown fractionated to the genus level in Figure 3.4.

Table 3.4. Results from the BLASTn searches with the sequences of the 16S rRNA gene fragments obtained from the screening of two clone libraries by sequence analysis

Target group	***Primers***	***Percentage of positive hits within mesocosm library (out of 20 clones tested)***	***Percentage of positive hits within North Atlantic gyre library (out of 50 clones tested)***
Alphaproteobacteria	28f/684r	100	100
Betaproteobacteria	359f/682r	100	30
Gammaproteobacteria	395f/871r	100	100
Bacteroidetes	555f/968r	100	100
Cyanobacteria	361f/785r	100	100
Planctomycetes	352f/920r	100	100
Firmicutes	350f/814r	70	0

3.3.2.1 *Alphaproteobacteria*

The forward primer (28f) used to amplify the *Alphaproteobacteria* was taken from a molecular probe designed by Ashelford *et al.* (2002). The sequence of the reverse primer (684r), coincidentally similar to a primer proposed by Blackwood *et al.* (2005), was homologous to that of a relatively large number (242) of bacteria outside the *Alphaproteobacteria*, mainly *Fusobacteria* and others belonging to the orders *Desulfovibrionales*, *Desulfobacterales* and *Desulfuromonadales* of the *Delta* subgroup of the *Proteobacteria*. In contrast, the sequence of forward primer 28f was homologous only to 13 16S rRNA gene sequences outside the target group, equally distributed among the *Gamma* and *Delta* subgroups of the *Proteobacteria,* and the *Verrucomicrobia*. Due to the high specificity of the forward primer and the high AT, the application of both primers as a primer pair in a PCR resulted in the amplification of 16S rRNA gene fragments from organisms that all belonged to the target group (Table 3.4). This primer pair therefore demonstrated the advantage of both primers of a primer pair being biased towards a particular bacterial group of interest, rather than one group-specific primer combined with a second, *Bacteria* or universal 16S rRNA gene primer.

Most of the sequences that were detected in either of the two clone libraries belong to members of the *Roseobacter* clade and the genus *Sphingomonas*, but several were also members of the genera *Rhodobium* and *Brucella* (Figure 3.4a). This demonstrates the value of the group-specific primer approach in detecting specifically a wide genetic diversity within each taxonomic group.

3.3.2.2 *Betaproteobacteria*

Primers 359f and 682r (Table 3.1) were identical to two of those suggested by Ashelford *et al.* (2002). All sequences analysed from the mesocosm sample belonged to the *Betaproteobacteria* (Table 3.4). In contrast, despite the high specificity of the primers, only one-third of the sequences from the North Atlantic gyre clone library were from *Betaproteobacteria* (Table 3.4 and Figure 3.4b). Interestingly, three of the clones detected in the clone library from the North Atlantic gyre show sequence similarities (at 92–95% sequence similarity) to 16S rRNA gene sequences of *Burkholderia* sp. also prevalent in the Sargasso Sea data set (Venter *et al.*, 2004).

To test whether this low specificity could be improved using a nested re-amplification PCR approach, a second clone library was screened. This library used

an aliquot of the PCR product obtained with primers 9bfm/1512uR as template for nested re-amplification with primers 359f/682r in a second PCR. Again, only 28 % of sequences screened from this clone library were from *Betaproteobacteria*. The reason for this low yield of positive hits from this oligotrophic environment is not known. The use of primer 682r generally results in the group-specific amplification of 16S rRNA gene fragments as the last 3'-end (G) base of the primer is highly specific for the vast majority of *Betaproteobacteria* sequences in the ROSE database. This is supported by the fact that the clone library prepared from the coastal sample proved to be composed entirely of target sequences (Table 3.4). However, the low specificity at the 3'-end of primer 359f (most 16S rRNA gene sequences have three guanosines at *E. coli* positions 376–378), combined with the potentially low abundance of *Betaproteobacteria* in the North Atlantic gyre sample may have led to the low percentage of target hits in this clone library. Even specific primers have difficulty amplifying their targets when there is a background high abundance of non-specific targets (J.E.M. Stach, personal communication).

3.3.2.3 *Gammaproteobacteria*

As in the case of the *Alphaproteobacteria*, one of the primers (359f) appears to have a high number (203) of matches outside the target group (Table 3.1). However, more than 160 of these sequence hits are due to a large number of sequences for a small number (five) of particular bacterial species. For example, 45 non-target hits are due to homologous sequences within the 16S rRNA gene of *Acetobacter* sp., *Gluconobacter* sp. and *Gluconacetobacter* sp.. Furthermore, 118 hits are due to *Ralstonia* sp. (54) and *Burkholderia* sp. (64) sequence hits. The reason for the high number of sequences from these bacteria in the databases is because they are important pathogens. Given this, and the higher specificity of the reverse primer 871r (Table 3.1), it is not surprising that the specificity found in both clone libraries is 100% (Table 3.4). This is further confirmed by the fact that the application of the *Betaproteobacteria*-specific primers demonstrated the presence of sequences with 92–95% sequence similarity to 16S rRNA gene sequences of *Burkholderia* sp. as the next nearest isolated species, but no such sequences were detected within the clone libraries produced with the *Gammaproteobacteria*-specific primers.

Many of the sequences obtained from the mesocosm sample have a high sequence similarity to 16S rRNA gene sequences that fall within phylogenetic clades

that belong to the oligotrophic marine *Gammaproteobacteria* (OMG) group (Figure 3.3). This group of *Gammaproteobacteria* was introduced by Cho and Giovannoni (2004) to indicate that all isolates were able to grow only in low-nutrient (oligotrophic) media. Phylogenetic analysis of their 16S rRNA gene sequences further confirmed that they form independent phylogenetic clades, which together comprise the OMG group of *Gammaproteobacteria* (Cho and Giovannoni, 2004). The fact that *Gammaproteobacteria* 16S rRNA gene sequences from the mesocosm cluster within two (SAR92 and OM182) of the five OMG clades identified by Cho and Giovannoni (2004) may indicate that this phylogenetic clade represents a group within the *Bacteria* that is genetically diverse and occurs also in coastal nutrient-rich environments. In fact, Stingl *et al.* (2007) state for the SAR92 clade, into which several of the 16S rRNA gene sequences from the mesocosm sample cluster, that "the peak of abundance correlates with the relatively high nutrient concentrations found in an upwelling region off the Oregon coast. In the lower nutrient regions farther off the coast, the abundance of the SAR92 was low, close to the limit of detection."

The phylogenetic tree shown in Figure 3.3 is a strict consensus of three trees, obtained using maximum likelihood, neighbour-joining and maximum parsimony (Section 3.2.6), that was achieved by collapsing the tree on those branching points where there were differences between the three different trees. Also, the sets of sequences that form the individual clades and the branching within these clades were identical in all of the three different trees (individual trees not shown), thus confirming the robustness of the consensus tree.

The success of this *Gammaproteobacteria*-specific primer pair (395f/871r) is demonstrated by the high specificity achieved in environmental PCR and the resulting discovery of 16S rRNA gene sequences in a nutrient-rich environment that appear to belong to the OMG group of *Gammaproteobacteria*. In addition, a relatively small sample (70 sequences) from the open ocean and the mesocosm clone libraries revealed sequences that cluster in a wide range of *Gammaproteobacteria* clades (Figures 3.3 and 3.4c).

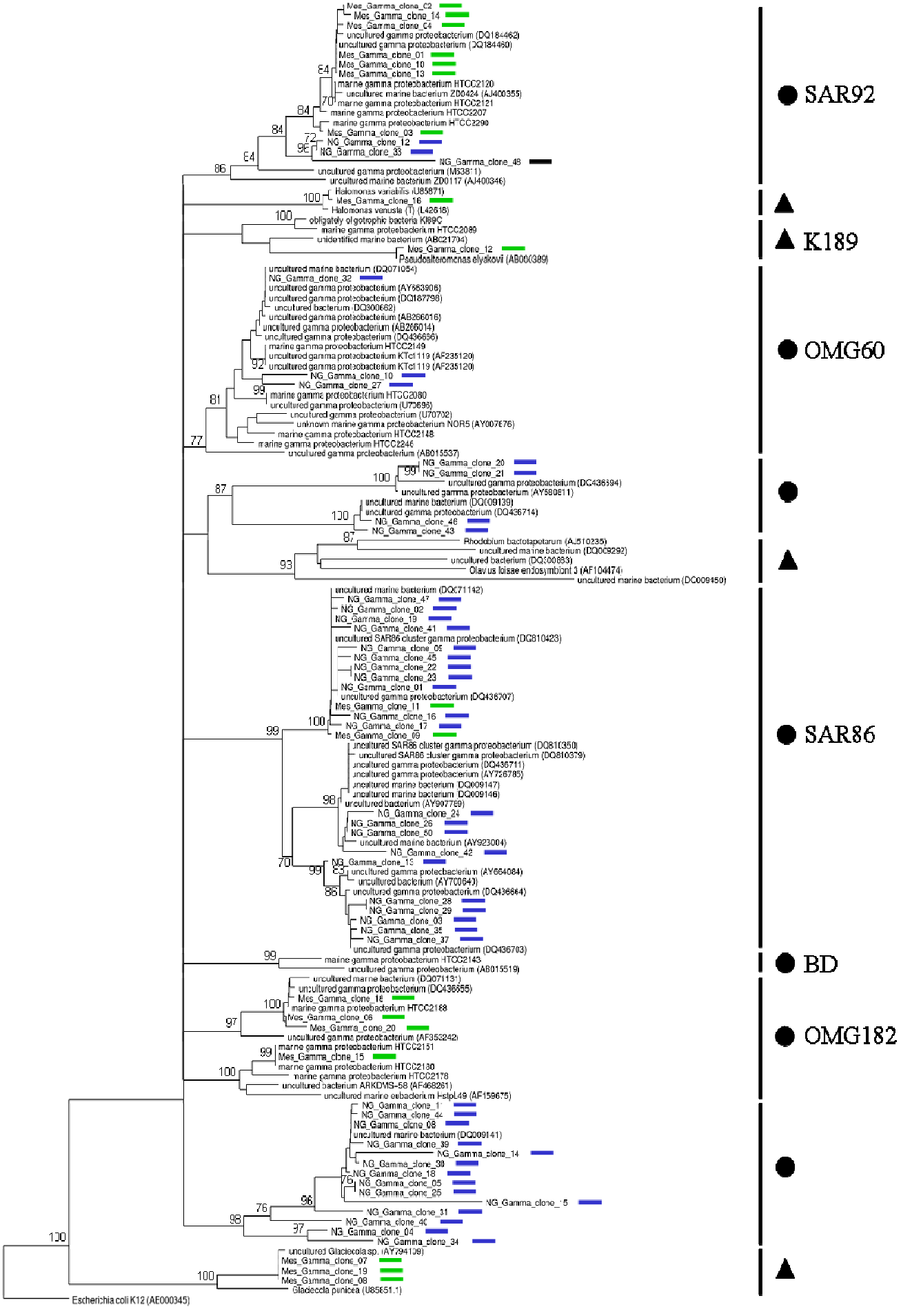

Mes_Gamma_clone_02
Mes_Gamma_clone_14
Mes_Gamma_clone_04
uncultured gamma proteobacterium (DQ184462)
uncultured gamma proteobacterium (DQ184460)
Mes_Gamma_clone_01
Mes_Gamma_clone_10
Mes_Gamma_clone_13
marine gamma proteobacterium HTCC2120
uncultured marine bacterium ZD0424 (AJ400355)
marine gamma proteobacterium HTCC2121
marine gamma proteobacterium HTCC2207
marine gamma proteobacterium HTCC2290
Mes_Gamma_clone_03
NG_Gamma_clone_12
NG_Gamma_clone_33
NG_Gamma_clone_48
uncultured gamma proteobacterium (M53811)
uncultured marine bacterium ZD0117 (AJ400346)
Halomonas variabilis (U85871)
Mes_Gamma_clone_16
Halomonas venusta (T) (L42618)
obligately oligotrophic bacteria KI89C
marine gamma proteobacterium HTCC2089
unidentified marine bacterium (ABC21704)
Mes_Gamma_clone_12
Pseudoalteromonas elyakovii (AB000389)
uncultured marine bacterium (DQ071054)
NG_Gamma_clone_32
uncultured gamma proteobacterium (AY663906)
uncultured gamma proteobacterium (DQ187798)
uncultured bacterium (DQ300662)
uncultured gamma proteobacterium (AB266016)
uncultured gamma proteobacterium (AB266014)
uncultured gamma proteobacterium (DQ436656)
marine gamma proteobacterium HTCC2149
uncultured gamma proteobacterium KTcf119 (AF235120)
uncultured gamma proteobacterium KTcf119 (AF235120)
NG_Gamma_clone_10
NG_Gamma_clone_27
marine gamma proteobacterium HTCC2080
uncultured gamma proteobacterium (U72695)
uncultured gamma proteobacterium (U70702)
unknown marine gamma proteobacterium NOR5 (AY007676)
marine gamma proteobacterium HTCC2148
marine gamma proteobacterium HTCC2246
uncultured gamma proteobacterium (AB015537)
NG_Gamma_clone_20
NG_Gamma_clone_21
uncultured gamma proteobacterium (DQ436694)
uncultured gamma proteobacterium (AY580811)
uncultured marine bacterium (DQ009139)
uncultured gamma proteobacterium (DQ436714)
NG_Gamma_clone_46
NG_Gamma_clone_43
Rhodobium bactotapetarum (AJ510235)
uncultured marine bacterium (DQ009292)
uncultured bacterium (DQ300853)
Olavius loisae endosymbiont 3 (AF104474)
uncultured marine bacterium (DQ009450)
uncultured marine bacterium (DQ071142)
NG_Gamma_clone_47
NG_Gamma_clone_02
NG_Gamma_clone_19
NG_Gamma_clone_41
uncultured SAR86 cluster gamma proteobacterium (DQ810423)
NG_Gamma_clone_05
NG_Gamma_clone_45
NG_Gamma_clone_22
NG_Gamma_clone_23
NG_Gamma_clone_01
uncultured gamma proteobacterium (DQ436707)
Mes_Gamma_clone_11
NG_Gamma_clone_16
NG_Gamma_clone_17
Mes_Gamma_clone_09
uncultured SAR86 cluster gamma proteobacterium (DQ810350)
uncultured SAR86 cluster gamma proteobacterium (DQ810379)
uncultured gamma proteobacterium (DQ436711)
uncultured gamma proteobacterium (AY726785)
uncultured marine bacterium (DQ009147)
uncultured marine bacterium (DQ009146)
uncultured bacterium (AY997759)
NG_Gamma_clone_24
NG_Gamma_clone_26
NG_Gamma_clone_50
uncultured marine bacterium (AY923004)
NG_Gamma_clone_42
NG_Gamma_clone_13
uncultured gamma proteobacterium (AY664084)
uncultured bacterium (AY700643)
uncultured gamma proteobacterium (DQ436664)
NG_Gamma_clone_28
NG_Gamma_clone_29
NG_Gamma_clone_03
NG_Gamma_clone_35
NG_Gamma_clone_37
uncultured gamma proteobacterium (DQ436703)
marine gamma proteobacterium HTCC2143
uncultured gamma proteobacterium (AB015519)
uncultured marine bacterium (DQ071131)
uncultured gamma proteobacterium (DQ436655)
Mes_Gamma_clone_18
marine gamma proteobacterium HTCC2188
Mes_Gamma_clone_06
Mes_Gamma_clone_20
uncultured gamma proteobacterium (AF353242)
marine gamma proteobacterium HTCC2151
Mes_Gamma_clone_15
marine gamma proteobacterium HTCC2130
marine gamma proteobacterium HTCC2178
uncultured bacterium ARKDMS-58 (AF468281)
uncultured marine eubacterium HstpL49 (AF158675)
NG_Gamma_clone_1
NG_Gamma_clone_44
NG_Gamma_clone_08
uncultured marine bacterium (DQ009141)
NG_Gamma_clone_39
NG_Gamma_clone_14
NG_Gamma_clone_30
NG_Gamma_clone_18
NG_Gamma_clone_05
NG_Gamma_clone_25
NG_Gamma_clone_15
NG_Gamma_clone_31
NG_Gamma_clone_40
NG_Gamma_clone_04
NG_Gamma_clone_34
uncultured Glaciecola sp. (AY794109)
Mes_Gamma_clone_07
Mes_Gamma_clone_19
Mes_Gamma_clone_08
Glaciecola punicea (U85851.1)
Escherichia coli K12 (AE000345)
0.10
SAR92
K189
OMG60
SAR86
BD
OMG182

Figure 3.3. (previous page) Phylogenetic analysis of representative 16S rRNA gene nucleotide sequences from the *Gammaproteobacteria* clone libraries prepared from samples from a Norwegian fjord mesocosm and the North Atlantic gyre, and representative sequences from the NCBI and ARB sequence database. The tree was calculated from a nucleotide alignment of 16S rRNA gene fragments (356 bases) using the neighbour-joining method within ARB, with Jukes–Cantor corrections and a maximum frequency filter (Ludwig *et al.*, 2004). *E. coli* (accession J01859) was used as an out group. The confidence of branch points was determined by three separate analyses (maximum likelihood, neighbour-joining, maximum parsimony), with multifurcations indicating branch points that were collapsed using a strict consensus rule until supported in all three analyses. Values of 100 bootstrap replicates (calculated using the neighbour-joining method) are given as numbers at branching points, but those <70 are omitted. Triangles indicate clades containing formally described species, and circles indicate clades or subclades for which no formally described species is available. Clones found in this study are colour-coded: blue, North Atlantic gyre library; green, mesocosm library.

3.3.2.4 *Bacteroidetes*

Primer 555f is highly specific with only two non-target hits (Table 3.1), while its sequence is identical to 84% of the *Bacteroidetes* sequences in the PRIMROSE database. In contrast, the reverse primer 968r has a relatively high number (305) of non-target hits, but the sequence is identical to over 90% of those in the database. However, 298 of the 305 non-target hits are due to identical sequences within the 16S rRNA gene of three members (*Borrelia* sp., *Spirochaeta* sp. and *Treponema* sp.) of the phylum *Spirochaeta*, and these pathogens are unlikely to be sufficiently abundant in the marine environment to be amplified by PCR. This, and the high specificity of the forward primer, seems to be responsible for the fact that all of the sequences in the clone libraries produced using this primer pair belong to the *Bacteroidetes*, according to BLASTn search.

3.3.2.5 *Planctomycetes*

The *Planctomycetes*, remarkable for their attached budding mode of replication and lack of peptidoglycan, are one of the least studied groups of bacteria, mainly due to the lack of successful laboratory cultivation. Only members of two (*Planctomyces* and *Pirellula*) of the four recognized genera have been grown in culture. The Sargasso Sea shotgun sequencing project (Venter *et al.*, 2004) revealed only a small number of molecular marker genes identifying genomic DNA clones as deriving from the *Planctomycetes*. FISH studies (Glöckner *et al.*, 1999; Brinkmeyer *et*

al., 2003) and 16S rRNA gene clone library-based approaches (Brinkmeyer *et al.*, 2003) have indicated low abundance (up to 3%) of *Planctomycetes* in a variety of environments. However, these studies may have underestimated *Planctomycetes* genetic diversity and abundance since the 16S rRNA gene sequences of the *Planctomycetes* may contain mismatches to several commonly used bacteria primers (Vergin *et al.*, 1998). The PCR primers developed here to be specific for the *Planctomycetes* have the potential to discover novel sequences and possibly phylogenetic lineages.

The oligonucleotide 920r was highly specific for the target group, while the sequence of the other primer, 350f, was homologous to 140 non-target sequences (Table 3.1). However, 130 of these 140 non-target hits belonged to the genera *Chlamydia* and *Chlamydophila*, which are pathogens that have not yet been detected and are unlikely to occur in the marine environment. The application of these oligonucleotides as a primer pair therefore led to only specific amplification of members of the *Planctomycetes* (Table 3.4), albeit with a large proportion of unclassified sequences from clones in the North Atlantic gyre library (Figure 3.4h).

3.3.2.6 *Cyanobacteria* (including chloroplasts of eukaryotic algae)

The screening of the 16S rRNA gene clone libraries demonstrated that these primers (361f/785r) were very selective and they specifically amplified only members of the target group. The two environment clone libraries showed contrasting results. The library from the mesocosm experiment was dominated by 16S rRNA gene clones from chloroplasts, but 96% of the clones from the North Atlantic gyre library had high sequence similarity with *Prochlorococcus* sp. isolates or environmental sequence clones (Figure 3.4f). *Prochlorococcus* sp. are known to dominate the oligotrophic regions of the oceans due to their moderate psychrophilia, high surface to volume ratio and efficient nutrient uptake mechanisms (Scanlan and West, 2002), and would not be expected to compete well in nutrient-rich Norwegian waters, as indicated by the results obtained here.

3.3.2.7 *Firmicutes*

The vast majority of marine bacteria detected routinely in *Bacteria*-specific 16S rRNA gene clone libraries are derived from Gram-negative bacteria, although a part (up to 0.05% by weight of clones) of the bacterial metagenome in the Sargasso Sea were derived from the *Firmicutes* (Venter *et al.*, 2004). Results from the screening of *Bacteria*-specific PCR-amplified 16S rRNA gene fragment clone libraries generally contain only a small fraction of 16S rRNA gene sequences from Gram-positive bacteria (for example, Zaballos *et al.*, 2006; Martín-Cuadrado *et al.*, 2007), suggesting that Gram-negative bacteria far outnumber Gram-positive bacteria in the sea. However, this large difference in the abundance of these two groups may be exacerbated by using a molecular approach. For example, the considerably thicker peptidoglycan layer in the *Firmicutes* and the higher GC content of the *Actinobacteria* are likely to result in lower genomic DNA yield after DNA isolation from environmental samples and lower yields during PCR amplification, respectively. In the future altering the DNA extraction protocol (Section 2.2.1) by either effecting non-enzymatic (chemical or mechanical) cell lysis (Purdy, 2005), or including the use of achromopeptidase, a bacteriolytic enzyme effective against many Gram-positive cells resistant to lysozyme (Ezaki and Suzuki, 1982), may increase the yield of Gram-positive genomic DNA.

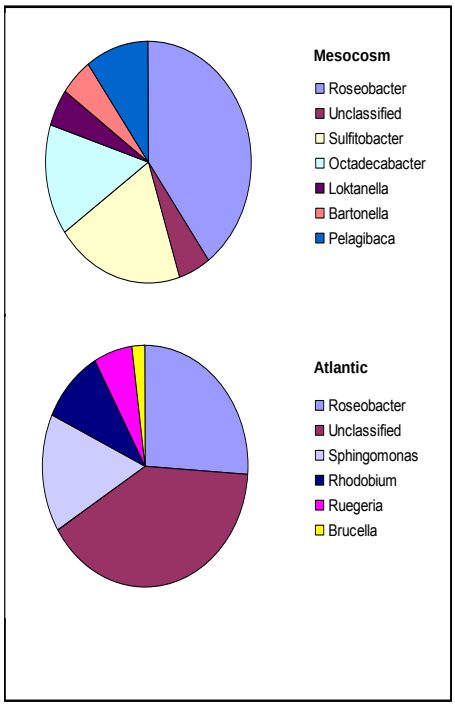
To investigate specifically the diversity of Gram-positive bacteria in the oceans, specific primers for the *Firmicutes* (350f/814r) were designed. *Actinobacteria*-specific primers are already available and proved to be successful for marine samples (Stach *et al.*, 2003). The phylum *Firmicutes* consists of three classes and 33 families (Garrity *et al.*, 2001). Given the huge genetic range among the *Firmicutes* and constrained by a low number of matches of the primers to 16S rRNA gene sequences outside the target group it is not surprising that the sequences of the primers are identical only to about 25% of the 2267 *Firmicutes* sequences within the PRIMROSE database (Table 3.1). However, despite the high specificity of the primers, 6 of the 20 sequences of the clone library prepared from the mesocosm DNA sample were from *Cyanobacteria* (two) or *Alphaproteobacteria* (four) illustrated in Figure 3.4g. Four of the seven non-target hits of primer 350f were to 16S rRNA gene sequences from chloroplasts and two were to *Alphaproteobacteria*. The sequence of the primer 814r is homologous to 17 non-target hits, 14 of which are members of the *Geobacter* group within the *Deltaproteobacteria* and only one from an

alphaproteobacterium (*Ehrlichia* sp.), but none to chloroplasts or *Cyanobacteria*. The clone library from the North Atlantic gyre sample consisted entirely of cyanobacterial (*Prochlorococcus* sp.) 16S rRNA gene fragments (Figure 3.4g). The precise reason for this is unknown. Perhaps the bias of the last 3'-end (C) base of primer 350f for a wide range of the *Firmicutes* sequences in the ROSE database was sufficient to amplify 16S rRNA gene fragments mainly from the target group in the case of the mesocosm sample. However, in the North Atlantic gyre sample, the potential low abundance of *Firmicutes* may have led to the absence of any target hits in this "*Firmicutes*" clone libraries (Table 3.4). This explanation is supported by the fact that the last three 3'-end bases of primer 350f are also present in all *Cyanobacteria*/chloroplast 16S rRNA gene sequences in the ROSE database, and all of the non-specific sequences derived from the cyanobacterium *Prochlorococcus*, which was the dominant photoautotroph in this sample (6400 cells mL^{-1}, Jameson *et al.*, 2007). The fact that a DNA extraction protocol specifically targetted at Gram-positive bacteria was not used, also may have exaggerated the low abundance of the *Firmicutes* in the gyre seawater samples.

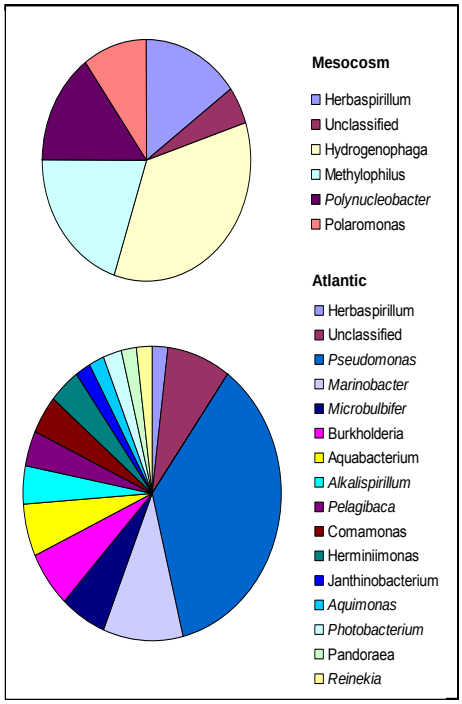
Their proven specificity when used in a PCR with a mesocosm seawater-derived template (Table 3.4) justifies the use of the *Firmicutes*-specific primers to amplify the 16S rRNA genes of these organisms where they are present. However, despite being *Firmicutes*-specific this primer pair (350f/814r), employing a PCR annealing temperature of 57^{o}C, will amplify 16S rRNA gene sequences from more distantly related bacteria, in the absence, or low abundance, of *Firmicutes*-derived DNA template. Similar non-specific amplification in the absence of specific template DNA will potentially occur with other group-specific primer pairs, although these employ variably higher annealing temperatures, restricting promiscuous base-pairing, in the PCR (Table 3.2), and also the other six target groups were demonstrably present in the two contrasting environments sampled.

3.3.3 Quantitative analysis of clone libraries

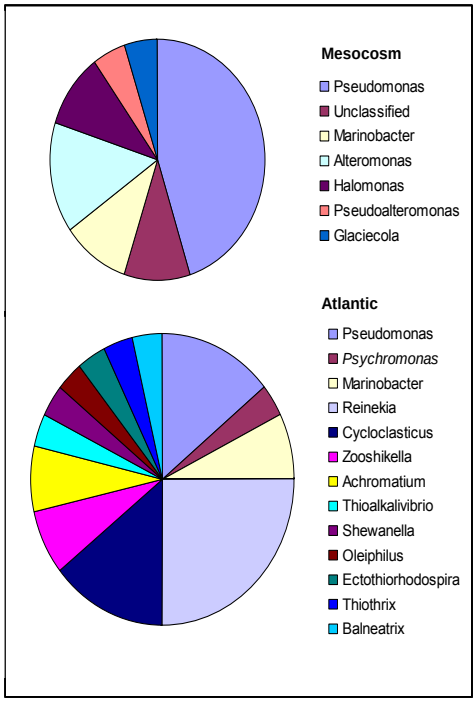
16S rRNA gene fragment sequences (ribotypes), from each pair of putatively group-specific clone libraries, were identified by BLASTn search. A genus label was applied to each ribotype by taking that from the highest scoring database homologue, from the list of the top 100 alignment hits. Where no hits were identified to the level of genus, these ribotype sequences were labelled "unclassified." Figure 3.4 shows a breakdown by genus for the ribotypes sampled using each group-specific PCR, and compares this breakdown between clone libraries constructed from two contrasting marine environments.



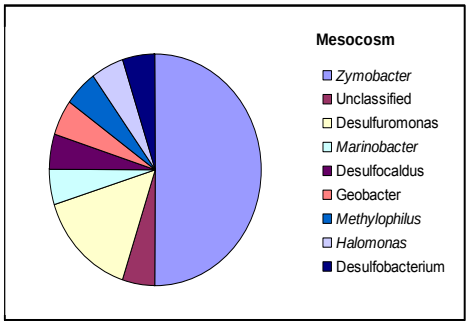
(a) class: *Alphaproteobacteria*



(b) class: *Betaproteobacteria*

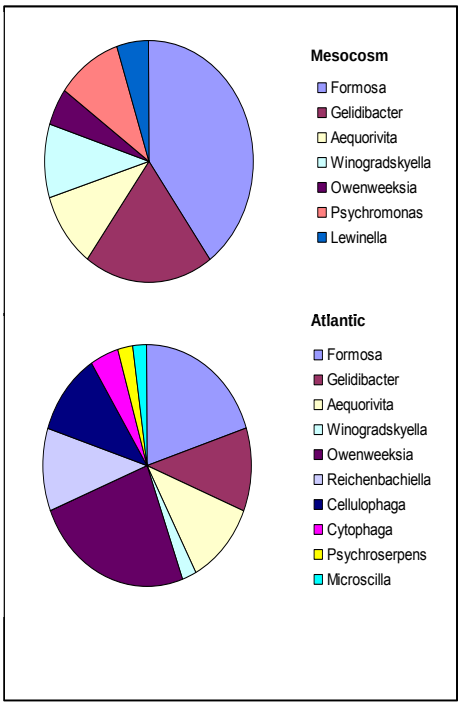


(c) class: *Gammaproteobacteria*

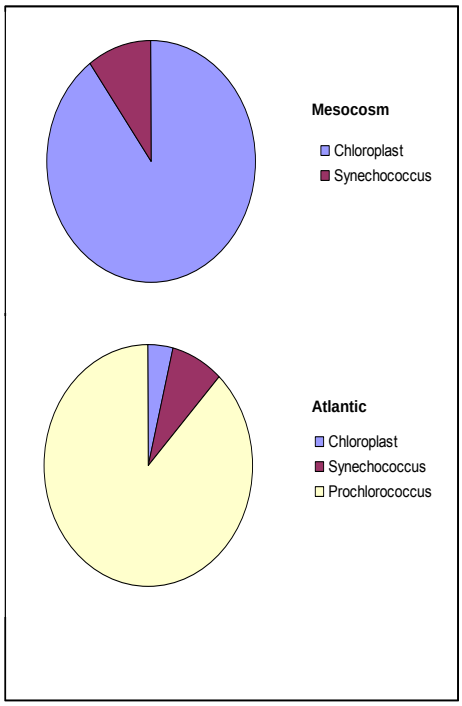


(d) class: *Deltaproteobacteria*

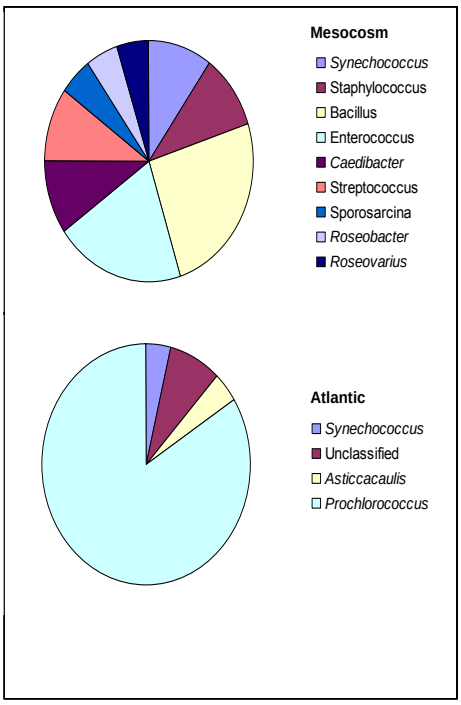
Figure 3.4 (Part I). Composition of clone libraries generated by group-specific PCR amplification of 16S rRNA gene fragments from environmental DNA extracted from a mesocosm in a Norwegian fjord ("Mesocosm," 20 clones per library) and 5 m depth in the subtropical North Atlantic gyre ("Atlantic," 50 clones per library). Genus identity was determined from the closest named homologue following a BLASTn search. Clones of italicised genera are the result of non-specific amplification.



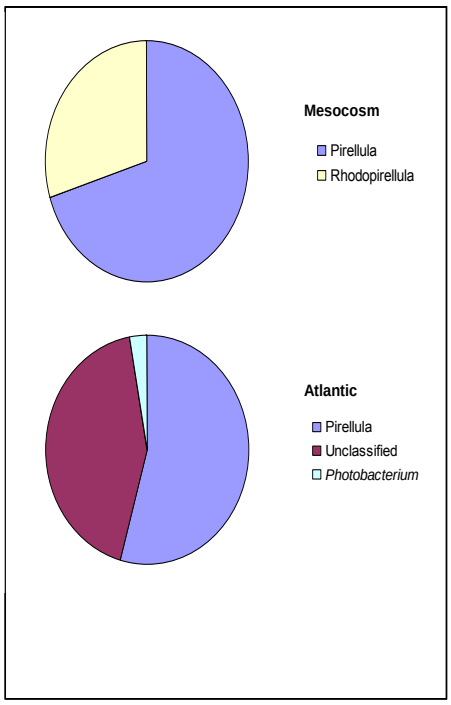
(e) phylum: *Bacteroidetes*



(f) phylum: *Cyanobacteria*



(g) phylum: *Firmicutes*



(h) phylum: *Planctomycetes*

Figure 3.4 (Part II). Composition of clone libraries generated by group-specific PCR amplification of 16S rRNA gene fragments from environmental DNA extracted from a mesocosm in a Norwegian fjord ("Mesocosm," 20 clones per library) and 5 m depth in the subtropical North Atlantic gyre ("Atlantic," 50 clones per library). Genus identity was determined from the closest named homologue following a BLASTn search. Clones of italicised genera are the result of non-specific amplification.

3.3.4 Denaturing gradient gel electrophoresis (DGGE)

The results of the DGGE analysis are summarised in Figure 3.5. In all cases, there were detectable differences in the microbial assemblages between the North Atlantic gyre and the mesocosm samples. However, the separation of the PCR fragments appeared to be better in those cases where a nested PCR was used as compared to a semi-nested PCR (that is for the *Betaproteobacteria*, and particularly for the *Bacteroidetes* and the *Cyanobacteria*/chloroplast; see Figure 3.5 and Section 3.2.5). In the case of the *Cyanobacteria* it may therefore be advisable to use the primer pair developed by Nübel *et al.* (1997) for DGGE analysis rather than the one used here, but to consider primers 361f/785r for the preparation of clone libraries.

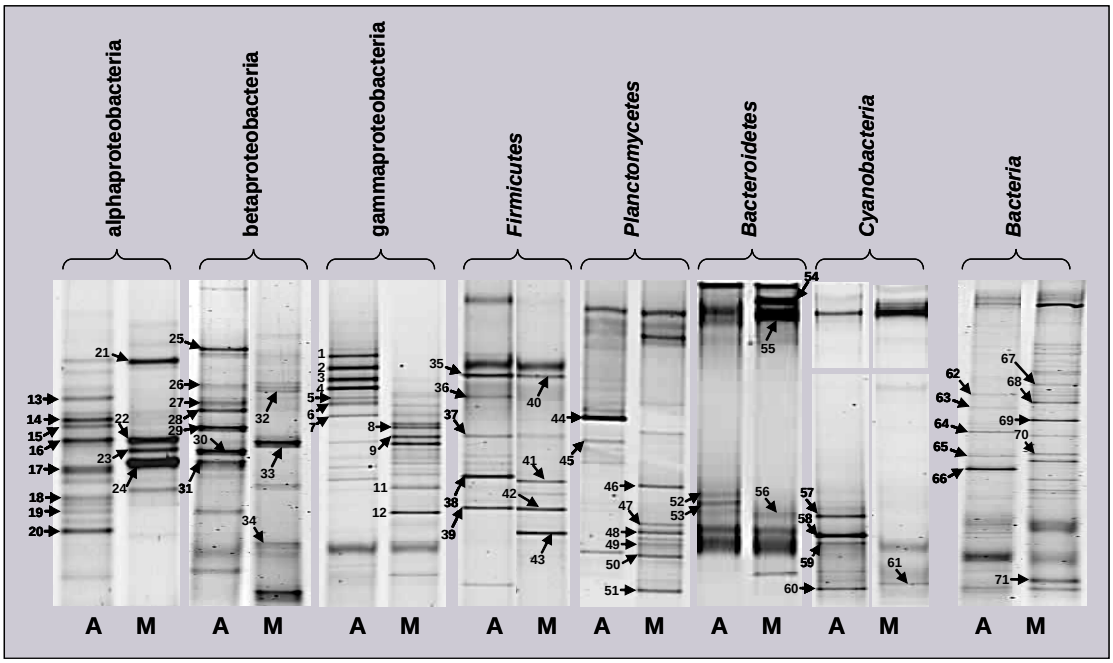


Figure 3.5. DGGE analyses of 16S rRNA gene fragments amplified from DNA samples from the North Atlantic gyre (A) and a Norwegian fjord mesocosm (M). Details of the primers used for the two- or three-step nested PCR approach are outlined in Sections 3.2.2, 3.2.5 and Tables 3.1 and 3.2. The sequence identities of the numbered bands, chosen as the most readily discernable under benchtop UV-illumination, are given in Table 3.5.

Table 3.5. Comparison of BLASTn analysis of sequences of 16S rRNA gene fragments obtained from DGGE bands excised from group-specific PCR-DGGE. Taxonomic names in bold indicate PCR amplicons obtained from bacteria that do not belong to the target group

Band Number	***Accession Number[a]***	***Nearest Match[b]***	***Max Ident.[c]***	***Nearest Isolate[d]***
Gammaproteobacteria				
1	AM747462	clone Surf1.29	98	None
2	AM747463	clone JL-ESNP-I14	98	None
3	AM747464	SAR86 clone ESP10-K9IV-26	95	None
4	AM747465	clone SPOTSAUG01 5m95	98	None
5	AM747466	clone T41 54	92	*Marinobacter* sp. 6-5/A9
6	AM747467	311 clone EBAC750-02H09	98	None
7	AM747468	clone ctg CGOCA18	97	None
8	AM747427	*Azotobacter armeniacus*	100	None
9	AM747428	clone ACE-1	97	*Pseudomonas rhodesiae*
10		n/s		
11	AM747430	clone ARKDMS-9	94	*Pseudomonas* sp. 6-SLY
12	AM747431	clone T31 155	99	None
Alphaproteobacteria				
13	AM747436	clone AS-PE-42	91	*Aeromonas* sp. MRF 600
14	AM747437	clone AS-PE-41	92	*Aeromonas* sp. MRF 600
15	AM747438	*Mesorhizobium* sp. str. M28	92	*Orientia tsutsugamushi*
16	AM747439	clone Fiji8-D10	93	*Rhizobium leguminosarum*
17	AM747440	clone V4.MO.18	93	*Ochrobactrum* sp. CA1
18	AM747441	*Sphingomonas parapaucimobilis*	95	n/a
19	AM747442	uncultured pBC16S3-43	95	None
20	AM747443	clone Toolik Shrub Organic 89	100	*Bradyrhizobiaceae* Unk-01
21	AM747401	clone F2C110	97	*Roseobacter* sp. ANT9276a
22	AM747402	*Rhodobacteraceae* HP47a	98	n/a
23	AM747403	*Rhodobacteraceae* HP47a	96	n/a
24	AM747404	*Rhodobacterales* clone PV2-27	98	*Roseobacter* sp. RED68
Betaproteobacteria				
25	**AM747394**	**clone HF70H6 P2**	**96**	***Aquimonas* sp. D11-34A**
26	AM747395	*Alcanivorax* sp. K3-3	95	n/a
27	**AM747396**	***Methylococcaceae* SIMO-1675**	**94**	***Aquimonas* sp. D11-34A**
28	**AM747397**	**clone HF70 H6 P2**	**97**	***Aquimonas* sp. D11-34A**
29	**AM747398**	**clone A714013**	**97**	***Aquimonas* sp. D11-34A**
30	**AM747399**	**clone A714013**	**98**	***Aquimonas* sp. D11-34A**
31	**AM747400**	**clone MSB-1E11**	**93**	***Aquimonas voraii***
32	AM747405	clone 224ds20 Beta	95	*Propionivibrio* sp.
33	AM747406	clone 224ds20	92	*Serratia marcescens*
34	**AM747407**	**clone 224ds20**	**91**	***Pseudomonas putida* PM1**
Firmicutes				
35	AM747444	clone aab26h05	92	*Staphylococcus* sp.
36	AM747445	clone aaa45g08	92	*Staphylococcus fleurettii*
37	AM747446	*Bacillus licheniformi*. RH104	93	n/a
38	AM747447	*Salinicoccus* sp. Y22	95	n/a
39	AM747448	clone BPC043	92	*Bacillus* sp. HZBN57
40	AM747408	*Bacillus* sp. MH-21	94	*Bacillus* sp. MH-21
41	AM747409	*Bacillus* sp. MH-21	99	*Bacillus* sp. MH-21
42	AM747410	clone TSB-E5	95	*Bacillus* sp. MH-21
43	AM747411	*Bacillus* sp. MH-21	95	*Bacillus* sp. MH-21
Planctomycetes				
44	AM747449	clone Flyn133b	98	*Pirellula* sp.
45	AM747450	clone FS266-28B-03	95	*Pirellula* sp.

Table 3.5. Comparison of BLASTn analysis of sequences of 16S rRNA gene fragments obtained from DGGE bands excised from group-specific PCR-DGGE. Taxonomic names in bold indicate PCR amplicons obtained from bacteria that do not belong to the target group

Band Number	***Accession Number***[a]	***Nearest Match***[b]	***Max Ident.***[c]	***Nearest Isolate***[d]
46	AM747412	clone DPC008	84	*Pirellula* sp.
47	AM747413	clone BSR3LB04	85	None
48	AM747414	OM190	84	None
49	AM747415	clone WCB136	91	None
50	AM747416	clone BSR3LB04	88	*Planctomyces maris*
51	AM747417	clone OM190	99	None
Bacteroidetes				
52	AM747451	clone PI 4e9c	93	*Flavobacteriaceae*. SW058
53	AM747452	uncultured *Bacteroidetes*	96	*Flavobacterium* sp.WED7.4
54	AM747418	clone MS056-2A	99	*Gelidibacter* sp. IMCC1914
55	AM747419	clone MS056-2A	99	*Gelidibacter* sp. IMCC1914
56	AM747420	clone MS056-2A	99	*Gelidibacter* sp. IMCC1914
Cyanobacteria				
57	AM747453	uncultured *Prochlorococcus* sp.	81	*Synechococcus* sp. CC9311
58	AM747454	chloroplast clone MC615-75	95	*Imantonia* sp. MBIC10497
59	AM747455	chloroplast clone MC622-82	93	*Imantonia* sp. MBIC10497
60	AM747456	chloroplast clone MC622-82	90	*Imantonia* sp. MBIC10497
61	AM747421	*Emiliania huxleyi* CCMP 373	99	n/a
Bacteria				
62	AM747457	clone ARDS108	92	*Methylobacterium* sp.
63	AM747458	clone R38A6	92	*Labrys ginsengisoli*
64	AM747459	DGGE band TS-BA4	92	*Sphingomonas* sp. PdSllD9
65	AM747460	*Rhizobium* sp. ST1	92	n/a
66	AM747461	clone T32 183	94	*Hyphomicrobium* sp.
67	AM747422	clone Phc4A	92	*Algibacter* sp. OB15
68	AM747423	*Flavobacteria* bacterium	92	*Sinorhizobium* sp. PRF215
69	AM747424	clone SanDiego3-E9	93	*Algibacter* sp. OB15
70	AM747425	clone Phc4A	90	*Algibacter* sp. OB15
71	AM747426	clone Emh3C	95	*Pibocella* sp. SL14

[a]Accession number reference for 16S rRNA gene sequences uploaded to the ncbi website (http://www.ncbi.nlm.nih.gov).

[b] Name of environmental clone, or cultured isolate, whose 16S rRNA gene sequence most closely matches that from the excised DGGE band, within the top 100 most closely matched sequences from the Genbank database, following alignment using the BLASTn tool.

[c] Degree of invariance between sample sequence and nearest matched sequence in the Genbank database, expressed as a percentage, where 100 % equals complete invariance.

[d] Nearest matched sequence from a laboratory cultured organism, within the top 100 most closely matched sequences from the Genbank database, following alignment using the BLASTn tool.

The identification by BLASTn search of 16S rRNA ribotypes from excised DGGE bands (shown in Figure 3.5 and selected as those most readily discerned and therefore excisable using a benchtop UV transilluminator) showed that the PCR fragments mostly belonged to the bacterial group that the primers specifically targeted. An exception was the DGGE analysis of *Betaproteobacteria*. As with the results from clone library screening, a large fraction (60%) of the bands on the DGGE gel made using the North Atlantic gyre DNA sample were from prokaryotes other than the *Betaproteobacteria*. In contrast to this, all of the bands on the *Firmicutes* DGGE gel that were identified by sequence analysis derived from strains belonging to the target group (Table 3.5 in contrast with clone library analysis in Figure 3.4). However, this may be an artifact due to the sequence analysis of only a relatively small number of DGGE band-derived sequences. Increasing this number may reveal non-target 16S rRNA gene fragments in *Firmicutes*-specific PCR-DGGE gels. Since not all bands in DGGE profiles were identified by sequencing, the data was not used to analyse the composition of DGGE profiles according to BLASTn-identified 16S rRNA gene fragment sequences. Instead profile comparison was restricted to using binary matrix representations of DGGE gels (Section 2.4.4).

3.4 Conclusions

This research has shown the clear benefits of using group-specific PCR primers. Much more information is obtainable than with general *Bacteria* primers as it is possible to target specific bacterial groups for more detailed investigation of diversity, either by screening clone libraries or by using these primers in association with the well-tried methodology of DGGE; there is greater discrimination on individual gels and more bands are visible and can be excised for sequencing.

Chapter 4

An investigation into the response of pelagic bacterial communities to changes in seawater pH equivalent to those predicted due to anthropogenic carbon emissions

4.1 Introduction

4.1.1 Effects of seawater acidification on marine organisms

Seawater acidification of the order of 0.35 to 0.8 pH units, the predicted range of change in surface waters due to human activity (Caldeira and Wickett, 2003; see Section 1.3.3), has been shown to have effects on a range of marine organisms. These include effects within taxonomically widely varying organisms and their populations, and effects on whole ecosystems. Demonstrations of organisms being affected by such a predicted degree of acidification include phytoplankton species such as coccolithophores *in vitro* (Zondervan *et al.*, 2002; Sciandra *et al.*, 2003), flagellates *in situ* (Rost *et al.*, 2003), diatoms *in situ* (Burkhardt *et al.*, 1999; Burkhardt *et al.*, 2001) and *in vitro* (Bucciarelli and Sunda, 2003), and metazoans such as echinoderms and burrowing invertebrates (reviewed in the report by Raven *et al.*, 2005). Physiological effects of increased pCO_2, within *in situ* mesocosms, in organisms of the euphotic zone include inhibited calcification in the coccolithophorid *Emiliania huxleyi* (Riebesell *et al.*, 2000; Zondervan *et al.*, 2001), and, following a 0.5 units fall in pH, an increase of up to 10 % in *in situ* photosynthetic production by the phytoplankton species *Emiliania huxleyi, Heterosigma carterae* and *Stichococcus bacillaris* in nutrient-rich coastal waters (Schippers *et al.*, 2004). Riebesell and co-workers showed raised pCO_2 causes increased DIC consumption by phytoplankton, and increased subsequent carbon drawdown through the water column in mesocosms (Riebesell *et al.*, 2007). This accelerated use of DIC constitutes an ecosystem effect of ocean acidification potentially felt throughout the world's oceans.

4.1.2 Effects of seawater acidification on bacteria

There has been little investigation into the effects of ocean acidification on pelagic bacteria. In 2005 a Royal Society report into anthropogenic ocean acidification stated, "Although their physiology suggests that they [non-photosynthetic microorgnisms] will respond to increased CO_2 in the surface oceans, and to decreased pH, there are few data available to demonstrate any responses and therefore impact that this might have on the functioning of the organisms in their natural environment" (Raven *et al.*, 2005). In the only study addressing heterotrophic bacterial dynamics in relation to elevated pCO_2, Grossart and colleagues (2006) found elevated CO_2 correlated with increased bacterial growth and protein production, especially in cells attached to POM.

4.1.3 Measuring the effects of elevated CO_2 using mesocosm enclosure studies

Work with bacterial cultures *in vitro* enables specific genetic, metabolic and physiological questions to be addressed. Stretton and colleagues (1996) detailed changes in gene expression pathways in a marine *Pseudomonas* sp. brought about by altered pCO_2. Genes non-essential for growth were mutated by fusion to the *LacZ* reporter gene from *E. coli*, the product of which, β-galactosidase, was expressed in a controlled manner by exposure of the cells to CO_2. The mechanism of promoter induction is reasoned to involve intracellular CO_2 partial pressure and HCO_3- concentration. Collins and Bell (2004) followed genotypic and phenotypic changes through 1000 generations of the eukaryotic alga *Chlamydomonas* sp. Artificially selected lines failed to evolve specific adaptation to pCO_2 of 1050 ppm. Other lines developed a phenotype of small cell size, increased chlorophyll concentration, increased photosynthesis and respiration. Existing neutral mutations in genes controlling carbon capture and concentration are probably at the heart of underlying genotypic selection. However extrapolating results from culture studies to mixed natural populations is problematic.

Studying enclosures of natural environments, with their contained communities, allows these communities to be studied while controlling a certain number of parameters. "Bottle" and other microcosm experiments have been used successfully to ascertain effects on marine microorganisms of anthropogenic environmental changes, including seawater acidification (for examples see Zondervan *et al.*, 2002; Schippers *et al.*, 2004). Limitations of these small enclosures, however, are unnatural physico-chemical and surface attachment effects. Significant changes to community composition and to rates of metabolism, independent of the treatment being tested, may occur within 24 hours, in what is known as the "bottle effect" (ZoBell and Anderson, 1936). Large mesocosms, especially located *in situ* rather than in the laboratory, can help reduce the bottle effect, allowing a pseudo-environmental investigation in which chosen parameters can be controlled. Bacterioplankton ecological studies have made use of mesocosms to assess: bacterial abundance and activity (Lebaron *et al.*, 2001), microbial community dynamics and changes in genetic diversity (Schäfer *et al.*, 2001), microbial organic matter production (Engel *et al.*, 2004); to follow bacteria-virus interaction in a bacterial community during the course of a viral infection (Hewson *et al.*, 2003), and effects of increased pCO_2 on bacterial communities *in situ* (Grossart *et al.*, 2006).

It is worth defining here the word "replicate" in the context of this experiment. Replicate mesocosms are replicated in their treatments. In this study (Section 4.2.1) mesocosms 1, 2 and 3, and mesocosms 4, 5 and 6 are triplicates sharing the same manipulated pCO_2 level and pH, and the same treatments of enclosure and bubbling. However, there is liability to a "founder effect" where each differs in the composition of the biota it contains, which results from an independent encapsulation of natural seawater, and in this way differs from other smaller scale experiments where the starting biota may be fully controlled; i.e. all components, biotic and abiotic, can be measured and put together wholly artificially. The two day delay between filling the mesocosms and beginning sampling, during which time CO_2 partial pressures were attained, may have produced also some degree of change in community composition due to containment effects. It follows that the bacterial composition, and thus the metagenome in each mesocosm will differ throughout the experiment, and not just owing to the experimental treatment. By comparing "replicate" mesocosms it is hoped to be able to account for this variation in initial community composition when looking for differences and similarities between treated and untreated mesocosms, using multivariate statistical analyses. Specifically, the aim is to distinguish effects on community composition due to increased pCO_2, from those due to other factors, such as founder effects. Unfortunately, due to restrictions on workload, only two replicates of each mesocosm treatment (high CO_2, mesocosms 1 and 2; ambient CO_2 mesocosms 5 and 6) were analysed, and this is below the minimum threshold of three replicates required to statistically validate the specific cause (e.g. CO_2) of any community change revealed.

4.1.4 Use of DGGE to study marine microbial communities

A small sample of previous work on marine microbial communities using DGGE includes community biogeographical distribution in the Arctic Ocean by Ferrari and Hollibaugh (1999), a study by Schäfer *et al*. (2002) of the structure of bacterial communities associated with pelagic diatoms, Rink *et al.*'s study of the effects on bacterial communities associated with a phytoplankton bloom (2007), and with a few studies employing mesocosms (for example, Schäfer *et al.*, 2001; Martínez Martínez *et al.*, 2007, and Chapter 3 of this study). A review of published experimental results did not find any studies reporting changes in marine bacterial community structure, due to ocean acidification or changes in DIC, analysed by

DGGE, a method identified as appropriate and potentially useful for assessing bacterial community dynamics (Ferrari and Hollibaugh, 1999).

It was therefore decided to study the effect on the bacterioplankton community of a pH decrease of up to 0.77 units, equivalent in a rise in pCO_2 to 750 ppm, akin to the IPCC scenario for the global surface ocean for 2100 (Houghton *et al.*, 2001), using 12,000 L *in situ* mesocosms as a proxy, and DGGE as a profiling technique.

4.1.5 Aims

The aim of this study was to use DGGE and *in situ* seawater mesocosms differentially altered to increase pCO_2 to investigate changes in the bacterial community contained therein. Further it was aimed to identify bacterial species whose contribution to the community profile was altered as a result of increased pCO_2. Finally it was aimed to describe these community changes using multivariate statistics, specifically to test the following hypothesis:

Hypothesis: 16S rRNA gene community profiles generated by PCR-DGGE using *Bacteria*-specific and group-specific primers will change due to increased pCO_2.

4.2 Method

4.2.1 Experimental setup

Six 12,000 L mesocosms were moored to an offshore raft, anchored in an 80 m water column, in the Raunefjord (60°20'N, 5°30'E), a branch of the Hardangerfjord, at the University of Bergen's marine biology field station at Esplenade, some ten kilometres southwest from Bergen. The mesocosms were filled with seawater from the surrounding fjord on 2nd May 2006. Prior to the experiment there had been blooms of *Phaeocystis* sp. and a diatom species, in March and April 2006, respectively. Transparent plastic lids were fitted over the mesocosms to allow the seawater and the airspace contained therein to equilibrate with respect to pCO_2, a

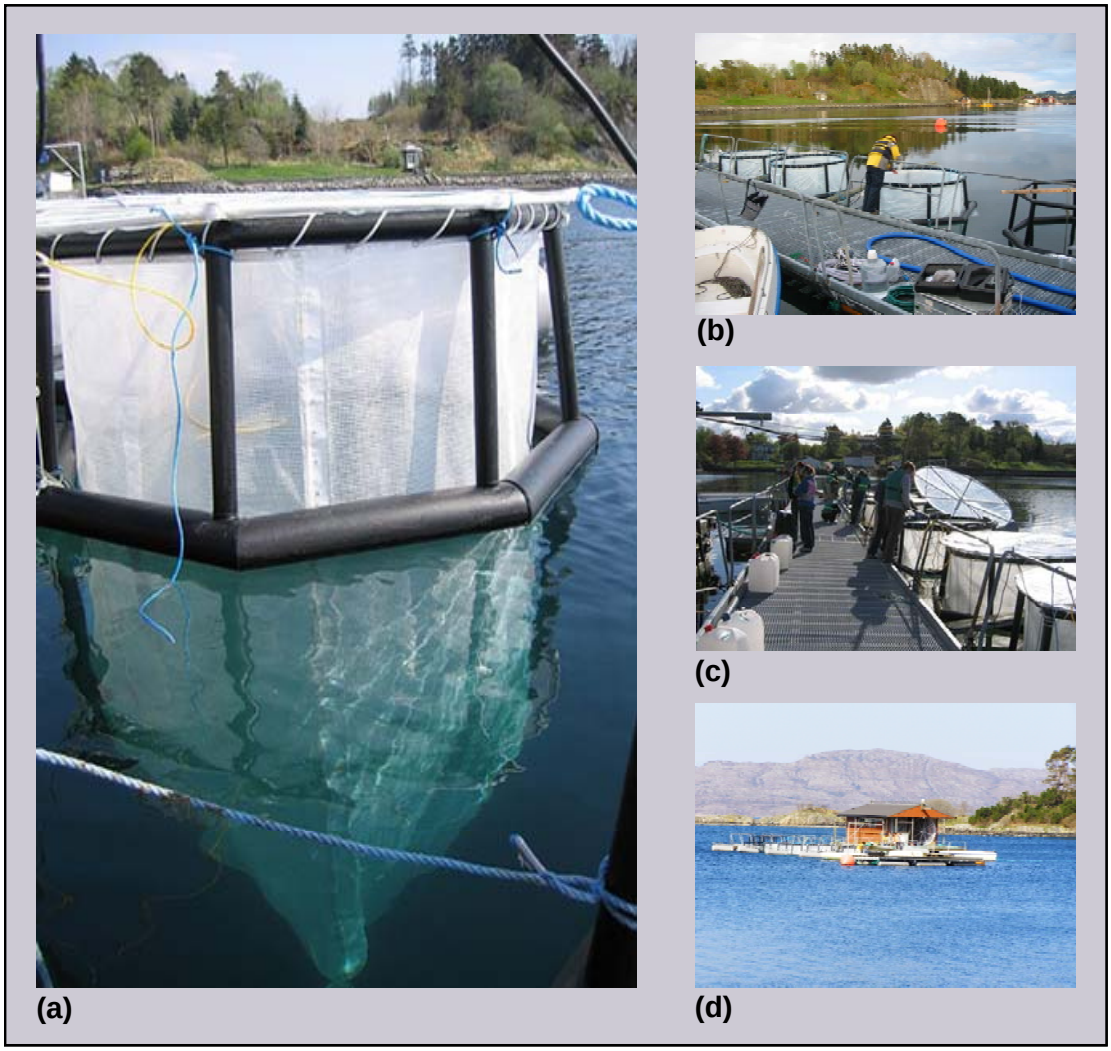


Figure 4.1. The Bergen mesocosm setup. Mesocosm with 12,000 L volume below the waterline (a), platform to which six mesocosms were moored (b, c), and mesocosm and platform complex viewed from shore (d).

supply of which was maintained to the airspace, above the acidified mesocosms. This minimised the loss of CO_2 to the atmosphere and therefore the rate of fall of pCO_2 (or rise of pH) within the enclosed water. Rainwater also was excluded, reducing potential effects of freshwater.

Three mesocosms were bubbled with CO_2 gas until the dissolved gas's partial pressure in the water was ~ 750 ppm, equivalent to that estimated for the global ocean surface for the year 2100 (Caldeira and Wickett, 2003). The remaining three mesocosms were bubbled with air to equilibrate for any effect of bubbling while maintaining ambient pCO_2 (~ 380 ppm). The six mesocosms, three replicates at elevated pCO_2 and three replicates at ambient pCO_2, labelled 1-6 were set up as detailed in Table 4.1.

Table 4.1. pCO_2 levels in experimental mesocosms

Mesocosm	***pCO_2 level (ppm)***
1	~ 750
2	~ 750
3	~ 750
4	~ 380
5	~ 380
6	~ 380

Sunny weather up until, and including, 11th May, resulting in 12 days of exposure of the mesocosms, caused stratification of the water in the mesocosms prior to the first sampling day, so it was decided to set up a continuous circulation in each with a small rotary pump.

Propeller-driven pumps, designed for use in aquaria, were inserted in the bottom of each mesocosm with the outflow attached to a 3 m pipe discharging in the upper 0.5 m. This setup provided vertical mixing, ensuring stratification did not have to be overcome when sampling. Nutrients, including nitrate (9 µmol L^{-1} NO_3^-) and phosphate (0.5 µmol L^{-1} PO_4^{3-}), were added on 5th May after attainment of desired pCO_2 levels, whereupon daily sampling began. The chronology of the experiment is illustrated in Figure 4.2.

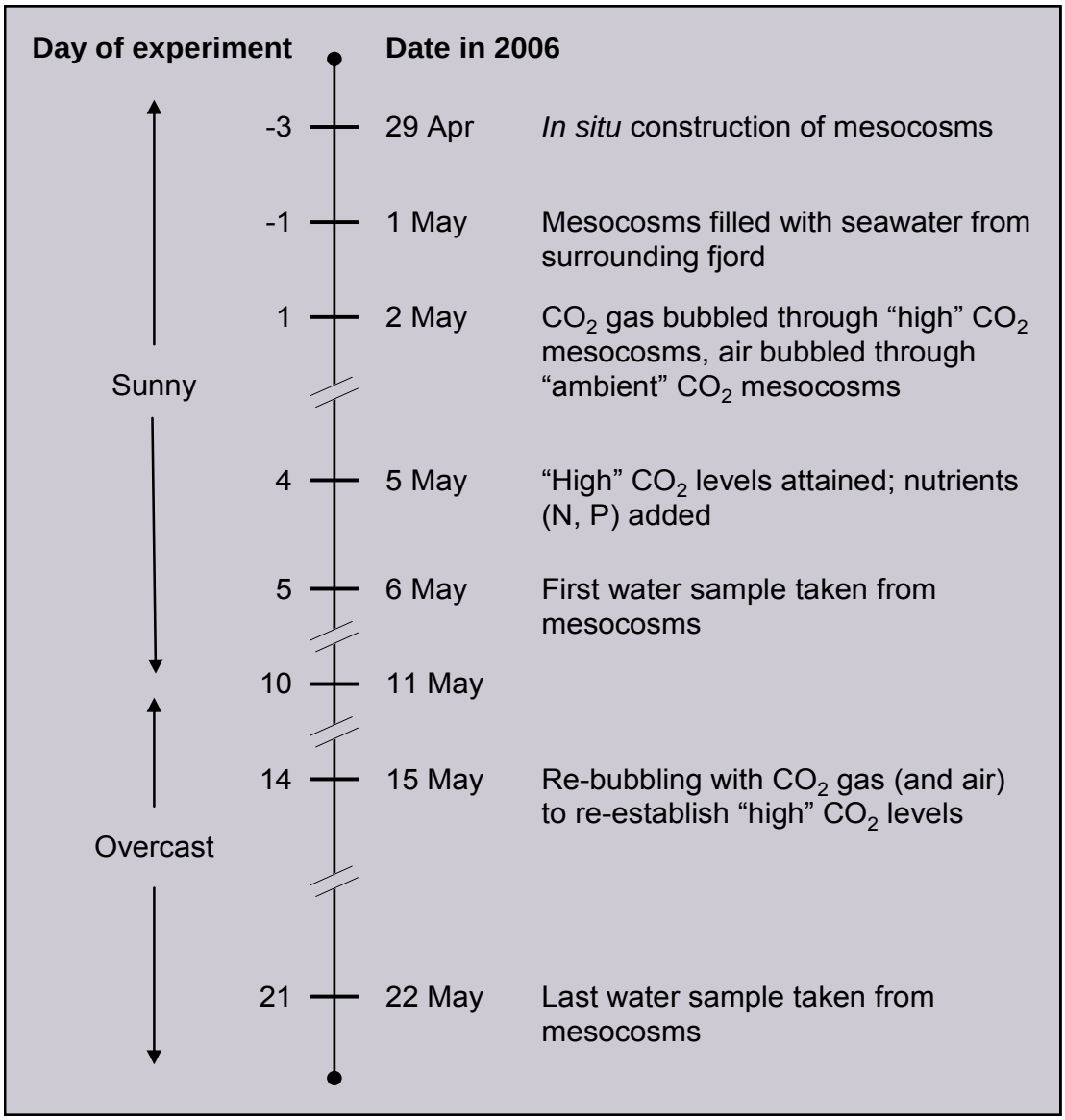


Figure 4.2. Timeline. Chronology of key events in the Bergen mesocosm experiment.

Mesocosms were re-bubbled on 15th May, the tenth sampling day of the experiment, with CO_2 (mesocosms 1-3) to regain pH levels which had risen to near-ambient, and with air (mesocosms 4-6) to avoid any bias due to bubbling effect.

The environment inside the mesocosms was controlled insofar as it was separated from the wider aquatic and atmospheric environments, and it was possible to add specific concentrations of nutrients and CO_2. Within the mesocosms all environmental parameters still varied temporally and spatially. Some intra-mesocosm biotic components, such as the nano- and picoplankton community which normally moves with the current, should vary less than their extra-mesocosm counterparts. Other abiotic components, such as temperature and salinity, may vary more inside

mesocosms relative to outside, owing to their increased susceptibility to the effects of containment (Munn, 2004).

4.2.2 Sampling and DNA extraction

Water samples from the surface 1 m were taken daily at 09:00 from all mesocosms. 16 L were retrieved using an acid-washed bucket, transferred to the site laboratory using an acid-washed carbouy, and 4 L filtered through each of four 0.22 μm Sterivex® filter cartridges, which were immediately frozen in liquid N_2 and stored at -80°C. DNA was extracted from the Sterivex filters according to the protocol in Section 2.2.1.

4.2.3 Optimisation of DGGE for mesocosm samples

Figure 4.3 (a, b, c) shows three DGGE gels consisting of three different denaturant gradients (42-53 %, 40-60 % and 32-68 % respectively). These progressively larger denaturant ranges (11% in Figure 4.3a, 20% in Figure 4.3b, and 36% in Figure 4.3c) have the effect of compressing the area on the gel within which the bands of the profiles are resolved. Each gel in Figure 4.3 contains the same series of 16S rRNA gene fragment profiles generated using the same PCR protocol using *Gammaproteobacteria*-specific primers (Table 3.1) with environmental template DNA, from individual experimental PCRs. Differences in gel profiles are due to the variation in denaturant gradient.

Sequencing of two resolved bands shows the same bands are present in each gel. Table 4.2 shows that bands 1, 3, and 5, and bands 2, 4, and 6, all have sequences which align in the Greengenes database with those of *Pseudomonas*, an aerobic rod common in the euphotic zone, and hence are likely to represent the same strain of *Gammaproteobacteria*. The six sequences are between 81 % and 88 % similar to *Pseudomonas* species sequences in the Greengenes database (http://greengenes.lbl.gov). The distance between these bands (1, 2, in Figure 4.3a) decreases in gels with higher denaturant ranges (bands 3, 4, and 5, 6, in Figures 4.3b and c respectively). Their presence in the different gels is testament to the reproducibility of results using PCR-DGGE. In gels (a) and (b) the higher initial denaturant (42 % and 40 % respectively) means some bands stop high in the gel. In gel (a) the relatively low final denaturant concentration of 53 % resulted in some bands electrophoresing through the gel completely. The compression of the

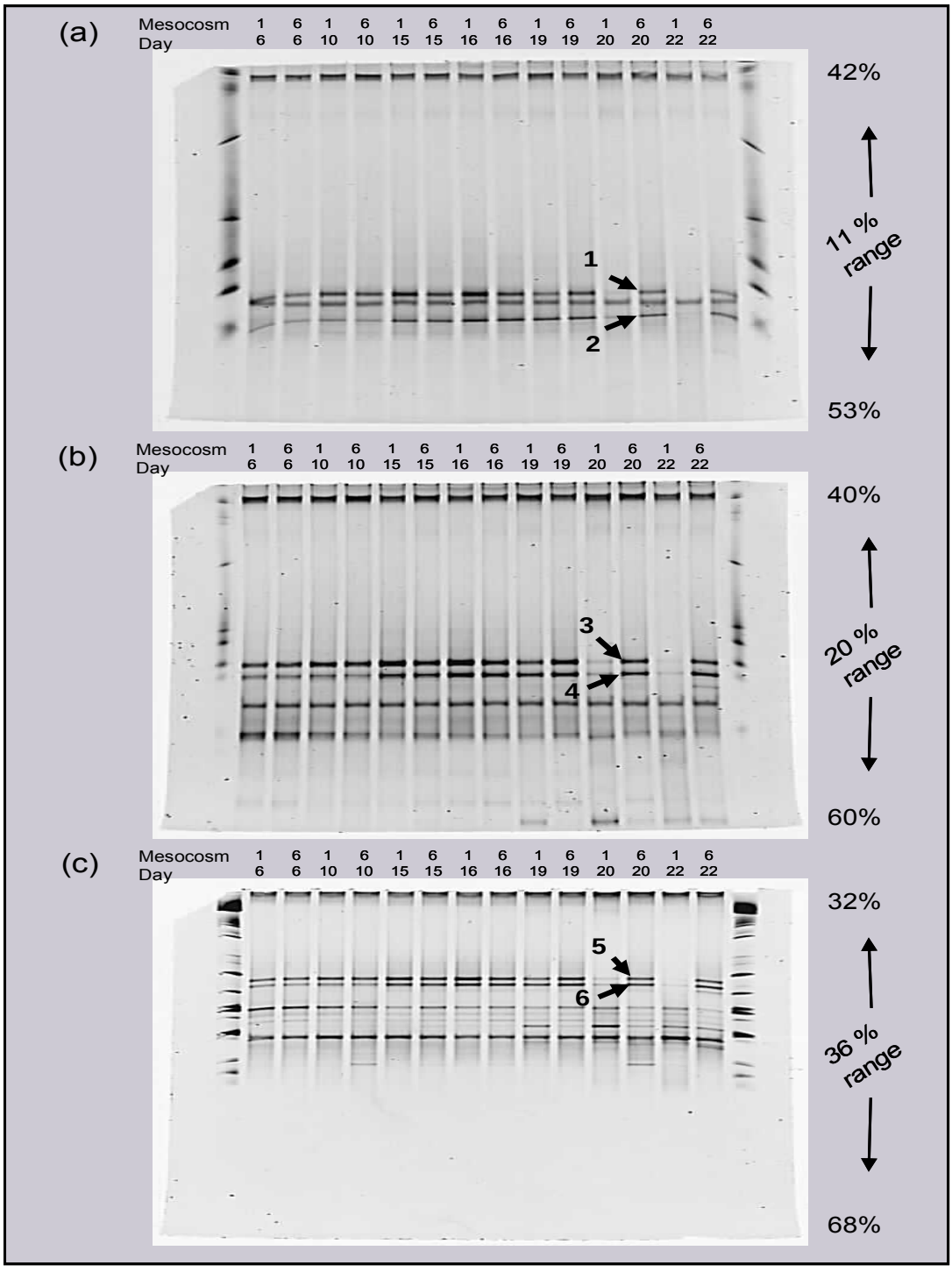


Figure 4.3. Varying band profiles are obtained using altered DGGE denaturant gradients. Three gels of varying denaturant gradients showing resolved banding profiles of PCR products obtained by a nested protocol using the *Gammaproteobacteria*-specific primer set 395f/871r and the *Bacteria* DGGE primers 518f-GC/785r, on three separate occasions using the same DNA as template for PCR. 500 ng DNA samples were loaded into each lane. Samples alternate from mesocosm 1 (high CO_2) and 6 (ambient CO_2) on 6th, 10th, 15th, 16th, 19th, 20th and 22nd May. Electrophoresis was for 17.5 hours at 60V. Numbered bands were excised and sequenced and the identities of their closest aligned database homologues are listed in Table 4.2.

Table 4.2. Phylogenetic classification of 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 4.3

Band Number	***Phylogenetic Alignment***[a]	***Similarity***[b]
1	*Gammaproteobacteria; Pseudomonadales; Pseudomonas*	82
2	*Gammaproteobacteria; Pseudomonadales; Pseudomonas*	81
3	*Gammaproteobacteria; Pseudomonadales; Pseudomonas*	84
4	*Gammaproteobacteria; Pseudomonadales; Pseudomonas*	83
5	*Gammaproteobacteria; Pseudomonadales; Pseudomonas*	88
6	*Gammaproteobacteria; Pseudomonadales; Pseudomonas*	85

[a] Phylogenetic classification, where possible to the level of genus. Sequences from excised bands were aligned with, and compared to, homologous 16S rRNA gene sequences in the Greengenes database (http://greengenes.lbl.gov) and classified accordingly.
[b] Degree of base similarity between an excised DGGE band sequence and that of its closest fully-aligned homologue in the Greengenes database (http://greengenes.lbl.gov).
* Base sequences for 16S rRNA gene fragments, amplified from excised DGGE bands, are listed in Appendix I.

band-resolving area within the gel, enabling more bands per profile to be visualised, as shown in Figure 4.3c, allows optimisation of the denaturant range. For the *Gammaproteobacteria*-specific PCR products used here the optimum is somewhere between that for the gel in Figure 4.3c which likely shows all bands in a compressed profile, and Figure 4.3b which shows a more expanded profile but in which some bands may have been lost due to excessive electrophoresis. The comparability of *Gammaproteobacteria* DGGE gels was borne in mind when deciding all further experimental gels would possess a denaturant gradient range of 40 % to 55 %. The same was decided for the *Cyanobacteria* specific group. The remaining groups were assigned the range of 40 % to 60 %.

4.2.4 Analysis of DGGE bands, and band profiles

DGGE was performed, according to methods described (Section 2.4). 16S rRNA gene fragment sequences, taken from bands excised from DGGE gels, were identified as set out in Section 2.3.6.1, while DGGE band profiles were converted to digital binary matrices (using Phoretix 2D) and statistically analysed (Section 2.4.4.2) using the multivariate analysis software package PRIMER-E (Clarke and Gorley, 2006). In addition to MDS analysis (Section 2.4.4.2) of temporal change in each mesocosm's bacterial diversity, second stage MDS was performed to compare both inter-mesocosm development, and inter-group development within each mesocosm.

4.2.4.1 Second stage MDS

In Section 4.4.2.4 two second stage MDS analyses are performed using PRIMER-E (for method see Clarke and Warwick, 2001). Firstly the temporal development of each group-specific community, within the different mesocosms, is compressed to a single point on a two dimensional graph for each specific group. This allows easy comparison of the development of each specific group between high CO_2 and ambient CO_2 mesocosms (Figure 4.15). Secondly, for each mesocosm, specific group temporal profile development is compressed to a single point. A plot for each mesocosm allows comparison of group-specific profile development in relation to that for the other groups (Figure 4.16).

4.3 Results - "background" data from the mesocosms

Figure 4.4 charts the temporal changes in key physical parameters in high and ambient CO_2 mesocosms, represented by mesocosm 1 and mesocosm 6 respectively. Spatial variation was considered insignificant since all mesocosms were constantly mixed, after temperature stratification was discovered in the mesocosms prior to the first sampling day. All physical measurements plotted on the graphs in Figure 4.4 are means of readings on each day from the surface, middle and bottom water layers in each mesocosm.

Figure 4.4a shows the partial pressure of CO_2 in mesocosm 1 was highest on 5th May at ~ 750 ppm and at ~ 600 ppm on 15th May after the mesocosm was bubbled for a second time with pure CO_2 gas, following a fall in the partial pressure to near ambient level. These levels remained for seven to eight days in each case whereupon the levels began to fall steeply. In mesocosm 6 (control) pCO_2 level averaged around 300 ppm until 13th May then fell to below 200 ppm, after which it stabilised at ~ 200 ppm following the second bubbling with air to maintain the replication effect between mesocosms. Figure 4.4b shows that the pH values mirrored pCO_2 trends, being steady at ~ pH 8.05 in mesocosm 6 for seven days then rising to pH 8.3 prior to re-bubbling on 15th May, and then steadily falling back to pH 8.1 by 23rd May. In mesocosm 1 this pattern was copied at an average level of 0.5 pH units lower. pH averaged around 7.7 for 6 days then began to rise between 11th and 15th May to 7.95. After re-bubbling a pH of 7.7 was regained and maintained between 16th and 23rd May.

Salinity was constant at 36.1 ‰ and is not charted here. Water temperature was the other abiotic parameter measured which was not altered deliberately, and showed little variation between mesocosms, as shown in Figure 4.4c. In mesocosms 1 and 6 a temperature peak of 11°C was observed on 10th and 11th May following sunny weather. Ensuing overcast conditions saw temperatures fall to below 9°C by 15th May, and subsequently rise again to 10°C on 22nd May.

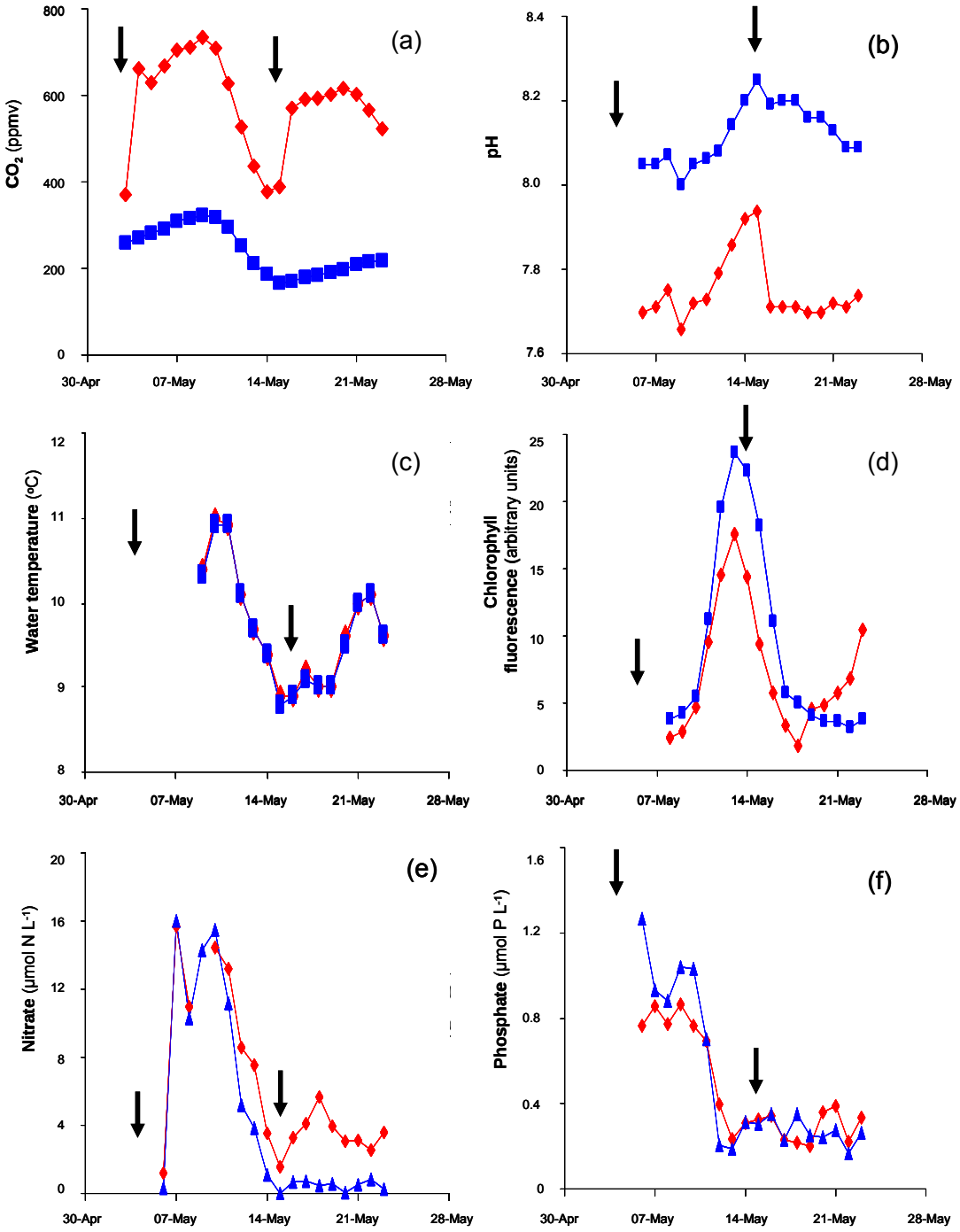


Figure 4.4. Abiotic environmental parameters, and chlorophyll fluorescence, measured in mesocosm 1 (high CO_2, red diamonds) and mesocosm 6 (ambient CO_2, blue squares/triangles) throughout the duration of the 2006 Norwegian mesocosm experiment. Bubbling with CO_2 gas to attain high pCO_2 levels, and with air in ambient CO_2 mesocosms, took place on 4th May and again on 15th May, indicated by black arrows. All mesocosms were fertilised with nitrate and phosphate on 5th May prior to sampling, which began on 6th May. Water temperatures (c) were the mean of readings from the surface, 1, 2 and 3 m. These data were supplied courtesy of I. Joint of PML.

Nitrate (9 μmol L^{-1} NO_3^- [final conc.]) and phosphate (0.5 μmol L^{-1} PO_4^{3-} [final conc.]) were added to the mesocosms on 6th May to stimulate phytoplankton and photosynthetic bacterioplankton blooms. Figure 4.4e shows nitrate levels rose from near zero to 16 μmol L^{-1} on 7th May. The level remained above 12 μmol L^{-1} for four days and then fell to near zero again by 15th May, where it remained in mesocosm 6 (ambient CO_2) for the rest of the experiment. Nitrate level in mesocosm 1 (high CO_2) did not fall as low as that in mesocosm 6, reaching only 2 μmol L^{-1} on 15th May and rising slightly to around 4 μmol L^{-1} for the rest of the experiment. Phosphate levels, illustrated in Figure 4.4f, behaved in a similar way to nitrate levels, only at lower overall concentrations, phosphate being a micronutrient, rather than a macronutrient. Phosphate levels peaked at around 1.0 μmol L^{-1} on 7th May for four days then fell to around 0.2 μmol L^{-1} by 12th May, remaining at this level for the duration of the experiment in both mesocosms 1 and 6.

The biotic components of the environments within mesocosms 1 (high CO_2) and 6 (ambient CO_2) are shown in Figure 4.5, and Figure 4.4d. Figure 4.4d shows overall chlorophyll-containing photosynthetic microorganisms (phytoplankton, represented by coccolithophores and cryptophytes in Figure 4.5a and 4.5b respectively, and cyanobacteria represented by the *Synecchococcus* genus in Figure 4.5c) increased their concentration by a factor of 10 within seven days of nutrient addition in mesocosm 6 (ambient CO_2) and by a factor of 7 in mesocosm 1 (high CO_2) (Figure 4.4d). This was followed by a crash in numbers over the next six days to starting levels, illustrated especially well in Figures 4.5a, b and f. The photosynthetic eukaryotes *Cryptophtya* (Figure 4.5b) and *Coccolithophoridaceae* (division *Prymnesiophyta*) (Figure 4.5a), the large and small subdivisions of the picoeukaryotes (0.2 μm to 2.0 μm) (Figures 4.5e and f), and the cyanobacteria belonging to the genus *Syncechococcus* (Figure 4.5c) all increased in concentration by a factor of between 3 and 4 in the ambient CO_2 mesocosm 6 post-nutrient addition, but only by a factor of 2 or less in the high CO_2 mesocosm 1. Although non-specific the "picoeukaryotes" comprise largely photosynthetic autotrophs, and as shown in Figures 4.5e and f, their growth patterns, especially that of the "large" fraction (Figure 4.5f), mimic that of the larger photosynthetic autotrophs (Figures 4.5a and b). Re-bubbling mesocosms on 15th May led to a crash in all the eukaryotic populations. However, as shown in Figure 4.5c, *Synechococcus* experienced only a decline in the growth rate after bubbling and continued increasing in numbers to the end of the experiment, albeit at a ~ 50 %

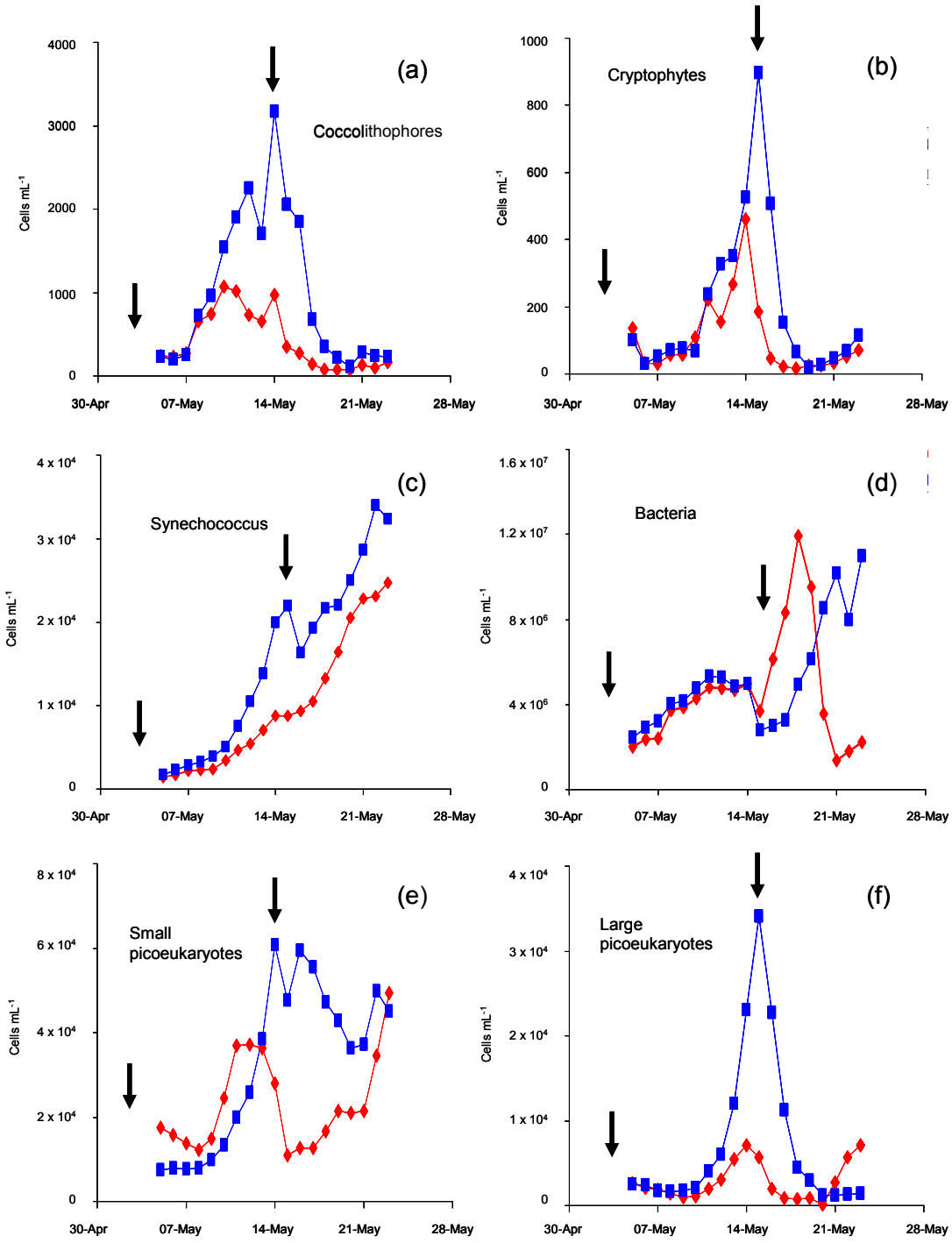


Figure 4.5. Biotic environmental parameters measured in mesocosm 1 (high CO_2, red diamonds) and mesocosm 6 (ambient CO_2, blue squares) throughout the duration of the 2006 Norwegian mesocosm experiment. Bubbling with CO_2 gas to attain high pCO_2 levels, and with air in ambient CO_2 mesocosms, indicated by black arrows, took place on 4th May and again on 15th May. All mesocosms were fertilised with nitrate and phosphate prior to sampling on 5th May. These data were supplied courtesy of I. Joint of PML.

higher level in mesocosm 6 (ambient CO_2) compared to mesocosm 1 (high CO_2). Figure 4.5d shows heterotrophic bacteria grew slowly from ~ 2 x 10^6 cells mL^{-1} to ~ 5 x 10^6 cells mL^{-1} by 15th May. After re-bubbling on 15th May the populations increased rapidly, initially in the high CO_2 mesocosm, which subsequently declined after three days. Following stability for three days between 15th May and 17th May at ~ 2 x 10^6 cells mL^{-1} there was a rise in cell numbers in mesocosm 6 (ambient CO_2) which continued to the end of the experiment, reaching 1.2 x 10^7 cells mL^{-1}.

There was a general pattern shown by all the growth curves illustrated in Figures 4.5 and 4.4d of slower growth and smaller final population attained in mesocosm 1, the high CO_2 environment. Exceptions to this were the heterotrophic bacteria (Figure 4.5d) which responded faster in mesocosm 1 (high CO_2), and the "small" picoeukaryotes (Figure 4.5e) which did likewise, but subsequently attained a greater final population size in mesocosm 6 (ambient CO_2) by over 50% compared to mesocosm 1 (high CO_2). Grossart *et al*. (2006) showed bacterial abundance declined as chlorophyll a (Chl *a*) rose. The inverse of this relationship is demonstrated here when bacterial abundance rises rapidly with the decline of Chl *a* following re-bubbling (Figures 4.4d and 4.5d). The implication is that when phytoplankton is in exponential growth the bacteria exhibit no net growth. They subsequently grow rapidly when microalgae die, feeding on the products of decay such as dissolved free fatty and amino acids. This is supported by observations that extracellular enzyme activity correlates with the amount of free organic substrate in the water (Ferrari and Hollibaugh, 1999; Grossart *et al.*, 2006). Similar increases in abundance of total bacteria in high CO_2 and ambient CO_2 mesocosms up to re-bubbling on the 15th May (Figure 4.5d) matches the finding of Grossart et al. (2006) using a similar mesocosm setup, that total bacterial abundance is independent of pCO_2.

4.4 Results - bacterial diversity

4.4.1 Testing the degree of replication in similarly-treated mesocosms

The data from this experiment come from two replicates of each of two different mesocosm treatments; namely those with pCO_2 elevated to ~750 ppm (high CO_2) – mesocosms 1 and 2 – and CO_2 at an untreated ~380 ppm (ambient CO_2) – mesocosms 5 and 6. This number of replicates (two) is not enough to draw statistically significant conclusions. However, valid statements about the observations made here, based on two duplicate data sets, can be reasonably and logically supported.

To test the similarity within, and dissimilarity between, replicated and dissimilarly treated mesocosm communities respectively, cluster analysis was performed on 16S rRNA gene fragment profiles from four mesocosms. DGGE profiles for communities from high CO_2 mesocosms 1 and 2, produced on a single gel, were combined with those for communities from ambient CO_2 mesocosms 5 and 6, using the gel image analysis software Gelcompar II® (Applied Maths, Sint-Martens-Latem, Belgium) (Figure 4.6a). Profiles were analysed and normalised using visually obvious reference bands common to all gels.

Changes in the patterns of bands can be seen in the DGGE profiles in the different mesocosms. Three such temporal changes are indicated in Figure 4.6a. The band labelled "a" appears in profiles of all mesocosms at some point between 10th and 17th May. The band labelled "c" is present in all profiles on 10th May and subsequently disappears by 19th May. The band labelled "b" is present in profiles of mesocosms 1, 2 and 5, but not 6. DNA from bands "a", "b" and "c" was sequenced and all aligned with the Gram-positive family *Flavobacteraceae* within the phylum *Firmicutes* (see equivalent bands in Figure 4.8d).

The twenty DGGE profiles in Figure 4.6a were digitised and arranged into a similarity matrix using Gelcompar II®. Using PRIMER-E the relationships between them were then visualised in a dendrogram (Figure 4.6b) and MDS plot (Figure 4.6c).

The dendrogram shows all profiles from mesocosms 1 and 2 clustered together as would be expected since these mesocosms were both high CO_2. However profiles for 10th May (the start of the experiment) for mesocosms 5 and 6 (ambient CO_2) cluster together, but separately from the equivalent two profiles for mesocosms 1 and 2 (high CO_2). If the two gels were completely normalised, and therefore fully comparable, the four profiles from 10th May should cluster together, since the four

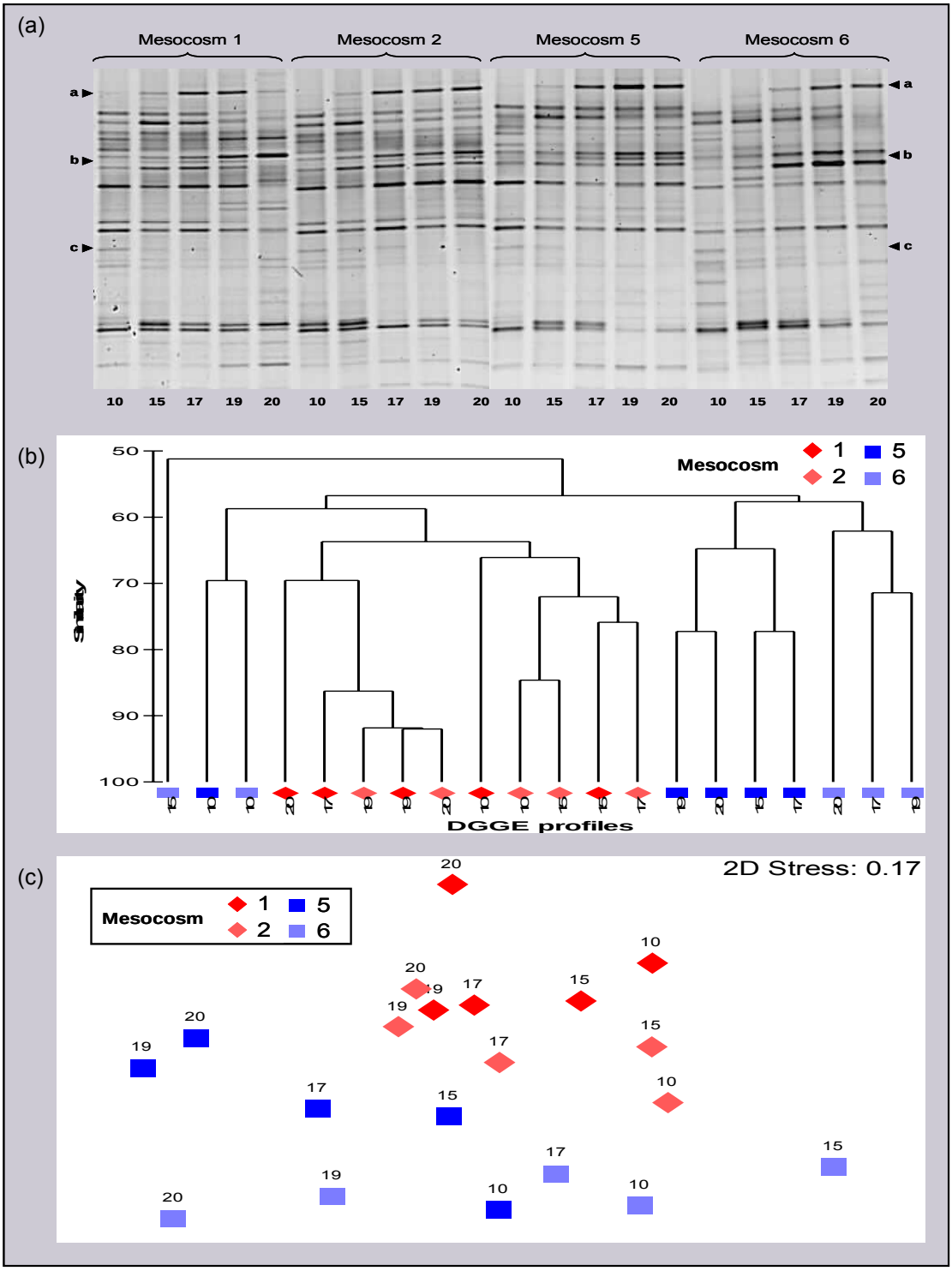


Figure 4.6. Analysis of profiles from separate DGGE gels**.** (a) DGGE profiles of 16S rRNA gene fragments amplified from DNA from four mesocosms using *Bacteria*-specific primers. DNA from mesocosms 1 and 2 was run on a single gel, as was DNA from mesocosms 5 and 6. The two gels were aligned and normalised using Gelcompar II®. Temporal changes to the band patterns in some or all mesocosm profiles are indicated a, b, and c. (b) A similarity matrix was made using all twenty profiles, and from this a cluster dendrogram, and (c) MDS plot, were constructed using PRIMER-E. Sample labels in a, b, and c refer to the date in May 2006.

communities would all be similar (at least this would be the case if all mesocosms were true replicates, and differed only in the alteration made to the pCO_2). So these early profiles are clustering due to the gel they are from, rather than due to the CO_2 treatment of the mesocosms.

4.4.2 Comparison of *Bacteria*-specific and seven group-specific DGGE profiles from high and ambient CO_2 mesocosms

4.4.2.1 Inspection of DGGE gels and taxonomic identification of certain bands (ribotypes)

16S rRNA gene fragment DGGE band profiles constitute a genetic fingerprint representative of the bacterial community's structure in the original water sample from which environmental DNA was extracted. This metagenomic community structure is hypothesised to change as a result of prolonged elevation of pCO_2. DGGE of 16S rRNA gene fragments was employed to test this hypothesis. Group-specific 16S rRNA gene fragments were amplified by nested PCR (Section 3.2.2) from total environmental DNA preparations from seven time points throughout the course of the experiment from both high and ambient CO_2 mesocosms. The resulting DGGE profiles were examined for temporal divergence in community structure between the high and low CO_2 mesocosms. For each specific bacterial group two gels were produced: one comparing mesocosms 1 and 6, and one comparing the "replicate" mesocosm pair 2 and 5. These DGGE gels are reproduced in Figures 4.7 and 4.8.

Observation of the comparative 16S rRNA gene fragment profiles from the high CO_2 and ambient CO_2 mesocosms in the DGGE gels in Figures 4.7 and 4.8 reveals varying degrees of temporal changes in bacterial community composition. Taxonomic identification of the 16S rRNA gene fragments of a selection of sequenced bands (Tables 4.3 and 4.4) added a qualitative aspect to these developments (see Section 2.4.3 for band sequencing method, and Section 2.3.6.1 for sequence identification methods).

Tables 4.3 and 4.4 show identification details of the organisms possessing the ribosomal genes, the fragments of which are numbered in the DGGE profiles in Figures 4.7 and 4.8 respectively. Phylogenetic classification was done by aligning sequences in the Greengenes 16S rRNA gene database (Section 2.3.6.1) and can be compared to the less robust identification by BLASTn (Section 2.3.6.1).

While disagreement between the two methods for the taxonomic identity of a sequence appears at first to be widespread in Tables 4.3 and 4.4, closer inspection reveals greater agreement. The phylogenetic (Greengenes) classification is often only to the level of family or order, and while the nearest BLASTn isolate is always a named or unnamed species, it is usually within the same clade specified by the phylogenetic alignment. Gross disparities between the two methods are rare and the result of poor similarity between the sequences comprising the closest alignment using either method (e.g.: Table 4.4, band 109).

Alphaproteobacteria. The DGGE gels showing bands resolved from a PCR product of "*Alphaproteobacteria*-specific" primers, reproduced in Figure 4.7a, show up to 15 distinguishable bands with no obvious temporal or inter-mesocosm profile variation. Alignment of sequences from excised bands showed them all to belong to the *Alphaproteobacteria*, except that of band 19 which weakly aligned with the gammaproteobacterial genus *Psychrobacter*,as detailed in Table 4.3. Most sequences excised from bands aligned within the *Rhodobacteraceae*.

Betaproteobacteria. The profiles generated with "*Betaproteobacteria*-specific" primers shown in the gels in Figure 4.7b consist of up to six prominent bands. Some bands, such as band 23, show a temporal decline in intensity in the high CO_2 mesocosm 1. This indicates a decrease in abundance of this ribotype. All sequences amplified from bands in Figure 4.7b aligned, using Greengenes, with those of *Betaproteobacteria.* The only exception was the sequence from band 23 (Table 4.3), the most intense band which, following BLASTn analysis, matched the *Gammaproteobacteria*, *Nitrosococcus* sp. This high specificity correlates well with that achieved by the primers in testing in coastal waters (100%), and contrasts with the poor specificity (30%) when used to amplify DNA from an ocean gyre (see Section 3.3.2). The most prominent bands (28, 29, and 30) closely aligned with the methanol-degrading genus *Methylophilus*.

Gammaproteobacteria. In the "*Gammaproteobacteria*-specific" DGGE gels in Figure 4.7c all bands matched to the *Gamma* subgroup of the *Proteobacteria*. There are up to 12 prominent bands, with one, represented by band 36, with 100 % sequence homology, following BLASTn analysis, to a SAR86 member (Pham *et al.*, 2008).

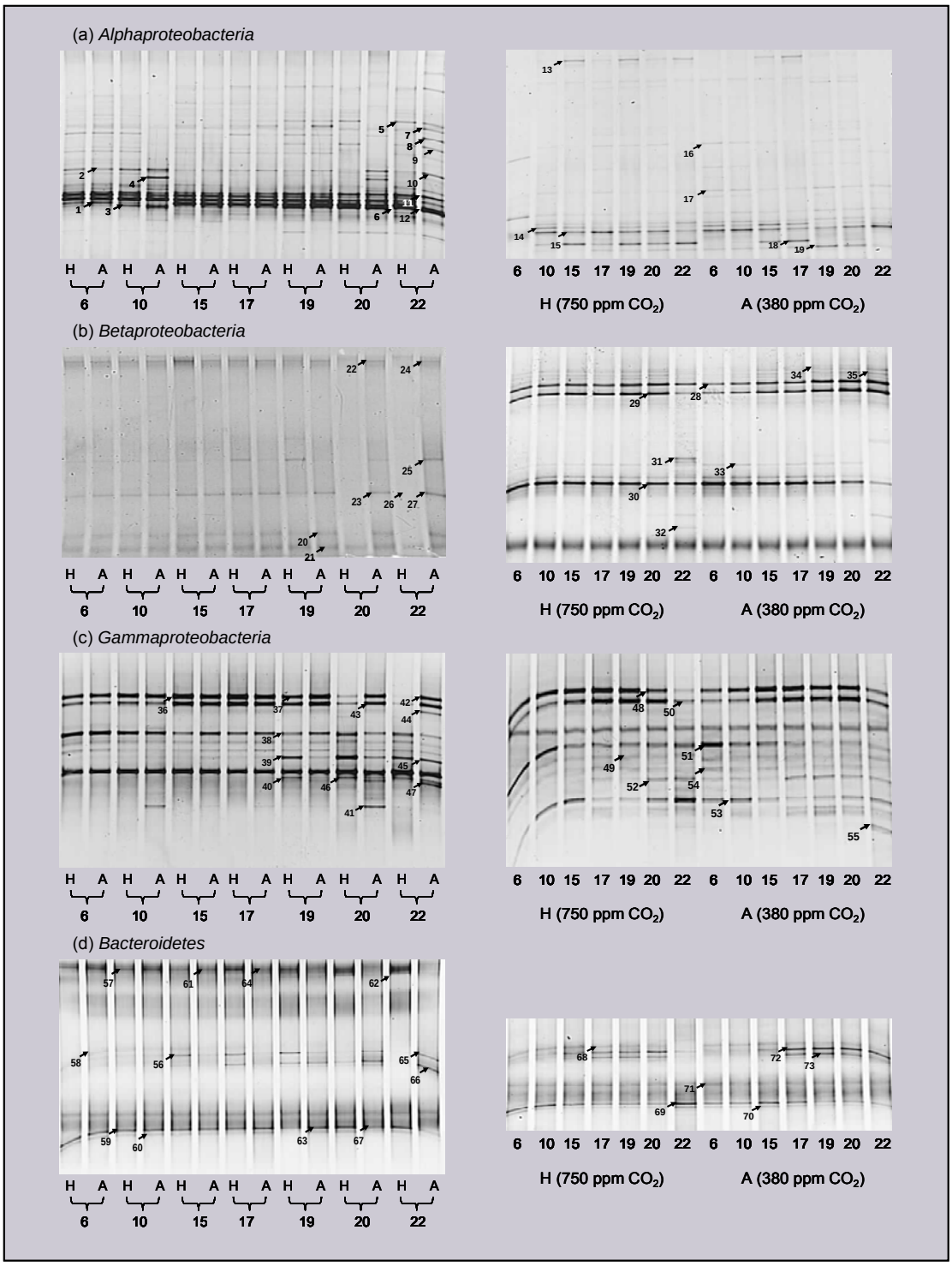


Figure 4.7. Comparison of group-specific bacterial community DGGE profiles I. Bands are homologous sequences of 16S rRNA gene fragments amplified by PCR (see Section 3.2.2 for details of group-specific PCR). Gels on the left hand side compare PCR products amplified from environmental DNA templates from mesocosms 1 and 6, those on the right hand side from mesocosms 2 and 5. Left and right-hand gel gradient ranges are not equal, resulting in differential spread of bands in equivalent date profiles. H, high CO_2; A, ambient CO_2. Identities of numbered bands were established by comparison with published sequences and are listed in Table 4.3.

Table 4.3. Phylogenetic classification of, and closest cultured isolate to, 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 4.7

Band Number	***Phylogenetic Alignment***[a]	***Similarity***[b]	***Nearest Isolate***[c]	***Max. Ident.***[d]
1	*Alphaproteobacteria; Roseovarius*	90	*Roseobacter* sp. CSQ-2	95
2	*Alphaproteobacteria;* unclassified *Rhodobacteraceae*	93	None	
3	*Alphaproteobacteria;* unclassified *Rhodobacteraceae*	94	None	
4	*Alphaproteobacteria; Dinoroseobacter*	98	*Jannaschia* sp. CSQ-9	98
5	None		*Rhizobium* sp. S133	83
6	None		*Ahrensia* sp. DFL-44	88
7	None		*Roseobacter denitrificans*	87
8	None		*Rhodobacteraceae* HP47a	92
9	None		*Rhodobacteraceae* HP47a	94
10	*Alphaproteobacteria;* unclassified *Rhodobacteraceae*	92	*Roseobacter* sp. Do-34	93
11	*Alphaproteobacteria;* unclassified *Rhodobacteraceae*	94	*Rhodobacteraceae* HP47a	92
12	*Alphaproteobacteria;* unclassified *Rhodobacteraceae*	95	*Roseobacter* sp. 3008	94
13	*Alphaproteobacteria; Acetobacter*	65	*Roseobacter* sp. LE17	81
14	*Alphaproteobacteria;* unclassified *Rhodobacteraceae*	96	*Roseobacter* sp. LE17	93
15	n/s			
16	*Alphaproteobacteria;* unclassified *Rhodobacteraceae*	78	*Roseobacter* sp. LE17	85
17	n/s			
18	n/s			
19	*Gammaproteobacteria; Psychrobacter*	62	*Roseobacter* sp. LE17	84
20	None		*Methylophilus* sp. EM8	97
21	None		*Janthinobacterium lividum*	93
22	n/s			
23	None		*Nitrosococcus* sp. XY-F	100
24	None		*Nitrosomonas* sp. Nm59	91
25	None		*Nitrosomonas* sp. Nm59	95
26	None		*Comamonas* sp. WW1	97
27	*Betaproteobacteria; Burkholderia*	85	*Methylophilus* sp. EM8	91

Table 4.3. Phylogenetic classification of, and closest cultured isolate to, 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 4.7

Band Number	***Phylogenetic Alignment***[a]	***Similarity***[b]	***Nearest Isolate***[c]	***Max. Ident.***[d]
28	*Betaproteobacteria; Methylophilus*	94	*Methylophilus* sp. EM8	97
29	*Betaproteobacteria; Methylophilus*	93	*Methylophilus* sp. EM8	96
30	*Betaproteobacteria; Methylophilus*	89	*Ralstonia* sp.	93
31	*Betaproteobacteria; Cupriavidus*	86	*Cupriavidus pauculus*	93
32	n/s			
33	*Betaproteobacteria; Cupriavidus*	88	*Cupriavidus campinensis*	91
34	*Betaproteobacteria; Methylophilus*	97	*Methylophilus* sp. EM8	97
35	*Betaproteobacteria; Methylophilus*	96	*Methylophilus* sp. EM8	98
36	*Gammaproteobacteria; Pseudomonas*	88	None	
37	*Gammaproteobacteria; Pseudomonas*	85	None	
38	Unclassified *Bacteria*	88	*Azotobacter chroococcum*	91
39	Unclassified *Bacteria*	93	None	
40	*Gammaproteobacteria; Pseudomonas*	89	*Pseudomonas libanensis*	92
41	*Gammaproteobacteria; Halomonas*	90	*Pseudomonas putida*	91
42	*Gammaproteobacteria; Pseudomonas*	85	None	
43	*Gammaproteobacteria; Pseudomonas*	86	None	
44	Unclassified *Bacteria*	88	None	
45	Unclassified *Bacteria*	90	None	
46	*Gammaproteobacteria; unclassified*	92	None	
47	*Gammaproteobacteria; Pseudomonas*	91	None	
48	*Gammaproteobacteria;* unclassified	82	None	
49	*Gammaproteobacteria; Pseudomonas*	82	None	
50	n/s			
51	n/s			
52	*Gammaproteobacteria;* unclassified	91	None	
53	n/s			
54	*Gammaproteobacteria;* unclassified	74	*Pseudomonas tuomuerense*	82

Table 4.3. Phylogenetic classification of, and closest cultured isolate to, 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 4.7

Band Number	*Phylogenetic Alignment*[a]	*Similarity*[b]	*Nearest Isolate*[c]	*Max. Ident.*[d]
55	*Gammaproteobacteria; Halomonas*	92	*Halomonas* sp. H21.2	92
56	n/s			
57	*Bacteroidetes;* unclassified *Flavobacteriaceae*	97	*Formosa algae*	99
58	*Bacteroidetes;* unclassified *Cryomorphaceae*	82	*Coccinistipes vermicola*	90
59	*Bacteroidetes;* unclassified *Cryomorphaceae*	82	*Coccinistipes vermicola*	82
60	*Bacteroidetes;* unclassified *Flavobacteriaceae*	77	*Coccinistipes vermicola*	84
61	*Bacteroidetes; Ulvibacter*	90	*Winogradskyella thalassocola*	91
62	*Bacteroidetes; Ulvibacter*	90	*Winogradskyella thalassocola*	91
63	*Bacteroidetes; Aquimarina*	83	*Marinoxanthomonas ophiurae*	86
64	*Bacteroidetes;* unclassified *Flavobacteriaceae*	97	*Formosa algae*	98
65	n/s			
66	n/s			
67	n/s			
68	*Bacteroidetes;* unclassified *Flavobacteriaceae*	93	*Olleya marilimosa*	94
69	*Bacteroidetes; Ulvibacter*	93	*Gilvibacter sediminis*	96
70	*Bacteroidetes; Zobellia*	81	*Coccinistipes vermicola*	89
71	*Bacteroidetes;* unclassified *Flavobacteriaceae*	97	*Olleya marilimosa*	96
72	*Bacteroidetes; Brumimicrobium*	82	*Marinoxanthomonas ophiurae*	85
73	*Bacteroidetes;* unclassified *Flavobacteriaceae*	95	*Olleya marilimosa*	95

[a] Phylogenetic classification, where possible to the level of genus. Sequences from excised bands were aligned with, and compared to, homologous 16S rRNA gene sequences in the Greengenes database (http://greengenes.lbl.gov) and classified accordingly.
[b] Degree of base similarity between an excised DGGE band sequence and that of its closest fully-aligned homologue in the Greengenes database (http://greengenes.lbl.gov).
[c] Nearest matched sequence from a laboratory cultured organism, within the top 100 most closely matched sequences from the Genbank database, following alignment using the BLASTn tool.
[d] Degree of invariance between sample sequence and nearest matched sequence in the Genbank database, expressed as a percentage, where 100 % equals complete invariance.
[*] Base sequences for 16S rRNA gene fragments, amplified from excised DGGE bands, are listed in Appendix I.

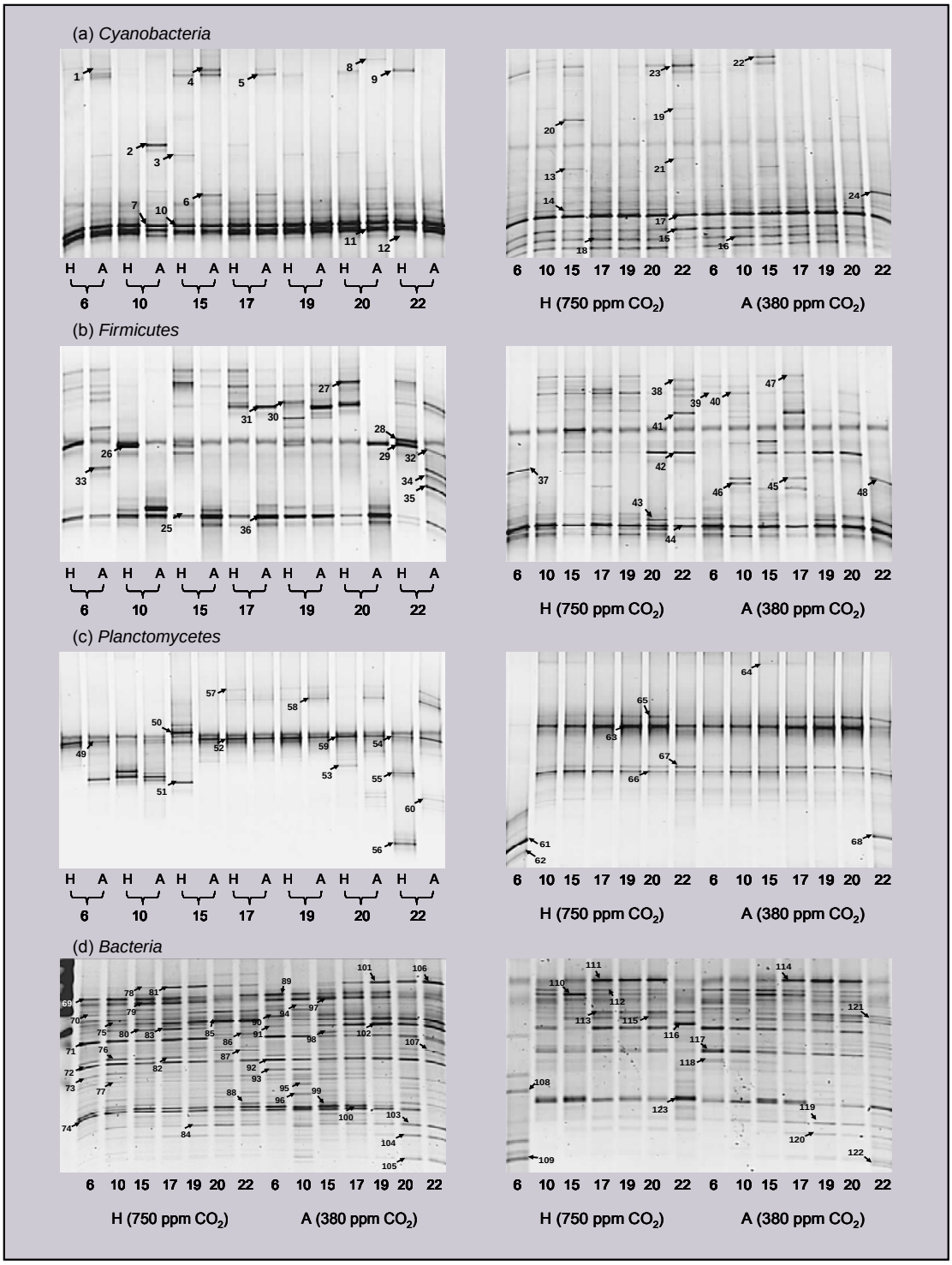


Figure 4.8. Comparison of group-specific bacterial community DGGE profiles II. Bands are homologous sequences of 16S rRNA gene fragments amplified by PCR (see Section 3.2.2 for details of group-specific PCR). Gels on the left hand side compare PCR products amplified from environmental DNA templates from mesocosms 1 and 6, those on the right hand side from mesocosms 2 and 5. Left and right-hand gel gradient ranges are not equal, resulting in differential spread of bands in equivalent date profiles.H, high CO_2; A, ambient CO_2. Identities of numbered bands were established by comparison with published sequences and are listed in Table 4.4.

Table 4.4. Phylogenetic classification of, and closest cultured isolate to, 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 4.8

Band Number	***Phylogenetic Alignment***[a]	***Similarity***[b]	***Nearest Isolate***[c]	***Max. Ident.***[d]
1	*Cyanobacteria*; Family II; GpIIa	99	*Synechococcus* sp. Almo3	88
2	*Cyanobacteria*; Family I; GpI	72	*Emiliania huxleyi* CCMP 373	89
3	*Cyanobacteria*; Family II; GpIIa	96	*Synechococcus* sp. Almo3	96
4	*Cyanobacteria*; unclassified	92	*Dinophyceae* sp. W5-1	95
5	*Cyanobacteria*; unclassified Chloroplast	95	*Chrysochromulina* sp. MBIC10518	97
6	*Cyanobacteria*; Family II; GpIIa	98	*Synechococcus* sp. Almo3	98
7	*Cyanobacteria*; Family VIII; GpVIII	79	*Synechococcus* sp. Almo3	98
8	*Cyanobacteria*; Family VIII; GpVIII	74	*Porphyra yezoensis*	84
9	*Cyanobacteria*; unclassified Chloroplast	96	*Imantonia* sp. MBIC10497	99
10	*Cyanobacteria*; Family II; GpIIa	96	*Synechococcus* sp. Almo3	99
11	*Cyanobacteria*; Family II; GpIIa	96	*Synechococcus* sp. Almo3	95
12	*Cyanobacteria*; Family II; GpIIa	98	*Synechococcus* sp. Almo3	98
13	*Cyanobacteria*; Chloroplast; *Cryptomonadaceae*	83	*Synechococcus* sp. M16.17	89
14	*Cyanobacteria*; Family II; GpIIa	95	*Synechococcus* sp. Almo3	98
15	*Cyanobacteria*; Family II; GpIIa	98	*Synechococcus* sp. Almo3	98
16	*Cyanobacteria*; Family II; GpIIa	98	*Synechococcus* sp. Almo3	98
17	*Cyanobacteria*; Family II; GpIIa	98	*Synechococcus* sp. Almo3	99
18	*Cyanobacteria*; Family II; GpIIa	98	*Synechococcus* sp. Almo3	99
19	*Cyanobacteria*; Family II; GpIIa	93	*Synechococcus* sp. RS9921	93
20	*Cyanobacteria*; Chloroplast; *Cryptomonadaceae*	90	*Cryptochloris* sp. C94	94
21	*Cyanobacteria*; Family II; GpIIa	89	*Synechococcus* sp. M16.17	89
22	*Cyanobacteria*; unclassified Chloroplast	92	*Phaeocystis antarctica* RS-24	94
23	*Cyanobacteria*; unclassified Chloroplast	94	*Imantonia* sp. MBIC10497	91
24	*Cyanobacteria*; Family I; GpI	81	*Microcoleus acremanii* UTCC 313	85
25	*Firmicutes; Bacillus*	89	*Bacillus firmus*	88
26	*Alphaproteobacteria; Roseobacter*	77	*Oceanobacillus* sp. UG03	87
27	*Alphaproteobacteria; Phaeobacter*	71	*Thalassobius* sp. MED612	89

Table 4.4. Phylogenetic classification of, and closest cultured isolate to, 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 4.8

Band Number	***Phylogenetic Alignment***[a]	***Similarity***[b]	***Nearest Isolate***[c]	***Max. Ident.***[d]
28	*Alphaproteobacteria*; unclassified *Rhodobacteraceae*	87	*Marinosulfonomonas methylotropha*	84
29	*Firmicutes; Bacillus*	88	*Bacillus endophyticus* SP31	94
30	*Firmicutes; Bacillus*	90	*Bacillus acidicola* CC-YY179	94
31	*Firmicutes; Bacillus*	93	*Bacillus isabelae* CVS-8T	94
32	*Firmicutes; Bacillus*	91	*Bacillus pallidus*	93
33	*Firmicutes*; unclassified	91	*Bacillus* sp. CNJ775 PL04	92
34	*Firmicutes; Marinibacillus*	89	*Bacillus* sp. CNJ958 PL04	93
35	*Firmicutes; Planococcus*	92	*Bacillus* sp. CNJ958 PL04	91
36	*Firmicutes; Bacillus*	98	*Bacillus pseudofirmus* 124-1	98
37	None			
38	*Firmicutes; Streptococcus*	88	*Streptococcus* sp. GM006	88
39	*Firmicutes; Streptococcus*	80	*Streptococcus* sp. 170702P1	80
40	*Firmicutes; Lactococcus*	90	*Pilibacter termitis*	86
41	None		*Enterococcus durans*	83
42	None		*Virgibacillus halodenitrificans*	84
43	n/s			
44	n/s			
45	n/s			
46	n/s			
47	n/s			
48	n/s			
49	n/s			
50	*Planctomycetes; Isosphaera*	77	None	n/a
51	*Planctomycetes; Isosphaera*	77	*Gemmata*-like str. CJuql4	90
52	*Alphaproteobacteria*; unclassified *Rhodobacteraceae*	97	None	n/a
53	*Planctomycetes*; unclassified *Planctomycetaceae*	87	*Pirellula* sp. 81	88
54	n/s			

Table 4.4. Phylogenetic classification of, and closest cultured isolate to, 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 4.8

Band Number	***Phylogenetic Alignment***[a]	***Similarity***[b]	***Nearest Isolate***[c]	***Max. Ident.***[d]
55	n/s			
56	*Planctomycetes; Planctomyces*	80	*Planctomyces* sp.	82
57	n/s			
58	n/s			
59	n/s			
60	n/s			
61	*Firmicutes; Bacillus*	100	*Geobacillus* sp. HM06-07	99
62	*Firmicutes; Bacillus*	99	*Bacillus* sp. HM06-06	98
63	*Bacteroidetes*; unclassified *Flavobacteriaceae*	89	*Olleya marilimosa*	90
64	*Bacteroidetes*; unclassified	77	*Jannaschia donghaensis*	90
65	*Bacteroidetes*; unclassified *Flavobacteriaceae*	83	*Mesoflavibacter zeaxanthinifaciens*	90
66	*Alphaproteobacteria; Phaeobacter*	97	*Roseobacter* sp. LE17	99
67	*Alphaproteobacteria; Rhizobium*	86	*Agrobacterium vitis*	89
68	*Firmicutes; Bacillus*	100	*Paenibacillus* sp. HM06-03	99
69	None		*Gelidibacter* sp. GSc-1	78
70	None		*Gillisia illustrilutea*	76
71	n/s			
72	None		*Salipiger mucescens* A3	82
73	n/s			
74	n/s			
75	n/s			
76	*Alphaproteobacteria*; unclassified *Rhodobacteraceae*	82	*Sulfitobacter donghicola*	88
77	None		*Antarctobacter* sp. TG22	78
78	n/s			
79	None		*Algibacter* sp. OB15	82
80	*Bacteroidetes*; unclassified *Prevotellaceae*	70	*Algibacter* sp. OB15	80
81	*Bacteroidetes; Sphingobacterium*	54	*Algibacter* sp. OB15	82

Table 4.4. Phylogenetic classification of, and closest cultured isolate to, 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 4.8

Band Number	***Phylogenetic Alignment***[a]	***Similarity***[b]	***Nearest Isolate***[c]	***Max. Ident.***[d]
82	*Alphaproteobacteria*; unclassified *Rhodobacteraceae*	85	*Thalassobius* sp. UST061013-004	92
83	None		*Flavobacterium* sp. SW254	76
84	*Alphaproteobacteria; Citreicella*	75	*Roseobacter* sp. GAI-109	82
85	None		None	n/a
86	None		*Flavobacterium* sp. 3008	74
87	*Alphaproteobacteria; Sphingomonas*	78	*Thalassobius* sp. 9PSW-3	89
88	*Alphaproteobacteria; Hyphomicrobium*	85	*Pleomorphomonas oryzae* F-4	90
89	*Bacteroidetes; Flavobacterium*	87	*Bizionia myxarmorum*	92
90	Unclassified *Bacteria*	21	*Algibacter* sp. Mac16 143.1	73
91	None		*Algibacter* sp. Mac16 143.1	77
92	*Alphaproteobacteria; Oceanicola*	86	None	n/a
93	None		*Rhizobium* sp. TG29	75
94	*Firmicutes*; unclassified	45	*Flavobacterium* sp. H7	81
95	*Alphaproteobacteria*; unclassified *Rhodobacteraceae*	85	*Roseobacter* sp. LE17	91
96	*Alphaproteobacteria*; unclassified *Rhodobacteraceae*	90	None	n/a
97	*Bacteroidetes*; unclassified *Flavobacteriaceae*	78	*Tenacibaculum maritimum*	83
98	*Alphaproteobacteria; Rhodothalassium*	79	*Roseobacter* sp. GAI-109	87
99	*Alphaproteobacteria*; unclassified *Rhodobacteraceae*	84	*Roseobacter* sp. LE17	92
100	*Firmicutes*; unclassified	48	*Rhodospirillum salexigens*	83
101	*Bacteroidetes; Polaribacter*	87	*Tenacibaculum* sp. HZBC22	87
102	*Bacteroidetes; Tenacibaculum*	92	*Aureimarina marisflavi*	91
103	*Alphaproteobacteria*; unclassified *Rhodobacteraceae*	86	*Antarctobacter* sp. TG22	93
104	*Alphaproteobacteria*; unclassified *Rhodobacteraceae*	85	*Thalassobius* sp. 9PSW-3	86
105	None		*Arthrobacter polychromogenes*	86
106	*Bacteroidetes*; unclassified *Flavobacteriales*	88	*Winogradskyella* sp. gap-f-41	90
107	*Alphaproteobacteria*; *Roseovarius*	86	*Roseobacter* sp. 49Xb1	92
108	*Firmicutes*; *Bacillus*	79	*Bacillus* sp. ADS9	85

Table 4.4. Phylogenetic classification of, and closest cultured isolate to, 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 4.8

Band Number	***Phylogenetic Alignment***[a]	***Similarity***[b]	***Nearest Isolate***[c]	***Max. Ident.***[d]
109	*Firmicutes; Paenibacillus*	47	*Rhodococcus* sp. ADC4	82
110	*Bacteroidetes; Elizabethkingia*	42	*Rhodothalassium salexigens*	73
111	*Bacteroidetes*; unclassified *Flavobacteriaceae*	91	*Algibacter* sp. OB15	92
112	*Bacteroidetes*; unclassified *Flavobacteriaceae*	91	*Flavobacterium* sp. S03	91
113	*Bacteroidetes*; unclassified *Flavobacteriaceae*	88	*Flavobacterium* sp. S03	90
114	*Bacteroidetes*; unclassified *Flavobacteriaceae*	86	*Flavobacterium* sp. S03	93
115	None		*Antarctobacter* sp. TG22	74
116	*Alphaproteobacteria; Ketogulonicigenium*	81	None	n/a
117	*Bacteroidetes*; unclassified *Flavobacteriaceae*	86	*Flavobacterium* sp. S03	87
118	*Bacteroidetes; Flavobacterium*	86	*Lacinutrix copepodicola*	91
119	*Alphaproteobacteria*; unclassified *Rhodobacteraceae*	78	*Roseobacter* sp. 49Xb1	88
120	*Alphaproteobacteria; Roseovarius*	83	*Roseobacter* sp. 49Xb1	90
121	*Firmicutes*; unclassified *Lachnospiraceae*	70	*Rhizobium* sp. Rf033	84
122	*Actinobacteria*; unclassified *Coriobacteriaceae*	79	*Rhodococcus* sp. ADC4	77
123	*Alphaproteobacteria*; unclassified	74	None	n/a

[a] Phylogenetic classification, where possible to the level of genus. Sequences from excised bands were aligned with, and compared to, homologous 16S rRNA gene sequences in the Greengenes® database (http://greengenes.lbl.gov) and classified accordingly.
[b] Degree of base similarity between an excised DGGE band sequence and that of its closest fully-aligned homologue in the Greengenes® database (http://greengenes.lbl.gov).
[c] Nearest matched sequence from a laboratory cultured organism, within the top 100 most closely matched sequences from the Genbank database, following alignment using the BLASTn tool.
[d] Degree of invariance between sample sequence and nearest matched sequence in the Genbank database, expressed as a percentage, where 100 % equals complete invariance.
* Base sequences for 16S rRNA gene fragments, amplified from excised DGGE bands, are listed in Appendix I.

Other bands appear to become more dominant during the course of the experiment, such as band 39 and band 45 in mesocosms 1 and 6, and band 52 in mesocosms 2 and 5. All band sequences align most closely with homologous sequences from marine or halophilic organisms. The alignment with highest
nucleotide base similarity was that of the sequence from band 55 which aligned in the genus *Halomonas*.

Bacteroidetes. Figure 4.7d shows "*Bacteroidetes*-specific" DGGE profiles. There are up to only five strong bands, indicating high abundance of these ribotypes, but all align well with cultured isolates from within the *Bacteroidetes* phylum. Certain bands appear only after bubbling on 15th May. These include band 72 (Table 4.3) which has a sequence with 85% BLASTn similarity to a *Marinomonas ophiurae* sequence, and band 73 (Table 4.3), whose sequence shares 95% BLASTn similarity with a sequence from *Olleya marilimosa*. The poorly-resolved bands at the top of the gel (Figure 4.7d) have weak BLASTn similarities to sequences from the north east Pacific Ocean, such as from *Formosa algae* (bands 57 and 64 in Table 4.3) and *Winogradskyella thalassocola* (bands 61 and 62 in Table 4.3). This contrasts with bands toward the bottom of the gels (bands 59, 60, and 70), which match isolates of *Coccinistipes vermicola* from Chinese coastal waters (Table 4.3), following BLASTn analysis.

In contrast to these BLASTn results, the same sequences aligned in the Greengenes database result in different classifications. While all aligned sequences fall within the *Bacteroidetes* phylum, only three are classified to genus level. However these three genera (*Ulvibacter*, *Aquimarina*, and *Zobellia*) failed to align using BLASTn.

Figure 4.8 illustrates 16S rRNA gene fragment DGGE profiles generated using the three remaining primer pairs, designed to be group-specific, and the pair targeting the domain *Bacteria*.

Cyanobacteria and chloroplasts. Sequences from all bands in the "*Cyanobacteria*-specific profiles" (Figure 4.8a) were identified, by both Greengenes database alignment and BLASTn search, as belonging to either members of the *Cyanobacteria* phylum or to eukaryotic phytoplankton chloroplasts. The majority of

BLASTn-aligned sequences matched members of the *Synecchococcus* genus, which falls within Group IIa as classified following alignment in the Greengenes database. The *Cyanobacteria* classification, by Greengenes, into numbered families and groups follows that of Rippka *et al*. (1979).

Firmicutes. Figure 4.8b shows gels produced using primers designed to be specific for the phylum *Firmicutes*, otherwise known as the low-GC content Gram-positive bacteria. Five bands are present in all profiles toward the bottom of the gels. Identification using BLASTn search showed these generally belong to the genus *Bacillus*, e.g. band 36 (Table 4.4) with 98% sequence similarity to a cultured *Bacillus pseudofirmus*, although organisms bearing this name may be polyphyletic (Takami and Krulwich, 2000). The gels exhibit up to 15 other bands but these are not present in all profiles in either mesocosm. Two sequences from bands most closely matched sequences outside the target phylum. These were bands 27 and 28 (Table 4.4) which showed BLASTn homology to *Alphaproteobacteria* isolated from the Arctic Ocean. All marine sequences aligning with bands were from sediment-dwelling organisms (*Oceanobacillus*, 87% with band 26) or from hypersaline environments (*Bacillus endophyticus*, 94% with band 29; and *B. isabelae*, 93% with band 31) according to BLASTn. No matches of *Firmicutes* bands in these gels were to documented pelagic bacteria (Table 4.4). There was strong classification agreement between results from BLASTn and Greengenes alignment. While only twelve percent of alignments fell outside the *Firmicutes* (in the *Alphaproteobacteria*), the highest similarities were consistent between the two methods. For instance the most prominent band in Figure 4.8b is band 36, classified as a *Bacillus* with 98 % similarity using Greengenes, and as *Bacillus pseudofermus* with 98 % similarity by BLASTn.

Planctomycetes. The DGGE profiles in Figure 4.8c were produced using primers designed to be specific for the *Planctomycetes* phylum. In contrast to earlier work which showed these primers to be 100% specific for genomic DNA of the *Planctomycetes* amplified from both coastal and oligotrophic seas (see Section 3.3.2), here there was some non-target PCR amplification. Profiles consisted of up to 15 bands, only three of these being present in all profiles. Using both BLASTn, and the Greengenes database, sequences from only four bands, out of 13 sequenced, aligned with database sequences belonging to organisms within the *Planctomycetes* phylum.

Non-specific amplifications were from bacteria belonging to the *Alphaproteobacteria* and the phyla *Firmicutes* and *Bacteroidetes*. Of the four *Planctomycetes* sequences, two aligned using BLASTn with marine sequences (band 53, *Pirellula* sp., 95 %; band 56, *Planctomyces* sp., 87 %, Table 4.4), and two with freshwater sequences (band 50, no cultured match; band 51, *Gemmata* sp., 98 %, Table 4.4). These latter two sequences were classified by Greengenes database alignment as being 77 % similar to members of the *Isosphaera* genus. Interestingly the non-target alignments with members of the *Firmicutes* phylum were between 99 % and 100 % similarity indicating considerable specificity failure of the primer pair 352f/920r, which was designed to be *Planctomycetes*-specific.

Bacteria. Figure 4.8d shows the DGGE profile gels made using primers specific for the entire domain *Bacteria*. Profiles in these gels consist of 30 or more bands, more than in any of the group-specific profiles, with the majority present in all profiles throughout the experiment and in all mesocosms. However the profiles representing mesocosms 2 and 5 show fewer bands per profile than those from mesocosms 1 and 6. The profile from 6th May, high CO_2 mesocosm 2 appears anomalous, with its bands being asynchronous with those of other profiles, and its two sequenced bands aligning with a terrestrial alphaproteobacterium according to BLASTn and a *Paenibacillus* according to alignment in the Greengenes database, although either alignment is tenuous due to very low sequence similarity (band 109, Table 4.4), and a strain belonging to the phylum *Firmicutes* (band 108, Table 4.4). Fifty 16S rRNA gene fragment sequences were produced from these two gels. Thirty eight matched marine, and 11 matched non-marine (possibly resulting from terrestrial runoff) cultured isolates (Table 4.4). The group breakdown of the 50 sequences was thus: *Alphaproteobacteria* (Greengenes 46 %, BLASTn 50 %), *Bacteroidetes* (Greengenes 38 %, BLASTn 42 %), *Firmicutes* (Greengenes 14 %, BLASTn 4 %) and *Actinobacteria* (Greengenes 2 %, BLASTn 4 %). These ratios are similar to that found by Schäfer *et al*. (2002), as well as for clone libraries produced using these primers (see Section 3.2.2). *Alphaproteobacteria* and *Bacteroidetes* dominate the sequences amplified with the *Bacteria*-specific primers (9bfm/1512uR), with the highest similarity alignments being with the genera *Roseovarius* (86 %, band 107) and *Tenacibaculum* (92 %, band 102) respectively (Figure 4.8d and Table 4.4, bands 69-123).

Certain phyla and classes of bacteria known to inhabit coastal pelagic environments are absent from the list (Table 4.4, bands 69-123) of bacterial groups amplified by the *Bacteria*-specific primers, and sequenced from the bands of the gels in Figure 4.8d. Groups that are not represented, by either Greengenes database alignment or BLASTn alignment of the NCBI database, include the *Beta-* and *Gamma-* subgroups of the *Proteobacteria* (notably *Pseudomonas*), the *Cyanobacteria* (notably *Synecchococcus*), the *Planctomycetes*, as well as any marine representatives of the *Firmicutes*.

4.4.2.2 Community richness analysis

The richness of a community of bacteria may be measured by a count of the number of bands in its 16S rRNA gene DGGE profile. Each band, otherwise known as a ribotype or operational taxonomic unit (OTU), represents a member of the community when defined by a 16S rRNA phylogeny. This is a measure of richness. However other 16S rRNA ribotypes will almost certainly be present in the community, but have not been amplified in the PCR sufficiently to produce visible DGGE bands. Community richness, within that portion of the group amplified by the group-specific PCR, was assessed for the *Cyanobacteria*, the *Planctomycetes* and the *Bacteria*, based respectively on the DGGE gels in Figures 4.8a, c and d.

Cyanobacteria. Cyanobacterial community richness is low in high CO_2 and ambient CO_2 mesocosms, 16S rRNA gene fragment band total counts being stable at between eight and ten. Neither treatment supports a richer population (Figure 4.9).

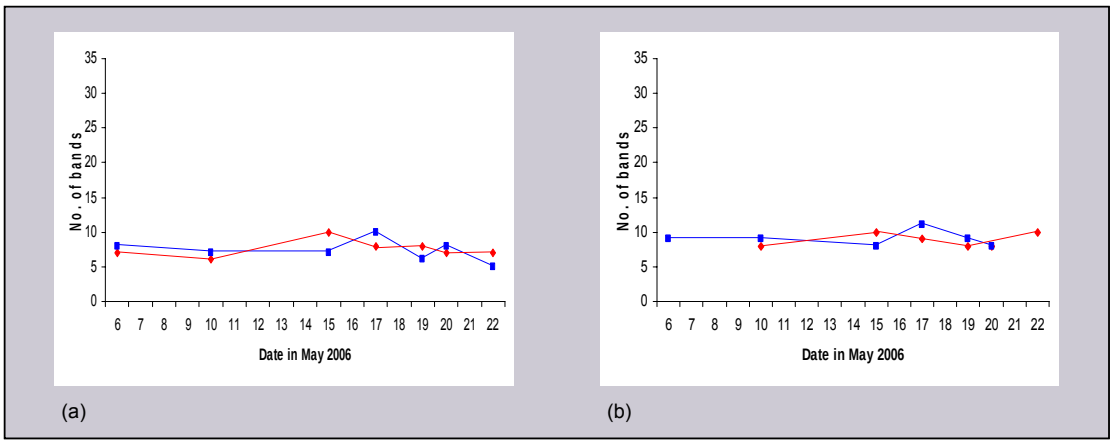


Figure 4.9. *Cyanobacteria* community richness analysis, assessed by number of bands in DGGE profiles, i.e. ribotypes (Figure 4.8a), throughout the course of the experiment (red – high CO_2, blue –

ambient CO_2). The left hand gel shows profiles from mesocosms 1 and 6 (a), and that on the right shows profiles from mesocosms 2 and 5 (b).

Planctomycetes. The *Planctomycetes* profiles in Figure 4.10 show a stable total number of bands of between six and eight. For the majority of the experiment the population in slightly richer in the ambient mesocosms (5 and 6), although this difference is already present in the first sample.

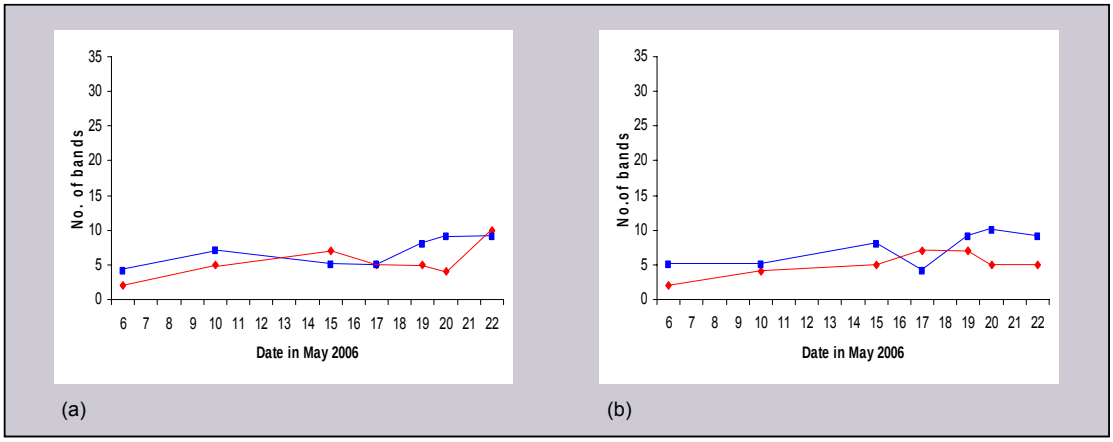


Figure 4.10. *Planctomycetes* community richness analysis, assessed by number of bands in DGGE profiles, i.e. ribotypes (Figure 4.8c), throughout the course of the experiment (red – high CO_2, blue – ambient CO_2). The left hand gel shows profiles from mesocosms 1 and 6 (a), and that on the right shows profiles from mesocosms 2 and 5 (b).

Bacteria. This DGGE analysis produces more bands than any of the group-specific analyses, even though the same nested primer pair was used, i.e. 341f-GC/518r. In both gels the ambient CO_2 community produces the highest ribotype counts at the start of the experiment, but by the end of the experiment these have been taken over by the ribotype counts in profiles from the high CO_2 mesocosm. This effect occurs sooner, with the crossover point on 15th May, and is more marked in mesocosms 1 and 6 (Figure 4.11a).

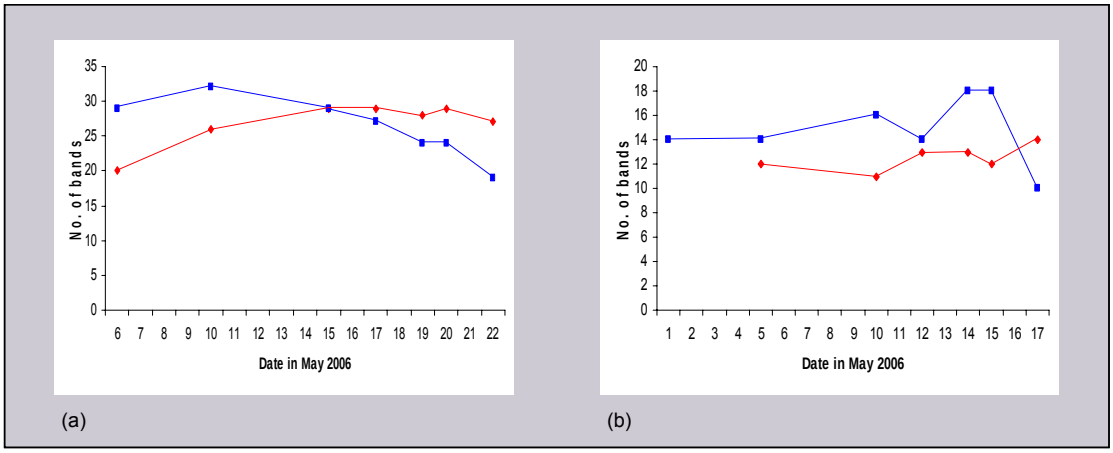


Figure 4.11. *Bacteria* community richness analysis, assessed by number of bands in DGGE profiles, i.e. ribotypes (Figure 4.8d), throughout the course of the experiment (red – high CO_2, blue – ambient CO_2). The left hand gel shows profiles from mesocosms 1 and 6 (a), and that on the right shows profiles from mesocosms 2 and 5 (b).

4.4.2.3 Multidimensional scaling analysis

DGGE profiles representative of bacterial communities in the high and ambient CO_2 mesocosms were analysed, and their patterns compared, using the multivariate statistical tool, multidimensional scaling (MDS). MDS reduces each DGGE profile to a point on a two dimensional graph, or plot. The method is described in Section 2.4.4.2. Briefly, bands from all the profiles of a DGGE gel are converted to a digital binary matrix, where "1" represents a band and "0" the absence of that band. The fidelity of the binary matrix depends on both the cleanliness and resolution of the original DGGE gel image, and the proofreading thoroughness of the investigator overseeing the derivation of the binary matrix using the specialist software (Phoretix 2D). It is important that the binary matrix is a true representation of the DGGE gel banding pattern, and that the same quality criteria are applied to all matrices. All subsequent analyses are performed under the assumption that the matrices are reliable. Binary profiles are related according to a similarity matrix. This is constructed using the Jaccard coefficient (Jaccard, 1908) by comparing each profile from the binary matrix in a stepped pair-wise manner. Each DGGE profile is displayed as a single point on an MDS plot.

Each point on the MDS plot clusters with other points according to their degree of similarity. Displaying all the profiles from a single DGGE gel on an MDS plot thus facilitates the visualisation of the relationships between them. The distances between profile points on MDS plots are real, and so subjective decisions concerning

clustering are not required. Community development patterns within mesocosms and potential divergence, in community structure, between them are revealed in the MDS plots. MDS plots in Figures 4.12 and 4.13 correspond to each of the bacterial group-specific DGGE gels comparing high and ambient CO_2 mesocosms, in Figures 4.7 and 4.8 respectively.

Observation of the specific group community MDS plots in Figures 4.12 and 4.13 reveals the varying extent that DGGE profiles change over time in mesocosms 1 and 6, and mesocosms 2 and 5. The *Alphaproteobacteria* (Figure 4.12a), *Bacteroidetes* (Figure 4.12d), and *Planctomycetes* (Figure 4.13c) profiles in mesocosms 1 and 6 behave as expected with profiles from different mesocosms becoming less similar over the time course of exposure. The *Firmicutes* (Figure 4.13b) and *Planctomycetes* (Figure 4.13c) profiles in mesocosms 2 and 5 appear to develop counter to this assumption, becoming more similar at the end of the experiment.

MDS plots were also constructed for DGGE gels comparing profiles from two different high CO_2 mesocosms, and two different ambient CO_2 mesocosms (Figure 4.6c and a, respectively). This was to show the level of similarity between the development of communities in mesocosms of replicate CO_2 treatment, in contrast to those from high and low CO_2 treatments.

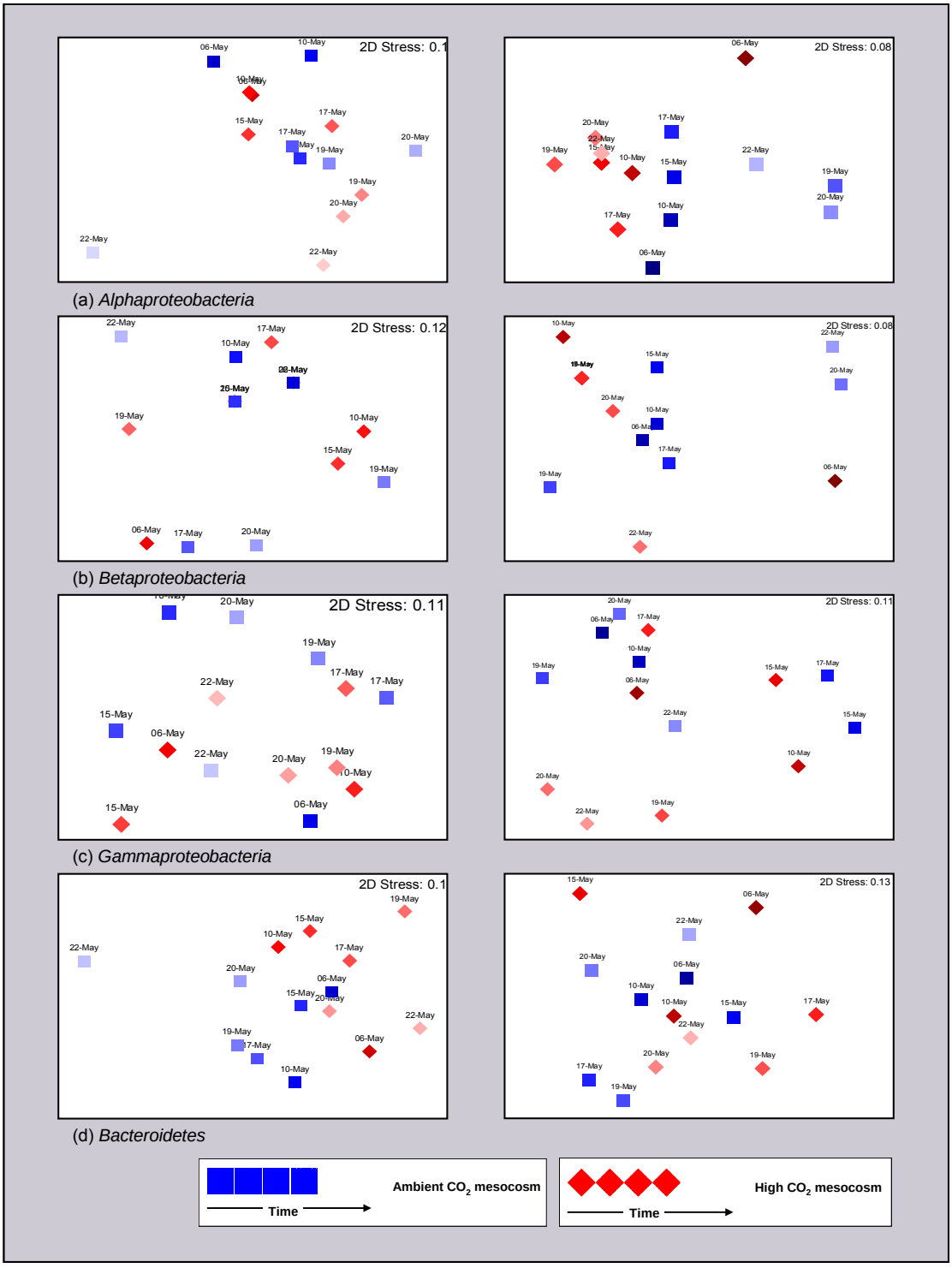


Figure 4.12. Multidimensional scaling comparison of community development in high CO_2 (red diamonds) and ambient CO_2 (blue squares) mesocosms. Mesocosm profiles from each mesocosm for successive sampling days are shown with decreasing colour intensity to aid visualisation of the timeline. All DGGE profiles for each specific group (CO_2 and ambient) were taken from the same gel (Figure 4.7). Lefthand plots display data from mesocosms 1 and 6, righthand plots display data from mesocosms 2 and 5.

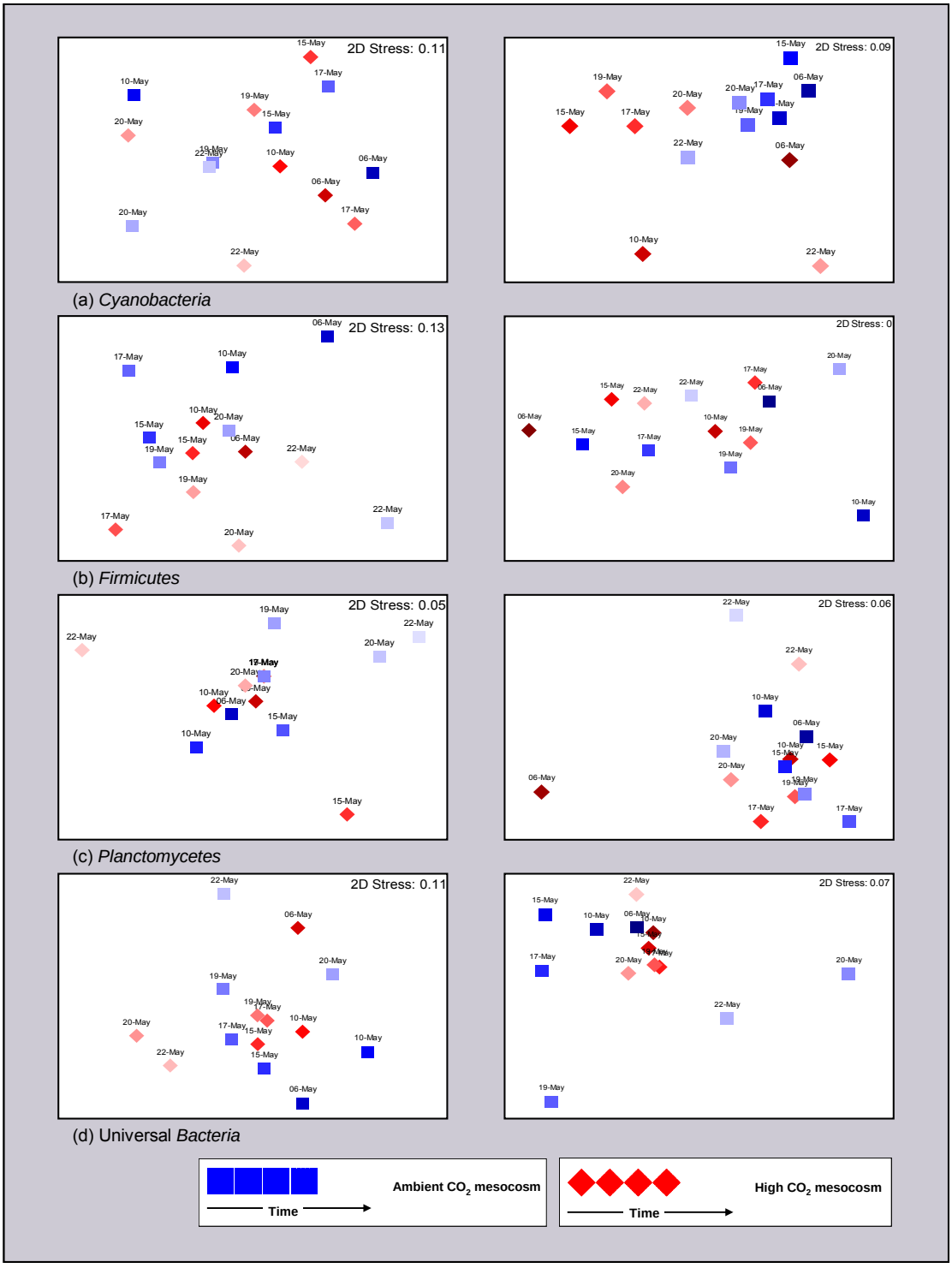


Figure 4.13. Multidimensional scaling comparison of community development in high CO_2 (red diamonds) and ambient CO_2 (blue squares) mesocosms. Mesocosm profiles from each mesocosm for successive sampling days are shown with decreasing colour intensity to aid visualisation of the timeline. All DGGE profiles for each specific group, including the *Bacteria* (CO_2 and ambient) were taken from the same gel (Figure 4.8). Left-hand plots display data from mesocosms 1 and 6, right-hand plots display data from mesocosms 2 and 5.

4.4.2.4 Second stage MDS analysis

A comparison of bacterial communities from high CO_2 and ambient CO_2 mesocosms has been done already using DGGE (see Figures 4.7 and 4.8) and MDS (Figures 4.12 and 4.13). Further comparison of these MDS plots with each other, or more specifically the similarity matrices of the profiles underlying them, allows comparison of the changes over time in the bacterial community in the different mesocosms. Such a second stage MDS was performed using PRIMER-E v.6. Analyses showed how each group-specific community compared in four different mesocosms (Figure 4.14), and how, in four different mesocosms, the seven specific groups and the *Bacteria* developed relative to each other (Figure 4.15).

Analysis by bacterial group. In the MDS plots in Figure 4.14, mesocosms with similar pCO_2 should cluster according to the hypothesis that increased pCO_2 has an effect on bacterial community composition. However neither mesocosms 1 and 2, nor mesocosms 5 and 6, cluster tightly, although the *Betaproteobacteria*, *Gammaproteobacteria* and *Planctomycetes* show loose clustering of these mesocosm pairs (Figure 4.14b, c and g, respectively). There is stronger correlation between profiles from the same DGGE gel, thus mesocosms 1 and 6, and mesocosms 2 and 5, show tighter clustering, than is seen for replicate treatment mesocosms. This same-gel clustering effect is strongest between mesocosms 2 and 5 in the *Bacteroidetes* (Figure 4.14d) and mesocosms 1 and 6 in the *Firmicutes* (Figure 4.14f) second stage MDS plots.

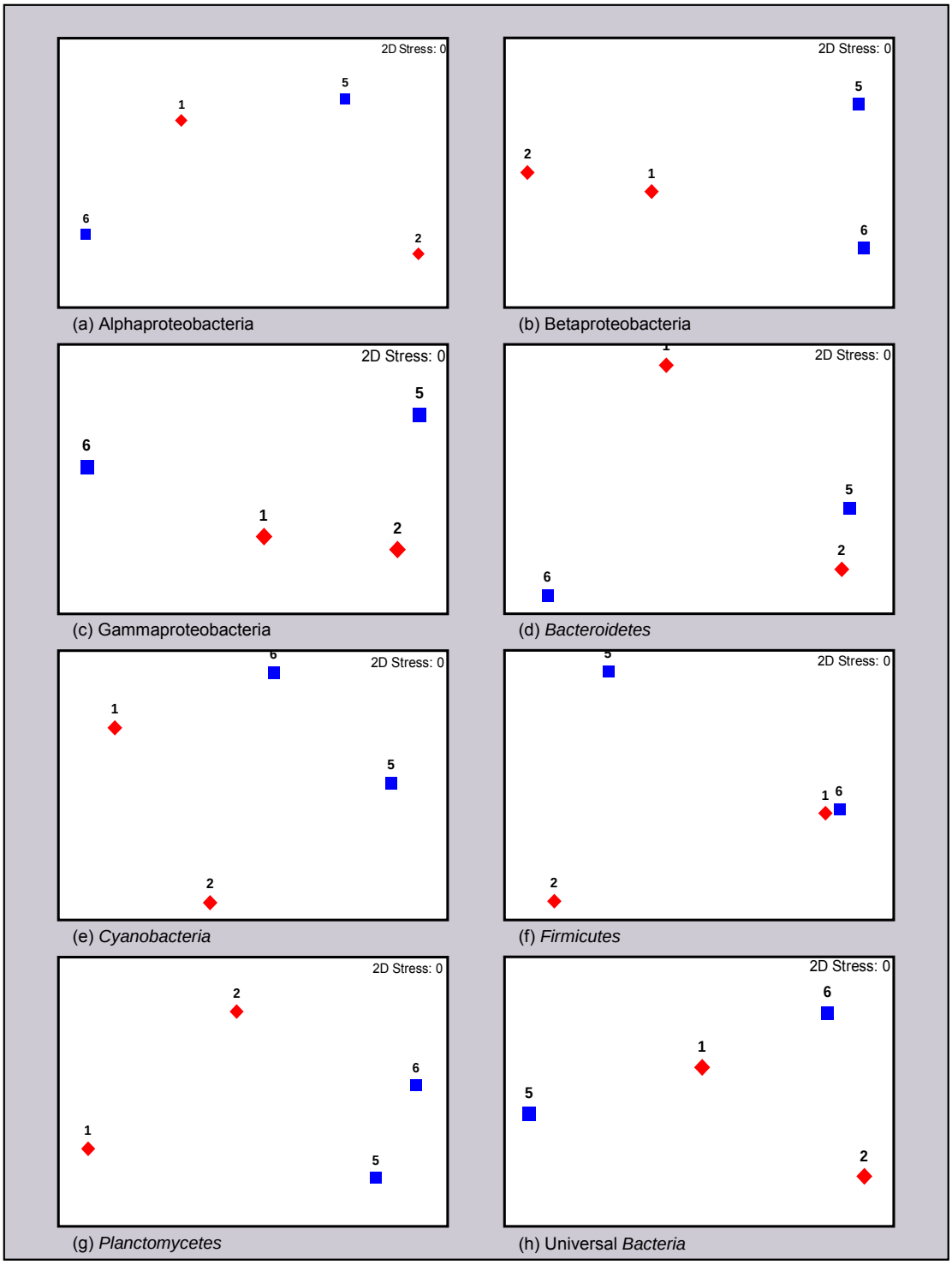


Figure 4.14. Comparison of overall community development through the course of the experiment in four different mesocosms, by second stage MDS, using a Kruskal Stress Formula and Spearman's Rank Correlation Coefficient mesocosms 1 and 2 (high CO_2), and mesocosms 5 and 6 (ambient CO_2), experienced replicate treatments. The entire temporal development of each group-specific community is reduced to a single point, which can be compared with that for the same group in all four mesocosms.

Analysis by mesocosm. The overall loose clustering in the MDS plots in Figure 4.15 may indicate that the different specific groups of bacteria develop community profiles generally independent of one another. In mesocosm 2 (high CO_2) the *Cyanobacteria* and the *Betaproteobacteria* cluster tightly (Figure 4.15c), but this pattern is not repeated in mesocosm 1, the other high CO_2 mesocosm (Figure 4.15a).

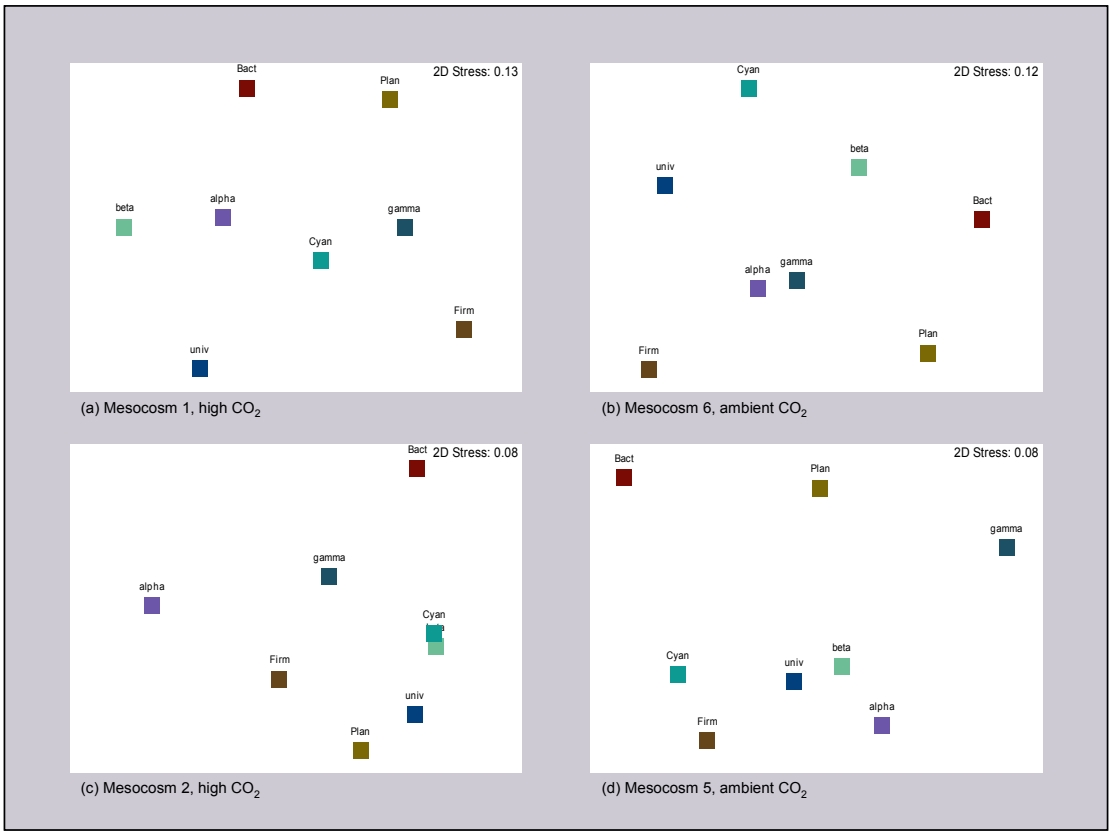


Figure 4.15 Second stage MDS showing how the group-specific community DGGE profiles develop in each mesocosm, relative to the other groups in the same mesocosm. Clustering indicates inter-group dependency, i.e. development of one group is linked to that of another. Key: alpha, *Alphaproteobacteria*; beta, *Betaproteobacteria*; gamma, *Gammaproteobacteria*; Bact, *Bacteroidetes*; Cyan, *Cyanobacteria*; Firm, *Firmicutes*; Plan, *Planctomycetes*; univ, *Bacteria*.

An interesting observation is that the community of the *Bacteria* (univ) clusters relatively closely with that of the *Alphaproteobacteria*, but is widely separated from the *Bacteroidetes* (Bact), in all mesocosms except mesocosm 2. A majority of *Bacteria*-specific DGGE profile bands belong to members of the *Alphaproteobacteria* and the *Bacteroidetes* according to BLASTn search. This fact alone leads to the expectation that the DGGE profiles of these three groups will not change independently of each other. In other words if the DGGE profiles of the

Alphaproteobacteria and the *Bacteroidetes* consist of the same ribotypes as found in the *Bacteria* (univ) DGGE profiles, this will make the DGGE profiles similar for these three groups, and cause their derived (compressed) data points to cluster in the second stage MDS (Figure 4.15). Therefore we would expect the *Alphaproteobacteria*, *Bacteroidetes* and *Bacteria* (univ) to cluster in all plots in Figure 4.15. However such clustering is not always clearly seen. This may be due to ribotypes in group-specific profiles being different from ribotypes (of the same group) in *Bacteria* (univ) profiles. Also ribotypes belonging to other groups within *Bacteria* (univ) profiles will prevent strong clustering with profiles of *Alphaproteobacteria* and *Bacteroides* groups.

4.4.2.5 Temporal change in bacterial diversity: a comparison of high CO_2 and ambient CO_2 mesocosms

Pair wise similarity comparisons between mesocosm community DGGE profiles was presented graphically in order to better illustrate the changes in community structure during the course of the experiment. Similarity between corresponding digital profiles from high CO_2 and ambient CO_2 mesocosms was calculated as a percentage. One hundred percent equates to complete similarity between two profiles, i.e. they share all their bands. Zero percent equates to complete dissimilarity, i.e. no band is common to both profiles. Percentages were calculated from corresponding profiles for mesocosms 1 and 6, and for mesocosms 2 and 5, at seven time point spread across the duration of the experiment. Figures 4.16a to k plot similarity between group-specific profiles from the two pairs of contrasting mesocosms. A decrease in similarity equates to a divergence in similarity between the two bacterial communities. Figures 4.16b-h show graphically the similarity between specific group 16S rRNA gene community profiles for either mesocosms 1 and 6, or 2 and 5. Figure 4.16i shows the divergence between mesocosms for the community amplified using *Bacteria* primers. The replicate mesocosms are also analysed in Figures 4.16j and k, using *Bacteria* profiles. According to the hypothesis that increased pCO_2 will drive change in diversity in the bacterial community, plots comparing high CO_2 and ambient CO_2 mesocosm will show a fall in similarity (%), while those comparing replicate mesocosms will show a high, and stable, similarity (%). These hypothetical relationships are illustrated in Figure 4.16a.

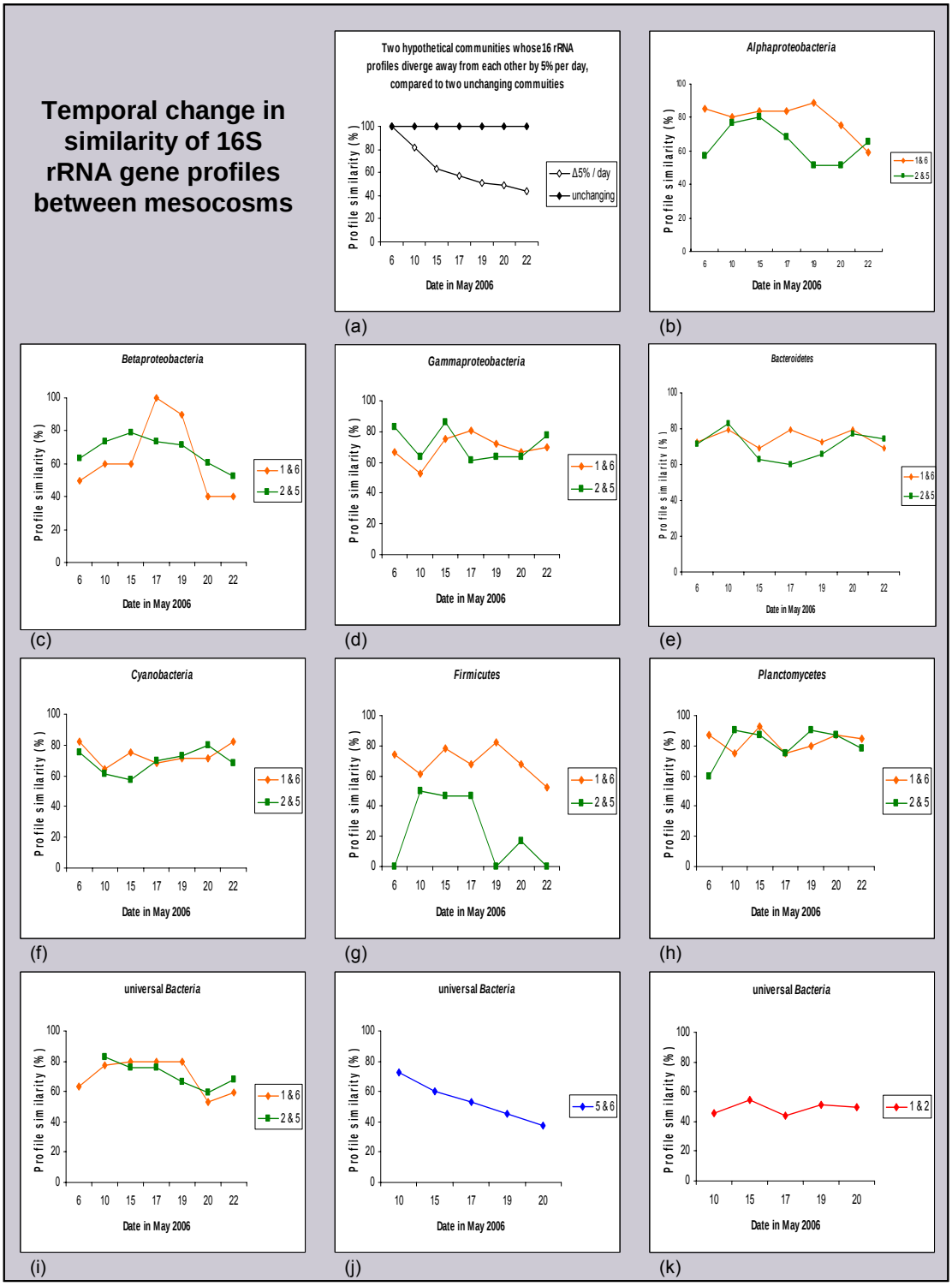


Figure 4.16. Similarity between corresponding DGGE profiles from different mesocosms. Similarity between mesocosms 1 and 6 (orange diamonds), mesocosms 2 and 5 (green squares), mesocosms 5 and 6 (blue diamonds), and mesocosms 1 and 2 (red diamonds). Bacterial communities analysed are hypothetical (a), group-specific (b-h), and *Bacteria* (i-k). See text for method.

For reference, the graph in Figure 4.16a illustrates two mesocosm communities not changing at all relative to each other (solid diamonds) and two

hypothetical communities diverging away from each other by 5% per day (empty diamonds). Some groups show little change in similarity between either mesocosm 1 and 6, or mesocosm 2 and 5 (*Gammaproteobacteria* – Figure 4.16d, *Bacteroidetes* - Figure 4.16e, *Cyanobacteria* - Figure 4.16f, *Planctomycetes* - Figure 4.16h). The *Alphaproteobacteria* (Figure 4.16b) communities show a more stable similarity between mesocosms 2 and 5, and a decreasing similarity between mesocosms 1 and 6 during the course of the experiment. All groups, including *Bacteria* (Figure 4.16i), exhibit apparent similarity fluctuation (both 1 versus 6, and 2 versus 5), whereby communities' profiles become less similar, and them more similar again! This pattern is seen most clearly in the *Betaproteobacteria* (Figure 4.16c) and the *Firmicutes* (figure 4.16g).

The *Bacteria* profiles in mesocosms 1 and 2 (high CO_2), in Figure 4.16k, maintain their degree of similarity to each other throughout the experiment, whereas the corresponding profiles in mesocosms 5 and 6 (low CO_2, Figure 4.16j) temporally become less similar.

4.4.2.6 Changes in bacterial diversity measured using ribotype accumulation curves

In an experiment aimed at describing temporal change in bacterial community diversity the total (accumulative) number of unique 16S rRNA bands (ribotypes) appearing in successive DGGE gel profiles may remain unaltered, but it is more likely to increase with time, owing to any banding variation contributing to an accumulative sum. Plotted as accumulation curves, this provides a measure of the occurrence and abundance of unique DGGE bands representing unique bacterial ribotypes. Curves were plotted for each group-specific community, and reproduced in Figure 4.17, from binary matrices using the Species Accumulation Curve tool in PRIMER-E, with the Jacknife 1 index, a non-parametric method for estimating the evenness of the sampling distribution, and 1000 permutations per analysis. Rates of accumulation for four separate mesocosm communities, are shown for each community amplified by group-specific, and *Bacteria*, primer pairs.

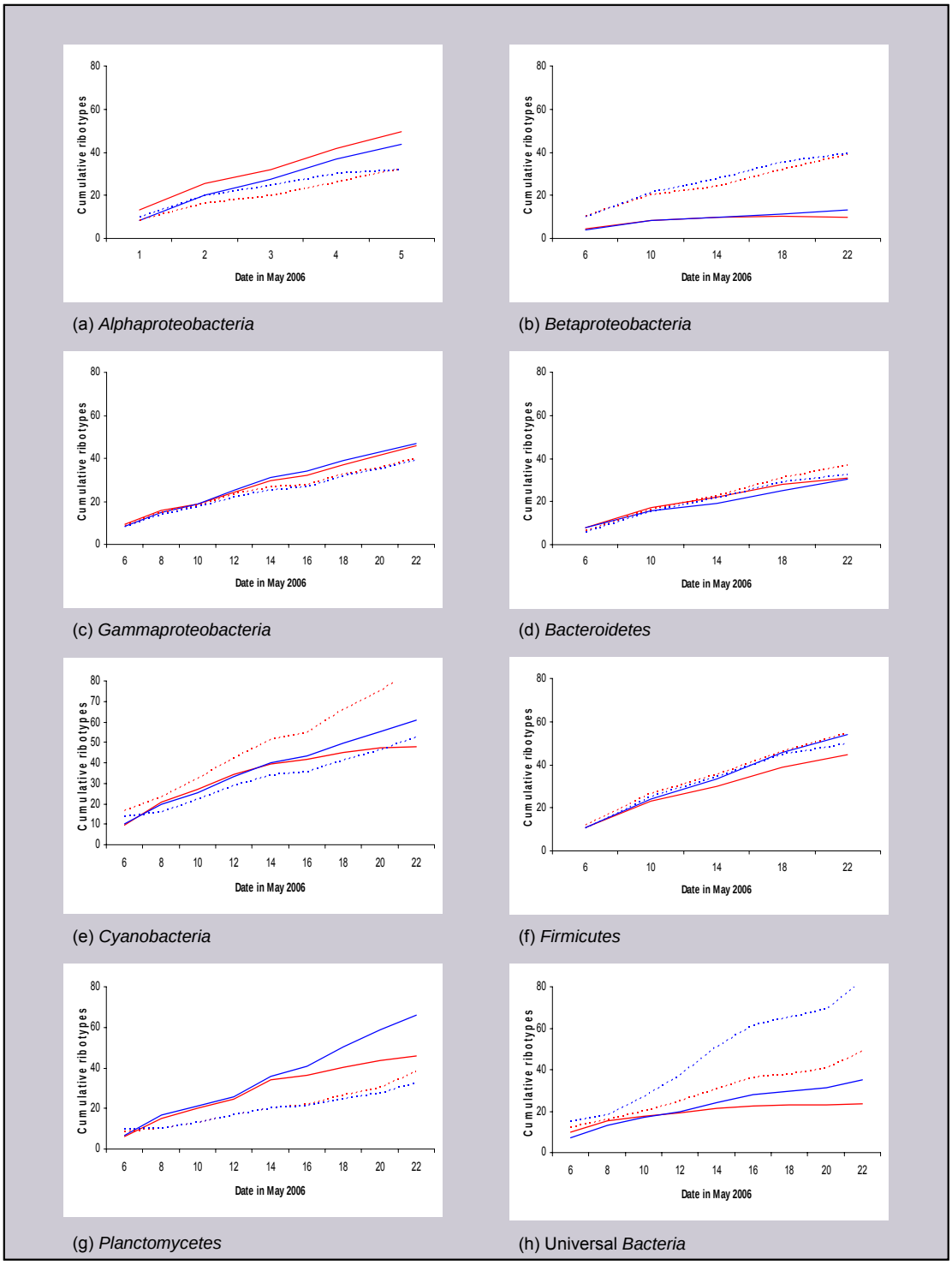


Figure 4.17. Unique ribotype accumulation curves for each specific bacterial group, for mesocosm 1 (red line), mesocosm 2 (red dashed line), mesocosm 6 (blue line) and mesocosm 5 (dashed blue line). The graphs chart the increase in number of unique DGGE bands (y-axis) in all cumulative profiles as the number of samples increases with time (x-axis).

The ribotype accumulation curves in Figure 4.17 show mixed results. In the *Aphaproteobacteria* and *Cyanobacteria*, the highest rates of unique ribotype

accumulation are in high CO_2 mesocosms. Contrasting this, the highest rates of unique ribotype accumulation are in ambient CO_2 mesocosms in the *Planctomycetes* and *Bacteria* communities. The *Alphaproteobacteria* and the *Planctomycetes* show a higher rate of unique ribotype accumulation in profiles from mesocosms 1 and 6, whereas the *Betaproteobacteria* and *Bacteria* profiles show higher unique ribotype accumulation in mesocosms 2 and 5. The three subgroups of the *Proteobacteria* and the *Bacteroidetes* have the lowest accumulation rates of unique ribotypes during the experiment. The highest rate of unique ribotype accumulation during the course of the experiment occurs in the *Bacteria*, where there is also the greatest difference in rates of unique ribotype accumulation between the four mesocosms. This variation occurs also to a lesser extent in the *Betaproteobacteria*, *Gammaproteobacteria*, *Cyanobacteria* and *Planctomycetes*. In this respect the four mesocosms are more similar for the *Alphaproteobacteria*, *Bacteroidetes* and *Firmicutes* than they are for the other groups.

4.4.2.7 Clustering and MDS analyses of the *Bacteroidetes* and the *Firmicutes* at increased temporal resolution

Ideally, to compare bacterial community profile development through time in two separate mesocosms, PCR products using samples from those mesocosms need to be run on the same DGGE gel. Such gels were produced to compare mesocosms 1 (high CO_2) and 6 (ambient CO_2), using seven group-specific, and *Bacteria*, primer sets. These gels are shown in Figures 4.7 and 4.8, and analysed in Section 4.4.2.1. However the analyses below allow a more detailed view of the changes in bacterial diversity over time, relative to other studies such as that of Grossart *et al*. (2006) who had only four post-bloom time points, and that above, as more samples from a single mesocosm can be analysed on one DGGE gel.

Figures 4.18 and 4.19 show the development of *Firmicutes* and *Bacteroidetes* community DGGE profiles, respectively, through the course of the experiment in both mesocosm 1 (750 ppm pCO_2) and mesocosm 6 (380 ppm pCO_2). There are 12 time points sampled throughout the course of the exposure, compared to the seven in the analyses in Section 4.4.2.1 (Figures 4.8b and 4.7d respectively). In addition to MDS analysis, a DGGE profile cluster analysis (Figures 4.18b and 4.19b) was performed on the binary matrices, using the Group Average Cluster program in PRIMER-E based on the UPGMA method (Clarke and Warwick, 2001).

Firmicutes. The profiles in the *Firmicutes*-specific gels vary markedly on a daily basis, with only two or three key bands appearing continuously throughout the exposure time (Figure 4.18a), examples of which are bands 2 and 6 from the high CO_2 gel, and bands 15, 16, and 19 from the ambient CO_2 gel. Other bands, such as band 24, closely matching a *Bacillus cohnii* isolate (Table 4.5) in the ambient CO_2 gel is present on 10th, 15th, 17th, 20th and 22nd of May, yet absent in profiles from the intervening days. This lack of profile continuity is reflected in the low significance of the branching in the related cluster analysis, where most profiles show less than 70 % similarity (Figure 4.18b). However, the MDS plots show a pattern of temporal development for the DGGE profile of the *Firmicutes* community (Figure 4.18c).

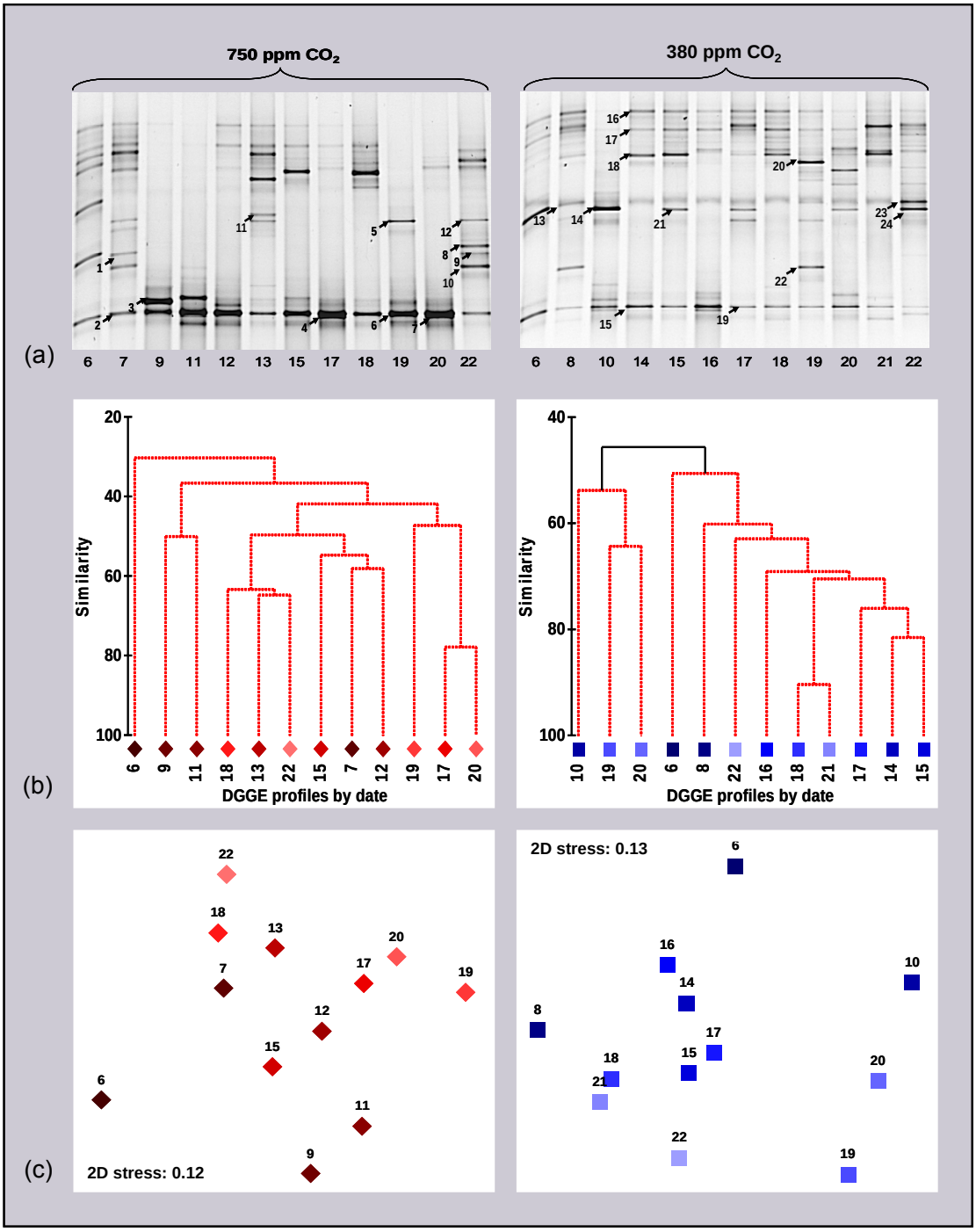


Figure 4.18. *Firmicutes* community development. **(a)** DGGE of PCR products amplified from DNA extracted from mesocosms 1 (high CO_2) and 6 (ambient CO_2). Profiles are labelled according to the sample date in May 2006, and numbered bands are detailed in Table 4.5. PCR protocol is detailed in Section 2.2.2 and the nested approach used was that for *Firmicutes*-specificity detailed in Table 3.1. Denaturant gradient was 30 % to 70 %. **(b)** UPGMA cluster analysis of DGGE profiles using PRIMER-E based on a binary matrix representation of the gel made using Phoretix-2D. **(c)** MDS plots indicate the relative development of profiles through time.

Table 4.5. Details of Genbank database 16S rRNA gene fragment sequences matching homologous sequences amplified using *Firmicutes*-specific PCR primers from DGGE bands shown in Figure 4.18

Band Number	***Nearest Match***[a]	***Max. Ident.***[b]	***Location***[c]	***Group***[d]	***Nearest Isolate***[e]	***Max. Ident.***[b]
1	*Thalassobacillus* sp. HS280	92	Houz-soltan Lake, Iran	*Firmicutes*	n/a	n/a
2	*Geobacillus* sp. HM06-07	97	Culture	*Firmicutes*	n/a	n/a
3	*Geobacillus* sp. HM06-07	95	Culture	*Firmicutes*	n/a	n/a
4	*Geobacillus pallidus*	92	Soil	*Firmicutes*	n/a	n/a
5	*Bacillus okuhidensis* GTC854	94	Okuhida spa area	*Firmicutes*	n/a	n/a
6	*Geobacillus* sp. HM06-07	99	Culture	*Firmicutes*	n/a	n/a
7	*Bacillus alveayuensis*	88	Ayu Trough sediment	*Firmicutes*	n/a	n/a
8	*Anoxybacillus kualawohkensis* KW 12	86	Malaysian Hot Spring	*Firmicutes*	n/a	n/a
9	*Geobacillus toebii* RH 127	97	Mud volcano	*Firmicutes*	n/a	n/a
10	*Geobacillus toebii* RH 127	97	Mud volcano	*Firmicutes*	n/a	n/a
11	*Bacillus okuhidensis* GTC854	96	Okuhida spa area	*Firmicutes*	n/a	n/a
12	*Bacillus okuhidensis* GTC854	96	Okuhida spa area	*Firmicutes*	n/a	n/a
13	clone P2D1-741	79	Human	*Firmicutes*	*Trichococcus flocculiformis*	77
14	*Bacillus halodurans*	99	Culture	*Firmicutes*	n/a	n/a
15	*Geobacillus toebii* RH 127	93	Mud volcano	*Firmicutes*	n/a	n/a
16	clone KG A3 120m2	88	Kerguelen Plateau	*Alphaproteo-bacteria*	None	n/a
17	*Streptococcus constellatus* CIP 105046	80	Culture	*Firmicutes*	n/a	n/a
18	*Streptococcus* sp. B00089B89	80	Human	*Firmicutes*	n/a	80
19	*Bacillus pallidus* Row2A1	98	Phenol-contaminated lake	*Firmicutes*	n/a	n/a
20	clone NR49	99	Soil	*Firmicutes*	*Bacillus* sp. SMS4	98
21	*Bacillus pseudofirmus* Mn6	93	Culture	*Firmicutes*	n/a	n/a
22	clone NR49	98	Soil	*Firmicutes*	*Bacillus* sp. SMS4	97
23	*Bacillus pseudofirmus* Mn6	94	Culture	*Firmicutes*	n/a	n/a
24	*Bacillus cohnii* US147	95	Culture	*Firmicutes*	n/a	n/a

Table 4.5

[a] Name of environmental clone, or cultured isolate, whose 16S rRNA gene sequence most closely matches that from the excised DGGE band, within the top 100 most closely matched sequences from the Genbank database, following alignment using the BLASTn tool.

[b] Degree of invariance between sample sequence and nearest matched sequence in the Genbank database, expressed as a percentage, where 100 % equals complete invariance.

[c] Sampling location from where the DNA of the nearest matched sequence was extracted.

[d] Group to which the nearest matched sequence belongs.

[e] Nearest matched sequence from a laboratory cultured organism, within the top 100 most closely matched sequences from the Genbank database, following alignment using the BLASTn tool.

Bacteroidetes. The DGGE gel images in Figure 4.19a show temporal development of *Bacteroidetes* band profiles. In mesocosm 1 (high CO_2) the bands labelled 6 and 7 are initially absent. Each subsequently is present in profiles but becomes absent again as the experiment progresses. Profile clustering in a UPGMA dendrogram correlates to some degree with time, as shown in Figure 4.19b. Profile development with time is most obvious in the MDS plots in Figure 4.19c. The *Bacteroidetes* profiles, albeit consisting of few bands, are relatively invariant as shown by the high profile similarity scores in Figure 4.19b. Temporal profile development is similar in both mesocosm 1 (high CO_2) and mesocosm 6 (ambient CO_2). This can be seen in the two gels in Figure 4.19a, and in the DGGE profile dendrograms and MDS plots derived from the two gels in Figures 4.19b and 4.19c respectively.

Profiles in DGGE gels made using *Bacteroidetes*-specific PCR primers (Figure 4.19a) maintain much more similarity from day to day, than those of the *Firmicutes*-specific gels (Figure 4.18a). Overall there are fewer bands in each profile. Obvious temporal changes include the disappearance of one band and the appearance of another as the experiment progressed (bands 6 and 7 in Figure 4.19a, 750 ppm CO_2). Sequences from these bands (6 and 7 in Table 4.6), possibly representing novel strains of *Bacteroidetes*, align most closely (94 % and 90 % homology respectively) to *Bacteroidetes* 16S rRNA gene sequences from environmental samples, and less closely (88 % and 89 % homology respectively) to the *Flavobacteria* isolates *Brumimicrobium mesophilum* (band 6) and *Bizionia paragorgiae* (band 7, Nedashkovskaya *et al.*, 2005). Such changes are less apparent in the mesocosm maintained at ambient (380 ppm) CO_2 (Figure 4.19a). Shorter branch lengths (Figure 4.19b) compared to those for the *Firmicutes* (Figure 4.18b) support the idea of temporally-stable DGGE profiles, with many *Bacteroidetes* profiles being over 90 %

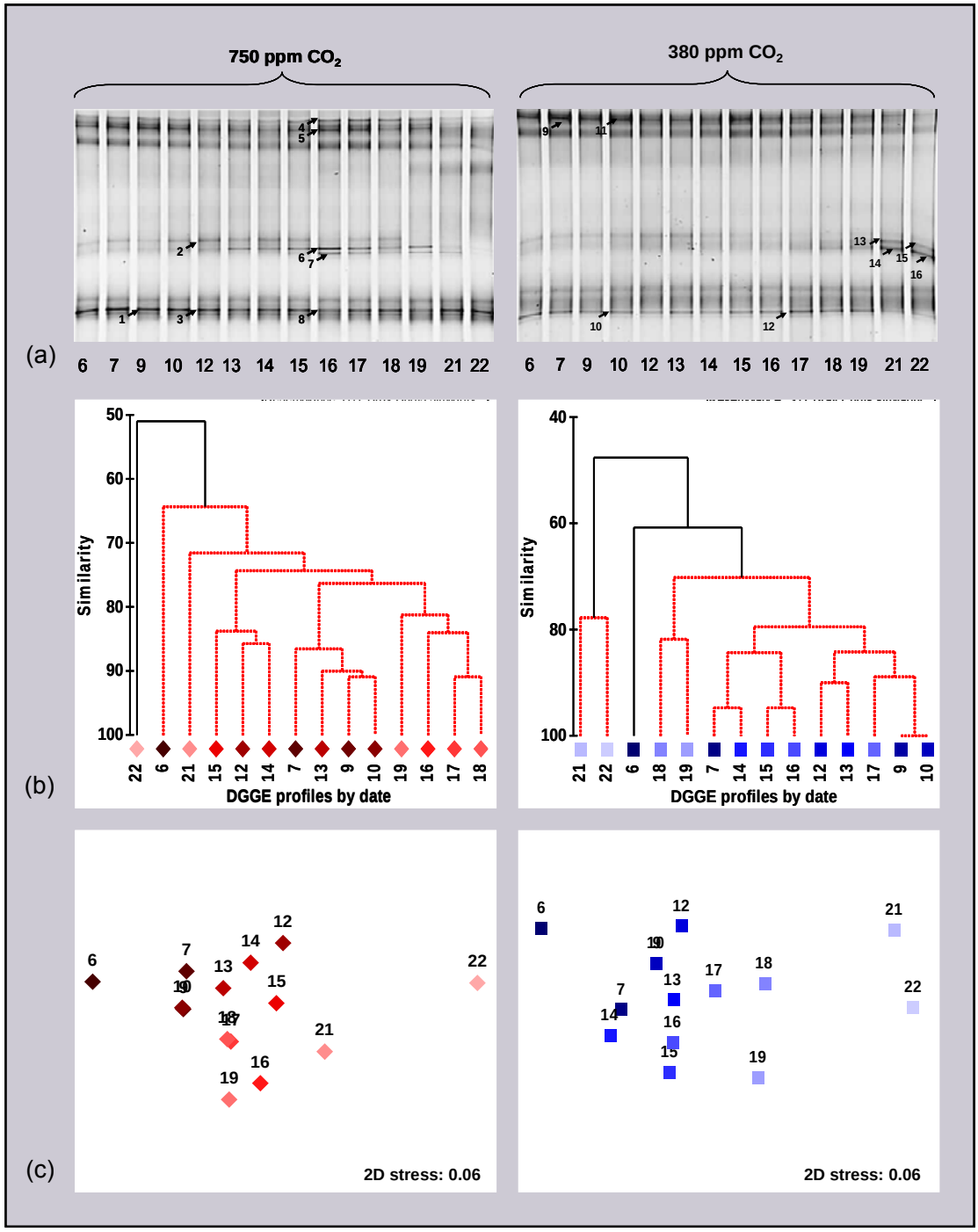


Figure 4.19. *Bacteroidetes* community development. Analysis of *Bacteroidetes*-specific 16S rRNA gene fragments from mesocosms 1 (high CO_2) and 6 (ambient CO_2). **(a)** DGGE, **(b)** DGGE profiles were subject to UPGMA cluster analysis using SIMPROF ($p = 0.05$), and **(c)** MDS. (b) and (c) are based on binary matrices generated from DGGE gels using Phoretix 2D, and constructed using Bray-Curtis similarity in PRIMER-E. Profiles are labelled according to the date of sampling in May 2006. PCR-DGGE used *Bacteroidetes*-specific primers, otherwise details are the same as for *Firmicutes* analysis in Figure 4.18.

Table 4.6. Details of Genbank database 16S rRNA gene fragment sequences matching homologous sequences amplified using *Bacteroidetes*-specific PCR primers from DGGE bands shown in Figure 4.19

Band Number	***Nearest Match***[a]	***Max. Ident.***[b]	***Location***[c]	***Group***[d]	***Nearest Isolate***[e]	***Max. Ident.***[b]
1	clone SW17	93	Antarctic sea ice	*Bacteroidetes*	*Psychroserpens mesophilus*	91
2	clone PRE-S47	84	Pearl River estuary	*Bacteroidetes*	None	n/a
3	clone PRE-S30	94	Pearl River estuary	*Bacteroidetes*	*Psychroserpens mesophilus*	92
4	clone ESP10-K9III-56	99	OMZ off Chile	*Bacteroidetes*	*Formosa algae*	98
5	clone 06	99	Norwegian fjord	*Bacteroidetes*	*Formosa algae*	98
6	DGGE band BP7	94	Baltic Sea	*Bacteroidetes*	*Brumimicrobium mesophilum*	88
7	clone KG A3 120m110	90	Kerguelen Plateau	*Bacteroidetes*	*Bizionia paragorgiae*	89
8	clone SW17	95	Antarctic sea ice	*Bacteroidetes*	*Olleya marilimosa*	94
9	clone 06	99	Norwegian fjord	*Bacteroidetes*	*Formosa algae*	98
10	clone PLY-P1-44	93	Plymouth Sound	*Bacteroidetes*	*Coccinistipes vermicola*	83
11	clone 06	98	Norwegian fjord	*Bacteroidetes*	*Formosa algae*	97
12	clone YRE-Q38	99	Yangtze River plume	*Bacteroidetes*	None	n/a
13	clone ESP10-K9III-51	89	OMZ off Chile	*Bacteroidetes*	*Mesoflavibacter zeaxanthinifaciens*	87
14	clone S26-56	92	Arctic Ocean sediment	*Bacteroidetes*	*Bizionia* sp. J69	90
15	clone S25 1015	87	Carribean Sea inshore	*Bacteroidetes*	*Bizionia* sp. J69	86
16	clone BS-E44	92	Bering Sea	*Bacteroidetes*	*Psychroserpens mesophilus*	90

[a] Name of environmental clone, or cultured isolate, whose 16S rRNA gene sequence most closely matches that from the excised DGGE band, within the top 100 most closely matched sequences from the Genbank database, following alignment using the BLASTn tool.
[b] Degree of invariance between sample sequence and nearest matched sequence in the Genbank database, expressed as a percentage, where 100 % equals complete invariance.
[c] Sampling location from where the DNA of the nearest matched sequence was extracted.
[d] Group to which the nearest matched sequence belongs.
[e] Nearest matched sequence from a laboratory cultured organism, within the top 100 most closely matched sequences from the Genbank database, following alignment using the BLASTn tool.

similar (Figure 4.19b). In Figure 4.19c *Bacteroidetes* profiles are represented by red and blue symbols in MDS plots of the high CO_2 and ambient CO_2 mesocosms respectively. These symbols reveal an overall progressive development of the *Bacteroidetes* profile. This contrasts with the equivalent MDS plots for *Firmicutes* profiles, illustrated in Figure 4.18c, where there is less order to the development of the profile. These results corroborate visual inspections of the representative gels in which the *Bacteroidetes* DGGE profile is seen to develop progressively with time (Figure 4.19a), but such progressive development is much harder to discern in the *Firmicutes* DGGE profiles (Figure 4.18a).

4.4.2.8 MDS analysis of *Planctomycetes* and *Cyanobacteria* using "replicate" mesocosms at increased temporal resolution

Planctomycetes. Figures 4.20 and 4.21 show MDS plots constructed from four independent DGGE gels for the *Planctomycetes* and *Cyanobacteria* respectively. The two sets of gels were made using same round PCR products, and matching DGGE parameters. The increased number of time points, analysed from mesocosms 1 and 6, gives a more robust interpretation of the temporal development of the mesocosms' 16S rRNA gene profiles.

We can see from Figure 4.20 that the community of *Planctomycetes* develops in a similar way in all four mesocosms. There is a small degree of "retrogressive" profile change, where a profile changes and then changes again to be more similar to the first profile than the second, such as between 10th and 17th May in mesocosm 2, 9th and 13th May in mesocosm 1, and 14th and 15th May in mesocosm 6. Overall there is a comparable degree of change in *Planctomycetes* community structure in the ambient CO_2 mesocosms (5 and 6) and in the high CO_2 mesocosms (1 and 2), indicated by the poor clustering of profiles from all four mesocosms. All mesocosm *Planctomycetes* communities show less change in diversity up to the re-bubbling point (15th May) than after this point.

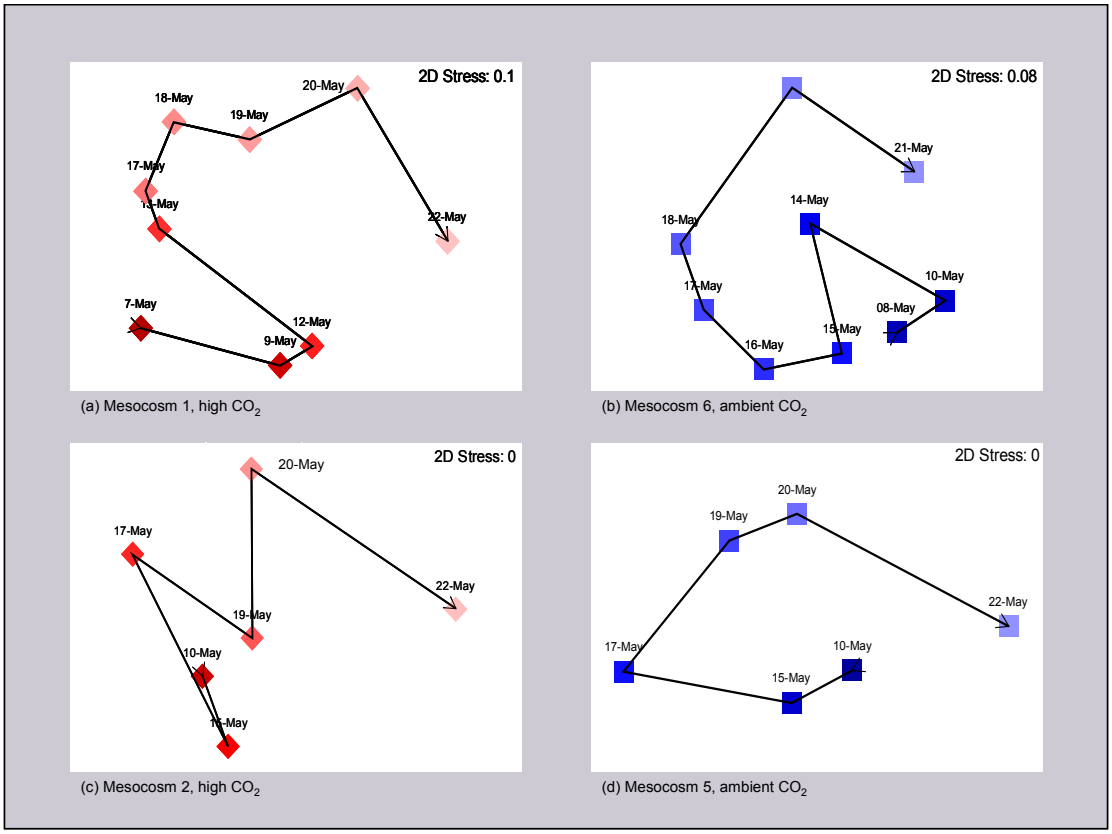


Figure 4.20. *Planctomycetes*. Comparison of community development in "replicate" high CO_2 and ambient CO_2 mesocosms. Dates differ for mesocosms 1 and 6 according to which DGGE profiles were available.

Cyanobacteria. The *Cyanobacteria* also show overall similar profile development, as seen in the temporal patterns traced on the MDS plots in Figure 4.21. As for the *Planctomycetes*, there is "retrogressive" profile development, especially in the high CO_2 mesocosms (1 and 2).

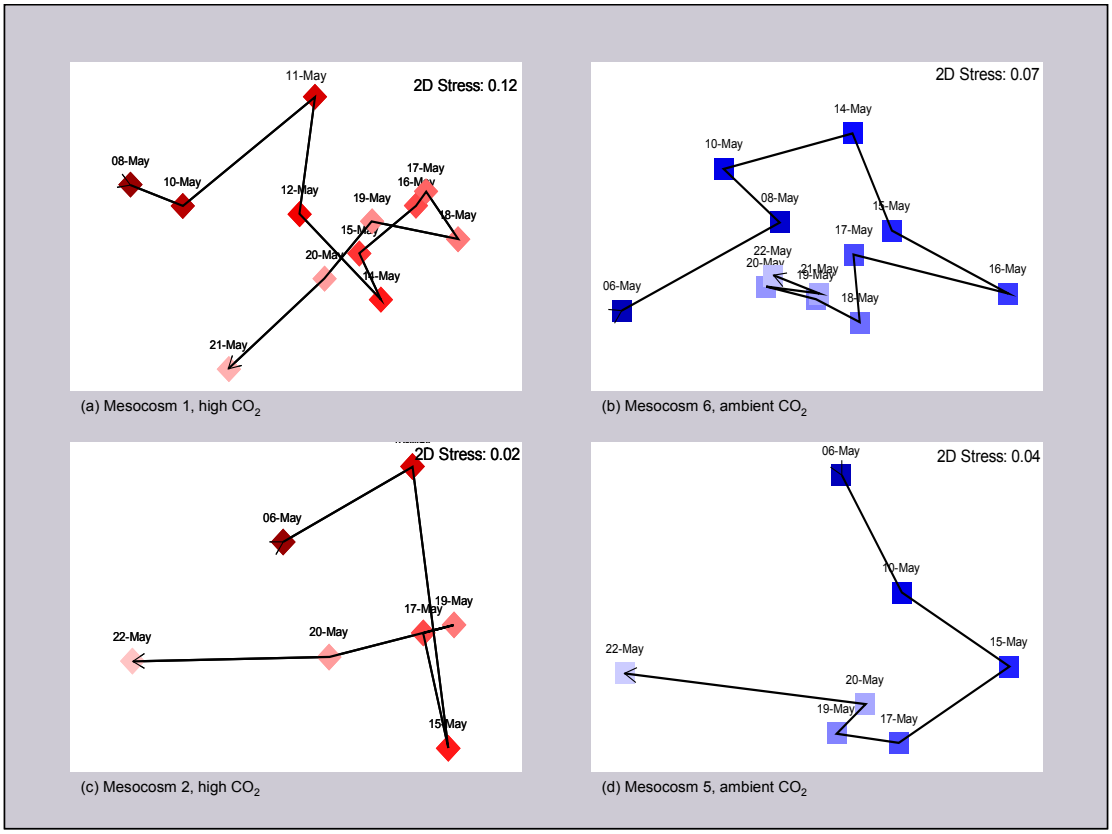


Figure 4.21. *Cyanobacteria*. Comparison of community development in "replicate" high CO_2 and ambient CO_2 mesocosms. DGGE profiles for mesocosms 1 and 6 were taken from separate gels; those for mesocosms 2 and 5 were from the same gel.

In all mesocosms the patterns described by the MDS plots show a tendency of returning towards the starting community profile. This is evident for both the *Planctomycetes* (Figure 4.20) and the *Cyanobacteria* (Figure 4.21). This indicates an initial change in bacterial diversity, for example DGGE profiles in Figures 4.8a and c, followed by a "regression" back towards a profile more similar to that of the sample from the starting community. This effect is demonstrated also by the shorter distance in the MDS plots between the first and last profiles, compared with others at intermediate time points.

4.5 Discussion

The experiment presented in this chapter aimed to evaluate whether predicted increases in the pCO_2 of the ocean (Caldeira and Wickett, 2003) will have an effect on the diversity of pelagic bacterial populations. DGGE was employed to profile 16S rRNA ribotypes from the bacterial metagenome of four *in situ* seawater mesocosms, two of which contained seawater with a pCO_2 artificially elevated to 750 ppm, the level predicted for the world's surface ocean in 2100.

4.5.1 Experimental Setting

The Large Scale Facility located in the Raunefjord used here, belonging to the University of Bergen, has been used before for large-scale mesocosm (> 12,000 L) studies (for example Grossart *et al.*, 2006; Riebesell *et al.*, 2007). In order to assess changes in bacterial diversity in different mesocosm populations it is desirable to begin with replicate populations or, as this is difficult to achieve, knowledge of the differences in bacterial diversity between mesocosms at the start of an experiment. Such differences translate into the distance between starting DGGE profiles for different mesocosms in MDS plots (Sections 2.4.4.2 and 4.4.2.3). The closer together these are, the more similar the starting bacterial communities.

Inter-mesocosm differences in bacterial diversity changes, inferred from DGGE band profiles, may be associated with high pCO_2 (*Bacteroidetes*, Figure 4.7d). Alternatively such differences in change in diversity between mesocosms, as shown in ribotype-richness analysis between high CO_2 mesocosms 1 and 2, which have less than 50 % ribotype profile similarity throughout the experiment (Figure 4.16k), may be present at the start of the experiment. Bacterial community diversity has been demonstrated to change according to the community diversity of the eukaryotic phytoplankton (Grossart *et al.*, 2005; Rink *et al.*, 2007). However this cause of change could not be assessed here, as analysis of the eukaryotic community was restricted to abundance measures of coccolithophores and cryptophytes (Figure 4.5), and species composition was not otherwise accounted for. In similar mesocosm experiments substantial differences in algal species composition have been recorded between mesocosms treated with CO_2 (Grossart *et al.*, 2006). However Rink *et al*. (2007) concluded that changes in phytoplankton abundance and composition are not reflected in the strongest bacterial ribotypes in DGGE profiles. Distinct components of a temporary DOM/POM "windfall" originating from specific species of dead

microalgae nourish specialised heterotrophic bacteria, whereas the bulk of bacterial beneficiaries are ubiquitous generalists (Hutchins *et al.*, 2001a).

Potential limitations of this mesocosm experiment, especially concerning extrapolation of results to the wider marine environment, include the within-mesocosm eutrophication (Section 4.2.1) method of stimulating a phytoplankton bloom, the short duration (17 consecutive sampling days), and gas bubbling in mid-experiment which may be seen to alter the experiment fundamentally. The subsequent rapid increase in organic matter in the water column made this the primary change to the bacterioplankton environment, possibly masking any causal effect on bacterial diversity on the part of CO_2. Witness all the major changes to the biotic components of the water column in high CO_2 and ambient CO_2 mesocosms after re-bubbling on 15th May, in Figures 4.4d and 4.5a to f . Finally, insulation from the external environment, particularly the entrapment of the plankton, usually tied to the movement of currents and tides, will have an effect on the microbial community compositions within the mesocosms. However, the limitations listed above, while affecting community composition inside mesocosms, relative to surrounding waters, should affect all mesocosms equally and so not mask any effect due to pCO_2 differences.

Mid-way through the experiment the demands of primary production led to a fall in the pCO_2 in the high CO_2 mesocosms to an ambient level. It was decided to re-establish 750 ppm CO_2 in high CO_2 mesocsoms, as the benefits of extending the exposure time would outweigh the disturbance caused by bubbles in the water column. Re-bubbling with CO_2 or air on 15th May had an impact on all the measured biotic parameters within the high CO_2 and ambient CO_2 mesocosms respectively, with eukaryotes and photosynthetic bacteria declining in abundance, and heterotrophic bacteria becoming more abundant (Figure 4.5). The experiment was thus split into two halves; firstly from the point of nutrient addition to immediately prior to re-bubbling (4th to 15th May), and secondly following re-bubbling up to the last sampling day (15th to 22nd May). It is important to understand the overall physical and biological dynamics in each mesocosm for these two time periods in order to properly interrogate the data and results obtained (see Figures 4.4 and 4.5).

Photosynthetic eukaryotic microalgae, including coccolithophores and cryptophytes, did not utilise the additional inorganic C to accelerate their primary production (Figures 4.5a and b), and photosynthetic bacteria (Figure 4.5c) exhibited

also slowed growth in the presence of elevated CO_2, relative to their counterparts in ambient CO_2 mesocosms. This suggests inorganic C availability is not limiting primary production in the mesocosm system. More likely another nutrient such as N or P is limiting growth since photosynthetic eukaryotes respond by growing strongly in all mesocosms. The slightly lower rates of primary production in high CO_2 mesocosms compared to ambient CO_2 mesocosms (Figure 4.5a, b and c) may therefore be due to a detrimental effect on microbial metabolism of the lowered pH (Figure 4.4b).

4.5.2 DGGE profile assessment of bacterial diversity

4.5.2.1 General considerations

The data from this study are basically comprised of band profiles on DGGE gels. These profiles show consistency in community diversity in different mesocosms at the start of the experiment (Figure 4.6a), and only slight change in diversity through the duration of exposure in "replicate" mesocosms (Figure 4.6a). The most conspicuous changes in bacterial diversity correlated to CO_2 treatment were in the *Alpha* and *Gamma* subgroups of the *Proteobacteria* (Figures 4.7a and c). On the other hand the clearest examples of ribotypes appearing and subsequently disappearing in profiles were ribotypes belonging to the *Bacteroidetes* in *Bacteroidetes*-specific DGGE gels (Figure 4.7d bands 56, 65, 66, 72 and 73). However presence of these bands correlates more strongly with the gas-bubbling event on 15th May, than with pCO_2. These three groups, the *Alpha*- and *Gammaproteobacteria* and the *Bacteroidetes*, have repeatedly been shown to include the most active and abundant heterotrophic members of coastal pelagic bacterioplankton (Pinhassi and Hagström, 2000; Schäfer *et al.*, 2002; Hamasaki *et al.*, 2007), as well as comprising key ribotypes associated with changes in bacterial diversity following experimentally induced phytoplankton blooms (Fandino *et al.*, 2001; Lebaron *et al.*, 2001; Schäfer *et al.*, 2001; Rink *et al.*, 2007). Grossart *et al*. (2005) further refined this conclusion, noting the *Bacteroidetes* associated with particulate matter and the *Alpha*- and *Gamma*proteobacteria were free-living.

Changes in bacterial diversity were inferred from changes in DGGE profiles, so it is important to understand what these profiles can, and cannot, inform. Departure from accurate representation of bacterial diversity begins with sampling. Methods of filtering and storage can have a selective influence on which cells persist at the point

of DNA extraction (Wintzingerode *et al.*, 1997). Coverage and specificity of the PCR primers, both those used for group-specific amplification, and those used for nested universal DGGE fragment amplification need to be considered. The implicit assumption that dominant community members are represented by strong DGGE bands is not always true (Sipos *et al.*, 2007). Huws *et al*. (2007) demonstrated amplification of *Archaea* ribotypes with subsequent misleading bands in DGGE profiles, using many "*Bacteria*-specific" primers including 341f, 518r, and 907r. These primers were used in this study for nested PCR using group-specific PCR templates, so they should not contribute to Archaeal ribotypes in DGGE profiles. *In silico* appraisal of primer pairs by these two criteria (group specificity and coverage) were conducted prior to PCR-DGGE. However analysis of experimental amplicons shows different figures for coverage and specificity for many primers compared to *in silico* analysis (compare Tables 3.3 and 3.1, respectively). Conclusions from ribotype profiles about bacterial groups must be made in light of the proportion of the group actually amplified (coverage, Table 3.4), as well as the proportion of amplicons from DNA template originating in organisms from other groups (specificity, Table 3.4).

The number of specific bands in *Bacteria* gel profiles may be the same as that in group-specific profiles. In this case the group-specific primers are no more specific than the *Bacteria* primers for that group. This is seen to a large extent with ribotypes belonging to the *Alphaproteobacteria* and the *Bacteroidetes*, but much less so with ribotypes belonging to the *Planctomycetes* and the *Beta* and *Gamma* subgroups of the *Proteobacteria* (Figures 4.7 and 4.8). There were no sequenced ribotypes from the *Bacteria* DGGE gel that belonged to these latter groups, when assessed by BLASTn search, demonstrating the utility of the primer pairs specific for these groups (*Planctomycetes* and *Gammaproteobacteria*).

Using rDNA, i.e. the 16S rRNA gene, rather than rRNA, to amplify and identify ribotypes in the community has both advantages (it involves simpler methodology and is less prone to contamination) and disadvantages (it reveals limited information about activity). Neither gives insight into linking bacterial identity with function. If 16S rRNA genes are present in a water sample irrespective of the activity of their hosts, then, assuming the continued healthy presence of those hosts, all bands in subsequent DGGE profiles should be persistently present for the duration of sampling. This is indeed the case for many bands in group-specific gels (Figures 4.7 and 4.8). It is harder to explain the presence of a band in a gel profile that was not

present in the previous profile. Increased abundance during the experiment of rare strains may allow the amplification of their 16S rRNA genes, by out-competing other genes for primers in a PCR, which did not occur before (Mühling *et al.*, 2007). Hence there may exist a threshold density for any specific 16S rRNA gene sequence in a water sample that enables its PCR amplification.

Anomalous ribotypes occur in several gel profiles, for example they are prominent in the *Planctomycetes* (Figure 4.8c) and *Cyanobacteria*/chloroplast-specific (Figure 4.8a) gels. This may be the result of high rates of diversity change, particularly through changing abundances amongst the rare ribotypes, within these populations.

Bias in the PCR can distort the DGGE ribotype profile relative to the bacterial community structure in the environment (Mary *et al.*, 2006) resulting in misleading amplification of some ribotypes and exclusion of others. Artefacts include chimaeric sequences, composed of sequence fragments from different genomes (Neufeld and Mohn, 2006), the likelihood of which forming is increased with greater fragmentation of template DNA, particularly in a first stage PCR performed with environmental DNA (Wintzingerode *et al.*, 1997). Such amplicons form bands on gels but are not representative of actual environmental ribotypes, and will give false results from BLASTn searches (Ashelford *et al.*, 2005; Sipos *et al.*, 2007).

The difficulties demonstrated in comparing data from separate gels, even when all controllable parameters are equal, means experiments of this type are limited in their analysis by the number of DNA samples that can be run on a single DGGE gel. This maximum figure is 40 using the narrowest lanes possible with the apparatus used here. So if it were feasible to perform this experiment with an unlimited number of mesocosm replications, another method of analysing the microbial diversity would need to be employed.

Accepting the above limitations however, and the partial nature of the information held in DGGE ribotypes, DGGE profiles remain a useful measure for the rapid comparison of the diversity of different bacterial communities.

4.5.2.2 Identity of bacterial ribotypes

A 16S rRNA gene fragment amplicon possessing a unique base sequence has been defined as a "ribotype" (Section 1.2.3) and in this study it is the basic unit of taxonomy, phylogenetically indivisible for the purposes of studying bacterial diversity

(Ferrari and Hollibaugh, 1999). In this sense a ribotype differs from an operational taxonomic unit (OTU) which is commonly defined as a grouping of ribotypes sharing 97 % or more sequence homology (Giovannoni and Rappé, 2000; Pommier *et al.*, 2006).

There are several ribotypes, equating to discrete DGGE gel bands, that responded to increased pCO_2 in mesocosms 1 and 2, and did not respond likewise in ambient CO_2 mesocosms 5 and 6. Those bands showing a decline in intensity in successive samples in high CO_2 mesocosms include band 10 in the "*Alphaproteobacteria*" gel (*Rhodobacteraceae*, Figure 4.7a), band 23 (*Nitrosococcus* sp., 100 % BLASTn similarity); and band 27 in the "*Betaproteobacteria*" gel (*Burkholderia*, Figure 4.7b). Also temporally increasing in intensity were band 36 (SAR86 clone, 100 % BLASTn similarity, Pham *et al.*, 2008), band 37 (Acinas *et al.*, 2004), band 42 (SAR86 clone, 99 % BLASTn similarity) and band 43 in the "*Gammaproteobacteria*" gel (Figure 4.7c). No isolates were matched to these sequences by BLAST, and in the Greengenes database they were all aligned with the genus *Pseudomonas*. Band 76 aligned with the *Rhodobacteraceae*, and with 88 % similarity by BLASTn to *Sulfitobacter*, in the "*Bacteria*" gel (Figure 4.8d). The organisms possessing all these ribotypes may be becoming less abundant, within the mesocosm, relative to others better able to respond positively to the increased pCO_2.

Other ribotypes increase in abundance following the mid-term gas bubbling. Examples of these include those excised from the "*Bacteroidetes*" gel (Figure 4.7d): band 72 (aligned with *Brumimicrobium* in the Greengenes database, and with *Marinoxanthimonas ophiurae* by BLASTn) and band 73 (*Flavobacteriaceae*, specifically *Olleya marilimosa*, 96 % BLASTn similarity). Band 65 - *Mesoflavibacter zeaxanthinifaciens* in the phylum *Bacteroidetes* with 90 % BLASTn similarity - from the "*Planctomycetes*" gel (Figure 4.8c). Bands increasing in intensity, or becoming present, with time include three from the "*Bacteria*" gel (Figure 4.8d): band 81 (weakly aligned with the *Sphingobacterium* genus, but having greater similarity to an *Algibacter* according to BLASTn), band 101 (*Polaribacter*), and 106 (the *Flavobacteriales Winogradskyella* sp., 90 % BLASTn similarity), band 111 (*Algibacter* sp., 93 % BLASTn similarity) and 114 (*Flavobacterium* sp., 94 % BLASTn similarity). These last two are only classified as far as the family *Flavobacteriaceae* by Greengenes database alignment. The organisms possessing these ribotypes may increase in abundance, within the high CO_2 mesocosm, due to

being more refractory than others to the fall in pH. Alternatively they may have responded well to the organic carbon "windfall" caused by gas bubbling.

There are other ribotypes in each group that are dominant throughout the experiment, unaffected by either elevated CO_2, or gas bubbling. These include those from bands 11, 12, 15, 16, 17 and 18 (all with > 98 % sequence homology to *Synechococcus* sp., *Cyanobacteria*, Figure 4.8, Table 4.4), 25 and 36 (*Bacillus pseudofirmus*, 98 %, *Firmicutes*, Table 4.4). *Synechococcus* sp. may be small enough to resist the cavitational shock of bubbling and, as already mentioned, they may be growth-limited in this setting by N or P, rather than inorganic C.

There are limitations to identifying the organism from which a ribotype originates by means of a BLASTn search (Ashelford *et al.*, 2005). Firstly the BLASTn results are defined by the query base sequence, the ribotype, which, as well as varying in length, is subject to the biases of the PCR listed above in Section 4.5.3.1 and discussed fully in Chapter 6. The longer a ribotype sequence is the more taxonomic information it can yield about the organism from which it was amplified. The ribotype sequences used here for BLASTn queries were generally short (~ 180 bp). However, the group-specific primer pairs were originally designed to produce BLASTn-appropriate length DGGE-ready PCR products with GC-clamps. Subsequent performance of a nested PCR on this template necessarily reduced the size of ribotype sequences for DGGE. However there is a balance to be struck, as both PCR and DGGE are compromised with longer sequences (Neufeld and Mohn, 2006).

The results of a BLASTn search are also dependent on the quantity and quality of the homologous sequences residing in the database. Submission to public repositories is subject to limited quality control, with the result that many sequences may be of low fidelity, for example those containing primer sequences, and those that are chimaeric constructs (Ashelford *et al.*, 2005; Neufeld and Mohn, 2006). The homology between a query sequence and its closest aligned database equivalent is an aspect of BLASTn data which needs to be considered carefully. Characteristics attributed to the database ribotype are not necessarily attributable to the query ribotype, especially below the different arbitrary sequence homology cut-off points for bacterial "species." These are 94 % sequence similarity, equivalent to 70 % DNA-DNA hybridisation (Wayne *et al.*, 1987), the more widely accepted 97 % sequence similarity (Venter *et al.*, 2004), and the strict 99 % sequence similarity unless niche-cohabitation can be demonstrated (Konstantinidis and Tiedje, 2005).

The alignment of sequences in the Greengenes database, performed for sequences in Tables 4.2, 4.3, and 4.4, is more reliable than a simple association with a sequence that aligns closest following BLASTn analysis. With Greengenes, sequences are placed in a phylogenetic tree according to maximum parsimony. The classification obtained is only given to the lowest attributable taxonomic level (e.g.: class, order, family, or genus). The nearest cultured isolate imparts a weighting to the classification proportional to its phylogenetic distance from the query sequence. The results are thus more robust than those sometimes obtained through BLASTn alignment.

Although many taxonomic demarcations in 16S rRNA gene bacterial phylogenetics are arbitrarily based on sequence differences, and there is clearly a need for quality policing of databased sequences, the system is currently the only basis for widespread application of bacterial phylogenetics. For the time being the approach used here to acquire 16S rRNA gene sequences, and to infer from their analysis patterns and developments within *in situ* bacterial communities, is valid.

4.5.2.3 Multivariate statistical analysis

MDS analysis reveals temporal patterns in the development of the bacterial community that may not be apparent in DGGE gels. Fundamental to this approach is a high-fidelity digital representation of profiles from DGGE gels – so-called binary matrices. The high pCO_2-associated diversity changes revealed by MDS analysis included the *Alphaproteobacteria* and *Bacteroidetes* communities in mesocosms 1 and 6. Separation of the two communities can be seen in MDS plots (Figure 4.12a and d) along the time course of the experiment.

Changing abundance of only a few ribotypes may occur as a result of increased pCO_2. In this case such changes will be masked in MDS based on changes in whole band (ribotype) profiles, such as the case with the *Gammaproteobacteria*-specific gel (Figure 4.7c) in which bands 42 and 43 become weaker in the high CO_2 mesocosm, but this is not translated into change in the corresponding MDS plot (Figure 4.12c). This limitation will apply to all DGGE analyses of bacterial diversity where the majority of ribotypes are not affected, in terms of presence or abundance, by increased pCO_2.

The phenomenon of "retrogressive" change in diversity, whereby ribotype profiles become more similar to a previous profile, is evident in many of the MDS plots in Figures 4.12 and 4.13. This may be a manifestation of community stability.

Second stage MDS analysis showed little clustering between the ribotypes of the *Alphaproteobacteria* and *Bacteroidetes* communities, and the ribotypes amplified by *Bacteria*-specific PCR. The fact that the majority of ribotypes amplified by *Bacteria* primers belong to these two groups make this an unexpected result. The *Bacteroidetes* were revealed to some degree to change independently of all other groups in all mesososms (Figure 4.15). This fits with the rise in abundance of the *Bacteroidetes* following the DOM/POM "windfall" after re-bubbling on 15th May (Figure 4.7d).

Early DGGE gel profiles (Figure 4.6), originating from the same gel, indicate good community replication between mesocosms at the start of the experiment. However the lack of clustering of these early profiles between gels, and their tendency to cluster when located on the same gel shows that comparison of profiles between gels cannot be trusted. This is certainly the case where complex profiles such as these are concerned, even when sophisticated image analysis software is employed.

Importantly from the point of view of the assessment of the impact of CO_2 on bacterial diversity, second stage MDS showed clustering of mesocosm communities according to shared DGGE gel (illustrated for the *Bacteroidetes* and *Firmicutes* in Figures 4.14d and f respectively), rather than according to pCO_2 level; i.e. mesocosms 1 and 6, and 2 and 5 (contrasting CO_2), clustered, rather than 1 with 2, and 5 with 6 (similar CO_2). This has probably more to do with experimental bias, than with bacterial responses to elevated pCO_2 in mesocosms. Without evidence to support Hypothesis II (Section 1.5), the null-hypothesis, that increased pCO_2 and decreased pH will not impact on bacterial diversity, is accepted.

4.5.2.4 Divergence of high CO_2 and ambient CO_2 bacterial communities

Community similarity between high CO_2 and ambient CO_2 mesocosms did not generally decrease during the time course of the experiment (Figure 4.16b to h), as would be expected if increased pCO_2 was a driver for community change, as in the hypothetical scenario of a daily 5 % change in diversity illustrated in Figure 4.16a. Non-specific ribotypes in group-specific DGGE gels may adversely affect this analysis, by falsely maintaining similarity when occurring in divergent communities from high and ambient CO_2 mesocosms, as is likely for the *Betaproteobacteria* and *Firmicutes* (Figures 4.16c and g). Nonetheless analysis of a PCR which still amplifies a specific target group, if not quite matching the intended one, revealing maintenance

of similarity between high-CO_2 and ambient-CO_2 communities, can be read as a true reflection of a lack of divergence between the two sets of communities.

4.5.3 **Further research**

Although not statistically significant, the changes in bacterioplankton diversity due to elevated pCO_2 demonstrated in this work make a good starting point for further investigation. Future studies of bacterioplankton response to increasing pCO_2, or ocean acidification, would benefit from focussing on the *Gamma* subgroup of the *Proteobacteria*, and by increasing greatly the time span of the experiment. The latter point however presents logistical and financial difficulties when working *in situ* in the marine environment with mesocosm volumes exceeding 10,000 L (Riebesell, 2004). Making century-scale predictions for the global ocean based on results from three weeks in 12,000 L mesocosms, even with the predicted pCO_2 level being appropriate, is difficult.

Future work may also benefit from sampling, reverse transcribing and amplifying 16S rRNA as opposed to its gene - crDNA (copy rDNA) instead of rDNA - and also amplifying longer ribotype sequences, possibly even moving to whole genome analysis, thus facilitating (and surpassing) the functional annotation of ribotypes, as pioneered by Venter *et al*. (2004).

Phylogenetic comparisons will give robustness to results, especially using near full-length 16S rRNA or 16S rDNA sequences. The advantage of sequence based phylogenetic studies is that each sequence is directly related to other samples, and the depth of coverage of a community is proportional to the number of sequences used (Neufeld and Mohn, 2006).

Functional annotation of a diversity inventory is always desirable. As well as reverse transcription of 16S rRNA, methods using labelled nucleotide incorporation into genomic DNA, such as DNA-based (Friedrich, 2006) or RNA-based (Manefield *et al.*, 2002; Whiteley *et al.*, 2006) stable isotope probing (Dumont *et al.*, 2006), and bromodeoxyuridine immunocapture (Hamasaki *et al.*, 2007), may be employed to identify actively-growing members of the community. While whole genome sequencing and comparison remains relatively costly and computationally-demanding, assessment of bacterial communities using 16S rRNA, or its gene, is still the most practical and informative strategy.

It will be interesting to use the group-specific primer pairs used in this study to analyse a larger number of clones per clone library than has been possible here, and from a variety of different geographic origins and seasons; a more extensive screening program has the potential to discover novel phylogenetic clades, particularly of the *Gammaproteobacteria*.

Finally it is not recommended to re-adjust pCO_2 levels in mesocosms after initial setting.

4.6 Conclusion

Temporal changes in 16S rRNA gene DGGE profiles were observed for the pelagic bacterial communities inhabiting control and CO_2-enriched seawater mesocosms. Notably the *Planctomycetes* and the *Firmicutes* showed marked daily fluctuations in their genetic diversity based on 16S rRNA gene fragment analysis using DGGE. *Gammaproteobacteria*-specific DGGE gels showed declining intensity of two bands towards the end of treatment with high CO_2. The ribotypes from these bands aligned phylogenetically with the genus *Pseudomonas*, and matched, but only with low BLASTn similarity (< 90 % 16S rRNA gene sequence identity), to ribotypes belonging to the SAR86 clade (Figure 4.7c). Overall changes to profiles correlated more to mesocosm containment, and especially to gas bubbling, than with elevated CO_2. No change in bacterial diversity was conclusively linked to increased pCO_2. With little difference detected in microbial diversity between two pairs of mesocosms it is likely further replicate mesocosm pairs would reveal any further degree of difference in microbial diversity between treatments. The experiment could have benefitted from more time, to reveal the results of adaptive evolution; and from extraction of rRNA, and its reverse transcription into a crDNA, to provide a template for first stage PCR with the *Bacteria*-specific primers 9bfm/1512uR, thus facilitating identification of the active bacteria and not necessarily just those present. In conclusion, large scale mesocosm experiments such as this, even if many more replicate treatments could be set up and sampled over a much longer time (which would almost certainly financially prohibit the work), are not the best way to determine the reaction of the marine microbial world to increased pCO_2.

Laboratory-based, small-scale experiments, in which many more variables can be constrained and controlled, and from which results can be extrapolated and up-scaled, have the potential to yield more meaningful data more efficiently.

Chapter 5

Assessment of the impact of large-scale ocean fertilisation with iron and phosphorus on the structure of pelagic bacterial communities using DGGE

5.1 Introduction

5.1.1 Fertilising the ocean with growth-limiting nutrients

Martin and Fitzwater (1988) first put forward the theory that previous global warming and cooling, during the Pleistocene epoch, was driven by Fe-rich dust. Wind-borne dust may be deposited far from its terrestrial source, as illustrated in the satellite image of Saharan dust clouds over the North Atlantic Ocean (Figure 5.1). Such Fe-rich dust falling on the oceans was hypothesised to have stimulated phytoplankton growth, and consumption and drawdown of CO_2, which in the form of DOC has been shown to be a more efficient transporter of C than POC, as suggested by C:N:P ratios (Hopkinson and Vallino, 2005). The increased flux of CO_2 from the atmosphere to the ocean would have resulted in a lessening of the gas's greenhouse effect. It was further hypothesised that when Fe-rich dust was scarce there would be a counteracting warming effect. The same group conceived the idea of large-scale ocean iron

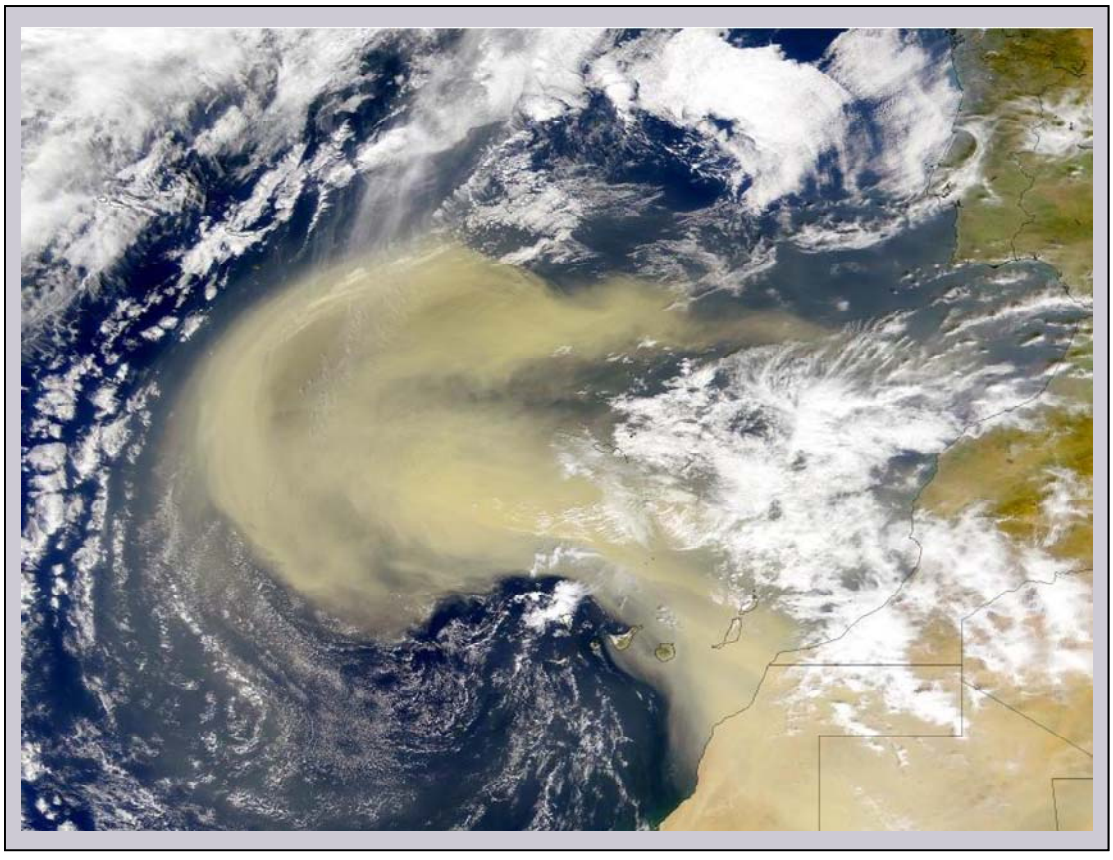


Figure 5. 1. SeaWIFS satellite image of the northeast Atlantic Ocean, taken in February 2000, showing a dust cloud from the Sahara Desert travelling well beyond the longitude of the FeeP experimental location. Image courtesy of Peter Miller, PML.

fertilisation (OIF) experiments in which biological and chemical fluxes would be monitored (Martin *et al.*, 1990; Munn, 2004). While the Fe-rich dust theory remains unproven, Fe-limitation of oceanic primary production was shown to be true for high nutrient, low chlorophyll (HNLC) regions (Coale *et al.*, 1996).

Some experiments have been carried out to observe these effects inside and outside of a naturally iron-fertilised area at the junction of the Indian and Southern Oceans, near the Kerguelen Plateau (Christaki *et al.*, 2008; Lefevre *et al.*, 2008; West *et al.*, 2008). There have been several large-scale artificial OIF experiments (for example Coale *et al.*, 1996; Boyd *et al.*, 2000; Arrieta *et al.*, 2004; Tsuda and Takeda, 2005), and these and others are reviewed by de Baar *et al.* (2005) and Boyd *et al*. (2007).

Primary production in the oceans may be also limited by the micronutrient phosphorus (P). Cotner *et al*. (1997) showed this to be the case for bacterioplankton production in the oligotrophic Sargasso Sea. Evolutionary adaptations in bacteria to this P-limitation include minimal genome sizes, for example 1.3 Mbp of the alphaproteobacterium *Pelagibacter ubique* (Bertilsson *et al.*, 2003; Giovannoni *et al.*, 2005), and sulphur-based lipids (S-lipids) in cellular membranes where otherwise phospholipids (P-lipids) would be found (Van Mooy *et al.*, 2006). Although receiving less attention than OIF, ocean phosphorus fertilisation (OPF) experiments have also taken place (Thingstad *et al.*, 1998; Thingstad *et al.*, 2005).

While OIF as a strategy to mitigate global warming has its advocates (Johnson and Karl, 2002), particularly those with commercial ambitions (see carbon offset credit trading in Munn, 2004), others fear both its failure to draw down significant quantities of CO_2 (Fuhrman and Capone, 1991; Chisholm *et al.*, 2001), and the unknown geochemical and ecological side effects of artificial fertilisation with Fe (Fuhrman and Capone, 1991). It has been postulated that increased primary production may increase ocean efflux of the greenhouse gases nitrous oxide (N_2O), and methane (CH_4) (Fuhrman and Capone, 1991), while ecological changes recorded to date include those to the community structures of both diatoms (Hutchins *et al.*, 2001a), and heterotrophic bacteria (Hutchins *et al.*, 2001a; West *et al.*, 2008).

5.1.2 Ocean iron fertilisation experiments

There have been several OIF experiments broadly aimed at testing the hypothesis that oceanic carbon sinks will be enhanced (Ducklow *et al.*, 2001), and at assessing the nature and extent of the ecological knock-on effects of OIF.

In 1995 IronEx II seeded an HNLC region in the eastern tropical Pacific Ocean with Fe (Coale *et al.*, 1996). "High nutrient (HN)" here refers to organic N and C. A fivefold increase in primary production in the mixed layer was matched by a threefold increase in bacterial production and a 1.7-fold increase in bacterial cell abundance (Cochlan, 2001). The biomass of *Synechococcus* and other photoautotrophic bacteria, determined by epifluorescence microscopy, only increased two-fold, compared to an 85-fold increase in diatom biomass (Coale *et al.*, 1996).

Although principally designed to study effects on phytoplankton in the sub-Antarctic Southern Ocean, the 1998 Southern Ocean sub-Antarctic Zone (SAZ) study's results are indirectly applicable to bacterioplankton (Hutchins *et al.*, 2001b). In an HNLC region low in both Fe (< 0.05 nM) and Si (< 1 μM) shipboard bottle experiments showed Fe alone increased community growth rates. The SAZ study concluded that Fe is the proximate limiting nutrient for chlorophyll production, photosynthetic efficiency, nitrate drawdown and diatom growth (Hutchins *et al.*, 2001a; Hutchins *et al.*, 2001b).

In 1999 the mesoscale OIF experiment SOIREE targeted the HNLC region of the polar Southern Ocean, looking at all aspects of the composition and dynamics of the microbial food web (Boyd *et al.*, 2000). Results showed bacterial numbers remained constant while bacterial production increased threefold following Fe fertilisation. They also indicated a direct link between primary production and Fe, which was only indirectly linked to heterotrophic bacterial (secondary) production, which resulted directly from an Fe-induced increase in DOM. Bacterial (secondary) production was more closely associated with available organic C and N derived from phytoplankton, than to Fe availability. Expanding bacterial populations were kept in check by zooplankton grazing (Hall and Safi, 2001), hence the stasis in bacterial abundance.

The Southern Ocean was again the site of Fe fertilisation in the 2000 EisenEx experiment (Gervais *et al.*, 2002; Arrieta *et al.*, 2004). Bacterial abundance was shown to double, and bacterial production peak at three times the level in surrounding HNLC waters. However Arrieta and colleagues (2004) concluded, in contrast to the

findings of SOIREE, that bacterial growth was directly limited by Fe. Significantly for the present study, they also reported no major phylogenetic changes to the structure of the bacterioplankton community as revealed by T-RFLP of 16S rRNA genes (Arrieta *et al.*, 2004). This indicated a phenotypic acclimation of the existing community, i.e. altered gene expression, rather than selection of genotypes extant but rare in the existing community; a conclusion backed up by large changes in enzyme activity (Arrieta *et al.*, 2004).

In 2001 an OIF experiment, SEEDS, was conducted in the northwest sub Arctic Pacific Ocean primarily to investigate phytoplankton response (Tsuda and Takeda, 2005). Results showed dramatic effects on both community structure and abundance of phytoplankton, but only a ~ 1.5-fold increase in bacterial abundance. Suzuki and co-workers (2005) concluded that "heterotrophic bacteria abundance was little respondent to Fe enrichment."

Again in the Southern Ocean HNLC region the SOFeX experiment in 2002 (Coale *et al.*, 2004) found a close correlation between bacterial abundance and production, and total primary production, with bacterial production amounting to < 10 % of the total production. However bacterial growth was found ultimately to be limited by organic carbon supply, and hence only indirectly by Fe (Oliver *et al.*, 2004).

5.1.3 Ocean phosphorus fertilisation experiments

Several studies have shown a lack of P to limit the growth of oceanic phytoplankton (Krom *et al.*, 1991; Mills *et al.*, 2004) and bacterioplankton (Dyhrman *et al.*, 2002). A literature search found only one report (Thingstad *et al.*, 2005) of an experiment involving large scale *in situ* OPF prior to that reported in this study. Thingstad *et al*. (2005) added phosphate to a final concentration of 110 nM across 16 km^2 in the oligotrophic Cyprus Eddy in the eastern Mediterranean Sea. Subsequent observations included a decline in chlorophyll, a rise in bacterial production, and a rise in copepod egg abundance. The authors (Thingstad *et al.*, 2005) speculated N had become the growth-limiting element for phytoplankton, leaving certain bacteria able to capitalise on the excess P, with copepods able to switch to direct grazing of bacteria.

Microcosm and mesocosm P addition experiments have attempted to answer questions surrounding the growth-limiting effects of P supply to heterotrophic and phototrophic bacteria in the open ocean. Cotner and colleagues (1997) found bacteria were P-limited in spring in the oligotrophic Sargasso Sea, with growth rates highly responsive to added inorganic phosphate. Similarly Thingstad and co-workers (1998), working with predator-free water from the oligotrophic northwest Mediterranean Sea in summer reported both phytoplankton and heterotrophic bacteria responded with a pulse uptake and subsequent accelerated growth, on addition of orthophosphate. No such effects were evident after C or N addition, indicating in contrast to their later conclusion (Thingstad *et al.*, 2005, above) that N had not become the limiting nutrient for phytoplankton in the northwest Mediterranean Sea. Both these studies help explain the build up of DOC, not utilised by microorganisms whose growth is P-limited, in these waters in spring and early summer.

5.1.4 Knowledge of the impact of ocean fertilisation, with either iron or phosphorus, or both together, on bacterioplankton communities is limited

Investigations into the effects of OIF and OPF on the pelagic bacterioplankton lag behind investigations into the effects on phytoplankton. This discrepancy was highlighted in the 1990s by Tortell and co-workers (1996) who pointed out that bacteria comprise up to 50 % by weight of pelagic POC, and through remineralisation of DOC, produce CO_2 by respiration, thus potentially nullifying the principal aim of ocean fertilisation. A bacterial reaction to Fe addition seemed reasonable due to bacteria having a higher Fe content per unit biomass than phytoplankton. Tortell *et al*. (1996) showed bacteria took up 20 % to 45 % of bio-available iron. This indicated also a significant role for bacteria in global oceanic Fe cycling, and carbon flux with the atmosphere. In another study Pakulski *et al*. (1996) showed added Fe induced a doubling of bacterial abundance and a four-fold increase in growth rate (μ). Bacteria were likely to respond to Fe addition, faster than phytoplankton, due to their higher surface area to volume ratios, greater abundance, and production of extracellular Fe-chelating siderophores (Pakulski *et al.*, 1996).

Most studies of bacterial response to OIF experiments have measured gross ecological parameters such as changes to bacterial abundance and production. Within OIF patches, relative to HNLC areas outside, the former has risen between three-fold (Cochlan, 2001) and 12-fold (Christaki *et al.*, 2008), while the latter has risen by 1.5-

fold (Suzuki *et al.*, 2005) to 2.8-fold (Christaki *et al.*, 2008). Bacterial abundance tended to correlate with chlorophyll *a* intensity, while bacterial production tended to correlate with total net primary production (Cochlan, 2001; Hall and Safi, 2001; Arrieta *et al.*, 2004; Oliver *et al.*, 2004; Suzuki *et al.*, 2005). Indeed Suzuki *et al*. (2005) concluded that bacterial abundance was linked directly with phytoplankton abundance, and only indirectly with Fe fertilisation. Meanwhile Hutchins *et al*. (2001a; 2001b) concluded that the structure of the bacterioplankton community changed in accordance with the phytoplankton community, which started out predominantly as mixed nano- and pico-phytoplankton, and became dominated by large diatoms. This and other investigations into effects of OIF on the structure of the bacterioplankton community, i.e. higher resolution effects, are limited to a few studies based on PCR of 16S rRNA, and its encoding genes, using primers broadly specific for the entire domain *Bacteria* (Hutchins *et al.*, 2001a; West *et al.*, 2008).

Hutchins *et al*. (2001a) used DGGE and universal *Bacteria* primers 338f-GC/517r (in this study called 341f-GC/518r and used in varying combinations listed in Table 3.2) to study community changes in Fe-fertilised shipboard mesocosms. Changes, associated with those of phytoplankton community composition, were largely confined to several ribotypes of the *Gammaproteobacteria*, and one of the *Alphaproteobacteria*. In another study West *et al*. (2008) employed single strand conformation polymorphism (SSCP) (see Section 1.2.3 for explanation of method) to assess changes in bacterial diversity measured by amplification of both 16S rRNA, and its encoding genes. Analysing clone libraries, community changes were largely ascribed to *Alphaproteobacteria*, *Gammaproteobacteria* and *Bacteroidetes* ribotypes, with the most significant changes due to switches in dominance between *Roseobacter* clusters (*Alphaproteobacteria*). Within the naturally-occurring OIF patch *Roseobacter* NAC11-7, SAR92, and a *Bacteroidetes* cluster related to agg58 dominated, while outside in the HNLC region these gave way to *Roseobacter* RCA, SAR11, and *Polaribacter* clusters.

On the other hand documented effects of OPF experiments on pelagic bacterial community structures could not be found, although bacteria are known to be P-limited in oligotrophic environments (Cotner *et al.*, 1997; Thingstad *et al.*, 1998; Thingstad *et al.*, 2005). Analysis of metabolic activity concluded that heterotrophic bacteria were well able to utilise a windfall of P, and that their direct predation by flagellates lead to a trophic by-pass of phytoplankton, and a subsequent increase in

phytoplankton predation due to the increase in abundance of zooplankton predators (Thingstad *et al.*, 2005).

5.1.5 Background to the 2004 FeeP experiment

The 2004 Plymouth Marine Laboratory (PML) core cruise between 24th April

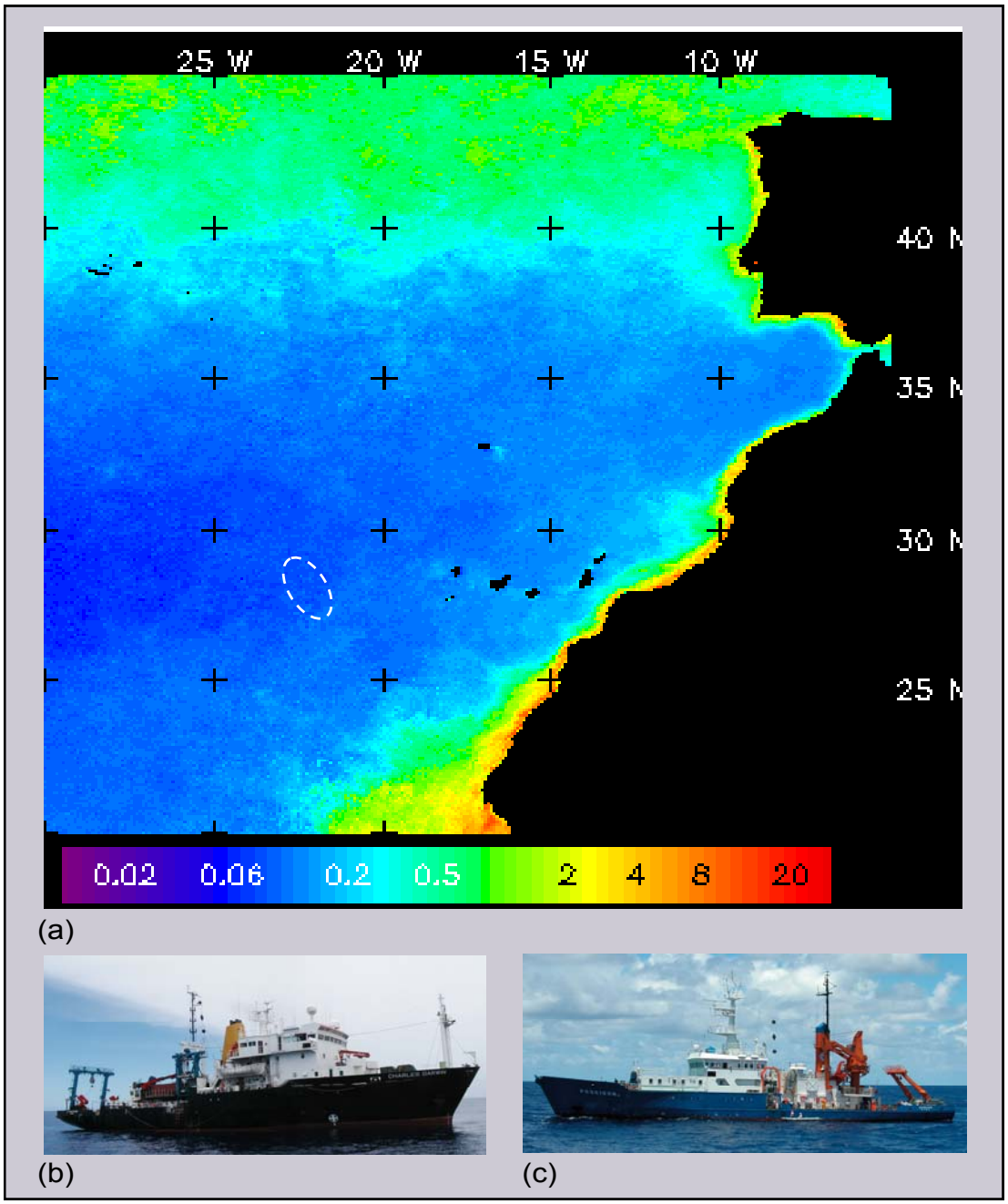


Figure 5.2. Location of the FeeP study (dotted circle) on a map showing chlorophyll *a* concentration, measured in µg chl L^{-1}, from the SeaWIFS satellite, averaged over the month of May 2004, courtesy of Peter Miller at PML (a), and the two research vessels taking part; the RRS Charles Darwin (b) and MV Poseidon (c).

and 26th May hosted a joint exercise between scientists from PML (UK), the University of East Anglia (UK), Laboratoire Arago (France), and the National Institute of Water and Atmospheric Research (New Zealand). The location of the experiment, shown in Figure 5.2, was the oligotrophic northeast Atlantic Ocean, 200 nautical miles west of the Canary Islands.

The RRS Charles Darwin and the MV Poseidon (Figure 5.2b and c) were used to test the hypothesis that the supply, separately and together, of Fe and P control biological activity and fluxes of macro nutrients (C and N) and organic matter in the euphotic zone. If these micro-nutrients (Fe, P) proved to be limiting factors in the level of total primary production, then using them as fertilisers in areas containing macro-nutrients (N, C) but exhibiting low primary production, could increase carbon dioxide fixation and export of C to the deep ocean. Two experiments were carried out, the first involving fertilisation with P alone, and the second with Fe and P together. Components of the plankton, including the bacterioplankton, were sampled throughout both experiments. These were analysed to assess ecological knock-on effects of the two treatments. 16S rRNA gene analysis of the component of the plankton filterable at 1.6 μm but retained at 0.22 μm was employed to assess community structural responses within the bacterial population.

5.1.6 Aims

It was decided to test the effects of *in situ* ocean fertilisation with the candidate primary production-limiting nutrients Fe and P, on pelagic bacterial diversity and community structure (method and results detailed in Sections 5.2 and 5.3 respectively). Changes in community profiles at 15 m depth throughout the experiment were ascertained by PCR-DGGE of 16S rRNA genes using *Bacterial* and group-specific primer sets previously designed (Mühling *et al.*, 2007) and validated (see Section 3.2.2).

Using samples from both a P-fertilised area, a P and Fe-fertilised area, and a control station outside of either treated area, the resulting DGGE profiles would be used to test the following hypotheses:

1) Both fertilisation treatments will lead to a change in abundance of heterotrophic bacteria.

2) Fertilisation with both P and Fe leads to changes in the structure of the bacterial community (i.e. bacterial diversity) and abundance of specific bacterial groups and ribotypes.

3) Fertilisation with P alone leads to changes in the structure of the bacterial community (i.e. bacterial diversity) and abundance of specific bacterial groups and ribotypes.

4) The changes observed resulting from the two different treatments (Hypotheses 2 and 3) will be distinct.

5.2 Method

5.2.1 Patch tracing using SF6

Sulphur hexafluoride (SF6) was used to locate regions of water amended by the addition of phosphate, or Fe and phosphate, and to provide a proxy for these nutrients when they were no longer detectable. 500 mL volumes were collected using glass-stoppered bottles on a sampling rosette. SF6 was removed by sparge-cryotrapping, and isolated chromatographically using an electron capture detector (method explained in Watson *et al.*, 1994). Inter-experiment calibrations (RRS Charles Darwin, Experiment One, fertilisation with P; MV Poseidon, Experiment Two, fertilisation with P and Fe), and calibrations against laboratory standards were undertaken.

5.2.2 Sampling

Water samples were collected using a conductivity, temperature and depth (CTD) recorder at various times during every day of the cruise from the RRS Charles Darwin, from stations within the P-seeded patch, and others outside the patch, and from various depths at each station, including one at 15 m. A similar sampling regime was undertaken aboard the MV Poseidon with the P and Fe experiment. Each 9 L sample of water was filtered at 1.6 μm and then through a 0.22 μm Sterivex filter cartridge (Millipore, Molsheim, France), to which lysis buffer (50 mM Tris-HCl pH 8.3, 40 mM EDTA, 0.75 M sucrose) was added. Filters were frozen using liquid-N_2 and stored at -20^{o}C. 33 samples from a total of 195, all from a depth of 15 m, were chosen for bacterial community structure analysis in this study. Details of these samples are listed in Tables 5.1 and 5.2, and their relative locations displayed in Figure 5.3. The following sections detail sampling protocols particular to each of the two fertilisation experiments, hereafter called Experiment One (P) and Experiment Two (Fe and P).

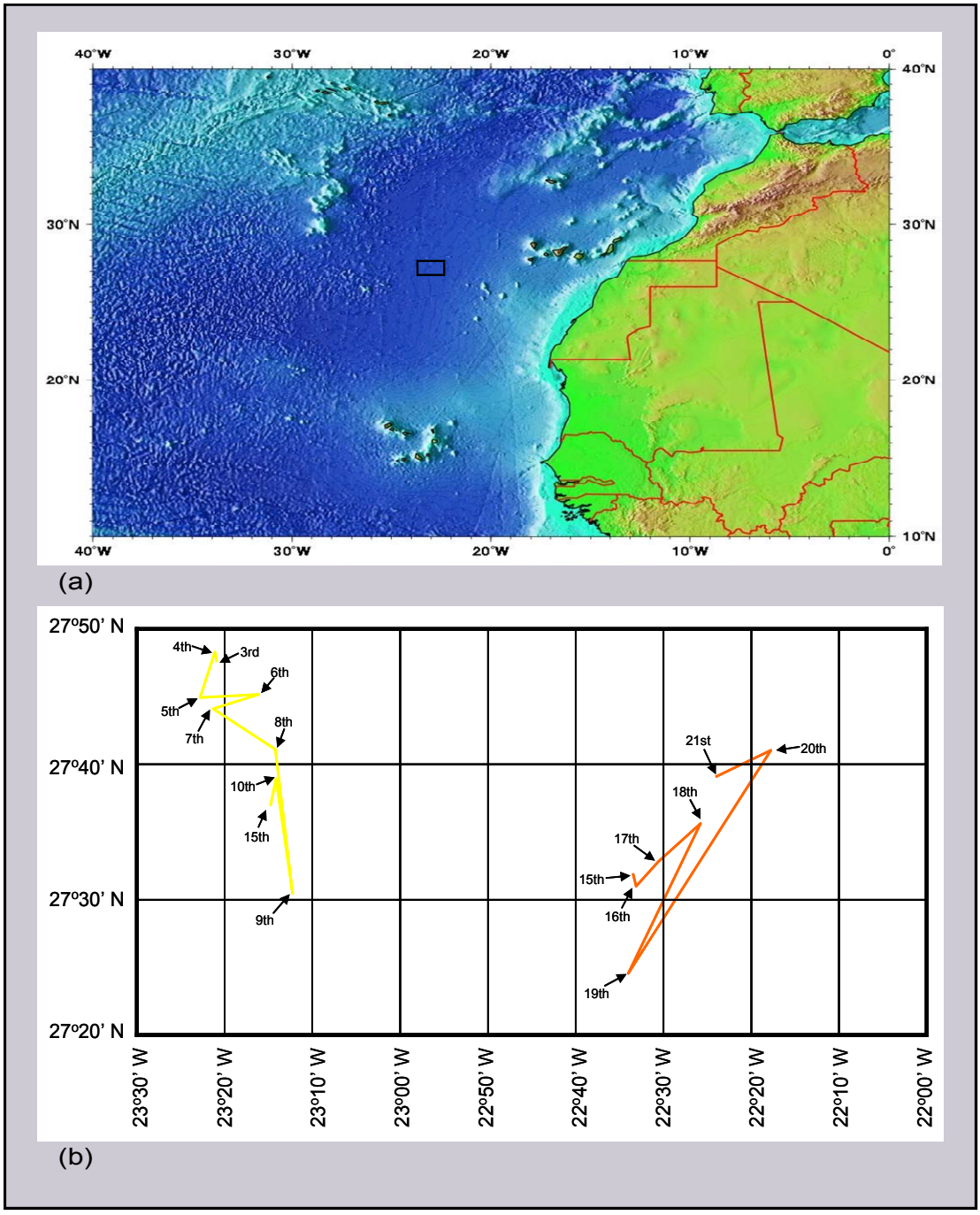


Figure 5.3. FeeP experimental location showing bathymetry (a), enlarged to display ships' courses and in-patch sampling stations (b). Experiment One, RRS Charles Darwin, yellow track; Experiment Two, MV Poseidon, orange track. Courses follow the fertilised patches by proxy tracking of SF6. Out-of-patch samples were taken where there was no SF6 signal.

5.2.3 Experiment One (ocean fertilisation with phosphorus alone)

Following horizontal and vertical mapping, 20 tonnes of anhydrous monosodium phosphate was added at 10 m depth over an area of approximately 25 km^2. The area centred on 27° 48' N, 23° 18' W and SF6 was used to trace amended

waters. Daily water samples were taken, from 5 m, 15 m, and 40 m, between the 5th and 15th May 2004 (Section 5.2.2). These were filtered and the microbial DNA extracted according to the method detailed in Sections 2.2.1 and 5.2.5.

Table 5.1. Environmental DNA samples, from Experiment One (P), used for PCR-DGGE. All samples were 9 L water taken from 15 m depth.

Date[a]	*Hours*[b]	*Time*[c]	*In/Out*[d]
3rd	0	0300	P
4th	0	0300	P
5th	5	0400	Being laid
5th	10	1200	I
5th	10	1200	I
6th	15	0300	I
6th	20	1000	I
6th	30	1500	O
7th	40	0300	O
7th	50	1000	O
7th	50	1500	I
8th	60	0300	I
8th	70	1000	I
8th	70	1400	O
9th	90	0300	O
9th	100	1400	I
10th	110	0300	I
15th	230	0000	I
15th	240	1500	O

[a] Date in May 2004
[b] Hours elapsed since deployment of P patch
[c] Time of day that samples were taken
[d] Sampling position relative to the P patch; P, prior to patch being laid; I, in patch; O, out of patch

5.2.4 Experiment Two (ocean fertilisation with phosphorus and iron)

Following horizontal and vertical mapping, 5 tonnes of acidified iron sulphate were added at 10 m depth, over an area of approximately 25 km^2 centred at 27° 30' N, 22° 30' W. 12 hours later 20 tonnes of anhydrous monosodium phosphate were added along the same track over the iron. Water samples were taken daily from 5 m, 15 m, and 40 m, between 16th and 22nd May 2004, and their handling and treatment was the same as for those from Experiment One.

Table 5.2. Environmental DNA samples, from Experiment Two (Fe, subsequently overlaid with P), used for PCR-DGGE. All samples were 9 L seawater taken from 15 m depth.

Date[a]	***Hours***[b]	***Time***[c]	***In/Out***[d]
	(First addition, Fe only)		
15th	0	0000	P
16th	5	1600	Being laid
17th	30	0300	I
17th	40	1000	O
17th	40	1400	I
	(Second addition, overlay of P)		
18th	50 (10)	0300	I
18th	60 (20)	1400	O
19th	80 (40)	0300	O
19th	90 (50)	1400	I
20th	100 (60)	0400	I
20th	110 (70)	1000	I
21st	110 (70)	0300	O
21st	120 (80)	1000	O
21st	130 (90)	1600	I

[a] Date in May 2004
[b] Hours elapsed since deployment of Fe patch, and, in brackets, since deployment of subsequent P overlay
[c] Time of day that samples were taken
[d] Sampling position relative to the P and Fe patch; P, prior to patch being laid; I, in patch; O, out of patch

5.2.5 DNA extraction

DNA was extracted from the Sterivex filters, by Christopher Bellas of PML, using the protocol outlined in Section 2.2.1, with the following minor modifications. Incubations with lysozyme and proteinase K/SDS were for 45 minutes and 1 hour respectively. Rinsing of lysate residue was carried out with 700 µL SET buffer for 15 minutes. DNA precipitation involved centrifugation for 10 minutes at 3000 x g. Precipitated DNA was concentrated, using a concentrator (Centricon), from 3 mL to 200 µL at 1000 x g for 30 minutes, resuspended in 1 mL H_2O^{MQ} and concentrated again to 100 µL, again at 1000 x g for 30 minutes. DNA concentration was repeated three times to give triplicate samples from each filter.

5.2.6 PCR-DGGE

16S rRNA gene fragments were amplified from environmental template DNA by PCR according to the protocol detailed in Section 2.2.2 and using the same primer sets as those listed in Table 3.1. Environmental amplicon mixtures were resolved into community-specific banding profiles using DGGE as outlined in Section 2.4. Profile patterns were visualised (Section 2.4.2) and digitised (Section 2.4.4) in preparation for visual and statistical analyses. Bands integral to changes in profiles were excised from gels and sequenced (Section 2.4.3). 16S rRNA partial gene sequences were used for *in silico* identification of interesting members of the community, by the method detailed in Section 2.3.6.1.

5.2.7 Statistical analysis

In Section 5.3.2 DGGE profiles derived from the microbial communities from within both amended water patches and from outside either patch, using 16S rRNA genes amplified with each pair of group-specific primers (Table 2.1), are compared using multidimensional scaling (MDS). This analysis reduces each DGGE profile to a single point in two-dimensional space, thus helping the visualisation of the relationships between different profiles. The method is explained in Section 2.4.4.2.

5.3 Results

5.3.1 Background results from the FeeP cruise

5.3.1.1 Monitoring distribution and duration of fertilised patches

Nutrient additions were tracked using the inert marker sulphur hexafluoride (SF6). Following this chemical, the area and depth extents of each nutrient fertilisation patch were monitored, affording clear confirmation of whether or not a

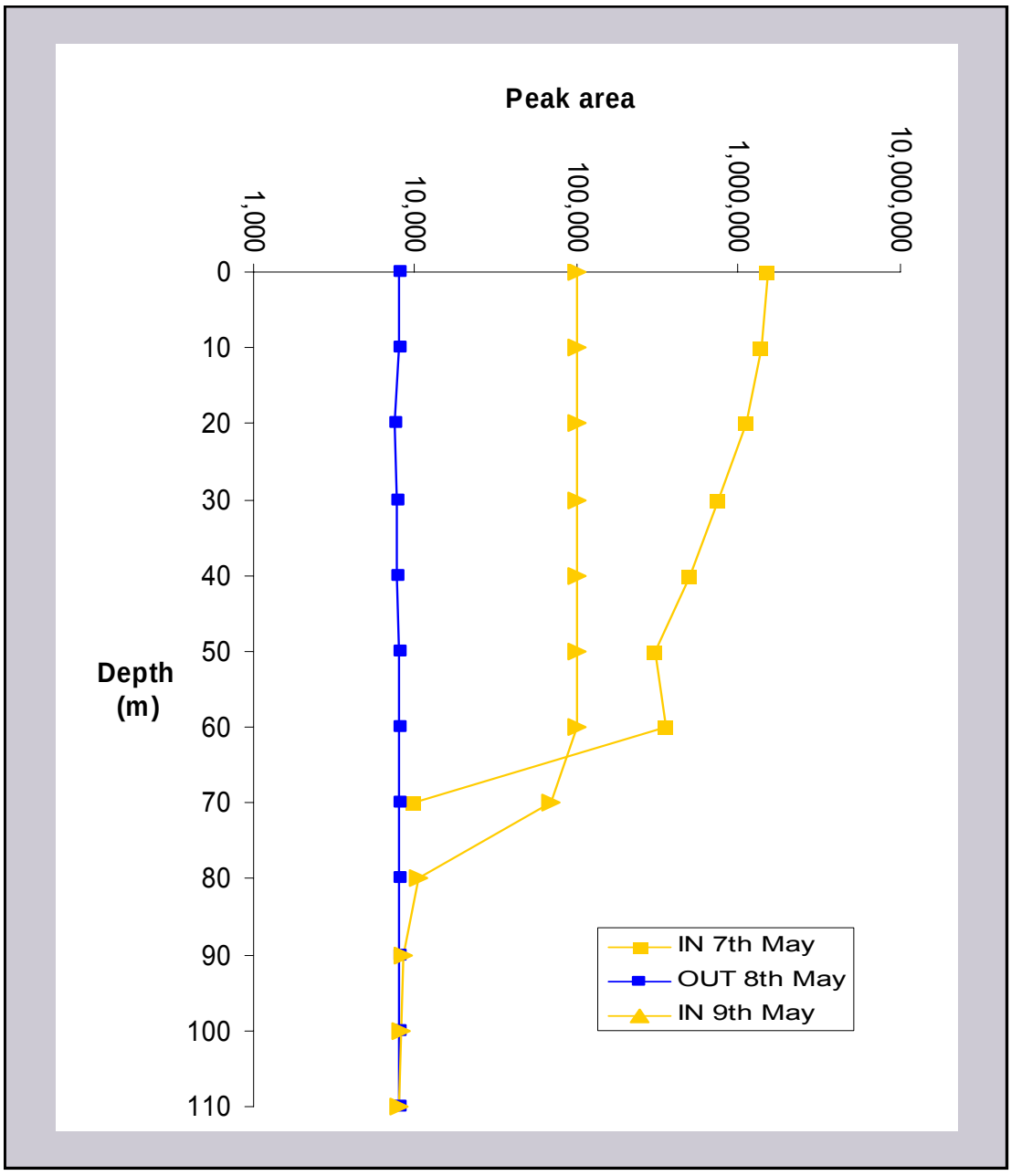


Figure 5.4. SF6, released simultaneously with P, measured as a proxy for depth (m) and concentration (peak area) of added P in Experiment One. Measurements were taken aboard MV Poseidon.

sampling station was in, or out, of patch. Figures 5.4 and 5.5 chart the distribution of SF6 deployed with P in Experiment One, and with Fe and P in Experiment Two, respectively.

During the course of Experiment One SF6 distribution (Figure 5.4) deepened from 40 m to 80 m, consistent with the surface mixed layer.

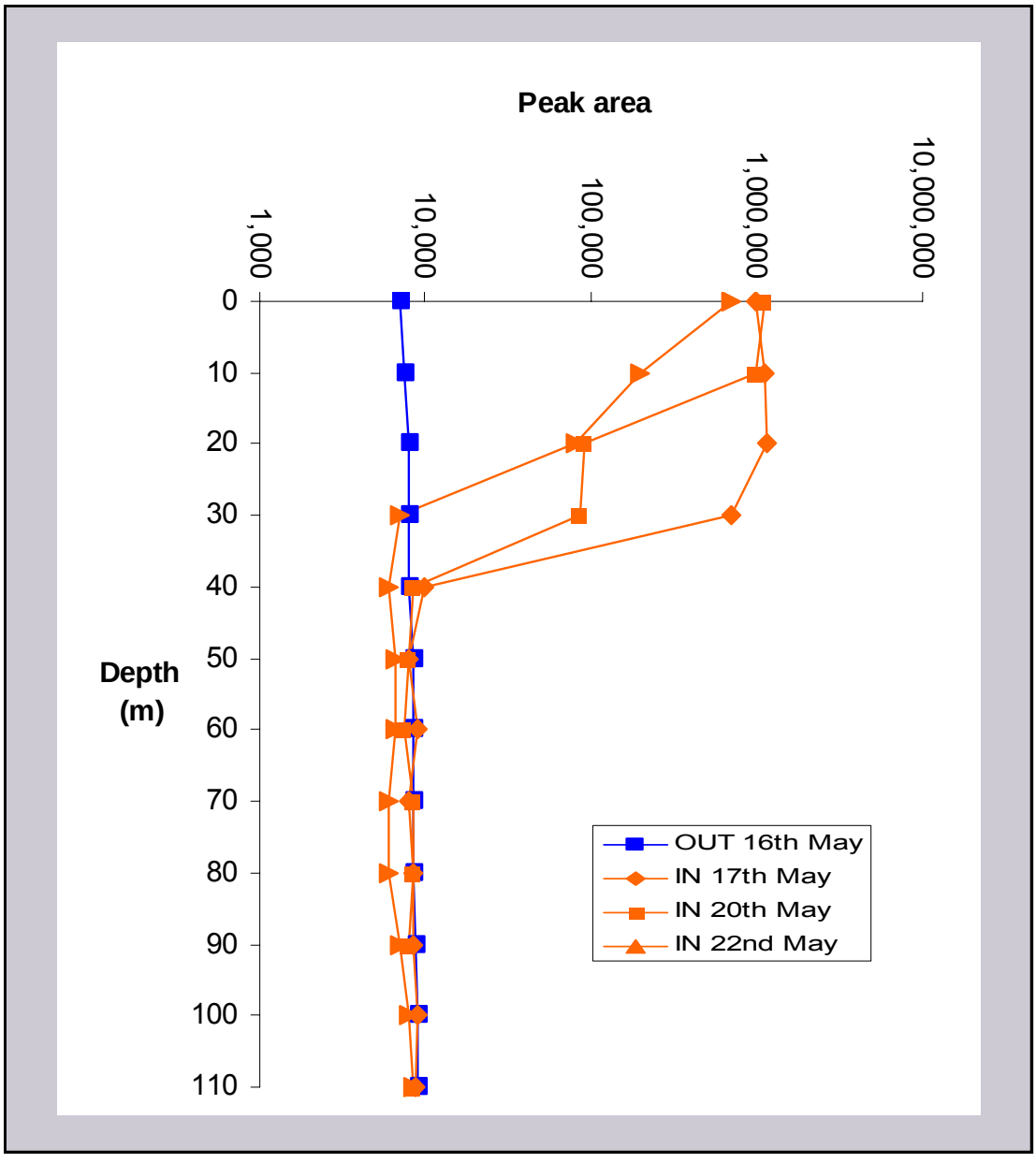


Figure 5.5. SF6, released simultaneously with the initial Fe deployment, measured as a proxy for depth (m) and concentration (peak area) of added P and Fe, during Experiment Two. Measurements were taken aboard MV Poseidon.

During the course of Experiment Two SF6 depth distribution (Figure 5.5), and the mixed layer depth, remained constant between 28 m and 35 m.

5.3.1.2 Preliminary shipboard analyses

N-fixation rates were low, but rose slightly toward the end of both experiments. Highest rates were associated with increases in abundance of the diazotrophic bacterium *Trichodesmium* sp. (C. Law, and A. Rees, unpublished).

Dark community respiration exceeded gross production, making the in-patch water columns, in both experiments, net heterotrophic (C. Robinson, unpublished). There was no evidence of a diel trend in the daily respired 0.8 mmol O_2 m^{-3}, which was less than compensated for by the 0.4 mmol O_2 m^{-3} photosynthetically produced. This implies a possible net increase in inorganic C within the system. *Prochlorococcus* sp. was the most abundant autotroph (Dixon, 2008). The majority of the respiration was accounted for by organisms under 0.8 μm, the micro- and nanoheterotrophs.

5.3.1.3 Microzooplankton

Possibly in keeping with the findings of Thingstad *et al*. (2005, see Section 5.1.4 above), a significant increase in growth rate (μ) of the cyanobacterium *Synechococcus* sp. was observed, along with predation of this organism by ciliates and heterotrophic dinoflagellates, following fertilisation with P in Experiment One (S. Kimmance, and E. Fileman, pers. comm.). This study was conducted in shipboard microcosms, containing serial dilutions of CTD-retrieved samples. A similar effect was not observed in samples from the Fe and P patch in Experiment Two.

5.3.1.4 Microbial productivity

Net primary production and chlorophyll *a* biomass are limited by N and P in the oligotrophic northeast Atlantic Ocean, both increasing as a result of the addition of Fe (Dixon, 2008). Bacterial production increased following P fertilisation in Experiment One, but again this was not mirrored following Fe and P fertilisation in Experiment Two.

In shipboard bottle experiments for both experiments, heterotrophic bacterial abundance and production, as well as *Prochlorococcus* sp. and *Synechococcus* sp. abundance, showed the greatest increases following *in vitro* addition of either cobalt or zinc to CTD-retrieved mesoscale P-enriched, and P and Fe-enriched, patch water. This suggests microbial, including bacterial, growth rates were multiple nutrient limited (Dixon, 2008), prior to being N-limited.

5.3.2 DGGE profile analyses of different bacterial groups

5.3.2.1 *Bacteria*

DNA samples were selected from both experimental patches, as well as from outside either fertilised area, and subjected to a nested PCR with two sets of primers universal for the domain *Bacteria* (see protocol in Section 3.2.2). The products were used to produce the DGGE gel shown in Figure 5.6. The times of samples taken in Experiment One and Experiment Two do not match exactly. This is because they were taken independently aboard different ships. The Experiment Two sample taken at 5 hours was after fertilisation with Fe only, and that at 50 hours was 50 hours after Fe fertilisation and 10 hours after additional P fertilisation. It must also be remembered the two experiments did not overlap in either time or space.

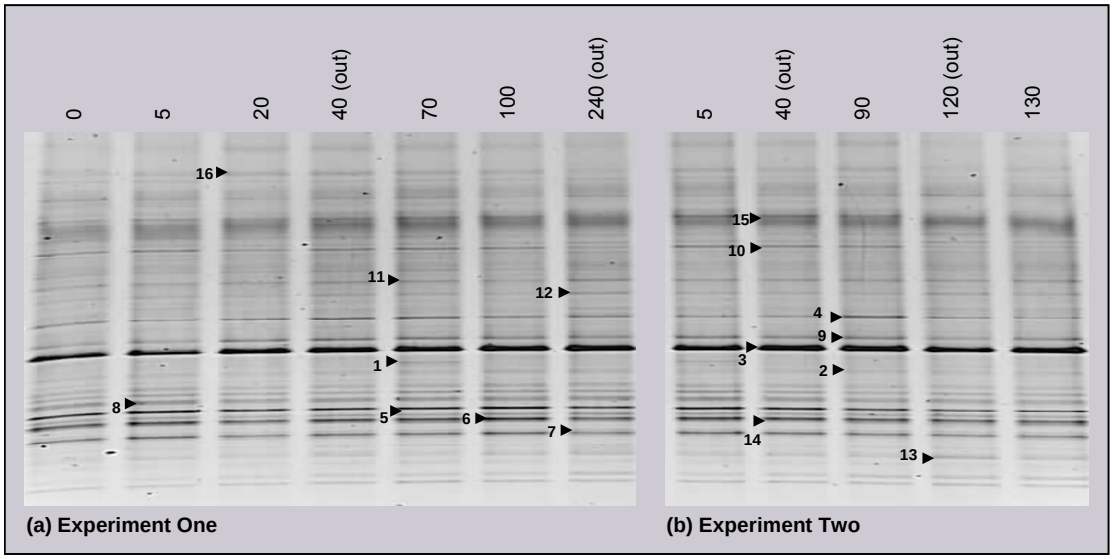


Figure 5.6. DGGE profiles of 16S rRNA partial genes amplified using primers designed to be specific for *Bacteria*, detailed in Table 3.1. DGGE gradient was 30 % to 60 %. Two water patches were sampled, the first amended with P (a), and the second with Fe and P (b).

DGGE profiles, shown in Figure 5.6, were uniform throughout for samples from both experimental patches and from the surrounding, HNLC, environment. The majority of sequenced bands aligned with *Alphaproteobacteria* (Table 5.3). The most prominent band (3) however aligned in all profiles with the Group IIa *Cyanobacteria*.

In keeping with the lack of variation between the *Bacteria* DGGE profiles (Figure 5.6), the corresponding MDS plot (Figure 5.7) shows little evidence of treatment-specific clustering.

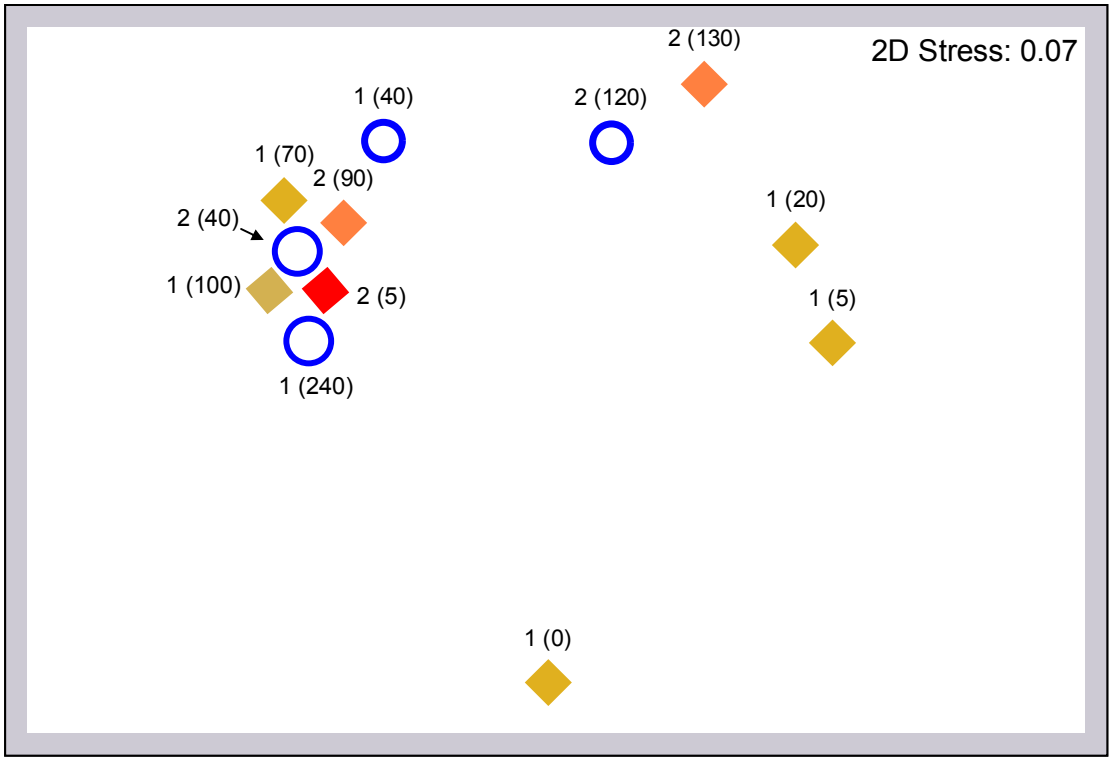


Figure 5.7. MDS plot showing relationships between putatively *Bacteria*-specific DGGE profiles shown in Figure 5.6. Non-metric scaling using Bray Curtis measure of similarity between profiles in a binary matrix based on DGGE band profiles in Figure 5.6. Experiment One (P, gold diamonds), Experiment Two (Fe and P, orange diamonds; Fe, red diamond), out-of-patch samples (blue circles). Numbers in brackets refer to hours elapsed since fertilisation.

Nucleotide sequences of the bands excised from the gel in Figure 5.6 were phylogenetically aligned to 16S rRNA gene sequences in the Greengenes database. The majority of sequences aligned within the *Alphaproteobacteria*, with four being identified to the level of genus. Three of these, *Bosea* (band 4), *Methylosinus* (band 8), and *Devosia* (band 11), are from members of the nitrogen-fixing *Rhizobiales*, although none are known to be marine.

Table 5.3. Phylogenetic classification of 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 5.6

Band Number	***PhylogeneticAlignment*** [a]	***Similarity*** [b]
1	*Alphaproteobacteria*; unclassified *Alphaproteobacteria*	85
2	*Alphaproteobacteria*; unclassified *Alphaproteobacteria*	77
3	*Cyanobacteria*; Family II; GpIIa	78
4	*Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bosea*	86
5	*Alphaproteobacteria*; unclassified *Alphaproteobacteria*	77
6	*Alphaproteobacteria*; unclassified *Alphaproteobacteria*	78
7	*Alphaproteobacteria*; unclassified *Alphaproteobacteria*	75
8	*Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylosinus*	78
9	*Alphaproteobacteria; Rickettsiales; Odyssella*	66
10	n/s	
11	*Alphaproteobacteria; Rhizobiales; Hyphomicrobiaceae; Devosia*	85
12	n/s	
13	n/s	
14	n/s	
15	*Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bosea*	83
16	n/s	

[a] Phylogenetic classification, where possible to the level of genus. Sequences from excised bands were aligned with, and compared to, homologous 16S rRNA gene sequences in the Greengenes database (http://greengenes.lbl.gov) and classified accordingly.
[b] Degree of base similarity between an excised DGGE band sequence and that of its closest fully-aligned homologue in the Greengenes database (http://greengenes.lbl.gov).
* Base sequences for 16S rRNA gene fragments, amplified from DGGE bands, are listed in Appendix I.

5.3.2.2 *Alphaproteobacteria*

The DGGE gel shown in Figure 5.8 was produced following the nested PCR protocol detailed in Section 3.2.2, using primers designed to be *Alphaproteobacteria*-specific.

The profiles of in-patch samples from Experiment One and Experiment Two are mostly invariant. There are distinct variations in the profiles from out of patch samples, especially evident in profiles from the two different samples taken at 40

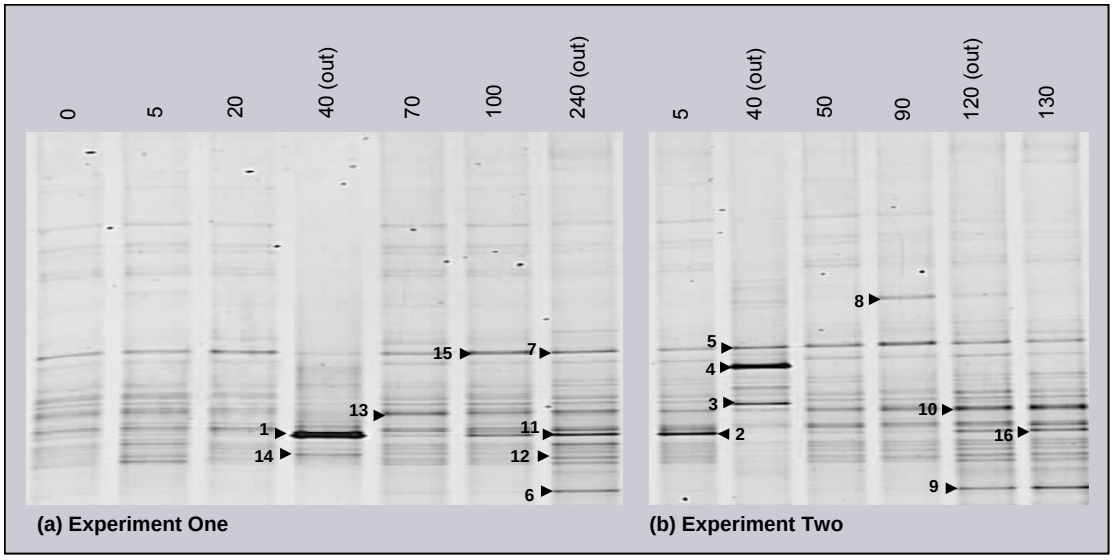


Figure 5.8. DGGE profiles of 16S rRNA partial genes amplified using primers designed to be specific for *Alphaproteobacteria*, detailed in Table 3.1. DGGE gradient was 30 % to 60 %. Two water patches were sampled, the first amended with P (a), and the second with Fe and P (b).

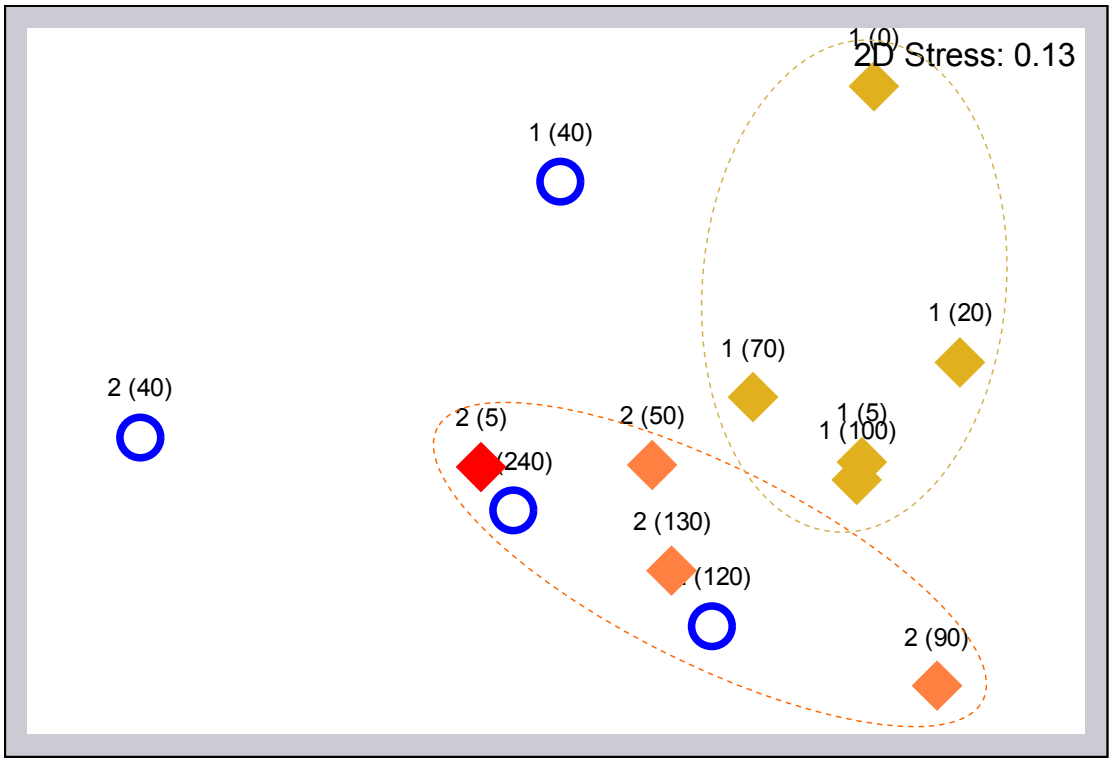


Figure 5.9. MDS plot of putatively *Alphaproteobacteria*-specific DGGE profiles shown in Figure 5.8, produced by non-metric analysis using Bray Curtis similarity measure. Experiment One (P, gold diamonds), Experiment Two (Fe and P, orange diamonds; Fe, red diamond), out-of-patch samples (blue circles). Numbers in brackets refer to hours elapsed since fertilisation. Dotted circles emphasise clustering of profiles according to location.

hours.

All but two of the sequenced bands are classified as *Alphaproteobacteria*. The sequence from band 11 aligns with the mammalian gut-residing *Firmicute*, *Ruminococcus*, and the sequence from band 13 with the epsilonproteobacterium *Sulfurovum*. However both alignments are with low sequence similarity values of 66 % and 50 %, respectively. The two highest similarity alignments are the sequence from band 14 having 95 % similarity to the methanol-degrader *Methylopila*, and the sequence from band 1 having 91 % similarity to the ethanol-oxidising *Saccharibacter*. This genus is associated with natural fermentation such as occurs in flowers.

MDS analysis (Figure 5.9) reveals the community profiles within the amended patch of Experiment One cluster apart from those profiles generated from out of patch samples. The profiles from the amended patch of Experiment Two are less distinct.

Table 5.4. Phylogenetic classification of 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 5.8

Band Number	***Phylogenetic Alignment*** [a]	***Similarity*** [b]
1	*Alphaproteobacteria; Rhodospirillales; Saccharibacter*	91
2	*Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bosea*	74
3	n/s	
4	*Alphaproteobacteria;* unclassified	85
5	n/s	
6	*Alphaproteobacteria; Rickettsiales; Odyssella*	48
7	*Alphaproteobacteria; Rhizobiales; Mesorhizobium*	83
8	n/s	
9	n/s	
10	*Alphaproteobacteria;* unclassified	60
11	*Firmicutes; Clostridia; Clostridiales; Ruminococcus*	66
12	*Alphaproteobacteria; Rickettsiales; Odyssella*	70
13	*Epsilonproteobacteria; Campylobacterales; Sulfurovum*	50
14	*Alphaproteobacteria; Rhizobiales; Methylocystaceae; Methylopila*	95
15	*Alphaproteobacteria; Rhodospirillales; Acetobacteraceae; Stella*	62
16	n/s	

[a] Phylogenetic classification, where possible to the level of genus. Sequences from excised bands were aligned with, and compared to, homologous 16S rRNA gene sequences in the Greengenes database (http://greengenes.lbl.gov) and classified accordingly.

[b] Degree of base similarity between an excised DGGE band sequence and that of its closest fully-aligned homologue in the Greengenes database (http://greengenes.lbl.gov).

* Base sequences for 16S rRNA gene fragments, amplified from DGGE bands, are listed in Appendix I.

Nucleotide sequence alignment results listed in Table 5.4 show 90 % of ribotypes were amplified from alphaproteobacterial 16S rRNA gene sequences present in the DNA extracts.

5.3.2.3 *Betaproteobacteria*

The *Betaproteobacteria*-specific PCR-DGGE protocol listed in Section 3.2.2 was employed to produce the set of community 16S rRNA gene profiles shown in Figure 5.10. As discussed in Section 3.3.2.2 the primers proved to be slightly non-specific.

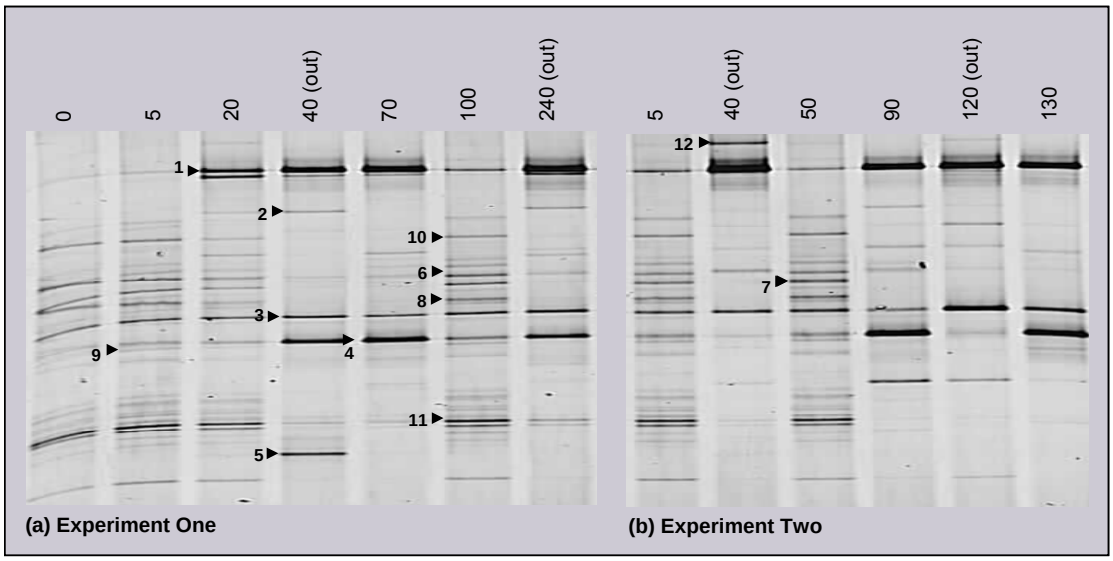


Figure 5.10. DGGE profiles of 16S rRNA partial genes amplified using primers designed to be specific for *Betaproteobacteria*, detailed in Table 3.1. Community PCR products were generated using protocol in Section 3.2.2. DGGE gradient was 30 % to 60 %. Two water patches were sampled, the first amended with P (a), and the second with Fe and P (b).

Profiles for 70 hours in Experiment One and 90 and 130 hours in Experiment Two, together with all out-of-patch profiles appear similar and differ markedly from the remaining profiles within the two patches. These in-patch profiles (Experiment One, 0, 5, 20, 100 hours; Experiment Two, 5, 50 hours) are highly invariant. Bands 8 and 9 have sequences aligning with the *Alphaproteobacteria,* and *Gammaproteobacteria* respectively. All other sequences align within the *Betaproteobacteria*. Band 4 is consistently present but varies in intensity, and its sequence matches the genus *Janthinobacterium,* which contains many endosymbionts of metazoans.

With the same exceptional profiles (1-70, 2-90, 2-130) as for the DGGE gel analysis (above) the profiles from Experiment One cluster together with those from Experiment Two, and separately from the out of patch profiles, which lack the *Gammaproteobacteria* bands, which resulted from mis-directed amplification during PCR using primer pair 59f/682r.

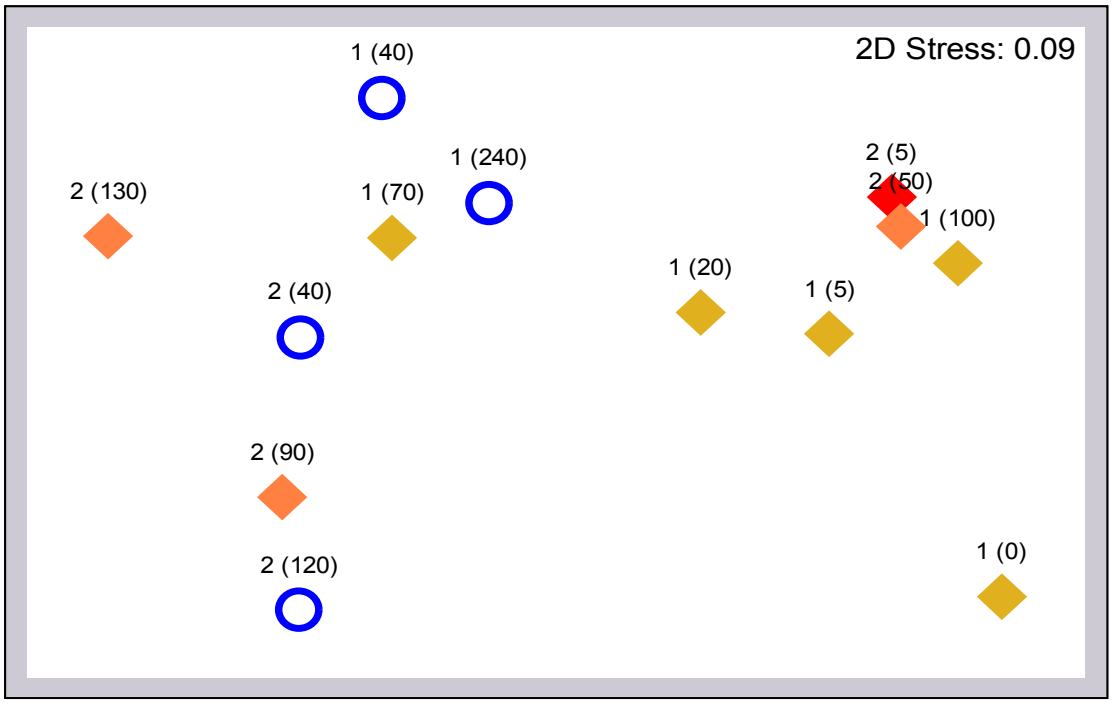


Figure 5.11. Non-metric MDS plot using Bray Curtis similarity of putatively *Betaproteobacteria*-specific DGGE profiles (Figure 5.10) produced with putatively *Betaproteobacteria*-specific PCR primers (Table 3.1). Experiment One (P, gold diamonds), Experiment Two (Fe and P, orange diamonds, out-of-patch samples (blue circles). Numbers in brackets refer to hours elapsed since fertilisation.

Table 5.5. Phylogenetic classification of 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 5.10

Band Number	***Phylogenetic Alignment*** [a]	***Similarity*** [b]
1	*Betaproteobacteria; Burkholderiales; Oxalobacteraceae; Oxalobacter*	81
2	*Betaproteobacteria; Burkholderiales;* unclassified *Oxalobacteraceae*	89
3	*Betaproteobacteria; Burkholderiales;* unclassified *Oxalobacteraceae*	96
4	*Betaproteobacteria; Burkholderiales; Janthinobacterium*	84
5	*Betaproteobacteria; Burkholderiales; Oxalobacteraceae; Herbaspirillum*	92
6	*Betaproteobacteria; Nitrosomonadales; Gallionellaceae; Gallionella*	72
7	n/s	
8	*Alphaproteobacteria; Sphingomonadales; Erythromicrobium*	73
9	*Gammaproteobacteria;* unclassified *Gammaproteobacteria*	77
10	*Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria*	81
11	*Betaproteobacteria; Methylophilales; Methylobacillus*	79
12	*Betaproteobacteria; Burkholderiales; Oxalobacteraceae; Oxalobacter*	81

[a] Phylogenetic classification, where possible to the level of genus. Sequences from excised bands were aligned with, and compared to, homologous 16S rRNA gene sequences in the Greengenes database (http://greengenes.lbl.gov) and classified accordingly.
[b] Degree of base similarity between an excised DGGE band sequence and that of its closest fully-aligned homologue in the Greengenes database (http://greengenes.lbl.gov).
* Base sequences for 16S rRNA gene fragments, amplified from DGGE bands, are listed in Appendix I.

Eighty two percent of sequences from bands excised from the *Betaproteobacteria*-specific DGGE gel (Figure 5.10) aligned most closely with betaproteobacterial 16S rRNA genes (Table 5.5). This contrasts with the low specificity (30 %) of the *Betaproteobacteria*-specific primer pair (359f/682r) achieved in amplifying 16S rRNA genes from another region of the oligotrophic North Atlantic Ocean, and is more consistent with the good specificity (100 %) achieved with samples from a Norwegian fjord (see Section 3.3.2 and Table 3.4). The most strongly aligned sequences, from bands 3 and 5, align with members of the *Oxalobacteracerae*, which include the nitrogen-fixing *Herbaspirillum*.

5.3.2.4 *Gammaproteobacteria*

Figure 5.12 shows DGGE community profiles made using primers designed to be specific for the *Gamma* subgroup of the *Proteobacteria*, obtained following the protocol in Sections 2.2.2 and 3.2.2.

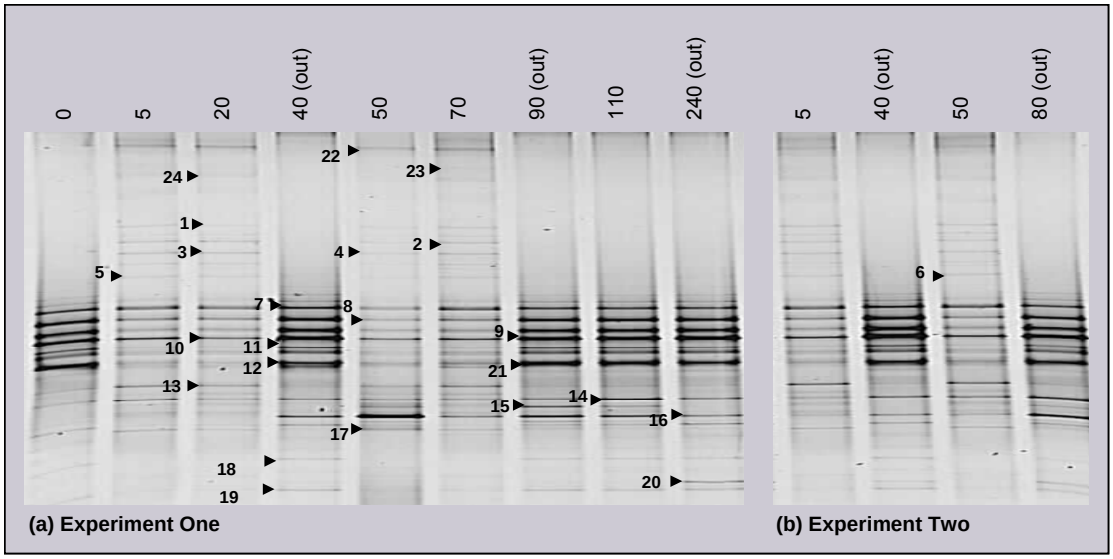


Figure 5.12. DGGE profiles of 16S rRNA partial genes amplified using primers designed to be specific for *Gammaproteobacteria*, detailed in Table 3.1. Community PCR products were generated using the protocol in Table 3.2. DGGE gradient was 30 % to 60 %. Two water patches were sampled, the first amended with P (a), and the second with Fe and P (b).

With the exception of the profile representing the community at 110 hours in the first experimental patch, which matches out-of-patch profiles, there are clear profile differences between the *Gammaproteobacteria* communities inside and outside the treated areas. Bands 1 to 6 and 17 are restricted to amended water. On the other hand bands 18 to 20 are only found in out-of-patch profiles. Figure 5.21 illustrates the proportion of *Gammaproteobacteria* ribotypes not being universal to all profiles, i.e. they are restricted to treated or untreated water. There is only one sequence (from band 22) aligning with sequences from outside of the target group. This sequence aligns well with *Winogradskyella*, a member of the *Bacteroidetes.*

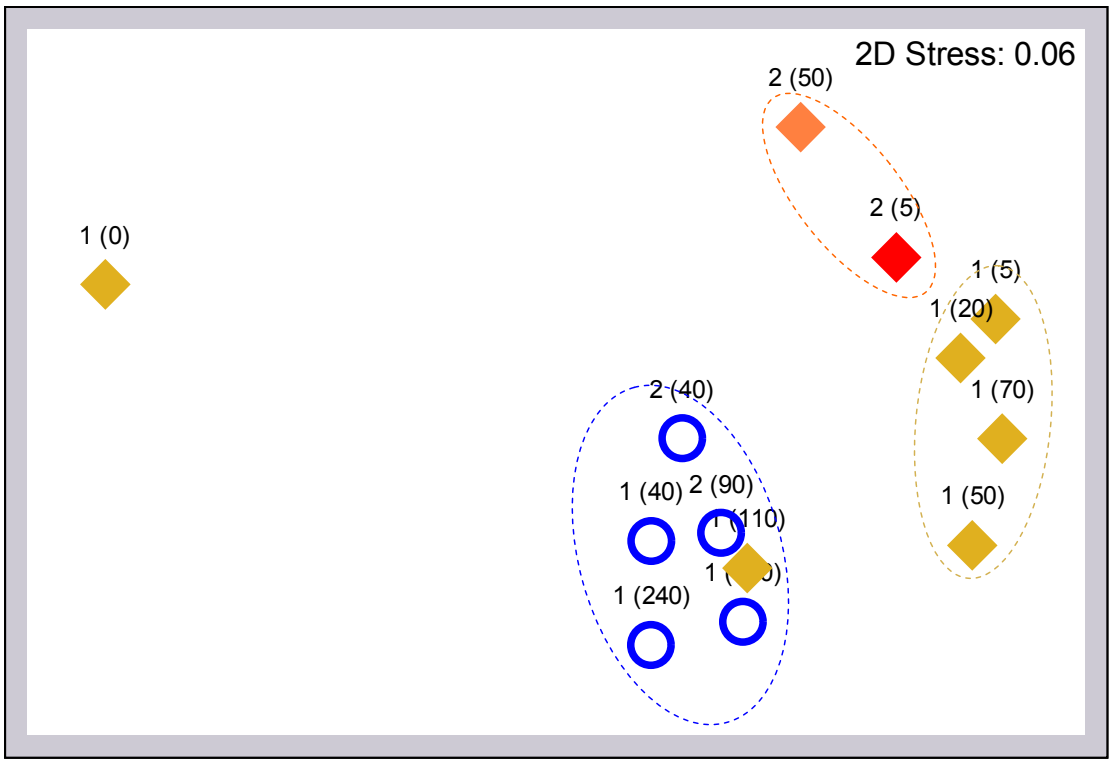


Figure 5.13. Non-metric MDS plot (Bray Curtis similarity) showing clear clustering of putatively *Gammaproteobacteria*-specific DGGE profiles (Figure 5.12) according to origin of environmental DNA (dotted circles). Experiment One (P, gold diamonds), Experiment Two (Fe and P, orange diamonds, out-of-patch samples (blue circles). Numbers in brackets refer to hours elapsed since fertilisation.

The MDS representation (Figure 5.13) of the profiles from Figure 5.12 shows distinct clustering according to treatment. Profiles from the first and second experimental patches group separately, and both cluster away from the out-of-patch profiles. The Experiment One profile from 110 hours differs from all other in-patch profiles by lacking bands 1 to 6 (see discussion in Section 5.4). Profile 1 (0 hours) also lacks these bands, but this profile represents pre-fertilisation and therefore clusters more closely to out-of-patch profiles than to in-patch profiles. The 130 hour template DNA failed to yield a PCR product so there are only two in-patch time points for Experiment Two.

Table 5.6. Phylogenetic classification of 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 5.12

Band Number	***Phylogenetic Alignment*** [a]	***Similarity*** [b]
1	n/s	
2	n/s	
3	*Gammaproteobacteria;* unclassified	63
4	*Gammaproteobacteria; Alteromonadales; Marinobacter*	71
5	*Proteobacteria;* unclassified	68
6	n/s	
7	*Gammaproteobacteria; Oceanospirillales;* unclassified *Halomonadaceae*	67
8	*Gammaproteobacteria; Pseudomonadales; Pseudomonas*	75
9	*Gammaproteobacteria; Pseudomonadales; Pseudomonas*	77
10	*Gammaproteobacteria;* unclassified *Oceanospirillales*	78
11	*Gammaproteobacteria; Pseudomonadales; Pseudomonas*	79
12	*Gammaproteobacteria; Enterobacteriales; Buchnera*	79
13	*Gammaproteobacteria;* unclassified	77
14	*Gammaproteobacteria; Alteromonadales; Glaciecola*	74
15	*Gammaproteobacteria; Oceanospirillales; Oceanospirillum*	71
16	*Gammaproteobacteria;* unclassified	74
17	*Gammaproteobacteria; Alteromonadales; Alteromonas*	90
18	*Gammaproteobacteria; Alteromonadales; Alteromonas*	96
19	*Gammaproteobacteria; Alteromonadales; Alteromonas*	99
20	*Gammaproteobacteria;* unclassified	85
21	*Gammaproteobacteria;* unclassified	77
22	*Bacteroidetes; Flavobacteria; Flavobacteriales; Winogradskyella*	93
23	n/s	
24	*Gammaproteobacteria; Alteromonadales; Teredinibacter*	54

[a] Phylogenetic classification, where possible to the level of genus. Sequences from excised bands were aligned with, and compared to, homologous 16S rRNA gene sequences in the Greengenes database (http://greengenes.lbl.gov) and classified accordingly.

[b] Degree of base similarity between an excised DGGE band sequence and that of its closest fully-aligned homologue in the Greengenes database (http://greengenes.lbl.gov).

* Base sequences for 16S rRNA gene fragments, amplified from DGGE bands, are listed in Appendix I.

Ninety five percent of alignments, for sequences obtained from excised bands from the "*Gammaproteobacteria*-specific" gel (Figure 5.12) are for 16S rRNA gene clones from that same phylum. Of those *Gammaproteobacteria* matches to genus level, all are with marine bacteria. The strongest alignments were to the flagellated *Alteromonas* (bands 17-19). Weaker alignments included to the lipid-degrading *Marinobacter* (band 4), *Pseudomonas* (bands 8 and 9), and *Glaciecola* (band 14). The sequences from bands 10 and 15 align weakly with the genus *Oceanospirillum*, some

of which are endosymbionts of deep-sea polychaete worms which feed on fish and cetacean carcasses.

5.3.2.5 *Firmicutes*

The DGGE profiles in Figure 5.14 were obtained using *Firmicutes*-specific primers, and the PCR protocols detailed in Sections 2.2.2 and 3.2.2.

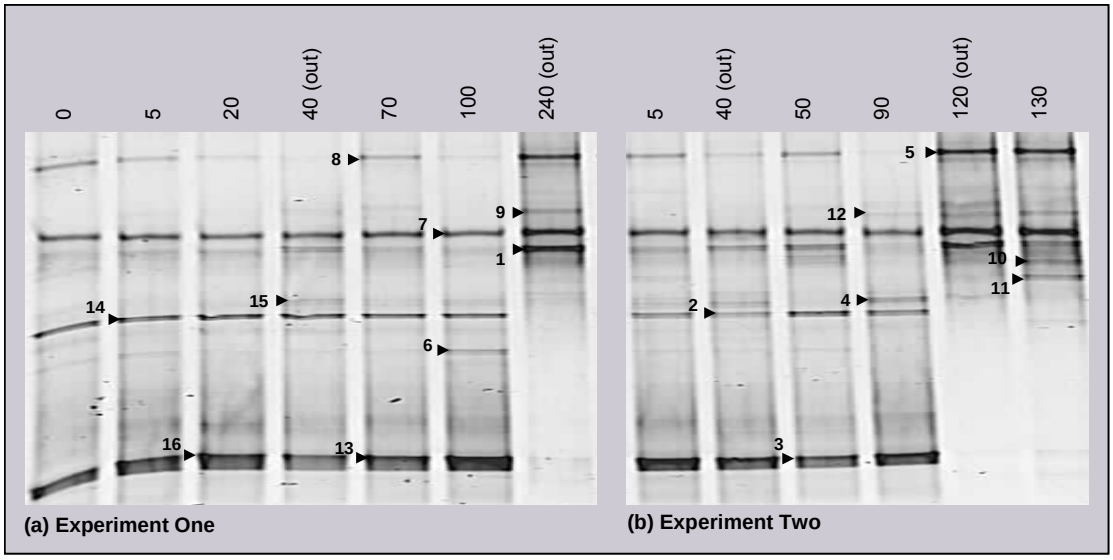


Figure 5.14. DGGE profiles of 16S rRNA partial genes amplified using primers designed to be specific for *Firmicutes*, detailed in Table 3.1. Group-specific PCR products were generated using the protocol in Table 3.2. DGGE gradient was 30 % to 60 %. Two water patches were sampled, the first amended with P (a), and the second with Fe and P (b).

In Figure 5.14 there is no obvious *Firmicutes*-specific profile change correlating with treatment. Profiles 1 (240), 2 (120) and 2 (130) appear different from the rest, in which the nucleotide sequence of band 2 matched the 16S rRNA gene sequence from the halophilic denitrifying firmicute, *Virgibacillus haldenitrificans* (Denariaz *et al*., 1989), following a BLASTn analysis.

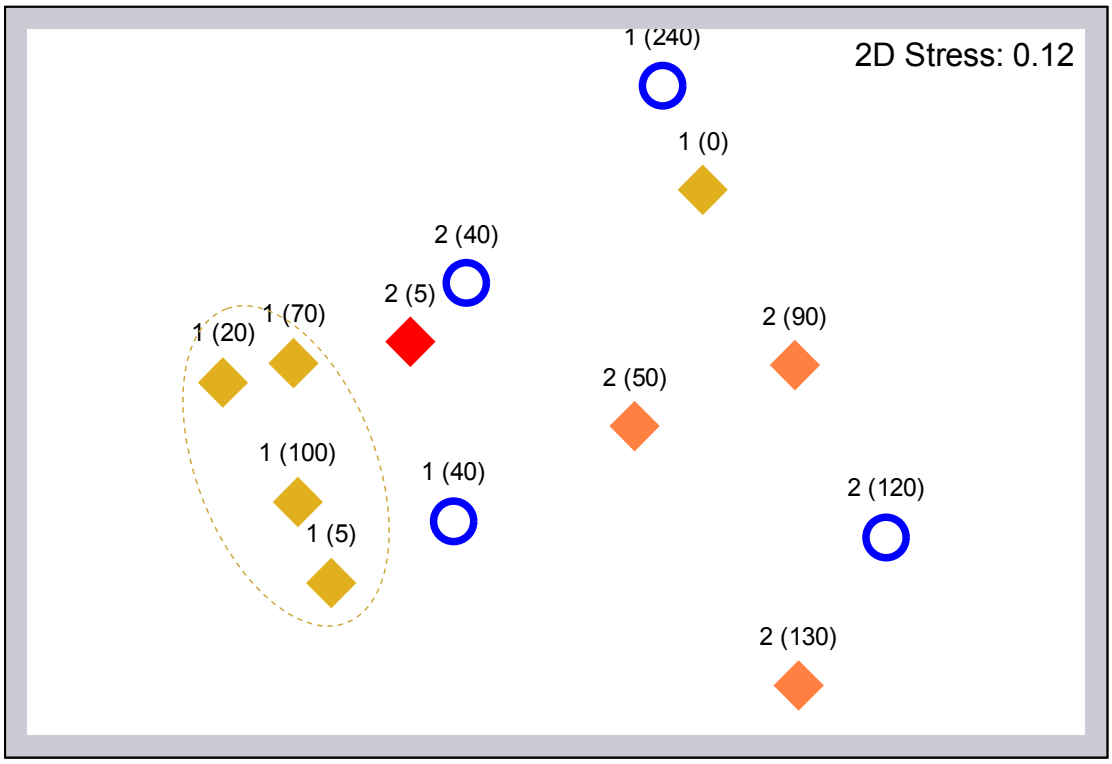


Figure 5.15. MDS plot showing relationships between DGGE profiles of putatively *Firmicutes*-specific 16S rRNA gene fragments. Experiment One (P, gold diamonds), Experiment Two (Fe and P, orange diamonds, out-of-patch samples (blue circles). Numbers in brackets refer to hours elapsed since fertilisation. Dotted circle highlights clustering of DGGE profiles consisting of gene fragments amplified from P-amended waters.

An MDS plot (Figure 5.15) however reveals an extent of clustering of profiles not visually apparent in the DGGE gel (Figure 5.14). Profiles from within Experiment One P-amended water (circled in Figure 5.15) clearly group separately from out-of-patch profiles. Statistical analysis of DGGE profiles shows that the bacterial diversity differs also between in-patches of Experiments One and Two, but not between in- and out-patches of Experiment Two.

Fifty six percent of successful sequence alignments were with organisms from the Gram-positive *Firmicutes* phylum (Table 5.7). However these were mostly with soil-associated organisms, with one ribotype (band 8) aligning with the *Clostridiaceae*, ubiquitous in marine sediments. Other alignments (bands 5, 10, and 11) were to marine representatives of the *Alphaproteobacteria* and *Gammaproteobacteria*. The sequence from band 1, common to all DGGE profiles, aligns with the phylum *Chloroflexi*: phototrophic filamentous bacteria formerly members of the green non-sulphur bacteria. The low specificity of amplified 16S

rRNA gene sequences is in keeping with the original testing of the *Firmicutes*-specific primers (350f/814r) by production of clone libraries from coastal and open ocean environmental DNA samples (Section 3.2.4). Those test results showed no *Firmicutes*

Table 5.7. Phylogenetic classification of 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 5.14

Band Number	***Phylogenetic Alignment*** [a]	***Similarity*** [b]
1	*Chloroflexi; Anaerolineae; Caldilineae;* unclassified *Caldilineacea*	78
2	n/s	
3	n/s	
4	*Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus*	54
5	*Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Thalassobius*	79
6	n/s	
7	*Firmicutes; Thermoanaerobacterales; Caldanaerobacter*	73
8	*Firmicutes; Clostridiales;* unclassified *Clostridiaceae*	70
9	n/s	
10	*Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Sulfitobacter*	67
11	*Gammaproteobacteria; Chromatiales; Chromatiaceae; Thiocystis*	72
12	*Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus*	78
13	n/s	
14	*Firmicutes; Bacillales; Bacillaceae; Geobacillus*	81
15	n/s	
16	n/s	

[a] Phylogenetic classification, where possible to the level of genus. Sequences from excised bands were aligned with, and compared to, homologous 16S rRNA gene sequences in the Greengenes database (http://greengenes.lbl.gov) and classified accordingly.
[b] Degree of base similarity between an excised DGGE band sequence and that of its closest fully-aligned homologue in the Greengenes database (http://greengenes.lbl.gov).
* Base sequences for 16S rRNA gene fragments, amplified from DGGE bands, are listed in Appendix I.

16S rRNA genes were amplified from the open ocean, while 70 % of coastal seawater amplicons matched with *Firmicutes* sequences, these being mostly associated with terrestrial bacteria.

All Greengenes database alignments showed poor nucleotide sequence similarity, all classifications being below 82 % (Table 5.7). It is generally agreed that 97 % homology between two 16S rRNA gene sequences is needed to group two ribotypes at the species level (Munn, 2004). The strongest sequence alignment was band 14 with *Geobacillus*: thermophilic rods abundant in geothermally heated marine sediments.

5.3.2.6 *Planctomycetes*

DGGE 16S rRNA gene fragment profiles (Figure 5.16) were obtained for the *Planctomycetes* using the specific protocol listed in Sections 2.2.2 and 3.2.2.

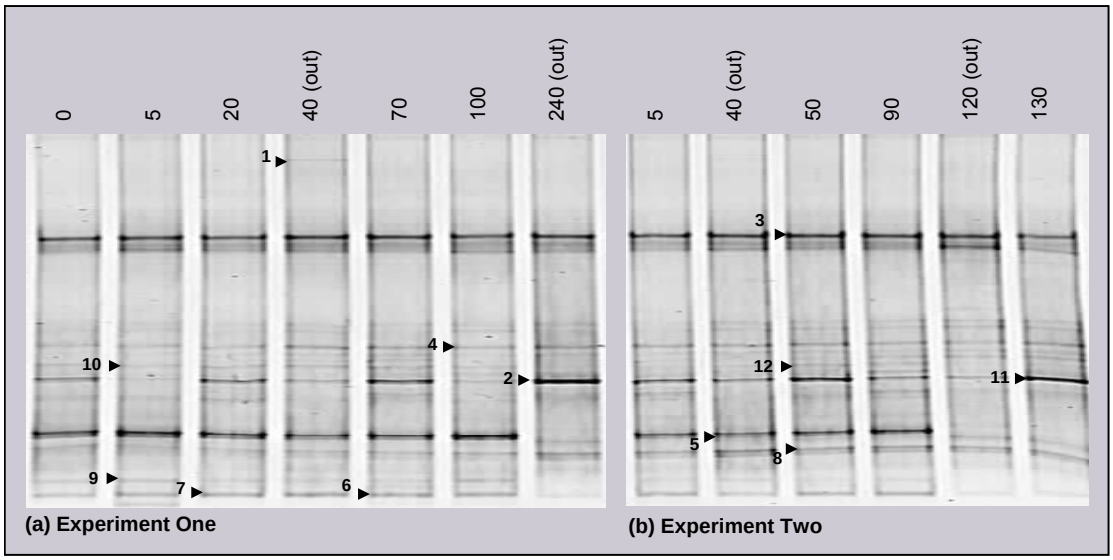


Figure 5.16. 16S rRNA partial gene DGGE profiles, from inside and outside of amended water, putatively specific for *Planctomycetes.* PCR products were generated using protocol in Table 3.2 and primers designed to be *Planctomycetes*-specific listed in Table 3.1. DGGE gradient was 30 % to 60 %. Two water patches were sampled, the first amended with P (a), and the second with Fe and P (b).

In terms of presence and absence of bands, there is little dissimilarity between the profiles in Figure 5.16. Sequences from 25 % of bands (2, 4, and 11) aligned with the *Planctomycetaceae*. The remaining bands yielded 16S rRNA partial gene sequences which aligned more closely to members of the *Firmicutes* and *Cyanobacteria* phyla. These non-target alignments are strong. Sequences from bands 1 and 8 align with 97 % and 98% similarity respectively to cyanobacterial sequences, while the sequence from band 5 has 99 % similarity to the genus *Bacillus*, within the *Firmicutes*.

Supporting the lack of DGGE profile differentiation in Figure 5.16, the corresponding MDS plot (Figure 5.17) also shows little evidence of any clustering of profile points. There is a marked overlap in the distribution of profiles from both experiments and out-of-patch profiles.

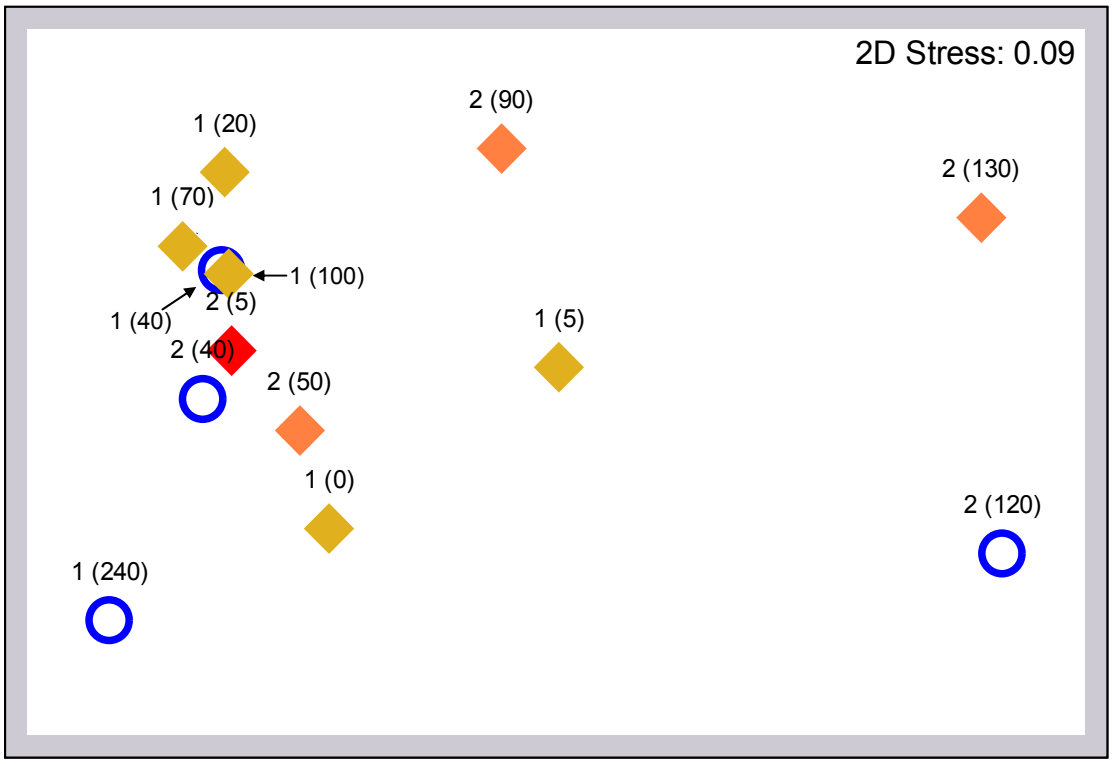


Figure 5.17. MDS plot of DGGE profiles generated using putatively *Planctomycetes*-specific PCR primers. Experiment One (P, gold diamonds), Experiment Two (Fe and P, orange diamonds, out-of-patch samples (blue circles). Numbers in brackets refer to hours elapsed since fertilisation.

Table 5.8. Phylogenetic classification of 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 5.16

Band Number	***Phylogenetic Alignment*** [a]	***Similarity*** [b]
1	*Cyanobacteria;* Family II; GpIIa	97
2	*Planctomycetes; Planctomycetacia;* unclassified *Planctomycetaceae*	86
3	*Actinobacteria; Actinobacteridae; Corynebacterineae; Skermania*	77
4	*Planctomycetes; Planctomycetacia;* unclassified *Planctomycetaceae*	86
5	*Firmicutes; Bacillales; Bacillaceae; Bacillus*	99
6	Unclassified *Bacteria*	71
7	*Firmicutes; Clostridiales; Veillonellaceae; Mitsuokella*	76
8	*Cyanobacteria;* Family II; GpIIa	98
9	*Firmicutes; Clostridiales; Veillonellaceae; Mitsuokella*	68
10	*Gammaproteobacteria;* unclassified *Oceanospirillales*	84
11	*Planctomycetes; Planctomycetacia;* unclassified *Planctomycetaceae*	86
12	*Acidobacteria; Acidobacteria; Acidobacteriaceae;* Gp6	90

[a] Phylogenetic classification, where possible to the level of genus. Sequences from excised bands were aligned with, and compared to, homologous 16S rRNA gene sequences in the Greengenes database (http://greengenes.lbl.gov) and classified accordingly.

[b] Degree of base similarity between an excised DGGE band sequence and that of its closest fully-aligned homologue in the Greengenes database (http://greengenes.lbl.gov).

* Base sequences for 16S rRNA gene fragments, amplified from DGGE bands, are listed in Appendix I.

A BLASTn analysis of the sequence from band 10 matched it to the thermophilic *Bacillus alveayuensis*, a marine sediment dweller (Bae *et al.*, 2005).

5.3.2.7 *Cyanobacteria*

The DGGE 16S rRNA gene community profiles in Figure 5.18 were obtained using *Cyanobacteria* and chloroplast-specific primers and a nested PCR protocol, detailed in Table 3.2 and Section 2.2.2 respectively.

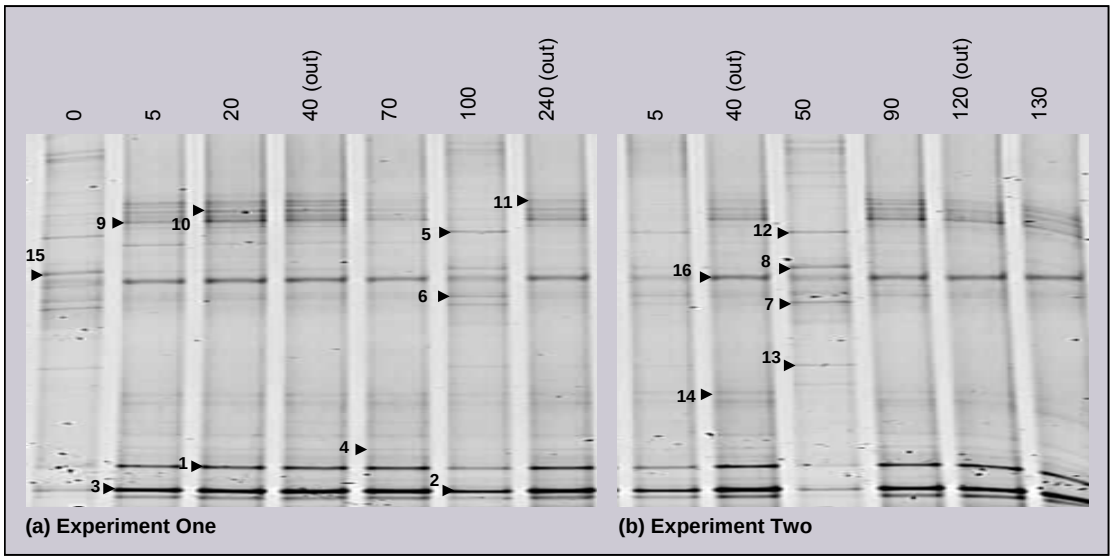


Figure 5.18. DGGE profiles of 16S rRNA gene fragments amplified, using primers designed to be specific for *Cyanobacteria* and chloroplasts, from environmental DNA taken from water amended with P (a) and Fe and P (b).

The DGGE profiles in Figure 5.18 show some variation, particularly 1 (0) and 2 (50), but there is no clear correlation of this variation with treatment. The majority of band sequences match most closely to members of the *Cyanobacteria* (Table 5.9). Non-target alignments were to 16S rRNA gene sequences belonging to *Gammaproteobacteria* and *Firmicutes.* The sequence from band 12 was unclassified by Greengenes alignment, but following a BLASTn analysis it most closely matched, at 88 % similarity, an endosymbiont of the Atlantic rift vent tubeworm *Ridgeia piscesae*, amongst other abyssal and hydrothermal-derived sequences.

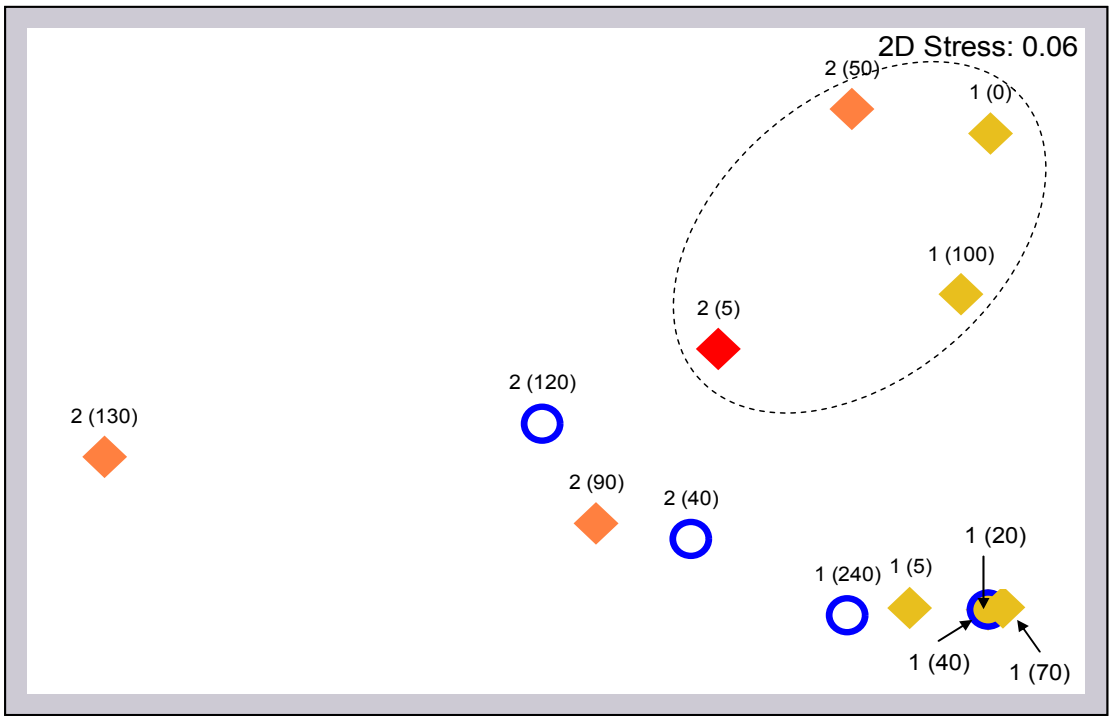


Figure 5.19. MDS plot of putatively *Cyanobacteria* 16S rRNA gene DGGE profiles from Figure 5.18. Experiment One (P, gold diamonds), Experiment Two (Fe and P, orange diamonds, out-of-patch samples (blue circles). Numbers in brackets refer to hours elapsed since fertilisation. See text for discussion concerning profiles clustered in dotted circle.

In Figure 5.19 DGGE profiles from Figure 5.18 do not cluster according to treatment. In fact the only four profiles displaying any degree of clustering (circled in Figure 5.19) stand out in the DGGE gel in Figure 5.18, but their origins have little in common. Discounting these four profiles, it becomes apparent there is little temporal change in the *Cyanobacteria* community 16S rRNA gene profile at 15 m in, or out of, either patch of amended water.

Table 5.9. Phylogenetic classification of 16S rRNA gene fragment sequences* obtained from bands excised from the DGGE gel in Figure 5.18

Band Number	***Phylogenetic Alignment*** [a]	***Similarity*** [b]
1	*Cyanobacteria;* Family II; GpIIa	98
2	*Cyanobacteria;* Family II; GpIIa	98
3	*Cyanobacteria;* Family II; GpIIa	97
4	*Cyanobacteria;* Family II; GpIIa	98
5	*Cyanobacteria;* Family II; GpIIa	77
6	*Firmicutes; Thermoanaerobacterales; Mahella*	51
7	Unclassified *Bacteria*	72
8	Unclassified *Bacteria*	64
9	*Cyanobacteria;* Family IV; GpIV	72
10	*Cyanobacteria;* Family I; GpI	74
11	*Cyanobacteria;* Family II; GpIIa	78
12	Unclassified *Bacteria*	82
13	*Gammaproteobacteria; Oceanospirillales; Chromohalobacter*	80
14	*Cyanobacteria;* Family II; GpIIa	97
15	*Firmicutes; Lactobacillales; Lactobacillaceae; Lactobacillus*	73
16	*Cyanobacteria;* Family II; GpIIa	86

[a] Phylogenetic classification, where possible to the level of genus. Sequences from excised bands were aligned with, and compared to, homologous 16S rRNA gene sequences in the Greengenes database (http://greengenes.lbl.gov) and classified accordingly.
[b] Degree of base similarity between an excised DGGE band sequence and that of its closest fully-aligned homologue in the Greengenes database (http://greengenes.lbl.gov).
* Base sequences for 16S rRNA gene fragments, amplified from DGGE bands, are listed in Appendix I.

Eighty one percent of 16S rRNA gene fragment sequences amplified from bands excised from the DGGE gel in Figure 5.18 aligned most closely with three families of the phylum *Cyanobacteria*. All of these families have marine representatives. Family I (band 10) divide by budding of binary fision. Family II (bands 1-5, 11, 14 and 16) divide by multiple fision, and Family IV (band 9) are filamentous (Rippka *et al*., 1979).

5.3.3 A comparison of ribosome accumulation rates in successional DGGE profiles of samples from P-fertilised, Fe and P-fertilised, and non-fertilised seawater

Accumulation of unique bands, or ribotypes, in a temporal series of DGGE profiles is an indication of the extent of community change in the environment being sampled. A static community will produce a uniform band profile at different sampling points in time, whereas species dying, or migrating into or out from a community, will cause bands to appear and disappear from DGGE profiles. Ribotype accumulation analysis, illustrated for the domain *Bacteria*, as well as specific groups in Figure 5.20, captures the former aspect of community profile change. Over the course of the two experiments DGGE profiles from samples taken from within the first patch (P) show a greater accumulation of unique ribotypes, than either those from in patch two (P + Fe) or from outside either patch. Exceptions to this trend were the *Cyanobacteria* (Figure 5.20f) where final band (ribotype) counts were higher for patch two than for patch one, and the *Planctomycetes* (Figure 5.20e) , where total ribotype counts in patch two and in out-of-patch profiles exceeded the count from patch one profiles.

The relatively flat curves in Figure 5.20a-g indicate ribotype accumulation rates were low during the experiments, especially after the first two profiles taken in each treatment location. The profiles made with 16S rRNA gene products universal for the domain *Bacteria* (Figure 5.6) showed the clearest distinction in accumulation rate of unique ribotypes (bands) between samples taken from amended water in Experiment One (P), and samples from amended water in Experiment Two (P + Fe). Figure 5.20g shows there is nearly a two-fold increase in the number of unique ribotypes occurring in successive DGGE profiles over the course of the two experiments in patch one, compared to those accumulating in patch two, or out of either patch.

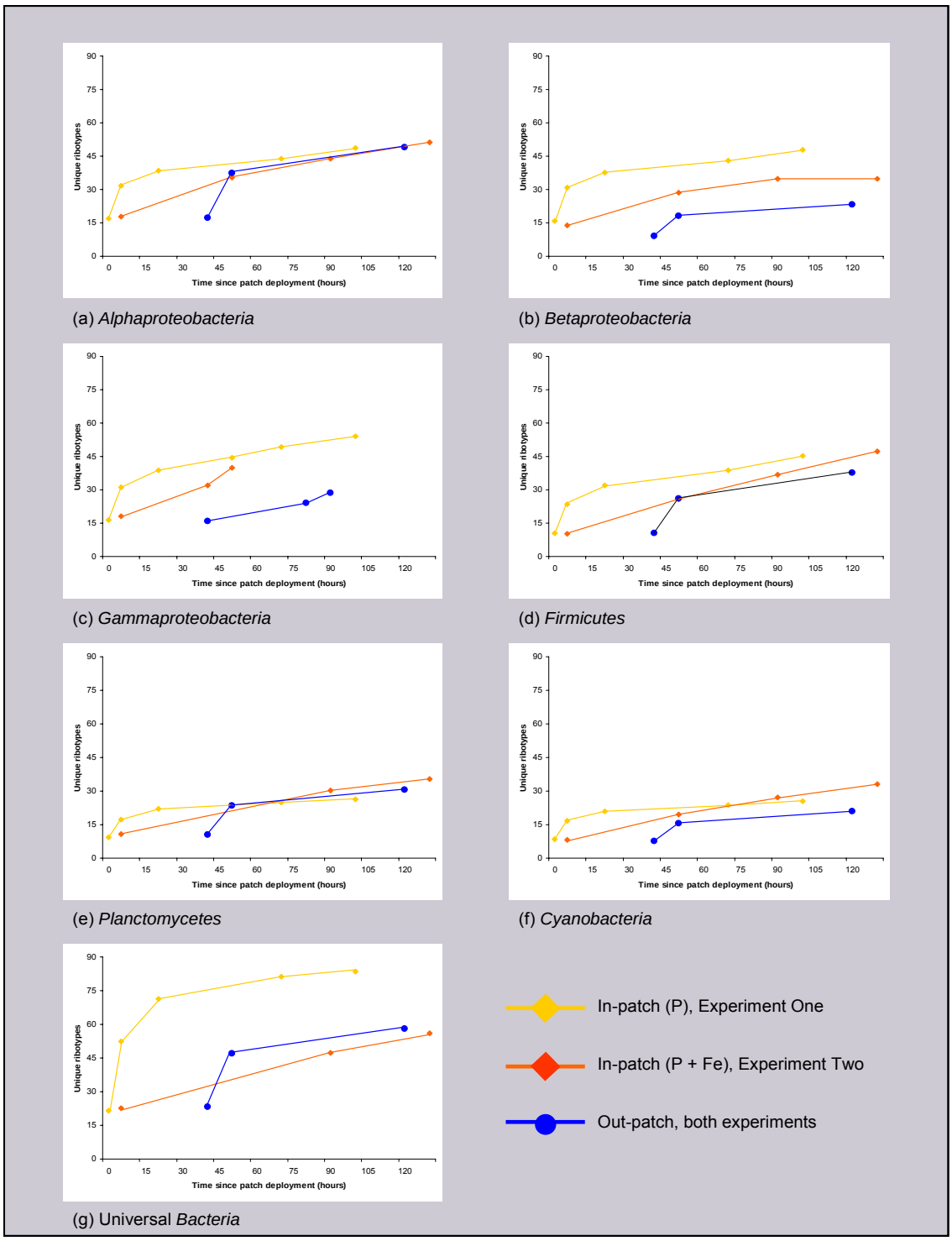


Figure 5.20. Ribotype accumulation curves. Distinct bands, representing unique 16S rRNA gene sequences (ribotypes), accumulate with successive DGGE profiles. Comparison of temporal ribotype accumulation between seven bacterial groups in water amended with P, with Fe and P, and non-amended water. In Experiment Two P was added after 40 hours, on top of Fe which was added at the beginning. Out-patch counts at 40 hours are separated on the graphs according to the actual order in which the samples were taken, that from Experiment One is first.

5.3.4 Ribotype composition of DGGE profiles of bacterial communities from nutrient-fertilised, and non-fertilised, seawater

The Venn diagram in Figure 5.21 illustrates the distribution of ribotypes, taken from DGGE profiles, according to location. Ribotypes are listed by group according

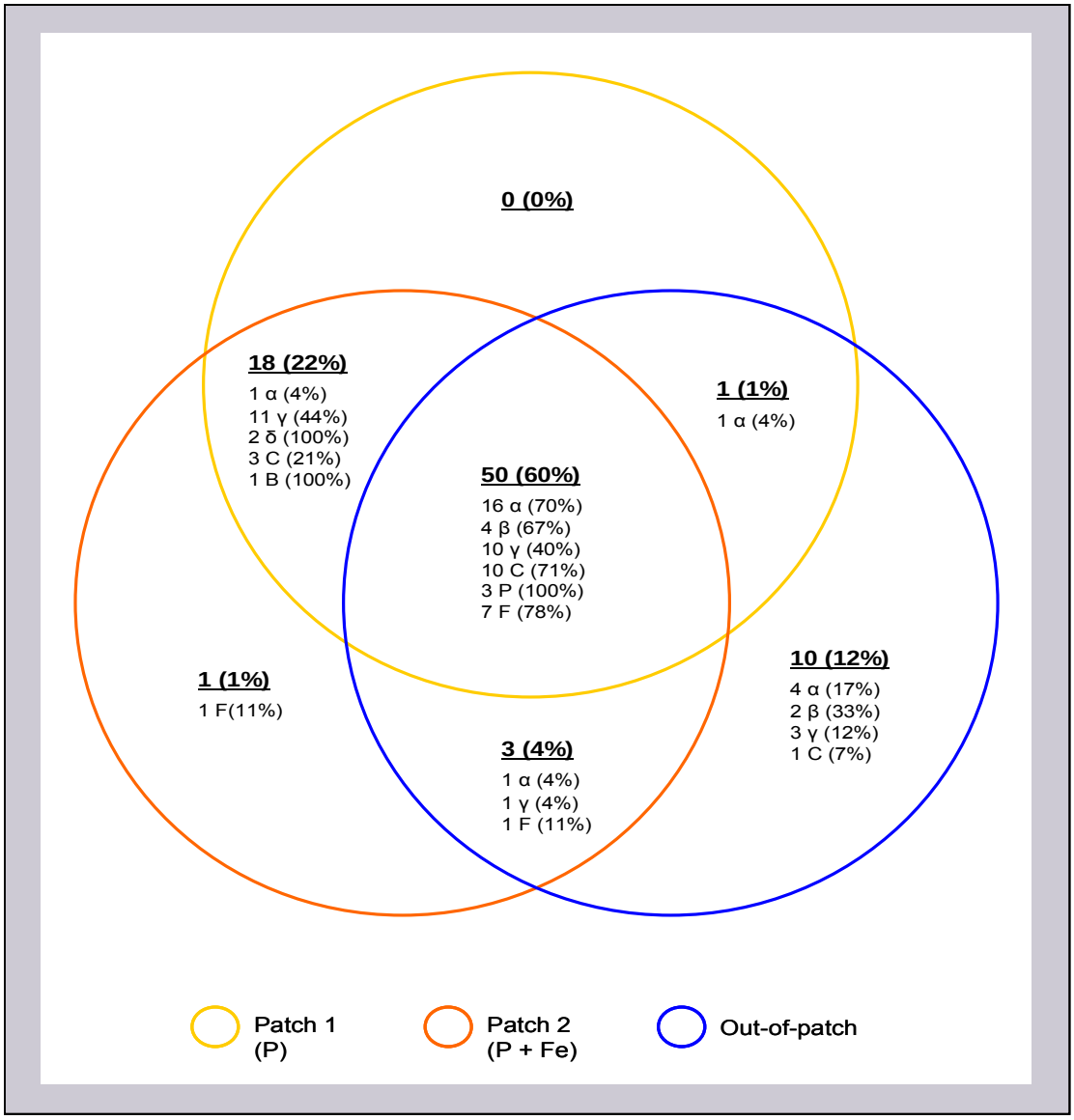


Figure 5.21. Distribution of 83 group-specific ribotypes (actual numbers of unique co-migrating sequences, and proportion of group-specific sequences originating from each of three different sampling stations). Group-specificity was ascertained by alignment with homologous NCBI database sequences. Distribution of ribotypes was measured by presence or absence of bands in DGGE profiles of 16S rRNA genes (see preceding figures in this chapter) from three locations: within P-amended water, within Fe and P-amended water, and non-amended water. Ribotypes were amplified by PCR with primers designed to be specific for the following groups: *Alphaproteobacteria* (α), *Betaproteobacteria* (β), *Gammaproteobacteria* (γ), *Firmicutes* (F), *Planctomycetes* (P) and *Cyanobacteria* (C). There was some non-specific amplification including that of ribotypes belonging to members of the *Bacteroidetes* (B) and *Deltaproteobacteria* (δ).

to that indicated by BLASTn analysis of the ribotype sequence. This allows for non-specific amplification by group-targeted primer pairs, as for example *Gammaproteobacteria* sequences which were amplified using "*Betaproteobacteria*-specific" primers (Table 5.5). Sixty percent of ribotypes are common to all DGGE profiles (both amended water patches and the out-of-patch station). Twenty three percent are restricted to the two amended water patches, 12 % are only found outside either patch, no ribotypes are only found in the P-treated area of Experiment One, and 1 % is restricted to the P and Fe amended water of Experiment Two.

The most noteworthy point on the Venn diagram is the 44 % of *Gammaproteobacteria* sequences which were amplified only from amended water samples. This illustrates the proportion of treatment-specific changes in 16S rRNA gene profiles attributable to the *Gammaproteobacteria*, visible in the profile bands in Figure 5.12.

5.4 Discussion

5.4.1 Experimental background

In oligotrophic pelagic marine environments, which constitute roughly 30 % of the area of the world's oceans, bacterial biomass often exceeds phytoplankton biomass (Cho and Azam, 1990). Thus the structure of the bacterial population has significant implications for the food-web structure, nutrient cycling pathways, and for the sinking into the mesopelagic zone of DOM and POM, in oligotrophic seas (Cho and Azam, 1990; DiTullio *et al.*, 1993).

The limited knowledge of the response of marine bacteria to ocean fertilisation with either Fe or P provided the impetus for the present study looking into responses of several bacterial groups at the class and phylum level. Environmental changes as a result of fertilisation with P, and Fe and P together, during these two experiments included a two-fold increase in the growth rate and abundance of the photoautotroph *Synechococcus* sp. More notable in the Fe-fertilised water (Experiment One), there was an associated increase in predation of *Synechococcus* sp. by microzooplankton ciliates and dinoflagellates (S. Kimmance, pers. comm.). This correlates with the findings of Thingstad *et al*. (1998, 2005) that increased bacterial production, resulting from P fertilisation, may bypass phytoplankton in the food chain and be consumed directly by metazoan predators.

It should be noted that these changes in the ribosomal gene content of the bacterial metagenome (the community of genomes) cannot be linked to changes in the physiological capabilities of the bacterioplankton community, except through information obtained from related cultured isolates.

Amplification of the 16S rRNA gene, rather than its product, the 16S rRNA component of cytoplasmic ribosomes, restricts the amount of information readable from DGGE profiles. Profiling 16S rRNA, or derived crDNA, allows the inference of metabolic activity in addition to genome presence in the environment. Increased band intensity on a gel would not only indicate increased abundance of a particular ribotype, but also increased metabolism, as active cells require greater numbers of ribosomes, and have higher ribosome turnover rates. However this would then make it difficult to dissect the proportions of a band's intensity due to abundance, from that due to activity. If the hypothesis concerning marine prokaryotes that "everything is everywhere" (Baas-Becking, 1934; Finlay, 2002) is true, then 16S rRNA gene profiling by DGGE will reveal very little. With PCR able to amplify the rarest

members of the community, changes to that community's structure should only be revealed by changing band intensities, a refinement to the DGGE technique not employed here. However in practise it has been shown here that rare ribotypes are not revealed by PCR-DGGE, but may subsequently form DGGE bands when the density of their genomes passes an unknown threshold in the sample DNA template (for example ribotypes absent and subsequently present in a *Bacteroidetes*-specific gel in Figure 4.7d).

5.4.2 Effects of iron and phosphorus fertilisation on pelagic members of the *Firmicutes*

The *Firmicutes*, or low G+C content Gram-positive bacteria, are well known components of soil bacterial communities, and both aerobic and anaerobic cells have been discovered to be abundant and diverse in marine sediments (Munn, 2004). However their presence in the upper water column of the open ocean remains poorly documented. A study of ten 16S rRNA gene clone libraries made from independent pelagic samples in 2000, showed the Gram-positive clones were exclusively from the *Actinobacteria*, or high G+C content Gram-positive bacteria (Rappé *et al.*, 2000). The low-scoring BLASTn matches in Table 5.7 to both *Firmicutes* and other 16S rRNA gene sequences, suggests members of the *Firmicutes* may indeed be absent from this environment. The 0 % of *Firmicutes* clones from the clone library constructed from a North Atlantic open ocean sample (Table 3.4) using "*Firmicutes*-specific" primers (350f/814r), may also be suggestive of an absence of *Firmicutes* in open water. The 70 % *Firmicutes*-specific clones in a similar library made from a coastal sample (Table 3.4) may be associated with the proximal land and/or benthic habitat. As the nearest matches are not very similar to the ribotypes amplified from the DGGE gel (Figure 5.14) then further investigation, along the line of that suggested for the *Bacteria*-specific *Cyanobium* sp. match (see below in Section 5.4.5), should be undertaken to elucidate the phylogeny of the bacteria whose 16S rRNA genes are amplified by these "*Firmicutes*-specific" primers (350f/814r). They may constitute a new, open water, lineage of Gram-positive bacteria.

5.4.3 Effects of iron and phosphorus fertilisation on pelagic members of the *Planctomycetes*

The parallel development of DGGE profiles amplified using *Planctomycetes*-specific primers (352f/920r), from samples within amended waters of Experiment Two (Figure 5.16) and out-of-patch suggest there is little effect of P and Fe fertilisation on the diversity of the pelagic *Planctomycetes* community. It is possible the addition of P (Experiment One) leads to a change in the diversity of the population, which then changes less than a similar population which has not been exposed to increased P. *Planctomycetes* also may be very scarce in this environment, as evidenced by the number of non-specific ribotype amplifications using primers (352f/920r) previously demonstrated to be 100 % specific, and the high similarity (98 % to 99 %) of the Planctomycete 16S rRNA gene homologues that were amplified (Table 5.8).

Another example, this one of a band which is non-specific and uncharacteristic for an otherwise highly-specific PCR-DGGE analysis is band 8, the sequence from which aligned to the Group IIa *Cyanobacteria*, and by BLASTn to *Prochlorococcus*. This band occurs faintly in every profile of the *Planctomycetes* gel in Figure 5.16. *Prochlorococcus* sp. cells make up the bulk of primary producing biomass in the oligotrophic subtropical oceans (Partensky *et al*. 1999). This fact together with a scarcity of *Planctomycetes* members in the same environment potentially combine to yield non-specific products (*Cyanobacteria* sequences) in a PCR, despite the use of primers of proven specificity (352f/920r, Table 3.1).

5.4.4 Effects of iron and phosphorus fertilisation on pelagic members of the *Cyanobacteria*

In the DGGE profiles of ribotypes amplified using *Bacteria*-specific primers a single ribotype band (band 3) stands out above all others in profiles from both experiments, and in samples from inside and outside of amended waters (Figure 5.6). The 114 bp fragment of a 16S rRNA gene sequenced was of good quality, with clear and unambiguous dye peaks in the chromatogram. However alignment in the Greengenes database was with the Group IIa *Cyanobacteria* with only 78 % sequence similarity. A BLASTn search found the most similar homologous sequence was only 84 % similar. This was to a cultivated freshwater cyanobacterium, *Cyanobium* sp. JJ12A2 (Jezberova, 2006, Accession AM710377), isolated from a freshwater

reservoir. The next most similar homologues belonged to *Prochlorococcus* sp. (83 % max. ident.) from a metagenomic trawl of the Pacific Ocean (Pham *et al.*, 2008, Accession EU361201). These values are well below the cut-off for species equity taken as 97 % 16S rRNA gene sequence similarity, equating to 70 % DNA-DNA hybridisation, which is accepted as the "species" delineation (Perry *et al.*, 2002), and therefore this ribotype could be representative of a new unicellular marine picophytoplankton. Further investigation should attempt to sequence the full-length 16S rRNA gene, with sequence purification using a clone library, perhaps with further genome sequencing to link function. The full length sequence should be applied to a phylogenetic analysis to confirm the relationship of the unknown cyanobacterium to other genera and families. The dominance of the band in the profile should make it readily obtainable from a clone library made with *Bacteria*-specific primers.

The sequence from band 12 in Figure 5.18, albeit only a single sequence from a single band, and therefore not a good basis for generalisation, is the product of "*Cyanobacteria*-specific" primers (361f/785CYAr) but according to BLASTn search matches 16S rRNA gene fragments from *Gammaproteobacteria* inhabiting the aphotic habitat around abyssal volcanic and hydrothermal vents (for example Accession EU491811 in Santelli *et al.*, 2008). At only 88 % sequence identity the unknown bacterium amplified here, may harbour a deep seated genomic linkage between the photoautotrophic *Cyanobacteria*, and the chemoautolithotrophic *Gammaproteobacteria* of the deep sea. Greengenes alignment did not permit classification of the organism to which this sequence belongs below the domain *Bacteria*. Again, more and longer sequences are needed to obtain more conclusive results, but it remains possible that these primers (361f/785CYAr) could be used to amplify genomic DNA from bacteria that belong outside of the current list of bacterial phyla.

5.4.5 Effects of iron and phosphorus fertilisation on pelagic members of the *beta* and *gamma* subgroups of the *Proteobacteria*

The 90 % group-specificity (Table 5.6) for the ribotypes sequenced from bands in Figure 5.12 does not compare favourably with the 100 % specificity of a clone library generated using the same primers (395f/871r) and environmental DNA from the oligotrophic North Atlantic Ocean (Table 3.4). However the non-specificity is due to two ribotypes only. The sequence from band 22 (Figure 5.12) aligns with

Winogradskyella, a member of the *Bacteroidetes;* while the most similar sequence according to BLASTn was amplified with primers specifically designed to amplify *Bacteroidetes* (Chen *et al.*, 2006). Such non-specificity is justification to review the design of the *Gammaproteobacteria*-specific primers (395f/871r) and improve their specificity.

Hutchins *et al*. demonstrated, using universal *Bacteria* PCR-DGGE, a response by the *Gammaproteobacteria* component of the bacterioplankton, to an OIF experiment (Hutchins *et al.*, 2001a). In the analysis presented here the greatest change to DGGE profiles also occurs in those profiles produced with *Gammaproteobacteria*-specific PCR primers. This is shown in the *Gammaproteobacteria* gel (Figure 5.12), and also in the Venn diagram in Figure 5.21, which shows 22 % of all ribotypes sequenced in this study were found to be treatment-associated and belong to the *Gammaproteobacteria* (reaffirming the findings of Hutchins *et al.*, 2001a). The *Gammaproteobacteria* include heterotophs that can obtain C from a variety of substrates, and some that obtain energy from anoxygenic phototrophy (Perry *et al.*, 2002). In the *Betaproteobacteria* gel (Figure 5.10) the majority of treatment-specific profile changes are also due to bands whose sequences align with the *Gammaproteobacteria*. The sequences from these bands (Figure 5.10: bands 6, 8, and 10, and Figure 5.12: bands 3, 4, 5, 18, and 19) match most closely to cloned sequences from uncultured organisms, according to BLASTn. The sequences from the two bands in the "*Gammaproteobacteria*-specific" gel restricted to profiles from outside either patch (bands 18, 19, Figure 5.12) both aligned with *Alteromonas*, the members of which possess large genomes (> 5 Mbp).

The amplification of these sequences from in-patch samples, but not from out-patch samples may be due to their growth in the amended water beyond a threshold, discussed elsewhere (Chapter 3), that permits their amplification in the PCR.

Results from previous studies (Hutchins *et al.*, 2001a; West *et al.*, 2008) complement and contrast with those from this study showing *Gammaproteobacteria*-specific ribotypes (Figure 5.12) related to the *Pseudomonas*, *Alteromonas*, and *Oceanospirillum* genera (some being members of the SAR86 clade), responding positively to fertilisation with P (Figures 5.12, 5.13, 5.21 and Table 5.6). Many species within these groups are photoheterotrophs with large genomes, for example *Rhodopseudomonas palustris* Bis A53 (5.5 Mbp), *Pseudoalteromonas atlantica* T6c (5.2 Mbp) and *Marinomonas* sp. MWYL1 (5.1 Mbp). It is possible that they contain a

wide array of genes conferring the ability to utilise P from many different substrates, such as phosphates and phosphonates, and are thus able to take advantage of an intermittent and unreliable supply of these substrates (Dyhrman *et al.*, 2007). Their large genomes may also mean their growth is directly limited by P-supply, nucleic acid making up a significant proportion of their demand for P.

An example of a profile which contradicts the temporal patterns shown in a series of gel profiles is that for the *Gammaproteobacteria* at 110 hours in Figure 5.12. This in-patch profile matches all the out-of-patch profiles rather than the other in-patch profiles.

5.4.6 Alignment of pelagic ribotypes with terrestrial homologues

There are a few 16S rRNA gene sequence matches to database homologues derived from terrestrial bacteria, notably *Alphaproteobacteria*, *Betaproteobacteria* and *Firmicutes* (Tables 5.4, 5.5 and 5.7 respectively). The experimental location was over the abyssal plain, 200 nautical miles from the Canary Islands, the nearest land, and 400 nautical miles from Africa, the nearest continent. Sequence alignment (BLASTn) scores for these matches tended to be low, around 90 % - Table 5.5 shows likely terrestrial *Betaproteobacteria* (the *Burkholderiales*) whose 16S rRNA genes aligned with sequences excised from the DGGE gel in Figure 5.10; while Table 5.7 shows likely terrestrial *Firmicutes* (*Thermoanaerobacterales* and *Clostridiales*) aligning with DGGE band-amplified sequences, while Table 5.4 details matches to terrestrial *Alphaproteobacteria* (*Rhizobiales* and *Rickettsiales*) - indicating the bacteria from which the sequences of the excised bands originated, are actually distantly related to the terrestrial organism possessing the 16S rRNA gene sequence in the database. At low levels of 16S rRNA gene sequence similarity, such as at or below a 90 % arbitrary cut-off, bacteria from different habitats and with different physiology may appear related, especially when they feature high in a list of BLASTn alignments.

A second explanation may account for high similarity between bacterial 16S rRNA gene sequences originating from marine and terrestrial habitats. Dust storms from the Sahara Desert are known to blow iron-rich dust and sand far out over the Atlantic Ocean, as seen in the satellite image in Figure 5.1. Indeed it was observed by Darwin aboard H.M.S. Beagle in January 1832 (Darwin, 1845). However Huxley looked out for, but failed to observe, the dust in the same locality aboard H.M.S

Rattlesnake in January 1847 (see page 22 of Huxley, 1935), thus demonstrating the intermittent nature of the phenomenon. It is not inconceivable that terrestrial bacteria are routinely transported from Africa to the open ocean by an airborne route. Ferrophilic members of these bacteria, or recipients of their metabolic genes via HGT, would only have to survive the cold, aquatic, saline environment in small numbers, or for short durations, in order to be stimulated into growth by iron from an OIF experiment, and thence to provide a DNA template for PCR-DGGE. Aeolian dust was also at the heart of the original theory of Fe-limitation of primary production (Martin and Fitzwater, 1988).

5.4.7 Overall effects of iron and phosphorus ocean fertilisation on the bacterioplankton and the wider pelagic ecosystem and geochemical cycling

The number of different ribotypes amplified by *Gammaproteobacteria*-specific (395f/871r) and *Bacteria*-specific (9bfm/1512uR) primers, indicative of the diversity present in the environment, increased upon addition of P to ocean water (Figures 5.20c and g). This may indicate a variety of P-limited heterotrophic bacteria that are normally scarce, but utilise the added P to grow to become abundant enough for their 16S rRNA genes to be amplified by PCR from a mixed environmental DNA template.

There were fewer *Alphaproteobacteria* ribotypes accounting for differences between DGGE profiles (Figure 5.21), and no treatment-associated profile differences discernible in the gels specific for the *Planctomycetes* (Figure 5.16), *Firmicutes* (Figure 5.14), *Cyanobacteria* (Figure 5.18), or for the entire domain *Bacteria* (Figure 5.6).

Assuming that everything is not everywhere, at least beyond a certain level of resolution of variation (after all no two bacterial genomes are likely to be 100 % identical in nucleotide sequence), the degree of variation uncovered will reflect the resolving capability of the tools and methods employed. The changes in profiles of the *Gammaproteobacteria* community shown in Figure 5.12, particularly the appearance of bands in the Experiment One (P-amended) patch, may be indicative of changes in the structure of the bacterioplankton, and increased abundance of certain of its members. These changes may have effects on the ecosystem's elemental cycling (Cho and Azam, 1990) and trophodynamics, such as increased abundance of

microzooplankton predators, having a knock-on negative effect on phytoplankton numbers. A theoretical increase in the system's existing net heterotrophy would lead to increased respiration, and so possibly to a net increase of pCO_2. Increasing the abundance and efficiency of bacteria utilising DOM and POM, such as members of the *Gammaproteobacteria* community, in the euphotic zone, could in addition, reduce the sink of carbon to the deeper ocean. Bacteria respond faster to nutrient fertilisation than eukaryotes, and constitute the majority of POC in these waters (Tortell *et al.*, 1996). With changes, resulting from a higher pCO_2, in the abundance and composition of the bacterioplankton in these oligotrophic waters remaining poorly constrained, it is hard to predict the biogeochemical consequences of wide-scale OIF or OPF beyond the supposition discussed above, and first touted over a decade and a half ago by Fuhrman and Capone (1991).

5.5 Conclusion

Phophorus fertilisation of the oligotrophic North Atlantic Ocean produced changes in the composition of the bacterial community, visualised in 16S rRNA gene profiles on DGGE gels. The most marked changes occurred to the *Gammaproteobacteria* fraction of that community, with different ribotypes appearing in profiles representative of amended water, compared to other ribotypes restricted to profiles from out-of-patch water (Section 5.3.2.4). Similar changes were observed in water amended with both P and Fe together. There is a limited amount and quality of information that can be taken from a 16S rRNA gene sequence alignment with homologous sequences in a database, and applied to the actual organisms at the heart of these changes. Sequences from those *Gammaproteobacteria* found only in non-amended water (bands 18 and 19 in Figure 5.12) align with the genus *Alteromonas* at 96-99 % similarity. BLASTn alignment agreed; there was 95 % similarity with sequences from members of the Alteromonas isolated from the world's most oligotrophic open ocean environment in the South Pacific Ocean (*Alteromonas* sp. MOLA 384, Accession No. AM990661, Lami *et al*. unpub.). Those ribotypes (bands 3, 4, and 5 in Figure 5.12) found only in profiles within the two fertilised areas (P and P + Fe), and not in profiles from the surrounding water, align within the same class (*Alteromonadales*) as *Alteromonas*, and match well by BLASTn to the SAR86 clade of the *Gammaproteobacteria* (Pham *et al.*, 2008). These organisms contain proteorhodopsins, light-powered transmembrane proton pumps. These bacteria are more distantly related to some methylophilic members of the *Betaproteobacteria*.

In summary the structure of the *Gammaproteobacteria* component of the bacterioplankton community changes in response to OIF, and more markedly to OPF, although the changes to gammaproteobacterial diversity resulting from fertilisation with P, and P and Fe together, were not distinguished.

Chapter 6

General discussion

Firstly in a synopsis the key points and results of each experiment will be summarised. Secondly suggestions for future work will be made that should lead to a more comprehensive understanding of pelagic bacterial diversity. The importance of this for understanding the structure and dynamics of marine bacterial communities, and their role in a more complete ecosystem understanding will be discussed.

6.1 Synopsis

6.1.1. Design and validation of bacterial group-specific primers

PCR oligonucleotide primers were designed to amplify fragments of 16S rRNA genes from mixed microbial DNA template, specific for seven groups of bacteria (*alpha*, *beta* and *gamma* subgroups of the *Preoteobacteria*, the *Bacteroidetes*, *Cyanobacteria*, *Firmicutes* and the *Planctomycetes*). Primers were all tested *in silico* and found to be appropriately group-specific (Table 3.3).

In vitro validation of the primers involved the construction of two clone libraries of 16S rRNA gene fragments using environmental DNA from two contrasting pelagic habitats. Cloned sequences from a plasmid clone library made using environmental DNA from the oligotrophic North Atlantic gyre showed all amplified ribotypes to be correctly group-specific, except for the *Betaproteobacteria*, and the *Firmicutes*. Specificity of ribotypes sequenced from a similar clone library derived from shallow inshore water in a Norwegian fjord was even higher, the only non-specific amplifications were among the *Firmicutes* ribotypes.

The group-specific primers permitted greater resolution of bacterial 16S rRNA gene diversity than is obtainable with *Bacteria* domain-specific primers alone. This applied to DGGE profiling of complex microbial communities, as well as total diversity found in group-specific clone libraries compared to *Bacteria*-specific clone libraries. Analysis of 16S rRNA gene sequences from the clone libraries revealed members of the gammaproteobacterial OMG group to be present in coastal samples, showing they are not restricted to oligotrophic environments (Mühling *et al.*, 2007).

6.1.2. PCR-DGGE application of group-specific primers

The technique of PCR-DGGE (Muyzer *et al.*, 1993) was optimised to enable the 16S rRNA gene (ribotype) profiling of pelagic samples exhibiting temporal and spatial variation (Sections 3.2.5 and 4.2.3). Gel profiles were well resolved for all specific groups, enabling conversion to digital binary matrices and multivariate statistical analysis.

6.1.2.1. Mesocosm bacterial diversity response to increased CO_2

Overall there was no statistically demonstrable change in bacterial 16S rRNA gene diversity linked to increased pCO_2. This lack of statistical application was due to there being insufficient mesocosm treatment replications (two is below the minimum required to validate statistical analysis). It is likely any such change in bacterial diversity will remain elusive upon increasing the number of mesocosm "replicates," as doing so would probably reduce, rather than improve, the similarity between the baseline (early) 16S rRNA gene profiles of "replicate" mesocosm communities.

There were, however, changes to the *Gammaproteobacteria* ribotype profiles associated with high CO_2 (Figure 4.7c). The ribotypes in question were from DGGE bands that decreased in intensity toward the end of the 22 days of exposure to elevated CO_2. Repeatable in separate PCR-DGGE analyses, this effect did not show up in the multivariate analysis since this was based on only presence and absence of bands.

Changes in DGGE profiles of ribosomal genes from mesocosm bacterial communities correlated more clearly with the effects of CO_2 and air bubbling. The response of bacteria to increased DOM and POM correlates with the findings of Grossart *et al*. (2006a) who demonstrated increased abundance of bacteria under elevated pCO_2 following the demise of a phytoplankton bloom, when non-living organic matter becomes plentiful. Bacterioplankton compositon has been shown to be associated with both that of the phytoplankton (Schäfer *et al.*, 2002; Green *et al.*, 2004) and the zooplankton (Heidelberg *et al.*, 2002). Indeed it is possible the effect of increased CO_2 on the bacterioplankton is indirect. Pelagic bacteria may, through enzyme-determined trophic restrictions, be tightly associated with either specific phytoplankton species, or specific "species" of phytoplankton-derived organic matter. In this case increased pCO_2 will lead to changes in the diversity of the phytoplankton community, and by association changes in the diversity of the bacterial community both in the upper ocean, with heterotrophic bacteria oxidising organic substrates, and through the water column via the biological pump, following these substrates as they descend (Ducklow *et al.*, 2001; Riebesell, 2004; Grossart *et al.*, 2006a; Grossart *et al.*, 2006b).

6.1.2.2. Pelagic bacterial diversity response to ocean fertilisation with iron and phosphorus

While the effect of Fe could not be dissected from that of P in this experiment, there was a strong effect on the diversity of the bacterial population in the euphotic zone attributable to P fertilisation. This effect was most apparent in the *Gammaproteobacteria* component of the community. Ribotypes amplified from P-amended water samples, but not from out-of-patch samples, possessed 16S rRNA gene sequences which aligned with members of the genera *Psychromonas* and *Oceanospirillum*. Cells of these genera have large (> 5 Mbp) genomes. It is possible that organisms with small streamlined genomes, such as *Pelagibacter ubique*, adapted to low dissolved organic phosphorus (DOP) levels, cannot respond rapidly to increased levels of DOP. However large genomes may allow for trophic redundancy (the ability to utilise a range of organic substrates), and cells possessing them, such as *Oceanospirillum* sp. have the genetic elasticity to utilise "windfalls" of various substrates, in this case P in the form of monosodium phosphate, by switching the expression of metabolic gene pathways (Giovannoni *et al.*, 2005).

6.2. Future work

This study provides insight for future studies of marine bacterial diversity using PCR-DGGE of environmental 16S rRNA genes, and for the undertaking of large scale CO_2 enrichment or nutrient fertilisation experiments. It also presents intriguing results which are worthy of further investigation.

6.2.1. Experimental design

The design of the two experiments (CO_2 enrichment and OIF/OPF) would benefit from three fundamental changes: increased number of treatment replication, increased time course, and increased profile resolution.

Although it would permit the application of significant statistical analyses, increasing the number of replications of the differently-treated mesocosms would be logistically and financially challenging. Neither would it increase the probability of revealing an effect of raised pCO_2 on bacterial diversity, as already discussed in Section 6.1.2.1, and therefore unlikely to be worthwhile.

Longer time sampling regimes would give a more robust assessment of changes in bacterial diversity, and allow easier extrapolation of results to real time scales (years or decades) that the experiments are designed to reflect. However maintaining treatment levels for longer, such as of dissolved CO_2 in a mesocosm with a minimised "bottle effect", or of inorganic nutrients in a tracked patch of the open ocean, over such time periods is challenging.

Increasing the resolution at which bacterial profiles are obtained may reveal changes in diversity not revealed in the current ribotype profiles. This might involve looking beyond changes in ribotype patterns, to changes in gene expression within defined cell types, or simply within the entire metagenome. However this will require much greater knowledge of the genomes present in the community, as well as further advances in sequencing technology, allowing rapid and direct acquisition of entire genome sequences from environmental samples.

6.2.2. Following up results from this study

6.2.2.1. Members of the *Gammaproteobacteria* respond to both increased pCO_2 and ocean fertilisation with phosphorus

DGGE profiles in Figures 4.8c and 5.12 show *Gammaproteobacteria* ribotypes that are present only in profiles from samples of ambient and not elevated pCO_2, and increased inorganic phosphate, respectively. A future study into the effects of raised CO_2, or ocean fertilisation, on the diversity of the pelagic bacteria would benefit from focussing exclusively on the *Gammaproteobacteria*, for instance isolated species such as *Oceanospirillum* sp., which responded to fertilisation with P, and for a strain of which (*Oceanospirillum* sp. MED92) a complete genome is currently being sequenced (The Gordon and Betty Moore Foundation Marine Microbiology Initiative, Project ID: 13561).

New primers could be designed, subdividing the *Gammaproteobacteria* still further, and increasing the profiling resolution, as has been done here for the domain *Bacteria*. Candidate groups include the SAR86 gene clone cluster which has been shown to be well distinct phylogenetically, if not an actual clade (Rappé *et al*., 2000). Full length 16S rRNA gene sequences would provide , via clone libraries, data for robust phylogenetic analyses. DGGE profiling could still be achieved with a nested PCR product, and could also be used to pre-screen clone libraries for redundant sequences prior to sequencing (Neufeld and Mohn, 2006). The variable regions of the

16S rRNA gene amplified by the current primers (395f/871r) are V3 to V5 (Figure 3.1), shown to reveal the most diversity within 16S rRNA genes (Schmalenberger *et al.*, 2001). The cells shown to respond here, especially to increased P availability, may well possess greater intragenomic variety of 16S rRNA genes on account of their large genomes and higher growth rates (Wintzingerode *et al.*, 1997). Cells with larger genomes will also likely be physically larger, and therefore use more P when it is available as an anabolic substrate for the synthesis of membranes and nucleic acids.

Analysis of rRNA, rather than its encoding gene, would benefit this analysis, revealing which *Gammaproteobacteria* were increasingly active following P fertilisation. Smaller microcosm scale experiments would facilitate longer exposure times following fertilisation, as well as longer maintenance of increased P levels, and may also facilitate the isolation of these candidate P-limited members of the *Gammaproteobacteria*. Subsequent metabolic and physiological characterisation of members of the bacterioplankton will bring some "biology" to bear on a field that is becoming increasingly based on theoretical reductionism (metagenomes are theoretical constructs – they are communities of genes, not physically linked, present in an environment whose boundaries are arbitrarily drawn) and assumption (for example that trophic, physiological and metabolic interactions can be extrapolated from nucleotide sequence data).

6.2.2.2. Some 16S rRNA gene sequences are not well identified by BLASTn search

Sequence alignments in the Greengenes database (http://greengenes.lbl.gov) are generally regarded by phylogeneticists as more reliable than equivalent alignments, using BLASTn, in the Genbank database (http://www.ncbi.nlm.nih.gov/genbank), and even as more reliable than alignments in the ARB database (http://www.arb-home.de). This stems from Greengenes alignments only being given to low taxonomic levels, such as genus, when there is high sequence similarity to support the alignment.

Nucleotide sequences, amplified from DGGE bands, aligned in this study indicated there are putative novel bacterial lineages in the ocean. Low similarity levels between 16S rRNA gene sequences and homologous sequences in the Greengenes, Genbank and RDP-II databases, such as for the *Firmicutes* (Table 5.7), provide a starting point to investigate such novel lineages. Longer sequences of these

candidate 16S rRNA genes could be obtained by designing forward and reverse primers specific for a number of the sequences from the DGGE bands in Figure 5.14 and amplifying upstream and downstream along the 16S rRNA gene. The products of this second stage PCR using the newly-designed and *Bacteria*-specific primer 9bfm, or the universal 1512uR, could then be put together *in silico* to give near full-length 16S rRNA gene sequences.

6.2.2.3. Response of eukaryotes to environmental change may be of greater fundamental importance than that of bacteria

There is evidence to suggest that the diversity of the heterotrophic pelagic bacterioplankton is closely linked to the species composition of the phytoplankton (Fandino *et al.*, 2001; Hutchins *et al.*, 2001; Schäfer *et al.*, 2002; Grossart *et al.*, 2005; Grossart *et al.*, 2006a; Mary *et al.*, 2006; West *et al.*, 2008). In this scenario bacteria may be restricted to specific phytoplankton-derived carbon substrates for growth, or be associated with specific phytoplankton species. It makes sense for future investigations, into ocean acidification or its potential mitigating strategies, to study also the effects on the phytoplankton, as effects on bacterial diversity will largely be a consequence of these.

6.2.3. Advancing investigation into bacterial diversity using the 16S rRNA gene marker and DGGE

Sequencing whole genomes has the potential to combine phylogenetic and functional characterisation of cultured isolates, and even of entire microbial communities. While whole genome sequencing remains an expensive and technically demanding undertaking, the analysis of 16S rRNA gene sequences remains the most functional and practical basis for rapid phylogenetic and taxonomic assessment of complex microbial communities (Neufeld and Mohn, 2006). The 16S rRNA gene remains one of very few molecules to meet the requirements of a universal genetic marker: universal occurrence, functional constancy, with appropriate sequence conservation and variation, as well as having a large sequence database to help identification of new sequences.

Fifteen years after its introduction (Muyzer *et al.*, 1993) as a tool for profiling mixed microbial populations from environmental samples, DGGE remains a useful tool for this purpose. It has advantages over other profiling techniques such as

restriction fragment length polymorphism (RFLP, used for example in Henriques *et al.*, 2004) and terminal RFLP (tRFLP, Liu *et al.*, 1997) which do not allow identification of specific ribotypes, and it is quicker and cheaper than other genetic screening techniques, such as the manufacture and screening of clone libraries of specific marker genes. Indeed clone libraries may be screened using DGGE to reduce subsequent time and effort in sequence analysis.

The group-specific primers used in this study allow increased phylogenetic coverage as well as resolution of specific clades within the domain *Bacteria* using DGGE with the 16S rRNA gene as a taxonomic marker. With sequences in public repositories becoming better standardised and quality assessed, it should be possible to design primers to explore bacterial community diversity at the sensitivity limits of the 16S rRNA gene. While specific ribotype clusters such as *Synechococcus* (Fuller *et al.*, 2003) and ecotypes such as the high and low light adapted *Prochlorococcus* clades (West and Scanlan, 1999), and broad phylogenetic divisions, such as those groups analysed in this study, can be assessed, there is a range of phylogenetic resolutions, lying between genus and class, which is not amenable to assessment using 16S rRNA (Ludwig and Schleifer, 2005).

Future advance in microbial diversity analysis using DGGE may be dependent on improved inter-gel comparison (see Section 4.4.1). Gel production and analysis would have to be greatly standardised and much attention paid to standardising gel images using software such as Gelcompar II or Phoretix 2D. Alternatively the choice, and number, of samples run on individual gels could be optimised for downstream analysis.

Notwithstanding the limitations of inter-gel comparison, automation of 16S rRNA gene DGGE analysis may facilitate high throughput screening of environmental samples, to monitor temporal or spatial changes in bacterial diversity. The Ingeny Company (Goes, The Netherlands) already produces a machine which performs PCR and DGGE in a single capillary, with a graphic profile output (www.ingeny.com). A robotic system, analogous to the gel-based system of reading successive nucleotide sequences in DNA sequencing machines, while requiring manual gel loading, could be set to separate bands from gels ready for subsequent sequencing. This automated approach will facilitate both continual year-round monitoring programs, and large scale spatial analyses. The latter could investigate high resolution spatial changes in bacterial diversity in small areas, such as in close

proximity to surfaces (e.g.: biofilms), or larger scale biogeographical studies looking at differences in bacterial diversity, such as across ocean basins or vertically through the water column, as pioneered by Pommier *et al*. (2006) and DeLong *et al*. (2006), respectively.

References

Acinas, S.G., Anton, J., and Rodriguez-Valera, F. (1999) Diversity of free-living and attached bacteria in offshore western Mediterranean waters as depicted by analysis of genes encoding 16S rRNA. *Applied and Environmental Microbiology* **65**: 514-522.

Acinas, S.G., Klepac-Ceraj, V., Hunt, D.E., Pharino, C., Ceraj, I., Distel, D.L., and Polz, M.F. (2004) Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* **430**: 551-554.

Alonso-Sáez, L., Balagué, V., Sà, E.L., Sánchez, O., González, J.M., Pinhassi, J. et al. (2007) Seasonality in bacterial diversity in north-west Mediterranean coastal waters: assessment through clone libraries, fingerprinting and FISH. *FEMS Microbiology Ecology* **60**: 98-112.

Amann, R., and Ludwig, W. (2000) Ribosomal RNA-targeted nucleic acid probes for studies in microbial ecology. *FEMS Microbiology Reviews* **24**: 555-565.

Amann, R., Ludwig, W., and Schleifer, K.-H. (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiological Reviews* **59**: 143-169.

Archibald, J.M., and Keeling, P.J. (2005) On the origin and evolution of plastids. In *Microbial phylogeny and evolution: Concepts and controversies*. Sapp, J. (ed). Oxford: Oxford University Press, pp. 238-260.

Arrieta, J.M., Weinbauer, M.G., Lute, C., and Herndl, G.J. (2004) Response of bacterioplankton to iron fertilisation in the Southern Ocean. *Limnology and Oceanography* **49**: 799-808.

Ashelford, K.E., Chuzhanova, N.A., Fry, J.C., Jones, A.J., and Weightman, A.J. (2005) At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. *Applied and Environmental Microbiology* **71**: 7724-7736.

Ashelford, K.E., Weightman, A.J., and Fry, J.C. (2002) PRIMROSE: A computer program for generating and estimating the phylogenetic range of 16S rRNA oligonucleotide probes and primers in conjunction with the RDP-II database. *Nucleic Acids Research* **30**: 3481-3489.

Baas-Becking (1934) *Geobiologie of inleiding tot de milieukunde*. The Hague: W.P. van Stockum and N.V. Zoon.

Bae, S.S., Lee, J.H., and Kim, S.J. (2005) *Bacillus alveayuensis* sp. nov., a thermophilic bacterium isolated from deep-sea sediments of the Ayu Trough. *International Journal of Systematic and Evolutionary Microbiology* **55**: 1211-1215.

Baker, G.C., Smith, J.J., and Cowan, D.A. (2003) Review and re-analysis of domain-specific 16S primers. *Journal of Microbiological Methods* **55**: 541-555.

Bertilsson, S., Berglund, O., Karl, D.M., and Chisholm, S.W. (2003) Elemental composition of marine Prochlorococcus and Synechococcus: Implications for the ecological stoichiometry of the sea. *Limnology and Oceanography* **48**: 1721-1761.

Biegala, I.C., Kennaway, G., Alverca, E., Lennon, J.-F., Vaulot, D., and Simon, N. (2002) Identification of bacteria associated with dinoflagellates (Dinophyceae) *Alexandrium* spp. using tyramide signal amplification-fluorescent *in situ* hybridization and confocal microscopy. *Journal of Phycology* **38**: 404-411.

Blackwood, C.B., Oaks, A., and Buyer, J.S. (2005) Phylum- and class-specific PCR primers for general microbial community analysis. *Applied and Environmental Microbiology* **71**: 6193-6198.

Blümel, M., Süling, J., and Imhoff, J.F. (2007) Depth-specific distribution of Bacteroidetes in the oligotrophic eastern Mediterranean Sea. *Aquatic Microbial Ecology* **46**: 209-224.

Boon, N., Windt, W.D., Verstraete, W., and Top, E.M. (2002) Evaluation of nested PCR-DGGE (denaturing gradient gel electrophoresis) with group-specific 16S rRNA primers for the analysis of bacterial communities from different wastewater treatment plants. *FEMS Microbiology Ecology* **39**: 101-112.

Bourne, D.G., and Munn, C.B. (2005) Diversity of bacteria associated with the coral *Pocillopora Damicornis* from the Great Barrier Reef. *Environmental Microbiology* **7**: 1162-1174.

Boyd, P.W., Jickells, T., Law, C.S., Blain, S., Boyle, E.A., Buesseler, K.O. et al. (2007) Mesoscale iron enrichment experiments 1993-2005: Synthesis and future directions. *Science* **315**: 612-617.

Boyd, P.W., Watson, A., Law, C.S., Abraham, E.R., Trull, T., Murdoch, R. et al. (2000) A mesoscale phytoplankton bloom in the polar Southern Ocean stimulated by iron fertilisation. *Nature* **407**: 695-702.

Boyle, E.A. (1997) Characteristics of the deep ocean carbon system during the past 150,000 years: ΣCO_2 distributions, deep water flow patterns, and abrupt climate change. *Proceedings of the National Academy of Sciences of the United States of America* **94**: 8300-8307.

Brinkmeyer, R., Knittel, K., Jürgens, J., Weyland, H., Amann, R., and Helmke, E. (2003) Diversity and structure of bacterial communities in Arctic and Antarctic pack ice. *Applied and Environmental Microbiology* **69**: 6610-6619.

Brown, M.V., Schwalbach, M.S., Hewson, I., and Fuhrman, J.A. (2005) Coupling 16S-ITS rDNA clone libraries and automated ribosomal intergenic spacer analysis to show marine microbial diversity: development and application to a time series. *Environmental Microbiology* **7**: 1466-1479.

Bucciarelli, E., and Sunda, W.G. (2003) Influence of CO_2, nitrate, phosphate, and silicate limitation on intracellular dimethylsulfoniopropionate in batch cultures of the coastal diatom *Thalassiosira pseudonana*. *Limnology and Oceanography* **48**: 2256-2265.

Buchan, A., González, J.M., and Moran, M.A. (2005) Overview of the marine *Roseobacter* lineage. *Applied and Environmental Microbiology* **71**: 5665-5677.

Buesseler, K.O., Andrews, J.E., Pike, S.M., and Charette, M.A. (2004) The effect of iron fertilization on carbon sequestration in the Southern Ocean. *Science* **304**: 414-417.

Burkhardt, S., Zondervan, I., and Riebesell, U. (1999) Effect of CO_2 concentration on C:N:P ratio in marine phytoplankton: A species comparison. *Limnology and Oceanography* **44**: 683-690.

Burkhardt, S., Amoroso, G., Riebesell, U., and Sültemeyer, D. (2001) CO_2 and HCO_3^- uptake in marine diatoms acclimated to different CO_2 concentrations. *Limnology and Oceanography* **46**: 1378-1391.

Button, D.K., Robertson, B.R., Lepp, P.W., and Schmidt, T.M. (1998) A small, dilute-cytoplasm, high-affinity, novel bacterium isolated by extinction culture and having kinetic constants compatible with growth at ambient concentrations of dissolved nutrients in seawater. *Applied and Environmental Microbiology* **64**: 4467-4476.

Button, D.K., Schut, F., Quang, P., Martin, R., and Robertson, B.R. (1993) Viability and isolation of marine bacteria by dilution culture: theory, procedures, and initial results. *Applied and Environmental Microbiology* **59**: 881-891.

Caldeira, K., and Wickett, M.E. (2003) Anthropogenic carbon and ocean pH. *Nature* **425**: 365.

Calvo, L., Vila, X., Abella, C.A., and Garcia-Gil, L.J. (2004) Use of ammonia-oxidizing bacterial-specific phylogenetic probe Nso1225 as a primer for fingerprint analysis of ammonia-oxidizer communities. *Applied Microbiology and Biotechnology* **63**: 715-721.

Carlson, C.A., Giovannoni, S.J., Hansell, D.A., Goldberg, S.J., Parsons, R., and Vergin, K. (2004) Interactions among dissolved organic carbon, microbial processes, and community structure in the mesopelagic zone of the northwestern Sargasso Sea. *Limnology and Oceanography* **49**: 1073-1083.

Casamatta, D.A., Johansen, J.R., Vis, M.L., and Broadwater, S.T. (2005) Molecular and morphological characterisation of ten polar and near-polar strains within the *Oscillatoriales*. *Journal of Phycology* **41**: 421-438.

Casamayor, E.O., Pedrós-Alió, C., Muyzer, G., and Amann, R. (2002) Microheterogeneity in 16S ribosomal DNA-defined bacterial populations from a stratified planktonic environment is related to temporal changes and to ecological adaptations. *Applied and Environmental Microbiology* **68**: 1706-1714.

Case, R.J., Boucher, Y., Dahllöf, I., Holmström, C., Doolittle, W.F., and Kjelleberg, S. (2007) Use of 16S rRNA and *rpoB* genes as molecular markers for microbial ecology studies. *Applied and Environmental Microbiology* **73**: 278-288.

Castle, D., and Kirchman, D.L. (2004) Composition of estuarine bacterial communities assessed by denaturing gradient gel electrophoresis and fluorescence *in situ* hybridization. *Limnology and Oceanography: Methods* **2**: 303-314.

Chen, X., Zeng, Y., and Jiao, N. (2006) Development and evaluation of specific 16S rDNA primers for marine Cytophaga-Flavobacteria cluster. *Molecular Ecology Notes* **6**: 1278-1281.

Chisholm, S.W., Falkowski, P.G., and Cullen, J.J. (2001) Discrediting ocean fertilisation. *Science* **294**: 309-310.

Chisholm, S.W., Olsen, R.J., Zehler, E.R., Goericke, R., and Waterbury, J.B. (1988) A novel free-living prochlorophyte abundant in the oceanic euphotic zone. *Nature* **334**: 340-343.

Chisholm, S.W., Frankel, S.L., Goericke, R., Olsen, R.J., Palenik, B., Waterbury, J.B. et al. (1992) *Prochlorococcus marinus* nov. gen. nov. sp.: an oxyphototrophic marine prokaryote containing divinyl chlorophyll *a* and *b*. *Archives of Microbiology* **157**: 297-300.

Cho, B.C., and Azam, F. (1990) Biogeochemical significance of bacterial biomass in the ocean's euphotic zone. *Marine Ecology Progress Series* **63**: 253-259.

Cho, J.C., and Giovannoni, S.J. (2004) Cultivation and growth characteristics of a diverse group of oligotrophic marine gammaproteobacteria. *Applied and Environmental Microbiology* **70**: 432-440.

Chouari, R., Le Paslier, D., Daegelen, P., Ginestet, P., Weissenbach, J., and Sghir, A. (2005) Novel predominant archaeal and bacterial groups revealed by molecular analysis of an anaerobic sludge digester. *Environmental Microbiology* **7**: 1104-1115.

Christaki, U., Obernosterer, I., Van Wambeke, F., Veldhuis, M.J.W., Garcia, N., and Catala, P. (2008) Microbial food web structure in a naturally iron-fertilised area in the Southern Ocean (Kerguelen Plateau). *Deep-Sea Research. Part II - Topical Studies in Oceanography* **55**: (5-7) 706-719.

Clarke, K.R., and Gorley, R.N. (2006) PRIMER v6: User Manual / Tutorial, PRIMER-E, Plymouth.

Clarke, K.R., and Warwick, R.M. (2001) Change in marine communities: an approach to statistical analysis and interpretation, 2nd Edition, PRIMER-E, Plymouth.

Coale, K.H., Johnson, K.S., Chavez, F.P., Buesseler, K.O., Barber, R.T., Brzezinski, M.A. et al. (2004) Southern Ocean iron enrichment experiment: Carbon cycling in high- and low-Si waters. *Science* **304**: 408-414.

Coale, K.H., Johnson, K.S., Fitzwater, S.E., Gordon, R.M., Tanner, S., Chavez, F.P. et al. (1996) A massive phytoplankton bloom induced by an ecosystem-scale iron fertilisation experiment in the equatorial Pacific Ocean. *Nature* **383**: 495-501.

Cochlan, W.P. (2001) The heterotrophic bacterial response during a mesoscale iron enrichment experiment (IronEx II) in the eastern equitorial Pacific Ocean. *Limnology and Oceanography* **46**: 428-435.

Cohan, F.M. (2002) What are bacterial species? *Annual Review of Microbiology* **56**: 457-487.

Cohn, F. (1875) Studies on bacteria. *Beiträge zur Biologie der Pflanzen* **1**: 127-222.

Collins, S., and Bell, G. (2004) Phenotypic consequences of 1,000 generations of selection at elevated CO_2 in a green alga. *Nature* **431**: 566-569.

Connon, S.A., and Giovannoni, S.J. (2002) High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. *Applied and Environmental Microbiology* **68**: 3878-3885.

Cook, A.A., Bhadury, P., Debenham, N.J., Meldal, B.H.M., Blaxter, M.L., Smerdon, G.R. et al. (2005) Denaturing gradient gel electrophoresis (DGGE) as a tool for identification of marine nematodes. *Marine Ecology Progress Series* **291**: 103-113.

Conway Morris, S. (1998) *The Crucible of Creation*. Oxford, Oxford University Press.

Cotner, J.B., Ammerman, J.W., Peele, E.R., and Bentzen, E. (1997) Phosphorus-limited bacterioplankton growth in the Sargasso Sea. *Aquatic Microbial Ecology* **13**: 141-149.

Cottrell, M.T., and Kirchman, D. (2000) Community composition of marine bacterioplankton determined by 16S rRNA gene clone libraries and fluorescence *in situ* hybridization. *Applied and Environmental Microbiology* **66**: 5116-5122.

Curtis, T.P., Sloan, W.T., and Scannell, J.W. (2002) Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 10494-10499.

Dahllöf, I., Baillie, H., and Kjelleberg, S. (2000) *rpoB*-based microbial community analysis avoids limitations inherent in 16S rRNA gene intraspecies heterogeneity. *Applied and Environmental Microbiology* **66**: 3376-3380.

Dar, S.A., Kuenen, J.G., and Muyzer, G. (2005) Nested PCR-denaturing gradient gel electrophoresis approach to determine the diversity of sulfate-reducing bacteria in complex microbial communities. *Applied and Environmental Microbiology* **71**: 2325-2330.

Darwin, C. (1845) Journal of Researches into the Natural History and Geology of the Countries visited during the voyage of H.M.S. Beagle round the World, under the Command of Captain Fitz Roy, R.N.

De Baar, H.J.W., Boyd, P.W., Coale, K.H., Landry, M.R., Tsuda, A., Assmy, P. et al. (2005) Synthesis of iron fertilisation experiments: From the Iron Age in the Age of Enlightenment. *Journal of Geophysical Research* **110**: 1-24.

DeLong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J., Frigaard, N.-U. et al. (2006) Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311**: 496-503.

Denariaz, G., Payne, W.J., and Gall, J.L. (1989) A halophilic denitrifier, *Bacillus halodenitrificans* sp. nov. *International Journal of Systematic Bacteriology* **39**: 145-151.

Díez, B., Pedrós-Alió, C., Marsh, T.L., and Massana, R. (2001) Application of denaturing gradient gel electrophoresis (DGGE) to study the diversity of marine picoeukaryotic assemblages and comparison of DGGE with other molecular techniques. *Applied and Environmental Microbiology* **67**: 2942-2951.

DiTullio, G.R., Hutchins, D.A., and Bruland, K.W. (1993) Interaction of iron and major nutrients controls phytoplankton growth and species composition in the tropical Pacific Ocean. *Limnology and Oceanography* **38**: 495-508.

Dixon, J.L. (2008) Macro and micro nutrient limitation of microbial productivity in oligotrophic subtropical Atlantic waters. *Environmental Chemistry* **5**: 135-142.

Dore, J.E., Lukas, R., Sadler, D.W., and Karl, D.M. (2003) Climate-driven changes to the atmospheric CO_2 sink in the subtropical North Pacific Ocean. *Nature* **424**: 754-757.

Ducklow, H.W., Steinberg, K., and Buesseler, K.O. (2001) Upper ocean carbon export and the biological pump. *Oceanography* **14**: 50-58.

Dumont, M.G., Neufeld, J.D., and Murrell, J.C. (2006) Isotopes as tools for microbial ecologists. *Current Opinion in Biotechnology* **17**: 57-58.

Dyhrman, S.T., Ammerman, J.W., and van Mooy, B.A.S. (2007) Microbes and the marine phosphorus cycle. *Oceanography* **20**: 110-116.

Dyhrman, S.T., Webb, E.A., Anderson, D.M., and Moffett, J.W. (2002) Cell-specific detection of phosphorus stress in *Trichodesmium* from the western North Atlantic. *Limnology and Oceanography* **47**: 1832-1836.

Eiler, A., and Bertilsson, S. (2006) Detection and quantification of *Vibrio* populations using denaturant gradient gel electrophoresis. *Journal of Microbiological Methods*.

Eilers, H., Pernthaler, J., Peplies, J., Glockner, F.O., Gerdts, G., and Amann, R. (2001) Isolation of novel pelagic bacteria from the German Bight and their seasonal contributions to surface picoplankton. *Applied and Environmental Microbiology* **67**: 5134-5142.

Elser, J.J., Stabler, L.B., and Hasset, R.P. (1995) Nutrient limitation of bacterial growth and rates of bacterivory in lakes and oceans: A comparative study. *Aquatic Microbial Ecology* **9**: 105-110.

Engel, A., Delille, B., Jacquet, S., Riebesell, U., Rochelle-Newall, E., Terbrüggen, A., and Zondervan, I. (2004) Transparent exopolymer particles and dissolved organic production by Emiliania huxleyi exposed to different CO2 concentrations: A mesocosm experiment. *Aquatic Microbial Ecology* **34**: 93-104.

Ezaki, T., and Suzuki, S. (1982) Achromopeptidase for lysis of anaerobic Gram-positive cocci. *Journal of Clinical Microbiology* **16**: 844-846.

Fandino, L.B., Riemann, L., Steward, G.F., Long, R.A., and Azam, F. (2001) Variations in bacterial community structure during a dinoflagellate bloom analysed by DGGE and 16S rDNA sequencing. *Aquatic Microbial Ecology* **23**: 119-130.

Feely, R.A., Sabine, C.L., Lee, K., Berelson, W., Kleypas, J., Fabry, V.J., and Millero, F.J. (2004) Impact of anthropogenic CO_2 on the $CaCO_3$ system in the oceans. *Science* **305**: 362-366.

Ferrari, V.C., and Hollibaugh, J.T. (1999) Distribution of microbial assemblages in the Central Arctic Ocean Basin studied by PCR/DGGE: Analysis of a large data set. *Hydrobiologia* **401**: 55-68.

Field, K.G., Gordon, D., Wright, T., Rappé, M., Urbach, E., Vergin, K., and Giovannoni, S.J. (1997) Diversity and depth-specific distribution of SAR11 cluster rRNA genes from marine planktonic bacteria. *Applied and Environmental Microbiology* **63**: 63-70.

Fieseler, L., Horn, M., Wagner, M., and Hentschel, U. (2004) Discovery of the novel candidate phylum "Poribacteria" in marine sponges. *Applied and Environmental Microbiology* **70**: 3724-3732.

Finlay, B.J. (2002) Global dispersal of free-living microbial eukaryote species. *Science* **296**: 1061-1063.

Fischer, S.G., and Lerman, L.S. (1983) DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: Correspondence with melting theory. *Proceedings of the National Academy of Sciences of the United States of America* **80**: 1579-1583.

Fisher, M.M., and Triplett, E.W. (1999) Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. *Applied and Environmental Microbiology* **65**: 4630-4636.

Freitag, T.E., and Prosser, J.I. (2003) Community structure of ammonia-oxidizing bacteria within anoxic marine sediments. *Applied and Environmental Microbiology* **69**: 1359-1371.

Friedrich, M.W. (2006) Stable-isotope probing of DNA: insights into the function of uncultivated microorganisms from isotopically labeled metagenomes. *Current Opinion in Biotechnology* **17**: 59-66.

Fuhrman, J.A., and Capone, D.G. (1991) Possible biogeochemical consequenses of ocean fertilisation. *Limnology and Oceanography* **36**: 1951-1959.

Fuhrman, J.A., McCallum, K., and Davis, A.A. (1993) Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific Oceans. *Applied and Environmental Microbiology* **59**: 1294-1302.

Fuller, N.J., Marie, D., Partensky, F., Vaulot, D., Post, A.F., and Scanlan, D.J. (2003) Clade-specific 16S ribosomal DNA oligonucleotides reveal the predominance of a single marine *Synechococcus* clade throughout a stratified water column in the Red Sea. *Applied and Environmental Microbiology* **69**: 2430-2443.

Garrity, G.M., Bell, J.A., and Lilburn, T.G. (2003) Taxonomic outline of the procaryotes. In *Bergey's Manual of Systematic Bacteriology*. New York: Springer-Verlag, p. 395.

Garrity, G.M., Winters, M., and Searles, D.B. (2001) *Taxonomic outline of the prokaryotic genera: Bergey's manual of systematic bacteriology*. New York: Springer-Verlag.

Gervais, F., Riebesell, U., and Gorbunov, M.Y. (2002) Changes in primary productivity and chlorophyll a in response to iron fertilisation in the southern Polar Frontal Zone. *Limnology and Oceanography* **47**: 1324-1335.

Giovannoni, S., and Rappé, M. (2000) Evolution, diversity, and molecular ecology of marine prokaryotes. In *Microbial Ecology of the Oceans*. Kirchman, D.L. (ed). New York: John Wiley & Sons: 47-84.

Giovannoni, S., Rappé, M., Vergin, K., and Adair, N. (1996) 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the green non-sulfur bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **93**: 7979-7984.

Giovannoni, S., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D. et al. (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**: 1242-1245.

Giovannoni, S.J., Britschgi, T.B., Moyer, C.L., and Field, K.G. (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**: 60-63.

Girvan, M.S., Campbell, C.D., Killham, K., Prosser, J.I., and Glover, L.A. (2005) Bacterial diversity promotes community stability and functional resilience after perturbation. *Environmental Microbiology* **7**: (3) 301-313.

Glöckner, F.O., Fuchs, B.M., and Amann, R. (1999) Bacterioplankton compositions of lakes and oceans: a first comparison based on fluorescence in situ hybridization. *Applied and Environmental Microbiology* **65**: 3721-3726.

Gnanadesikan, A., Sarmiento, J., and Slater, R. (2000) Potential limits on the efficiency of ocean fertilisation as a carbon sequestration strategy: the importance of circulation. In *American Geophysical Union Fall Meeting*. San Francisco.

Goddard, V.J., Baker, A.C., Davy, J.E., Adams, D.G., de Ville, M.M., Thackeray, S.J. et al. (2005) Temporal distribution of viruses, bacteria and phytoplankton throughout the water column in a freshwater hypereutrophic lake. *Aquatic Microbial Ecology* **39**: 211-223.

Gonzalez, J.M., Ortiz-Martinez, A., Gonzalez-delValle, M.A., Laiz, L., and Saiz-Jimenez, C. (2003) An efficient strategy for screening large cloned libraries of amplified 16S rDNA sequences from complex environmental communities. *Journal of Microbiological Methods* **55**: 459-463.

Green, D.H., Llewellyn, L.E., Negri, A.P., Blackburn, S.I., and Bolch, C.J.S. (2004) Phylogenetic and functional diversity of the culturable bacterial community associated with the paralytic shellfish poisoning dinoflagellate *Gymnodinium catenatum*. *FEMS Microbiology Ecology* **47**: 345-357.

Grossart, H.P., and Simon, M. (2007) Interactions of planktonic algae and bacteria: effects on algal growth and organic matter dynamics. *Aquatic Microbial Ecology* **47**: 163-176.

Grossart, H.P., Allgaier, M., Passow, U., and Riebesell, U. (2006a) Testing the effect of CO_2 concentration on the dynamics of marine heterotrophic bacterioplankton. *Limnology and Oceanography* **51**: 1-11.

Grossart, H.P., Levold, F., Allgaier, M., Simon, M., and Brinkhoff, T. (2005) Marine diatom species harbour distinct bacterial communities. *Environmental Microbiology* **7**: 860-873.

Grossart, H.P., Kiørboe, T., Tang, K.W., Allgaier, M., Yam, E.M., and Ploug, H. (2006b) Interactions between marine snow and heterotrophic bacteria: aggregate formation and microbial dynamics. *Aquatic Microbial Ecology* **42**: 19-26.

Grüntzig, V., Nold, S.C., Zhou, J., and Tiedje, J.M. (2001) *Pseudomonas stutzeri* nitrate reductase gene abundance in environmental samples measured by real-time PCR. *Applied and Environmental Microbiology* **67**: 760-768.

Gunderson, L.H. (2000) Ecological resilience - in theory and application. *Annual Review of Ecology and Systematics* **31**: 425-439.

Gupta, R.S., and Griffiths, E. (2002) Critical issues in bacterial phylogeny. *Theoetical Population Biology* **61**: 423-434.

Hagström, Å., Pinhassi, J., and Zweifel, U.L. (2000) Biogeographical diversity among marine bacterioplankton. *Aquatic Microbial Ecology* **21**: 231-244.

Hagström, Å., Pommier, T., Rohwer, F., Simu, K., Stolte, W., Svensson, D., and Zweifel, U.L. (2002) Use of 16S ribosomal DNA for delineation of marine bacterioplankton species. *Applied and Environmental Microbiology* **68**: 3628-3633.

Hall, J.A., and Safi, K. (2001) The impact of in situ Fe fertilisation on the microbial food web in the Southern Ocean. *Deep-Sea Research. Part II - Topical Studies in Oceanography* **48**: 2591-2613.

Hamasaki, K., Taniguchi, A., Tada, Y., Long, R.A., and Azam, F. (2007) Actively growing bacteria in the Inland Sea of Japan, identified by combined bromodeoxyuridine immunocapture and denaturing gradient gel electrophoresis. *Applied and Environmental Microbiology* **73**: 2787-2798.

Handelsman, J. (2004) Metagenomics: Application of genomics to uncultured microorganisms. *Microbiology and Molecular Biology Reviews* **68**: 669-685.

Hardy, A. (1970) *The Open Sea: Its Natural History, Part 1: The World of Plankton*: Collins Fontana.

Hayashi, K. (1991) PCR-SSCP: a simple and sensitive method for the detection of mutations in genomic DNA. *PCR Methods and Applications* **1**: 34-38.

Heidelberg, J.F., Heidelberg, K.B., and Colwell, R.R. (2002) Bacteria of the γ-subclass *Proteobacteria* associated with zooplankton in Chesapeake Bay. *Applied and Environmental Microbiology* **68**: 5498-5507.

Henriques, I.S., Almeida, A., Cunha, A., and Correia, A. (2004) Molecular sequence analysis of prokaryotic diversity in the middle and outer sections of the Portuguese estuary *Ria de Aveiro*. *FEMS Microbiology Ecology* **49**: 269-279.

Herzog, H., Eliasson, B., and Kaarstad, O. (2000) Capturing greenhouse gases. *Scientific American* **282**: 72-77, 79.

Hewson, I., Capone, D.G., Steele, J.A., and Fuhrman, J.A. (2006) Influence of Amazon and Orinoco offshore surface water plumes on oligotrophic bacterioplankton diversity in the west tropical Atlantic. *Aquatic Microbial Ecology* **43**: 11-22.

Hewson, I., Vargo, G.A., and Fuhrman, J.A. (2003) Bacterial diversity in shallow oligotrophic marine benthos and overlying waters: Effects of virus infection, containment, and nutrient enrichment. *Microbial Ecology* **46**: 322-336.

Hicks, R.E., Amann, R., and Stahl, D.A. (1992) Dual staining of natural bacterioplankton with 4',6-diamidino-2-phenylindole and fluorescent oligonucleotide probes targeting kingdom-level 16S rRNA sequences. *Applied and Environmental Microbiology* **58**: 2158-2163.

Hoffmann, L., Peeken, I., Assmy, P., Veldhuis, M.J.W., and Lochte, K. (2005) Response of pico-, nano-, and microphytoplankton during the Southern Ocean iron fertilization experiment EIFEX. In *ASLO 2005 Summer Meeting*. Santiago de Compostela, Spain.

Holben, W.E., Feris, K.P., Kettunen, A., and Apajalahti, J.H.A. (2004) GC fractionation enhances microbial community diversity assessment and detection of minority populations of bacteria by denaturing gradient gel electrophoresis. *Applied and Environmental Microbiology* **70**: 2263-2270.

Hopkinson, C.S.J., and Vallino, J.J. (2005) Efficient export of carbon to the deep ocean through dissolved organic matter. *Nature* **433**: 142-145.

Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., Dai, X. et al. (2001) *Climate change 2001: the scientific basis. Contribution of Working Group I to the Third Assessment Report of the International Panel on Climate Change*. Cambridge: Cambridge University Press.

Huber, J.A., Welch, D.B.M., Morrison, H.G., Huse, S.M., Neal, P.R., Butterfield, D.A., and Sogin, M.L. (2007) Microbial population structures in the deep marine biosphere. *Science* **318**: 97-100.

Hutchins, D.A., Campbell, B.J., Cottrell, M.T., Takeda, S., and Cary, S.C. (2001a) Response of marine bacterial community composition to iron additions in three iron-limited regimes. *Limnology and Oceanography* **46**: 1535-1545.

Hutchins, D.A., Sedwick, P.N., DiTullio, G.R., Boyd, P.W., Quéguiner, B., Griffiths, F.B., and Crossley, C. (2001b) Control of phytoplankton growth by iron and silicic acid availability in the subantarctic Southern Ocean: Experimental results from the SAZ project. *Journal of Geophysical Research* **106**: 31559-31572.

Huws, S.A., Edwards, J.E., Kim, E.J., and Scollan, N.D. (2007) Specificity and sensitivity of eubacterial primers utilised for molecular profiling of bacteria within complex microbial ecosystems. *Journal of Microbiological Methods* **70**: 565-569.

Huxley, J. (1935) T.H. Huxley's diary of the voyage of H.M.S. Rattlesnake. London: Chatto and Windus.

Iglesias-Rodríguez, M.D., Brown, C.W., Doney, S.C., Kleypas, J., Kolber, D., Kolber, Z. et al. (2002) Representing key phytoplankton functional groups in ocean carbon cycle models: Coccolithophorids. *Global Biogeochemical Cycles* **16**: (47)41-20.

Isao, K., Hara, S., Terauchi, K., and Kogure, K. (1990) Role of sub-micrometre particles in the ocean. *Nature* **345**: 242-244.

Jaccard, P. (1908) Nouvelles recherches sur la distribution florale. *Bulletin de la Society Vaudoise des Sciences Naturale* **44**: 223-270.

Jameson, E., Joint, I., Mann, N.H., and Mühling, M. (2008) Application of a novel *rpoC1*-RFLP approach reveals that marine *Prochlorococcus* populations in the Atlantic gyres are composed of greater microdiversity than previously described. *Microbial Ecology* **55**: (1) 141-151.

Jannasch, H.W., and Jones, G.E. (1959) Bacterial populations in seawater as determined by different methods of enumeration. *Limnology and Oceanography* **4**: 128-139.

Jaspers, E., and Overmann, J. (2004) Ecological significance of microdiversity: Identical 16S rRNA gene sequences can be found in bacteria with highly divergent genomes and ecophysiologies. *Applied and Environmental Microbiology* **70**: 4831-4839.

Jezberova, J. (2006) Phenotypic diversity and phylogeny of picocyanobacteria in mesotrophic and eutrophic reservoirs investigated by a cultivation-dependent polyphasic approach. In *Depatment of Faculty of Biological Sciences*. Ceske Budejovice: University of South Bohemia.

Jiao, N., Yang, Y., Hong, N., Ma, Y., Harada, S., Koshikawa, H., and Watanabe, M. (2005) Dynamics of autotrophic picoplankton and heterotrophic bacteria in the East China Sea. *Continental Shelf Research* **25**: 1265-1279.

Johansen, J.E., Nielsen, P., and Sjoholm, C. (1999) Description of *Cellulophaga baltica* gen. nov., sp. nov. and *Cellulophaga fucicola* gen. nov., sp. nov. and reclassification of *Cytophaga lytica* to *Cellulophaga lytica* gen. nov., comb. nov. *International Journal of Systematic Bacteriology* **49**: 1231-1240.

Johnson, K.S., and Karl, D.M. (2002) Is ocean fertilisation credible and creditable? *Science* **296**: 467.

Johnson, P.W., and Sieburth, J.M. (1979) Chroococcoid cyanobacteria in the sea: A ubiquitous and diverse phototrophic biomass. *Limnology and Oceanography* **24**: 928-935.

Kaeberlein, T., Lewis, K., and Epstein, S.S. (2002) Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* **296**: 1127-1129.

Kasting, F., and Seifert, J.L. (2002) Life and the evolution of Earth's atmosphere. *Science* **296**: 1066-1068.

Kheshgi, H.S. (1995) Sequestering atmospheric carbon dioxide by increasing ocean alkalinity. *Energy - The International Journal* **20**: 915-922.

Kirchman, D. (2000) Uptake and regeneration of inorganic nutrients by marine heterotrophic bacteria. In *Microbial Ecology of the Oceans*. Kirchman, D. (ed). New York: John Wiley and Sons.

Kirchman, D., and Rich, J. (1997) Regulation of bacterial growth rates by dissolved organic carbon and temperature in the equatorial Pacific Ocean. *Microbial Ecology* **33**: 11-20.

Kita-Tsukamoto, K., Oyaizu, H., Nanba, K., and Simidu, U. (1993) Phylogenetic relationships of marine bacteria, mainly members of the family Vibrionaceae, determined on the basis of 16S rRNA sequences. *International Journal of Systematic Bacteriology* **43**: 8-19.

Koch, R. (1876) The etiology of anthrax, based on the life history of *Bacillus anthracis*. *Beiträge zur Biologie der Pflanzen* **2**: 277-310.

Konstantinidis, K.T., and Tiedje, J.M. (2005) Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 2567-2572.

Krempin, D.W., and Sullivan, C.W. (1981) The seasonal abundance, vertical distribution, and relative microbial biomass of chroococcoid cyanobacteria at a station in southern California coastal waters. *Canadian Journal of Microbiology* **27**: 1341-1344.

Krom, M.D., Kress, N., Brenner, S., and Gordon, L.I. (1991) Phosphorus limitation of primary productivity in the eastern Mediterranean Sea. *Limnology and Oceanography* **36**: 424-432.

Kysela, D.T., Palacios, C., and Sogin, M.L. (2005) Serial analysis of V6 ribosomal sequence tags (SARST-V6): A method for efficient, high-throughput analysis of microbial community composition. *Environmental Microbiology* **7**: 356-364.

Labrenz, M., Brettar, I., Christen, R., Flavier, S., Bötel, J., and Höfle, M.G. (2004) Development and application of a real-time PCR approach for quantification of uncultured bacteria in the central Baltic Sea. *Applied and Environmental Microbiology* **70**: 4971-4979.

Lebaron, P., Servais, P., Troussellier, M., Courties, C., Muyzer, G., Bernard, L. et al. (2001) Microbial community dynamics in Mediterranean nutrient-enriched seawater mesocosms: changes in abundances, activity and composition. *FEMS Microbiology Ecology* **34**: 255-266.

Lee, S.H., Malone, C., and Kemp, P.F. (1993) Use of 16S rRNA-targeted fluorescent probes to increase signal strength and measure cellular RNA from natural planktonic bacteria. *Marine Ecology Progress Series* **101**: 193-201.

Lefevre, D., Guigue, C., and Obernosterer, I. (2008) The metabolic balance at two contrasting sites in the Southern Ocean: The iron-fertilised Kerguelen area and HNLC waters. *Deep-Sea Research. Part II - Topical Studies in Oceanography* **55**: (5-7) 766-776.

Li, W.K.W., and Dickie, P.M. (1985) Growth of bacteria in seawater filtered through 0.2 m nucleopore membranes: implications for dilution experiments. *Marine Ecology Progress Series* **26**: 245-252.

Li, Y., Saxena, D., Barnes, V.M., Trivedi, H.M., Ge, Y., and Xu, T. (2006a) Polymerase chain reaction-based denaturing gradient gel electrophoresis in the evaluation of oral microbiota. *Oral Microbiology and Immunology* **21**: 333-339.

Li, Z.-Y., He, L.-M., Wu, J., and Jiang, Q. (2006b) Bacterial community diversity associated with four marine sponges from the South China Sea based on 16S rDNA-DGGE fingerprinting. *Journal of Experimental Marine Biology and Ecology* **329**: 75-85.

Liu, W.-T., Marsh, T.L., Cheng, H., and Forney, L.J. (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Applied and Environmental Microbiology* **63**: 4516-4522.

Lopez, I., Ruiz-Larrea, F.R., Cocolin, L., Orr, E., Phister, T., Marshall, M. et al. (2003) Design and evaluation of PCR primers for analysis of bacterial populations in wine by denaturing gradient gel electrophoresis. *Applied and Environmental Microbiology* **69**: 6801-6807.

Ludwig, W., and Schleifer, K.-H. (2005) Molecular phylogeny of bacteria based on comparative sequence analysis of conserved genes. In *Microbial Phylogeny and Evolution*. Sapp, J. (ed). Oxford: Oxford University Press.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar et al. (2004) ARB: A software environment for sequence data. *Nucleic Acids Research* **32**: 1363-1371.

Lunn, M., Sloan, W.T., and Curtis, T.P. (2004) Estimating bacterial diversity from clone libraries with flat rank abundance distributions. *Environmental Microbiology* **6** (10): 1081-1085.

Manefield, M., Whiteley, A.S., Griffiths, R.I., and Bailey, M.J. (2002) RNA stable isotope probing, a novel means of linking microbial community function to phylogeny. *Applied and Environmental Microbiology* **68**: 5367-5373.

Marchesi, J.R., Sato, T., Weightman, A.J., Martin, T.A., Fry, J.C., Hiom, S.J., and Wade, W.G. (1998) Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. *Applied and Environmental Microbiology* **64**: 795-799.

Martin, J.H., and Fitzwater, S.E. (1988) Iron deficiency limits phytoplankton growth in the northeast Pacific subarctic. *Nature* **331**: 341-343.

Martin, J.H., Gordon, R.M., and Fitzwater, S.E. (1990) Iron in Antarctic waters. *Nature* **345**: 156-158.

Martin, J.H., Other, A.N., and Other, A.N. (1994) Testing the iron hypothesis in ecosystems of the equatorial Pacific Ocean. *Nature* **371**: 123-129.

Martín-Cuadrado, A.-B., López-García, P., Alba, J.C., Moreira, D., Monticelli, L., and Strittmatter, A. (2007) Metagenomics of the deep Mediterranean, a warm bathypelagic habitat. *PLoS ONE* **2**: e914.

Martínez Martínez, J., Schroeder, D.C., Larsen, A., Bratbak, G., and Wilson, W.H. (2007) Molecular dynamics of *Emiliania huxleyi* and cooccurring viruses during two separate mesocosm studies. *Applied and Environmental Microbiology* **73**: 554-562.

Mary, I., Cummings, D.G., Biegala, I.C., Burkill, P.H., Archer, S.D., and Zubkov, M.V. (2006) Seasonal dynamics of bacterioplankton community structure at a coastal station in the western English Channel. *Aquatic Microbial Ecology* **42**: 119-126.

Mayrhofer, S., Mikoviny, T., Waldhuber, S., Wagner, A.O., Innerebner, G., Franke-Whittle, I.H. et al. (2006) Microbial community related to volatile organic compound (VOC) emission in household biowaste. *Environmental Microbiology* **8**: 1960-1974.

McCaig, A.E., Glover, L.A., and Prosser, J.I. (2001) Numerical analysis of grassland bacterial community structure under different land management regimens by using 16S ribosomal DNA sequence data and denaturing gradient gel electrophoresis banding patterns. *Applied and Environmental Microbiology* **67**: 4554-4559.

McCaig, A.E., Embley, T.M., and Prosser, J.I. (1994) Molecular analysis of enrichment cultures of marine ammonia oxidisers. *FEMS Microbiology Letters* **120**: 363-367.

Mehling, A., Wehmeier, U.F., and Piepersberg, W. (1995) Nucleotide sequences of streptomycete 16S ribosomal DNA—towards a specific identification system for streptomycetes using PCR. *Microbiology* **141**: 2139-2147.

Mills, M.M., Ridame, C., Davey, M., La Roche, J., and Geider, R.J. (2004) Iron and phosphorus co-limit nitrogen fixation in the eastern tropical North Atlantic. *Nature* **429**: 292-294, 232.

Miyamoto, H., Yamamoto, H., Arima, K., Fujii, J., Maruta, K., and Izu, K. (1997) Development of a new semi-nested PCR method for detection of Legionella species and its application to surveillance of legionellae in hospital cooling tower water. *Applied and Environmental Microbiology* **63**: 2489-2494.

Moeseneder, M.M., Arrieta, J.M., and Herndl, G.J. (2005) A comparison of DNA- and RNA-based clone libraries from the same marine bacterioplankton community. *FEMS Microbiology Ecology* **51**: 341-352.

Moeseneder, M.M., Arrieta, J.M., Muyzer, G., Winter, C., and Herndl, G.J. (1999) Optimization of terminal-restriction fragment length polymorphism analysis for complex marine bacterioplankton communities and comparison with denaturing gradient gel electrophoresis. *Applied and Environmental Microbiology* **65**: 3518-3525.

Moon-van der Staay, S.Y., de Wachter, R., and Vaulot, D. (2001) Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* **409**: 607-610.

Moore, L.R., Rocap, G., and Chisholm, S.W. (1998) Phisiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature* **393**: 464-467.

Moyer, C.L., Dobbs, F.C., and Karl, D.M. (1994) Estimation of diversity and community structure through restriction fragment length polymorphism distribution analysis of bacterial 16S rRNA genes from a microbial mat at an active, hydrothermal vent system, Loihi seamount, Hawaii. *Applied and Environmental Microbiology* **60**: 871-879.

Mühling, M., Fuller, M.J., Somerfield, P.J., Post, A.F., Wilson, W.H., Scanlan, D.J. et al. (2006) High resolution genetic diversity studies of marine *Synechococcus* isolates using *rpoC1*-based restriction fragment length polymorphism. *Aquatic Microbial Ecology* **43**: 263-275.

Mühling, M., Fuller, N.J., Millard, A., Somerfield, P.J., Marie, D., Wilson, W.H. et al. (2005) Genetic diversity of marine Synechococcus and co-occurring cyanophage communities: evidence for viral control of phytoplankton. *Environmental Microbiology* **7**: 499-508.

Mühling, M., Woolven-Allen, J., Murrell, J.C., and Joint, I. (2007) Improved group-specific PCR primers for denaturing gradient gel electrophoresis analysis of the genetic diversity of complex microbial communities. *The ISME Journal*. **2**: 379-392.

Mullis, K., and Faloona, F. (1987) Specific synthesis of DNA *in vitro* via a polymerase catalyzed chain reaction. *Methods in Enzymology* **55**: 335-350.

Munn, C.B. (2004) *Marine Microbiology, Ecology and Applications*. London: Bios Scientific Publishers.

Muyzer, G., Brinkhoff, T., Nubel, U., Santegoeds, C., Schäfer, H., and Waver, C. (1998) Denaturing gradient gel electrophoresis (DGGE) in microbial ecology. In

Molecular Microbial Ecology Manual. Akkermans, A.D.L., van Elsas, J.D., and de Bruijn, F.J. (eds). Dordrecht: Kluwer Academic Publishers, pp. 1-27.

Muyzer, G., de Waal, E.C., and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology* **59**: 695-700.

Nakatsu, C.H., Torsvik, V., and Øvreås, L. (2000) Soil community analysis using DGGE of 16S rDNA polymerase chain reaction products. *Soil Science Society of America Journal* **64**: 1382-1388.

Nedashkovskaya, O.I., Kim, S.B., Lysenko, A.M., Frolova, G.M., Mikhailov, V.V., and Bae, K.S. (2005) *Bizionia paragorgiae* gen. nov., sp. no., a novel member of the family Flavobacteriaceae isolated from the soft coral *Paragorgia arborea*. *International Journal of Systematic and Evolutionary Microbiology* **55**: 375-378.

Neef, A., Amann, R., Schlesner, H., and Schleifer, K.-H. (1998) Monitoring a widespread bacterial group: *in situ* detection of planctomycetes with 16S rRNA-targeted probes. *Microbiology* **144**: 3257-3266.

Neufeld, J.D., and Mohn, W.W. (2006) Assessment of microbial phylogenetic diversity based on environmental nucleic acids. In *Molecular Identification, Systematics, and Population Structure of Prokaryotes*. Stackebrandt, E. (ed). Berlin: Springer-Verlag, pp. 219-259.

Neufeld, J.D., Yu, Z., Lam, W., and Mohn, W.W. (2004) Serial analysis of ribosomal sequence tags (SARST): a high-throughput method for profiling complex microbial communities. *Environmental Microbiology* **6**: 131-144.

Nübel, U., Garcia-Pichel, F., and Muyzer, G. (1997) PCR primers to amplify 16S rRNA genes from cyanobacteria. *Applied and Environmental Microbiology* **63**: 3327-3332.

Ohmoto, H., Watanabe, Y., and Kumazawa, K. (2004) Evidence from massive siderite beds for a CO_2-rich atmosphere before ~1.8 billion years ago. *Nature* **429**: 395-399.

Oliver, J.L., Barber, R.T., Smith, W.O., Jr., and Ducklow, H.W. (2004) The heterotrophic bacterial response during the Southern Ocean Iron Experiment (SOFeX). *Limnology and Oceanography* **49**: 2129-2140.

Ormerod, W.G., Freund, P., and Smith, A. (2002) *Ocean storage of CO2, 2nd Edition*. Cheltenham.

Pace, N.R. (1997) A molecular view of microbial diversity and the biosphere. *Science* **276**: 734-740.

Pahlow, M., and Riebesell, U. (2000) Temporal trends in deep ocean Redfield ratios. *Science* **287**: 831-833.

Pakulski, J.D., Coffin, R.B., Kelley, C.A., Holder, S.L., Downer, R., Aas, P. et al. (1996) Iron stimulation of Antarctic bacteria. *Nature* **383**: 133-134.

Partensky, F., Hess, W.R., and Vaulot, D. (1999) *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiology and Molecular Biology Reviews* **63**: 106-127.

Pearson, P.N., and Palmer, M.R. (2000) Atmospheric carbon dioxide concentrations over the past 60 million years. *Nature* **406**: 695-699.

Pernthaler, J., and Amann, R. (2005) Fate of heterotrophic microbes in pelagic habitats: Focus on populations. *Microbiology and Molecular Biology Reviews* **69**: 440-461.

Pernthaler, J., Glöckner, F.-O., Schönhuber, W., and Amann, R. (2001) Fluorescence in situ hybridization (FISH) with rRNA-targeted oligonucleotide probes. In *Marine Microbiology*. Paul, J.H. (ed). San Diego: Academic Press pp. 207-226.

Perreault, N.N., Andersen, D.T., Pollard, W.H., Greer, C.W., and Whyte, L.G. (2007) Characterization of the prokaryotic diversity in cold saline perennial springs of the Canadian high Arctic. *Applied and Environmental Microbiology* **73**: 1532-1543.

Perry, J.J., Staley, J.T., and Lory, S. (2002) *Microbial Life*. Sunderland: Sinauer Associates.

Pham, V.D., Konstantinidis, K.T., Palden, T., and DeLong, E.F. (2008) Phylogenetic analyses of ribosomal DNA-containing bacterioplankton genome fragments from a 4000 m vertical profile in the North Pacific subtropical gyre. *Environmental Microbiology* **10**: 2313-2330.

Pinhassi, J., and Hagström, Å. (2000) Seasonal succession in marine bacterioplankton. *Aquatic Microbial Ecology* **21**: 245-256.

Plasinskl, S., Deel, D., Miller, L., and Kane, B. (2007) Carbon sequestration technology roadmap and program plan. In. Energy, U.S.D.o. (ed): National Energy Technology Laboratory.

Platt, T., Subba Rao, D.V., and Irwin, B. (1983) Photosynthesis of picoplankton in the oligotrophic ocean. *Nature* **316**: 747-749.

Pomeroy, L.R. (1974) The ocean's food web, a changing paradigm. *BioScience* **24**: 499-504.

Pomeroy, L.R., Sheldon, J.E., Sheldon, W.M.J., and Peters, F. (1995) Limits to growth and respiration of bacterioplankton in the Gulf of Mexico. *Marine Ecology Progress Series* **117**: 259-268.

Pomeroy, L.R., Williams, P.J.l., Azam, F., and Hobbie, J.E. (2007) The microbial loop. *Oceanography* **20**: 28-33.

Pommier, T., Canbäck, B., Riemann, L., Boström, K.H., Simu, K., Lundberg, P. et al. (2006) Global patterns of diversity and community structure in marine bacterioplankton. *Molecular Ecology* **16**: 867-880.

Prabagaran, S.R., Manorama, R., Delille, D., and Shivaji, S. (2007) Predominance of *Roseobacter*, *Sulfitobacter*, *Glaciecola* and *Psychrobacter* in seawater collected off Ushuaia, Argentina, sub-Antarctica. *FEMS Microbiology Ecology* **59**: 342-355.

Proujan, C. (1979) *Secrets of the Sea*. London: Reader's Digest Association Ltd.

Purdy, K.J. (2005) Nucleic acid recovery from complex environmental samples. *Methods in Enzymology* **397**: 271-292.

Rappé, M.S., and Giovannoni, S.J. (2003) The uncultured microbial majority. *Annual Review of Microbiology* **57**: 369-394.

Rappé, M.S., Vergin, K., and Giovannoni, S.J. (2000) Phylogenetic comparisons of a coastal bacterioplankton community with its counterparts in open ocean and freshwater systems. *FEMS Microbiology Ecology* **33**: 219-232.

Rau, G.H., and Caldeira, K. (1999) Enhanced carbonate dissolution: a means of sequestering waste CO_2 as ocean bicarbonate. *Energy Conversion and Management* **40**: 1803-1813.

Rau, G.H., and Caldeira, K. (2002) Minimising effects of CO_2 storage in oceans. *Science* **276**: 275-276.

Raven, J., Caldeira, K., Elderfield, H., Hoegh-Guldberg, O., Liss, P., Riebesell, U. et al. (2005) Ocean acidification due to increasing atmospheric carbon dioxide. In: The Royal Society.

Ridgwell, A.J., Kennedy, M.J., and Caldeira, K. (2003) Carbonate deposition, climate stability, and neoproterozoic ice ages. *Science* **302**: 859-862.

Riebesell, U. (2004) Effects of CO_2 enrichment on marine phytoplankton. *Journal of Oceanography* **60**: 719-729.

Riebesell, U., Schulz, K.G., Bellerby, R.G.J., Botros, M., Fritsche, P., Meyerhöfer, M. et al. (2007) Enhanced biological carbon consumption in a high CO_2 ocean. *Nature* **450**: 545-548.

Riebesell, U., Zondervan, I., Rost, B., Tortell, P.D., Zeebe, R.E., and Morel, F.M.M. (2000) Reduced calcification of marine plankton in response to increased atmospheric CO_2. *Nature* **407**: 364-367.

Riemann, L., Steward, G.F., Fandino, L.B., Campbell, L., Landry, M.R., and Azam, F. (1999) Bacterial community composition during two consecutive NE Monsoon periods in the Arabian Sea studied by denaturing gradient gel electrophoresis (DGGE) of rRNA genes. *Deep-Sea Research. Part II - Topical Studies in Oceanography* **46**: 1791-1811.

Rink, B., Seeberger, S., Martens, T., Duerselen, C.-D., Simon, M., and Brinkhoff, T. (2007) Effects of phytoplankton bloom in a coastal ecosystem on the composition of bacterial communities. *Aquatic Microbial Ecology* **48**: 47-60.

Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., and Stanier, R.Y. (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Journal of General Microbiology* **111**: 1-61.

Rivkin, R.B., and Anderson, M. (1997) Inorganic nutrient limitation of oceanic bacterioplankton. *Limnology and Oceanography* **42**: 730-740.

Robinson, C., Poulton, A.J., Holligan, P.M., Baker, A.R., Foster, G., and Gist, N. (2006) The Atlantic Meridional Transect Programme (AMT): a contextual view 1995–2005. *Deep-Sea Research. Part II - Topical Studies in Oceanography* **53**: 1485-1515.

Rocap, G., Distel, D.L., Waterbury, J.B., and Chisholm, S.W. (2002) Resolution of Prochlorococcus and Synechococcus ecotypes by using 16S-23S ribosomal DNA internal transcribed spacer sequences. *Applied and Environmental Microbiology* **68**: 1180-1191.

Rost, B., Riebesell, U., Burkhardt, S., and Sültemeyer, D. (2003) Carbon acquisition of bloom-forming marine phytoplankton. *Limnology and Oceanography* **48**: 55-67.

Sabehi, G., Béjà, O., Suzuki, M.T., Preston, C.M., and DeLong, E.F. (2004) Different SAR86 subgroups harbour divergent proteorhodopsins. *Environmental Microbiology* **6**: 903-910.

Sabine, C.L., Feely, R.A., Gruber, N., Key, R.M., Lee, K., Bullister, J.L. et al. (2004) The oceanic sink for anthropogenic CO_2. *Science* **305**: 367-371.

Saitou, N. and Nei, M. (1987) The neighbour-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4** (4): 406-425.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**: 5463-5467.

Santelli, C.M., Orcutt, B.N., Banning, E., Bach, W., Moyer, C.L., Sogin, M.L. et al. (2008) Abundance and diversity of microbial life in ocean crust. *Nature* **453**: 653-656.

Scanlan, D.J., and West, N.J. (2002) Molecular ecology of the marine cyanobacterial genera *Prochlorococcus* and *Synechococcus*. *FEMS Microbiology Ecology* **40**: 1-12.

Schäfer, H., and Muyzer, G. (2001) Denaturing gradient gel electrophoresis in marine microbial ecology. In *Marine Microbiology*. Paul, J.H. (ed). San Diego: Academic Press pp. 425-468.

Schäfer, H., Abbas, B., Witte, H., and Muyzer, G. (2002) Genetic diversity of "satellite" bacteria present in cultures of marine diatoms. *FEMS Microbiology Ecology* **42**: 25-35.

Schäfer, H., Bernard, L., Courties, C., Lebaron, P., Servais, P., Pukall, R. et al. (2001) Microbial community dynamics in Mediterranean nutrient-enriched seawater mesocosms: Changes in the genetic diversity of bacterial populations. *FEMS Microbiology Ecology* **34**: 243-253.

Schippers, P., Lürling, M., and Scheffer, M. (2004) Increase of atmospheric CO_2 promotes phytoplankton productivity. *Ecology Letters* **7**: 446-451.

Schmalenberger, A., Schwieger, F., and Tebbe, C.C. (2001) Effect of primers hybridising to different evolutionarily conserved regions of the small-subunit rRNA gene in PCR-based microbial community analyses and genetic profiling. *Applied and Environmental Microbiology* **67**: 3557-3563.

Schut, F., de Vries, E.J., Gottschal, J.C., Robertson, B.R., Harder, W., Prins, R.A., and Button, D.K. (1993) Isolation of typical marine bacteria by dilution culture: growth, maintenance, and characteristics of isolates under laboratory conditions. *Applied and Environmental Microbiology* **59**: 2150-2160.

Sciandra, A., Harlay, J., Lefèvre, D., Lemée, R., Rimmelin, P., Denis, M., and Gattuso, J.-P. (2003) Response of coccolithophorid *Emiliania huxleyi* to elevated partial pressure of CO_2 under nitrogen limitation. *Marine Ecology Progress Series* **261**: 111-122.

Sheffield, V.C., Cox, D.R., Lerman, L.S., and Myers, R.M. (1989) Attachment of a 40-base-pair G+C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. *Proceedings of the National Academy of Sciences of the United States of America* **86**: 232-236.

Sieburth, J.M. (1976) Bacterial substrates and productivity in marine ecosystems. *Annual Review of Ecology and Systematics* **7**: 259-285.

Simu, K., and Hagström, Å. (2004) Oligotrophic bacterioplankton with a novel single-cell life strategy. *Applied and Environmental Microbiology* **70**: 2445-2451.

Sipos, R., Szekely, A.J., Palatinszky, M., Revesz, S., Marialigeti, K., and Nikolausz, M. (2007) Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis. *FEMS Microbiology Ecology* **60**: 341-350.

Sneath, P.H. (1993) Evidence from *Aeromonas* for genetic crossing-over in ribosomal sequences. *International Journal of Systematic Bacteriology* **43**: 626-629.

Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R. et al. (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences of the United States of America* **103**: 12115-12120.

Stach, J.E.M., Bathe, S., Clapp, J.P., and Burns, R.G. (2001) PCR-SSCP comparison of 16S rDNA sequence diversity in soil DNA obtained using different isolation and purification methods. *FEMS Microbiology Ecology* **36**: 139-151.

Staley, J.T., and Konopka, A. (1985) Measurements of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annual Review of Microbiology* **39**: 321-346.

Stingl, U., Desiderio, R.A., Cho, J.C., Vergin, K., and Giovannoni, S. (2007) The SAR92 clade: an abundant coastal clade of culturable marine bacteria possessing proteorhodopsin. *Applied and Environmental Microbiology* **73**: 2290-2296.

Stretton, S., Marshall, K.C., Dawes, I.W., and Goodman, A.E. (1996) Characterisation of carbon dioxide-inducible genes of the marine bacterium, *Pseudomonas* sp. S91. *FEMS Microbiology Letters* **140**: 37-42.

Strous, M., Pelletier, E., Mangenot, S., Rattei, T. et al., (2006) Deciphering the evolution and metabolism of an anammox bacterium from a community genome. *Nature* **440**: 790-794.

Stubner, S. (2004) Quantification of Gram-negative sulphate-reducing bacteria in rice field soil by 16S rRNA gene-targeted real-time PCR. *Journal of Microbiological Methods* **57**: 219-230.

Suzuki, K., Hinuma, A., Saito, H., Kiyosawa, H., Liu, H., Saino, T., and Tsuda, A. (2005) Responses of phytoplankton and heterotrophic bacteria in the northwest subarctic Pacific to *in situ* iron fertilisation as estimated by HPLC pigmant analysis and flow cytometry. *Progress in Oceanography* **64**: 167-187.

Suzuki, M.T., Rappé, M., and Giovannoni, S. (1998) Kinetic bias in estimate of coastal picoplankton community structure obtained by measurements of small-subunit rRNA gene PCR amplicaon length heterogeneity. *Applied and Environmental Microbiology* **63**: 983-989.

Suzuki, M.T., Taylor, L.T., and DeLong, E.F. (2000) Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Applied and Environmental Microbiology* **66**: 4605-4614.

Suzuki, M.T., Preston, C.M., Béjà, O., de la Torre, J.R., Steward, G.F., and DeLong, E.F. (2004) Phylogenetic screening of ribosomal RNA gene-containing clones in bacterial artificial chromosome (BAC) libraries from different depths in Monterey Bay. *Microbial Ecology* **48**: 473-488.

Suzuki, M.T., Rappé, M.S., Haimberger, Z.W., Winfield, H., Adair, N., Ströbel, J., and Giovannoni, S.J. (1997) Bacterial diversity among small-subunit rRNA gene clones and cellular isolates from the same seawater sample. *Applied and Environmental Microbiology* **63**: 983-989.

Takahashi, T., Sutherland, S.C., Feely, R.A., and Cosca, C.E. (2003) Decadal variation of the surface water pCO_2 in the western and central equatorial Pacific. *Science* **302**: 852-856.

Takahashi, T., Feely, R.A., Weiss, R.F., Wanninkhof, R.H., Chipman, D.W., Sutherland, S.C., and Takahashi, T.T. (1997) Global air-sea flux of CO_2: An estimate

based on measurements of sea-air pCO_2 difference. *Proceedings of the National Academy of Sciences of the United States of America* **94**: 8292-8299.

Takami, H., and Krulwich, T.A. (2000) Reidentification of facultatively alkaliphilic *Bacillus firmus* OF4. *Extremophiles* **4**: 19-22.

Thingstad, T.F. (2000) Control of bacterial growth in idealised food webs. In *Microbial Ecology of the Oceans*. Kirchman, D. (ed). New York: John Wiley and Sons.

Thingstad, T.F., Krom, M.D., Mantoura, R.F.C., Flaten, G.A.F., Groom, S., Herut, B. et al. (2005) Nature of phosphorus limitation in the ultraoligotrophic eastern Mediterranean. *Science* **309**: 1068-1071.

Thingstad, T.F., Zweifel, U.L., and Rassoulzadegan, F. (1998) P-limitation of both phytoplankton and heterotrophic bacteria in NW Mediterranean summer surface waters. *Limnology and Oceanography* **43**: 88-94.

Thistle, D., Carman, K.R., Sedlacek, L., Brewer, P.G., Fleeger, J.W., and Barry, J.P. (2005) Deep-ocean, sediment-dwelling animals are sensitive to sequestered carbon dioxide. *Marine Ecology Progress Series* **289**: 1-4.

Thurman, H.V. (1990) *Essentials of Oceanography*. Columbus: Merrill Publishing Company.

Todorov, J.R., Chistoserdov, A.Y., and Aller, J.Y. (2000) Molecular analysis of microbial communities in mobile deltaic muds of southeastern Papua New Guinea. *FEMS Microbiology Ecology* **33**: 147-155.

Tortell, P.D., Maldonado, M.T., and Price, N.M. (1996) The role of heterotrophic bacteria in iron-limited ocean ecosystems. *Nature* **383**: 330-332.

Tsuda, A., and Takeda, S. (2005) Results from the sub Arctic Pacific iron experiment for ecosystem dynamics study (SEEDS). *Progress in Oceanography* **64**: 91-324.

Ueda, K., Seki, T., Kudo, T., Yoshida, T., and Kataoka, M. (1999) Two distinct mechanisms cause heterogeneity of 16S rRNA. *Journal of Bacteriology* **181**: 78-82.

Uphoff, H.U., Felske, A., Fehr, W., and Wagner-Döbler, I. (2001) The microbial diversity in picoplankton enrichment cultures: a molecular screening of marine isolates. *FEMS Microbiology Ecology* **35**: 249-258.

Urbach, E., Scanlan, D.J., Distel, D.L., Waterbury, J.B., and Chisholm, S.W. (1998) Rapid diversification of marine picophytoplankton with dissimilar light-harvesting structures inferred from sequences of *Prochlorococcus* and *Synechococcus* (Cyanobacteria). *Journal of Molecular Evolution* **46**: 188-201.

Vancanneyt, M., Schut, F., Snauwaert, C., Goris, J., Swings, J., and Gottschal, J.C. (2001) *Sphingomonas alaskensis* sp. nov., a dominant bacterium from a marine oligotrophic environment. *International Journal of Systematic and Evolutionary Microbiology* **51**: 73-79.

Van Mooy, B.A.S., Rocap, G., Fredricks, H., Evans, C.T., and Devol, A.H. (2006) Sulfolipids dramatically decrease phosphorus demand by picocyanobacteria in oligotrophic marine environments. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 8607-8612.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A. et al. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66-74.

Verdugo, P., Alldredge, A.L., Azam, F., Kirchman, D.L., Passow, U., and Santschi, P.H. (2004) The oceanic gel phase: a bridge in the DOM-POM continuum. *Marine Chemistry* **92**: 67-85.

Vergin, K.L., Urbach, E., Stein, J.L., DeLong, E.F., Lanoil, B.D., and Giovannoni, S.J. (1998) Screening of a fosmid library of marine environmental genomic DNA fragments reveals four clones related to members of the order *Planctomycetales*. *Applied and Environmental Microbiology* **64**: 3075-3078.

Wagner-Döbler, I., and Biebl, H. (2006) Environmental biology of the marine *Roseobacter* lineage. *Annual Review of Microbiology* **60**: 255-280.

Walker, N.J. (2002) A technique whose time has come. *Science* **296**: 557-559.

Ward, N., Rainey, F.A., Stackebrandt, E., and Schlesner, H. (1995) Unravelling the extent of diversity within the order *Planctomycetales*. *Applied and Environmental Microbiology* **61**: 2270-2275.

Waterbury, J.B., Watson, S.W., Guillard, R.R., and Brane, L.E. (1979) Widespread occurrence of a unicellular, marine, planktonic cyanobacterium. *Nature* **277**: 293-294.

Watson, A., Law, C.S., Van Scoy, K.A., Millero, F.J., Yao, W., Friederich, G. et al. (1994) Minimal effect of iron fertilisation on sea-surface carbon dioxide concentrations. *Nature* **371**: 143-145.

Wayne, L.G., Brenner, D.J., and Colwell, R.R. (1987) Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *International Journal of Systematic Bacteriology* **37**: 463-464.

Weisburg, W.G., Barns, S.M., Pelletier, D.A., and Lane, D.J. (1991) 16S ribosomal DNA amplification for phylogenetic study. *Journal of Bacteriology* **173**: 697-703.

West, N.J., and Scanlan, D.J. (1999) Niche-partitioning of *Prochlorococcus* populations in a stratified water column in the eastern North Atlantic Ocean. *Applied and Environmental Microbiology* **65**: 2585-2591.

West, N.J., Obernosterer, I., Zemb, O., and Lebaron, P. (2008) Major differences of bacterial diversity and activity inside and outside of a natural iron-fertilised phytoplankton bloom in the Southern Ocean. *Environmental Microbiology* **10**: 738-756.

Whiteley, A.S., Manefield, M., and Lueders, T. (2006) Unlocking the "microbial black box" using RNA-based stable isotope probing technologies. *Current Opinion in Biotechnology* **17**: 67-71.

Whitman, W.B., Coleman, D.C., and Wiebe, W.J. (1998) Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 6578-6583.

Wichels, A., Würtz, S., Döpke, H., Schütt, C., and Gerdts, G. (2006) Bacterial diversity in the breadcrumb sponge *Halichondria panicea* (Pallas). *FEMS Microbiology Ecology* **56**: 102-118.

Widmer, F., Seidler, R.J., Gillevet, P.M., Watrud, L.S., and Di Giovanni, G.D. (1998) A highly selective PCR protocol for detecting 16S rRNA genes of the genus *Pseudomonas* (sensu stricto) in environmental samples. *Applied and Environmental Microbiology* **64**: 2545-2553.

Wintzingerode, F.v., Göbel, U.B., and Stackebrandt, E. (1997) Determination of microbial diversity in environmental samples: Pitfalls of PCR-based rRNA analysis. *FEMS Microbiology Reviews* **21**: 213-229.

Woese, C.R. (1987) Bacterial evolution. *Microbiological Reviews* **51**: 221-271.

Wolf-Gladrow, D.A., Riebesell, U., Burkhardt, S., and Bijma, J. (1999) Direct effects of CO_2 concentration on growth and isotopic composition of marine plankton. *Tellus* **51**: 461-476.

Yoon, H.S., Hackett, J.D., Pinto, G., and Bhattacharya, D. (2002) The single, ancient origin of chromist plastids. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 15507-15512.

Young, E. (2007) A drop in the ocean: Can dumping tonnes of fertiliser into the oceans really turn back the tide of global warming? *New Scientist* 15 Sept. 2007: 42-45.

Zaballos, M., López-López, A., Ovreas, L., Bartual, S.G., D'Auria, G., Alba, J.C. et al. (2006) Comparison of prokaryotic diversity at offshore oceanic locations reveals a different microbiota in the Mediterranean Sea. *FEMS Microbiology Ecology* **56**: 389-405.

Zachos, J.C., Röhl, U., Schellenberg, S.A., Sluijs, A., Hodell, D.A., Kelly, D.C. et al. (2005) Rapid acidification of the ocean during the Paleocene-Eocene thermal maximum. *Science* **308**: 1611-1615.

Zeidner, G., and Béjà, O. (2004) The use of DGGE analyses to explore eastern Mediterranean and Red Sea marine picophytoplankton assemblages. *Environmental Microbiology* **6**: 528-534.

Zeidner, G., Preston, C.M., DeLong, E.F., Massana, R., Post, A.F., Scanlan, D.J., and Béjà, O. (2003) Molecular diversity among marine picophytoplankton as revealed by *psbA* analyses. *Environmental Microbiology* **5**: 212-216.

ZoBell, C.E., and Anderson, D.Q. (1936) Observations on the multiplication of bacteria in different volumes of stored sea water and the influence of oxygen tension and solid surfaces. *Biology Bulletin* **70-71**: 324-342.

Zondervan, I., Rost, B., and Riebesell, U. (2002) Effect of CO_2 concentration on the PIC/POC ratio in the coccolithophore *Emiliania huxleyi* grown under light-limiting conditions and different daylengths. *Journal of Experimental Marine Biology and Ecology* **272**: 55-70.

Zubkov, M.V., Fuchs, B.M., Tarran, G.A., Burkill, P.H., and Amann, R. (2003) High rate of uptake of organic nitrogen compounds by *Prochlorococcus* cyanobacteria as a key to their dominance in oligotrophic oceanic waters. *Applied and Environmental Microbiology* **69**: 1299-1304.

Zuckerkandl, E., and Pauling, L. (1965) Molecules as documents of evolutionary history. *Journal of Theoretical Biology* **8**: 357-366.

Zweifel, U.L., Norrman, B., and Hagström, Å. (1993) Consumption of dissolved organic carbon by marine bacteria and demand for inorgainc nutrients. *Marine Ecology Progress Series* **101**: 23-32.

Zondervan, I., Zeebe, R.E., Rost, B., and Riebesell, U. (2001) Decreasing marine biogenic calcification: a negative feedback on rising atmospheric pCO2. *Global Biogeochemical Cycles* **15**: 507-516.

Appendix I

Nucleotide sequences of 16S rRNA gene fragments re-amplified by PCR from numbered, and excised, bands following separation by DGGE

Table 4.2

1 AGCGCGCGTAGGTGGTTTGTATAAGTGAGATGTGAAGCCCTGGGCTCAACCTAGGAACTGCATCTTATACTGATTTG
CTAGAGTACGATAGAGGGGTGTGGAATTTACTGTGTAGCGGTGGAATGCGTAGATATTAGGAGAATACCAGTGGCG
AAGGCGAGCTACTGGGATCTGATACTGACACTGGAGGTGCGAAAGCGTGGGTGAGCGAACA

2 TGGATGTGAAGCCCTGGGCTCAACCTGGGAACTGCATCCTAGACTGATTCACTAGAGTACGATAGAGGGATGTAGA
ATTCACAGTGTAGCGGTGGAATGCGTAGATATTGTGAAGAATACCAATGGCGAAAGCGGCCTCCTGGATCTGATACT
GACACTGAGGTGCGAAAGCCGTGGGTAGCGAACA

3 AGGTGGTTTGTTAAGTTGGATGTGAAAGCCCTGGGCTCAACCTGGGAACTGCATCCAAAACTGACTCACTAGAGTAC
GATAGAGGGAGGTAGAATTCATAGTGTAGCGGTGGAATGCGTAGATATTATGAAGAATACCAGTGGCGAAGGCGGC
CTCCTGGATCTGTACTGACACTGAGGTGCGAAAGCGTGGGTAGCGAACA

4 AGCGCGCGCTAGGTGGTTTGTTAAGTTGGATGTGAAAGCCCTGGGCTCAACCTGGGAACTGCATCCCAAACTGATT
CACTAGAGTACGATAGAGGGAGGTAGAATTCACAGTGTAGCGGTGGAATGCGTAGATATTATGAAGAATACCAATGG
CGAAGGCGGCCTCCTGGATCTGTACTGACACTGGAGGTGCGAAAGCGTGGGTAGCGAACA

5 CTACACTATGAATTCTACCTCCCTCTATCGTACTCTAGTGAGTTAGTTTTGGATGCAGTTCCTAGGTTGAGCCCAGGG
CTTTCACATCCAACTTAACAAACCACCTACGCGCGCTTTACGCCCAGTAATTCCGATTAACGCTTGCACCTTCCGT

6 CCAGGaGGCTGCCTTCGCCACTGGTATTCTTCACATATATCTACGCATTCTCACCGCTACACTGTGAATTCTACCTCC
CTCTATCGTACTCTAGTGAGTCAGTTTTGAGCATGCAGTTCCCAGGTTGAGCCCGGGCATATCACATCCAACTTAAC
AAACCACCTACGCGCGCTTTACGCCCAGTAATTCCGATTAACGCTTGCACCTTCCGT

Table 4.3

1 GATACTGTCATTATCATCCTCATGGCGAGAGAGCTTTACGACCCTAGGGCCTTCATCATCACACGCGGCATGGCTAG
ATCAGGGTTGCCCCCATTGTCTAAGATTCCCTA

2 ATACTGTCATTATCATCTCTGGCGAGAGTGCTTTACGACCCTAGGGCCTTCACTCACACACGCGGCATGGCTAGATC
AGGGTTGCCCCCATTGTCTAAGATTCCCTA

3 GATACTGTCATTATCATCTCTGGCGAGAGAGCTTTACGACCCTAGGGCCTTCATCATACACGCGGCATGGCTAGATC
AGGGTTGCCCCCATTGTCTAAGATTCCCTA

4 GTCATTATCATCACTGGCGACAGTGCTTTACGACCCTAAGGCCTTCATCACACACGCGGCATGGCTAGATCAGGGTT
GCCCCCATTGTCTAAGATTCCCCA

5 TACTGTCATCATCATCACTGTCGACGTGCTGTACGACCTCTGACTTCATCACACACGCGGCATGCTGAGATCAGCTT
GCGCCATTGTCTATATTCCCA

6 TACTGTCATCATCTTCACTGTGAAGTGCTTTACGACCCTAGGCTTCATCACACACGCGGCATTGCTGGATCAGGCTT
GCGCCATTGTCTATATTCCCA

7 ACTGTCATCATCATCACTGGCGACGTAGCTCTACGAGCCTAGACTTCATCACTCACGCAGCATGCTAGATCAGGCTT
GCGCGCATTGTCTAGATCGCGA

8 TACTGTCATTATCATCACTGGCGAAGAGCTTTACGACCCTAAGGCCTTCACTCTTACACGACGGCATGGCTAGATCA
GGCTTGCGCCCATTGTCTAAGATTCCCCA

9 ACTGTCATTATCATCACTGGCGACGAGCTCTACGACCCTAAGACTTCATCACTCACGCAGGCATGGCTAGATCATGC
TTGCGCCCATTGTCTATATTCGCA

10 ACTGTCATTATCATCTCTGGTCGAAGTAGCTTTACGACCCTAAGGCCTTCATCACATCACGAAGGCATGGCTAGATCA
GGCTTGCGCCCATTGTCTAAGATTCCCCA

11 ACTGTCATTATCATCACTGGCGAAGAGCTCTACGACCCTAAGCTTCATCACTCACGCGGCATGGCTAGATCATGCTT
GCGCCCATTGTCTAAGATTCTCCA

12 TACTGTCATTATCATCACTGTGCGAAGTAGCTTTACGACCCTAAGGCCTTCTTCATATCACGCGGCATGGCTAGATCA
GGCTTGCGCCCATTGTCTAAGATTCCCCA

13 CTGATCTAGCTATGCGCGTGTGTGATGAGGCTAGGCTGTAAAGCTACTTTCGTCTAGATGATGATGATGACTGTATA
CTCGAGAGAGAGCCCGGCTAACTTCGTGCCAGCAGCCGCGAT

14 CTGATCTAGCATGCGCGTGTGTGATGCAGGCCTAGGATCGTAAAGCACTTTCGCAGAGATGATAATGACAGTATCTG
GTACAGAACCCCGGCTAACTCCGTGCCAGCAGCCGCGA

16 CTGATCTAGCTATGCGCGTGTGTGATGAATGCCTAGGCTCGTAGAGCTCTTTCTATAGAGATGATAATGACAGTATCT
GTGTAGAGATCACCGGATAACTCCGTGCCAGCAGCCGCTGATATATATGCCAGCAGCCGCGATAAT

19 ACCTGATCTAGCATGCGCGTGTGTGATGAGGCTAGGATCTGTAAAGCACTTTCGCTAGATGATGATGATGACTGTAT
CTGAGTAGAGAGCTCCGGCTAACTCGTGCCAGCAGCCGCGATaATA

20 TTCCCAAGTTGAGCTCGGGGATTTCACATCTGACTTACAAAACCGCCTGCGCACCCTTTACGCCCAGTAATTCCGAT
TAACGCTCGCACCCTATGT

21 GAGCCCGGGGATTTCGCATCAGACTTACAAAACCGCGTGCGCACGCTTTACGCCCAGGAATTCCGATTAACGCTCG
CACCCTATGT

23 CTGCTACACGTACGCATTTCACTGCTACACAGACCCGTGACACCGCATCTGACATACATAACACCGCGACACGCCTT
TACGCCCAGGAAATCCGATTAACGCTCGCAGCCTATGG

24 ACTTACAAAACCGCGTGCGCACCCTTTACGCCCAGGAATTCCGATTAACGCTCGCACCCTATGT

25 CGGGGATTTCGCCCTTGACATACAAAACCGCGTGCGCACCCTTTACGCCCAGTAATTCCGATTAACGCTCGCACCCT
ATGT

26 GAGCCCGGGGATTTCACCCCTGACTTACAAAACCGCCTGCGCACGCTTTACGCCCAGTAATTCCGATTAACGCTCG
CACCCTACGT

27 CTGGTACTTTCACGCGCAGTCACCCAAGTTGAGCTCGGGGATTTCGCATCTGACTTACAAAACCGCCTGCGCACCCT
TTACGCCCAGGAATTCCGATTAACGCTCGCACCCTATGT

28 CGATTACTGGGCGTAAAGGGTGCGCAGGCGGATTTGTAAGTCAGATGTGAAATCCCCGAGCTCAACTTGGGAACTG
CGTTTGAAACTATAAGACTAGAGTGTGTCAGAGGGGGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAC

29 ATACTGGGCGTAAGGGTGCGCAGGCGTTTTGTAAGTCAGATGTGAATCCCCGATGCTCAACTTGGGAACTGCGTTT
GAAACTACAAGACTAGAGTGTGTCAGAGGGGGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAC

30 ATTACTGGGCGTAAAGCGTGCGCAGGCGTTTTGTAAGACAGGATGTGAATCCCCGAGTCTCAACCTTGGGAACTGC
GCTTGTAGACTGCAAGACTAGAGTGTGTCAGAGGGGGGTAGAATTCCGCGTGTAGCAGTGAAATGCGTAC

31 ATACTGTGGCGTAAGCGTGCGCGGGCGGCTTTGTAAGACAGACGTGAATCCCCGGGCTCAACCTGGGAATTGCGC
TTGTGACTGCAAGGCCTAGAGTGCAGCAGAGGGGGGTAGAATTCCGCGTGTAGCAGTGAAATGCGTAC
33 TCATACTGGGCGTAAAGCGTGCGCAGCGTTTTGTAAGACAGGACTGTGAAATCCCCGGGTCTCAACCTGGGAATTG
CGCTTGTAGACTGCAAGGATAGAGTACGGTAGAGGGGGGTAGAATTCCGCGTGTAGCAGTGAAATGCGTA
34 TGCATTACTGGGCGTAAAGGGTGCGCAGGCGTTTTGTAAGTCAGATGTGAAATCCCCGAGCTCAACTTGGGAACTG
CGTTTGAAACTACAAGACTAGAGTGTGTCAGAGGGGGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAC
35 GCATACTGGGCGTAAAGGGTGCGCAGGCGGTTTTGTAAGTCAGATGTGAAATCCCCGAGCTCAACTTGGGAACTGC
GTTTGAAACTATAAGACTAGAGTGTGTCAGAGGGGGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAC
36 CTACACTATGAATTCTACCTCCCTCTATCGTACTCTAGTGAGTTAGTTTTGGATGCAGTTCCTAGGTTGAGCCCAGGG
CTTTCACATCCAACTTAACAAACCACCTACGCGCGCTTTACGCCCAGTAATTCCGATTAACGCTTGCACCTTCCGT
37 CCAGGaGGCTGCCTTCGCCACTGGTATTCTTCACATATATCTACGCATTCTCACCGCTACACTGTGAATTCTACCTCC
CTCTATCGTACTCTAGTGAGTCAGTTTTGAGCATGCAGTTCCCAGGTTGAGCCCGGGCATATCACATCCAACTTAAC
AAACCACCTACGCGCGCTTTACGCCCAGTAATTCCGATTAACGCTTGCACCTTCCGT
38 CTCAGCGTCAGTATCGAGCCAGGTGGCCGCCTTCGCCACTGGTATTCCTCCATATATCTACCCATTTCACCGCTACA
CAGGGAAATTCTACCTCCCTCTATCGTACTCTAGTCAGACAGTTTTGGATGCAGTTCCCAGGTTGAGCCCGGGGCTT
TCACATCCAACTTAACAAACCACCTACGCGCGCTTTACGCCCAGTAATTCCGATTAACGCTTGCACCTTCCGT
39 GTCAGTATCGAGCCAGGTGGCCGCCTTCGCCACTGGTATTCCTCACATATATCTACGCATTTCACCGCTACACATGG
AATTCTACCTCCCTCTCTCGTACTCTAGTCAGCCAGTATCGAATGCAGTTCCCAGGTTGAGCCCGGGGCTTTCACAT
CTGACTTAACAAACCGCCTACGCGCGCTTTACGCCCAGTAATTCCGATTAACGCTCGCACCCTCCGT
40 TCAGTATCGAGCAGGTGGCCGCCTTCGCCACTGGTATTCCTCCATATATCTACGTTATTTCACCGCTACACTGGGAA
ATTCTACCTCCCTCTATCGTACTCTAGTGAGTCAGTTTTGGATGCAGTTCCCAGGTTGAGCCCGGGGCTTTCACATC
CAACTTAACAAACCACCTACGCGCGCTTTTACGCCCAGTAATTCCGATTAACGCTTGCACCTTCCGT
41 ATCCAGGTGGCCGCCTTCGCCACTGGTATTCCTCCTGATATCTACGCATTTCACCGCTACACTGGGAATTCTACCTC
CCTCTATCGTACTCTAGCCTGACAGTTTTGGATGCAGTTCCCAGGTTGAGCCCGGGGCTTTCACATCCGACTTATCA
GACCACCTACGCGCGCTTTACGCCCAGTAATTCCGATTAACGCTCGCACCTTCCGT
42 ACAGTACAGATCCAGGTGGCCGCCTTCGCCACTGGTATTCTTCATAATATCTACGCATTCCACCGCTACACTATGAAT
TCTACCTCCCTCTATCGTACTCTAGTGAGTTAGTTTTGGATGCAGTTCCTAGGTTGAGCCCAGGGCTTTCACATCCAA
CTTAACAAACCACCTACGCGCGCTTTACGCCCAGTAATTCCGATTAACGCTTGCACCTTCCGT
43 GTCAGTACAGATCCAGGTGGCTGCCTTCGCCATTGGTATTCTTCACAATATCTACTGCATTCCACCGCTACACTGGG
AATTCTACCTCCCTCTATCGTACTCTAGTGAGTCAGTTTTGGATGCAGTTCCCAGGTTGAGCCCGGGGCATTTCACA
CCCAACTTAATAAACCACCTACGCGCGCTTTACGCCCAGTAATTCCGATTAACGCTTGCACCTTCCGT
44 CTCAGCGTCAGTATCGAGCCAGGTGGCCGCCTTCGCCACTGGTATTCCTCCATATATCTACCCATTTCACCGCTACA
CAGGGAAATTCTACCTCCCTCTATCGTACTCTAGTCAGACAGTTTTGGATGCAGTTCCCAGGTTGAGCCCGGGGCTT
TCACATCCAACTTAACAAACCACCTACGCGCGCTTTACGCCCAGTAATTCCGATTAACGCTTGCACCTTCCGT
45 ACAGTATCGAGCAGGTGGCCGCCTTCGCCACTGGTATTCCTCCCATATCTACGCATTTCACCGCTACACATGGAATT
CTACCTCCCTCTCTCGTACTCTAGTCAGCCAGTATTGGATGCAGTTCCCAGGTTGAGCCCGGGGCTTTCACATCCAA
CTTAACAAACCACCTACGCGCGCTTTACGCCCAGTAATTCCGATTAACGCTTGCACCTTCTGT
46 GTGGCCGCCTTCGCCACTGGTATTCCTCCATATATCTACGCATTTCACCGCTACACATGGAATTCTACCTCCCTCTAT
CGTACTCTAGTCAGCCAGTATCGGATGCAGTTCCCAGGTTGAGCCCGGGGCTTTCACATCCGACTTAACAAACCACC
TACGCGCGCTTTACGCCCAGTAATTCCGATTAACGCTTGCACCCTCCGT
47 ATCCAGGTGGCCGCCTTCGCCACTGGTATTCCTCCCATATCTACGCATTTCACCGCTACACATGGAATTCTACCTCC
CTCTATCGTACTCTAGTGAGCCAGTTTTGGATGCAGTTCCCAGGTTGAGCCCGGGGCTTTCACATCCAACTTAACAA
ACCACCTACGCGCGCTTTACGCCCAGTAATTCCGATTAACGCTTGCACCTTCCGT
48 ACGCGCGCGTAGGTGGTTTGTTAAAGTTGGATGTGAAGCCCTGGGTCTCAACCTAGGAACTGCATCCAATACTGACT
CACTAGAGTACGATAGAGGGAGGTAGAATTCATAGTGTAGCGGTGGAATGCGTAGATATTATGAAGAATACCAGTGG
CGAAGGCGGCCTCCTGGATCTGTACTGACACTGAGGTGCGAAAGCGTGGGTAGCGAACAGGATTAGATACCCTGGT
AGA
49 CTACTGGGCGTACGCGCGCGTAGGTGGTTGTTAAGTTGGATGTGAAGCCTGGGTCTCAACCTGGGAACTGCATCCA
GACTGACTCACTAGAGTACGATAGAGGGAGGTAGAATTCATAGTGTAGCGGTGGAATGCGTAGATATTGTGAAGAAT
ACCAGTGGCGAAGGCGGCCTCCTGGATCTGTACTGACACTGAAGTGCGAAAGCGTGGGTAGCGAACAGGATTAGAT
ACCCTGGTAGA
52 CTGGGCGTAAGCGCGCGTAGGCGGTTGTTAGTGTGATGTGAAGCACAGGGCTCAACCTGGGAACTGCATCACATAC
TGGCAAGCTAGAGTACGGTAGAGGGAGGTAGAATTCCATGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAATAC
CAGTGGCGAAGGCGGCCTCCTGGATCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATA
CCCCTGGTAGA
54 GATACTGGGCGTACGCGCGCGTAGGTGGTTGTATAGTCTAGATGTGAAGCCTGGGCTCAACCTAGGACTGCATCTA
GTACTGTCTCGACTAGAGTACTGAGTAGAGGGATGTAGAATCACATGTGTAGCGGTGCAATGCGTAGATATGAGGA
GCATCACTCAGTGCGAGGCGCCTACTGGGATCTGATACTGACACTGAGGTGCGAAAGCGTGGGGAGCAAACAGGA
TTAGATACCCCTGGTAGA
55 ACGCATCCGGACTGTCAGGCTACAGTGCAGGAGAGGAGGTAGACTTCCCGGTGTAGCGTTGAAATGCGTAGAGATC
GGGAGGAATACCAGTTGGCGAAGCGGGCCTCCTGGACTGACACTGACGCTGAAGGTGCGAACAGCGTGGGTAGCA
AAACAGGAATTAGATACCCCTGGTAGAA
57 ACGCTTTCGTCCATCAGTGTCAGTTGATTATTAGTAATCTGCCTTCGCAATTGGTATTCTATGTAATATCTATGCATTT
CACCGCTACACTACATATTCTAACTACTTCATAATAACTCAAGATAACCAGTTTCAAAGGCAATTTTACAGTTGAGCTG
CAAGATTTCACCTCTGACTTAATTATCCACCTACGGA
58 ACGGCGTGGACTACCAGGGTATCTAATCCTGTTCGCTCCCCACGCTTTCGTCCCTCAGCGTCAGTTACATGTTAGTA
AGCTGCCTTCGCAATCGGTGTTCTGAGTCATATCTATGCATTTCACCGCTACATGACTCATTCCGCCTACTTCATTTG
TACTCAAGGTCTCCAGTATCAATGGCAGTTCGACAGTTGAGCTGTCGGCTTTCACCACTGACTTAAAGACCCGCCTA
CGGA
59 TACGGCGTGGACTACCAGGGTATCTAATCCTGTTCGCTCCCCACGCTTTCGTCCCTCAGCGTCAGTTACATGTTAGT
AAGCTGCCTTCGCAATCGGTGTTCTGAGTCATATCTATGCATTTCACCGCTACATGACTCATTCCGCCTACTTCATTT
GTACTCAAGGTCTCCAGTATCAATGGCAGTTCGACAGTTGAGCTGTCGGCTTTCACCACTGACTTAAAGACCCGCCT
ACGGA

60 GTCCCTCAGCGTCAGTATACATGTTAGTAATCTGCCTTCGCAATCGGTGTTCTGAGTCATATCTATGTATTTCACCGC
TACATGACATCATTCCGCCTACTTCATTTGCACTCAAGGTCTCCAGTATCAATGGCAGTTTCGACAGTTGAGCTG

61 CTTATCACTTTCGCTTAGCCACTCAATCCGAAGACCGAACAGCTAGTATCCATCGTTTACGGCGTGGACTACCAGGG
TATCTAATCCTGTTCGCTCCCCACGCTTTCGTCCCTCAGCGTCAATATAGTGTTAGTGATCTGCCTTCGCAATTGGTA
TTCTGTGTAATATCTATGCATTTCACCGCTACACTACACATTCTAATCACTTCACACTAATTCAAGAAACCCAGTATCA
ATGGCAATTTTACAGTTGAGCTGTAAGATTTCACCACTGACTTAGATTCCCGCCTACGGA

62 CTTATCACTTTCGCTTAGCCACTCAATCCGAAGACCGAACAGCTAGTATCCATCGTTTACGGCGTGGACTACCAGGG
TATCTAATCCTGTTCGCTCCCCACGCTTTCGTCCCTCAGCGTCAATATAGTGTTAGTGATCTGCCTTCGCAATTGGTA
TTCTGTGTAATATCTATGCATTTCACCGCTACACTACACATTCTAATCACTTCACACTAATTCAAGAAACCCAGTATCA
ATGGCAATTTTACAGTTGAGCTGTAAGATTTCACCACTGACTTAGATTCCCGCCTACGGA

63 TTTACGGCGTGGACTAGCAGGGTATCTAATCCTGTTCGCTCGCACGCTTTCGTCCCTCAGCGTCAATATACATGTTA
GTGATCTGCCTTCGCAATCGGTGTTCTGAGTCATATCTATGCATTTCACCGCTACATGACTCATTCTAATTACTTCACA
CTAACTCAAGAAACCCAGTATCAATGGCAATTTTACAGTTGAGCTGTCGATTTCACCCACTGACTTAAATGACCCGCC
TACCGA

64 ATCGTTTACGGCGTGGACTACCAGGGTATCTAATCCTGTTCGCTACCCACGCTTTCGTCCATCAGTGTCAGTTGATTA
TTAGTGATCTGCCTTCGCAATTGGTATTCTATGTAATATCTATGCATTTCACCGCTACACTACATATTCTAACTACTTCA
TAATAACTCAAGATAACCAGTTTCAAAGGCAATTTTACAGTTGAGCTGCAAGATTTCACCTCTGACTTAATTATCCACC
TACGGA

68 GGTGAATCCTGCAGCTCACTGTAGAACTGCCTTTGATACTGGTTGTCTTGAGTTATTATGAAGTGGGTAGAATGTGTC
GTGTAGCGGTGAAATGCATAGATATTACTCAGAATACCGATTGCGAAGGCAGATTACTAATAATTTACTGACGCTGAG
GGACGAAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGTTACTAGCTGTT

69 AGTCAGTGGTGAATCTTACAGCTCAACTGTAGATTGCCATTGATACTGGGTGTCTTGAATTAGTGTGAAGTGATTAGA
ATGTGTAGTGTAGCGGTGAAATGCATAGATATTACTCAGAATACCAATTGCGAAGGCAGATCACTAACACTATATTGA
CGCTGAGGGACGAAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGATACTA
GCTGTTCGGTCTTCGGATTGAGTGGCTAAGCGAAAGTGATAAGTATCCCACCTGGGGAGTACGTTCGCAAGAATGA
AACTCAAAGGAATTGACGGGGA

70 TAGTCAGTGGTGATGCCGACAGCTCAACTGTCGATCTGCCATTGATACTGGATTCCTTGAGTTCTAGTGAAGTGAGT
AGAATGAGTCATGTAGCGGTGAAATGCATAGATATTACTCAGAACACCGATTGCGAAGGCAGATTACTAACATGATA
CTGACGCTGAGGGACGAAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGTC
ACTCGCTGTTGGCGATATACTGTGAGCGGCCAAGCGAAAGTATTAAGTGACCCACCTGGGGAGTACGCCCGCAACA
GTGAAACTCAAAGGAATTGACGGGGA

71 CTCAACTGTAGAATTGCCTTTGATACTGGTTGACTTGAGTTATTATGAAGTAGTTAGAATATGTAGTGTAGCGGTGAA
ATGCATAGATATTACTTAGAATACCGATTGCGAAGGCAGATTACTAATAATATACTGACGCTGATGGACGAAAGCGTG
GGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGATACTAGCTGTT

72 GGTGAATCTTTCAGGCTCTACTGTAGAACTGCCTTTGATACTGGACGTCTTGAATTCTTGTCGAAGTGGGTAGAATGT
GTCATGTAGCGGTGAAATGCATAGATATGACTTAGAACACCGATTGCGAAGGCAGATTACTAATATTTTAATGACGCT
GAGGGACGAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGATTACTCGCTGTT
GGAAATATACTGTCAGTGTCTAAGCGAAAGTGATAAGTAATCCACCTGGGGAGTACGATCGCAAGGTTGAAACTCAA
AGGAATTGACGGGG

73 GTGAATCCTGCAGCTCACTGTAGATTGCCTTTGATACTGGTTGTCTTGAGTTATTATGAAGTGGTTAGAATGTGTCGT
GTAGCGGTGAAATGCATAGATATTACACAGAATACCGATTGCGAAGGCAGATTACTAATAATATACTGACGCTGAGG
GACGAAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGATACTAGCTGTT

Table 4.4

1 CAGCAGAGCGCCTTCGCCACTGGTGTTCTTCCCGATATCTTACGCATTTCACCGCTACACCGGGAATTCCCTCTGCC
CCTACCACACTCTAGCCCAACAGTTTCCACTGCCATGATGGAGTTAAGCTCCACTTTTTAACAGCAGACTTGTTGGG
CCGCCTGCGGACGCTTTACGCCCAATAATTCCGGATAACGCTTGCCACTCCCGT

2 TCAGTGATGGTACAGCAGAGCGCTTTCGCAGCTGTGTTCTTCTCTGATATCTACGCATTTCACCGCTACACTGAGAAT
TCCTCTGACCCTGCCATACTCTAGTCTAGTAGTTTCCACTTGCTTTCACAGAGTTAAGCTCTGCTCTTTAACAACAGA
CTTACTAGACAACCTGACAGGCGCTTTACGCCCAATGATTCCGGATAACGCTTGCATCTCCCGT

3 TCATGAGCGTCAGTGATGGTACAGCAGAGCGCTTCGCCACTGGTGTTCTTCTCTGATACTCTACGCATTTCACCGCT
ACACCGGGAATTCCCTCTGCCCCTACCACACTCTAGCCCAGCAGTTTCCACTGCCATGTATGGAGTTAAGCTCCACT
TTTTAACAGCAGACTTGATTGGGCCGCCTGCGGACGCTTTACGCCCAATAATTCCGGATAACGCTTGCAACTCCCGT

4 CAGTAGAGCGCCTTCGCCACCTGGTGTTCTTTCTAATATCTCACGCATTTCACCGCTACACTAGAATTCCCTCTAGCC
CCTGTCATACTCAAGTCTAGTAGTTTCCATTGCTTTCCTAGGGTTAAGCCCTAGTCTTTAACAACAGACTTATTAGACA
ACCTACAGGCGCTTTACGCCCAATGATTCCGGATAACGCTTGCATCTCCCGT

5 AGAGCGCTTTCGCCACTGGTGTTCTTTCTAATATCTCACTGCATTTCACCGCTACACTAGAATTCCCTCTACCCCTGC
CATACTCTAGTCTAGTAGTTTCCATTGCTTTGTCTAGGGTTAAGCTCTAGTCTTTAACAACAGACTTATTAGACAACCT
ACAGGCGCTTTACGCCCAATGATTCCGGATAACGCTTGCATCCCCCGT

6 CAGCAGAGCGCCTTCGCCACTGGTGTTCTTCCCGATACTCTACGACATTTCACCGCTACACCGGGAATTCCCTCTGC
CCCTACCACACTCTAGCCCAACAGTTTCCACTGCCATGTATGGAGTTAAGCTCCACTTTTTAACAGCAGACTTGTTGG
GCCGCCTGCGGACGCTTTACGCCCAATAATTCCGGATAACGCTTGCCACTCCCGT

7 TCGCCATCTGGTGTTCTTCCGATCATCTACGTCATTTCACCGCTCACACTCGGAATTCCCTTCTAGCCCCCTACTCAT
CACTCTAGTCCAATCAGTTTCCACTGCCATGTATGAAGTTGAGCTCCACTATTTAACAGCTGACTTGTTGGGCCGCCT
GCGGACGCTTTACGCCCAATGATTCCGGATAACGCTTGCATCTCCCGT

8 GGATAGTAGATCGCCTTTCGCATCCTGGTGTTCTTCCGATAATCTACGTCATTTCACCGCTCACACTCGGGGAATTC
CCTTCTAGCCCTCCTACCATACTCTAGTCTAACAGTTTCCACTGCCATTATGGAGTTGAGCCCCACTATTTAACAGCT
GACTTATTGAGCCACCTGCGGACGCTTTTACGCCCAATGATTCCGGATAACGCTTGCCTCCTCCGT

9 GTCAGTGATGGTACAGTAGAGCGCTTTCGCCACCGGTGTTCTTTCTAATATCTACGCATTTCACCGCTACACTAGAAA
TTCCCTCTACCCCTGCCATACTCTAGTCTAGTAGTTTCCATTGCTTTTCTAGGGTTAAGCCCTAGTCTTTAACAACAGA
CTTATTAAACAACCTACAGGCGCTTTACGCCCAATGATTCCGGATAACGCTTGCATCCCCCGT

10 CAGCAGAGCGCCTTCGCCACTGGTGTTCTTCCTGAATATCTCACGCATTTCACCGCTACACCGGGAATTCCCTCTGC
CCCTACCACACTCTAGTCCAACAGTTTCCACTGCCGTGATGGAGTTGAGCTCCACTTTTTAACAGCAGACTTGTTGG
GCCGCCTGCGGACGCTTTACGCCCAATGATTCCGGATAACGCTTGCCACTCCCGT

11 AGCGCCTTCGCCACTGGTGTTCTTCCCGATATCTACGCATTTCACCGCTACACCGGGAATTCCCTCTGCCCCTACCA
CACTCTAGCCCAGCAGTTTCCACTGCCATGATGGAGTTAAGCTCCACTTTTTAACAGCAGACTTGTTGAGCCGCCTG
CGGACGCTTTACGCCCAATGATTCCGGATAACGCTTGCCTCTCCCGT

12 CAGCAGAGCGCCTTCGCCACTGAGTGTTCTTCCCGATATCTCACTGCATTTCACCGCTACACCGGGAATTCCCTCTG
CCCCTACCACACTCTAGCCCAACAGTTTCCACTGCCATGATGGAGTTAAGCTCCACTTTTTAACAGCAGACTTGTTGG
GCCGCCTGCGGACGCTTTACGCCCAATAATTCCGGATAACGCTTGCCACTCCCGT

13 ATCATTGGGCGTAAGCGTCTCGCAGCGGCCCAATAAGTCTGCTGTTAAGACTGGAGCTCAACTCTATCATGGCAGTG
GATCTGCTGGACTAGAGTGTGGTAGGGGTAGAGGGAATTCCTAGTGTAGCGGTGAGATGCGTAGATATCAGGAAGA
ACACCAGTGGCGAAGGCGCTCTGCTGGGCCATCACTGACACTCATAGACGACAGCTAGGGGAGCAAATGGGATTA
GATACCCCAGTAGTCA

14 TATTGGGCGTAAGCGTCCGCAGGCGGCCCAACAAGTCTGCTGTTAAAAAGTGGAGCTTAACTCCATCATGGCAGTG
GAAACTGTTGGGCTAGAGTGTGGTAGGGGCAGAGGGAATTCCCGGTGTAGCGGTGAAATGCGTAGATATCGGGAA
GAACACCAGTGGCGAAGGCGCTCTGCTGGGCCATCACTGACGCTCATGGACGAAAGCCAGGGGAGCGAAAGGGAT
TAGATACCCCAGTAGTCAACTTCGTTGACGCTATGGAAAGAGGCAACCAAACAAGAATTAATAATAAACCCCATATCA
CAAA

15 ATTATTGGGCGTAAGCGTCCGCAGGCGGCCCAACAAGTCTGCTGTTAAAAGTGGAGCTTAACTCCATCATGGCAGT
GGAAACTGTTGGGCTAGAGTGTGGTAGGGGCAGAGGGAATTCCCGGTGTAGCGGTGAAATGCGTAGATATCGGGA
AGAACACCAGTGGCGAAGGCGCTCTGCTGGGCCATCACTGACGCTCATGGACGAAAGCCAGGGGAGCGAAAGGGA
TTAGATACCCCAGTAGTCA

16 CATTATTGGGCGTAAGCGTCCGCAGGCGGCCCAACAAGTCTGCTGTTAAAAAGTGGAGCTTAACTCCATCATGGCA
GTGGAAACTGTTGGGCTAGAGTGTGGTAGGGGCAGAGGGAATTCCCGGTGTAGCGGTGAAATGCGTAGATATCGG
GAAGAACACCAGTGGCGAAGGCGCTCTGCTGGGCCATCACTGACGCTCATGGACGAAAGCCAGGGGAGCGAAAGG
GATTAGATACCCCAGTAGTCA

17 ATTATTGGGCGTAAGCGTCCGCAGGCGGCCCAACAAGTCTGCTGTTATAAAGTGGAGCTTAACTCCATCATGGCAGT
GGAACTGTTGGGCTAGAGTGTGGTAGGGGCAGAGGGAATTCCCGGTGTAGCGGTGAAATGCGTAGATATCGGGAA
GAACACCAGTGGCGAAGGCGCTCTGCTGGGCCATCACTGACGCTCATGGACGAAAGCCAGGGGAGCGAAAGGGAT
TAGATACCCCAGTAGTCA

18 ATATTGGGCGTAAGCGTCCGCAGGCGGCCCAACAAGTCTGCTGTTAAAAAGTGGAGCTTAACTCCATCATGGCAGT
GGAGACTGTTGGGCTAGAGTGTGGTAGGGGCAGAGGGAATTCCCGGTGTAGCGGTGAAATGCGTAGATATCGGGA
AGAACACCAGTGGCGAAGGCGCTCTGCTGGGCCATCACTGACGCTCATGGACGAAAGCCAGGGGAGCGAAAGGGA
TTAGATACCCCAGTAGTCA

19 ATTGGGCGTAAGCGTCTGCAGGCGGTCCATAAGTCTGCTGTTAAGACTGGAGCTTACTCCATCATGGCAGTGGATA
CTGCTGGACTAGAGTGTGGTAGGGGCAGAGGGAATTCCCGGTGTAGCGGTGAGATGCGTAGATATCGGGAAGAAC
ACCAGTGGCGAAGGCGCTCTGCTGGGCCATCACTGACACTCATAGACGAAAGCTAGGGGAGCGAAAGGGATTAGA
TACCCCAGTAGTCAC

20 AGTCTGCTGTTAAAGACTGGGGCTCAACCCCATCATGAGCAGTGGAAACTGCTAGACTTGAGTGTGGTAGAGGTAG
AGGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAAGAACACCAGTGGCGAAAGCGCTCTGCTGGGCCA
TCACTGACACTCATAGACGACAGCTAGGGGAGCAAATGGGATTAGATACACCAGTAGTCA

21 ATTGGGCGTAAGCGTCTGCAGGCTGCTTAATAAGTCTGCTGTTAAGACTGGAGCTCACTCTATCATGGCAGTGGAGA
CTGCTGGACTAGAGTGTGGTAGGGGCAGAGGGAATTCCCGGTGTAGCGGTGAGATGCGTAGATATCGGAAAGAGC
ACCAGTGGCGAAGGCGCTCTGCTGGACCATTCACTGACACTCATAGACGAAAGCTAGGGGAGCAAAAGGGATTAGA
TACCCCAGTAGTCA

22 TTGGGCGTAAGCGCCTGCAGGTTGTTTAATAAGTCTGTTGTTAAGACTGGGGCTTACCCTAGGAGGGCAATGGAACT
GCTAGACTTGAGTGTGATAGGGGTAGAGGGAATTTCTAGTGTAGCGGTGAAATGCGTAGATATTAGAAAGAACACCG
GTGGCGAAGGCGCTCTGCTGGATCATTACTGACACTCATAGGCGAAAGCTAGGGTAGCAAAAGGGATTAGATACCC
CAGTAGTCA

23 TTGGGCGTAAGCGCCTGCAGGTGTGTTCTAATAAGTCTGTTGTTAAGACTAGGGCTTAACCCTATGCATGAGCAATG
GACACTACTAGACTAGAGTATGGCAGGAGTAGAGGGAATTTCTAGTGTAGCGGTGATATGCGTAGATATTAGAAAGA
ACACCGGTGGCGAAAGCGCTCTACTGGACCATCACTGACACTCAGAGGACGAAAGCTAGGGTAGCAAAAGGGATTA
GATACCCCCAGTAGTCA

24 ATTGGGCGTAAGGGTTCGCAGGTGGTACTGCAAAGTCTGCTGTTAAAAATCGGAGCTCAACTCCGTAACCGGCGGT
GGAAAACTGCTCAGGCTAGAGTAGCTGGTAGGGGCAGAGGGAATTCCTGGTGTAGCGGTGAAATGCGTAGAGTCAT
CAGGGAAGAACACCGAGTGGCGAAGGCGCTCTGCTAGGAGCCTAGTTCACTGACAGCTGCATGGACGAAAGCTAG
GGGAGCGAATGGGATTAGATACCCCCTGTAGTCT

25 AGCCGCCTTCGCCACTGGTGTTCCTCCACATCATCTACGCAATTTCACCGTCTACACTTGGAATTCCACTCTCCTCTC
CTGCACTCAAGATTACCAGTTTCCAATGCACGTCCAGCGGTTGAGCCGTGCAGATTTCACATCAAACTTATATaAACC
GCCTGCGCGCGCTTTACGCCCAATAATTCCGaACAACGCTTGCCACCTACGT

26 GCGCCTTCGcCACTGGTGTTCTCCACATATCTACGCTAAATTACACCGTCTACACTTGGAATTCCACTCACCTCTCTC
TGCACTCAAGATTACCAGTTTCCAAGTGCACGTCCAGCGGTTGAGCCGTGAGATTTCACCTCAAACTTATATAATCCG
CCTACGTGCGCTTTACGCCCAGTAATTCCGAACAACGCTTGCCACCTACGT

27 CGCCTTCGCACTGGTGGCTCCAGATATCTAGCATATTCACACGTCTACACTTGGAATTCCACTTACCTCTCTCTGTCA
CTCAAGATTAGCAGTTTCAAATGAAGGTCCAGGGTTGAGCCCTGGGATTTCACCTCTAACTTAATAATACGCCTACGT
GCGCTTTACGCCCAGTAATTCCTGAACAACGCTTGCCACCTACGT

28 CTTCGcCACTGGTGTTCTCCATATCATCTGACGCTAATTTCACCGTCTACACTTGGAATTCCACTTACCTCTCTCGAAC
TCAAGACTAGCAGTATCCAATGACACCTCCACGGTTGAGCCCTGAGATTTCACCTCTGACTTACTAATACGCCTACGT
GCGCTTTACGCCCAGTAATTCCGGACAACGCTAGCCACCTACGT

29 ACCAGTGAGTCGCCTTCGCCACTGGTGTTCCTCCACATCATCTGACGTCATTTCACCGCTACACGTGGAATTCCGCT
CTCCTCTCTCTGCACTCAAGATCCCCAGTTTCCAATGACCCTCCACGGTTGAGCCGTGGGATATCACATCAGACTTA
ATGAACCGCCTGCGCGCGCTTTACGCCCAATAATTCCGAACAACGCTTGCCACCTACGT

30 ACCAGCTGAGTCGCCTTCGCCACTGGTGTTCCTCCACATCATCTACGCTATTACACACGTCTACACTTGGAATTCCA
CTCTCCTCTCTCTGCACTCAAGACTACCCAGTTTCCAATGACCCTCCACGGTTGAGCCGTGGGCATTTCACATCAGA
CTTAATAGACCGCCTGCGCGCGCTTTACGCCCAATAATTCCGGACAACGCTTGCCACCTACGT

31 GACCAGTGAGTCGCCTTCGCCACTGGTGTTCCTCCACATCATCTACGCATTTCACCGTCTACACGTGGAATTCCACT
CTCCTCTCCTGTACTCAAGTCTCGCAGTTTCCAATGACCCTCCGCGGTTGAGCCGTGCGATATCACATCAGACTTAA

AGAACCGCCTGCGCGCGCTTTACGCCCAATAATTCCGGACAACGCTTGCCACCTACGT

32 GCCAGTGAGTCGCCTTCGCCACTGGTGTTCCTCCACATCATCTGACGACTACTTTCACCGCTACACGTGGAATTCCA
GCTCTCCTCTCCTGCACTCAAGTTCACCAGTTTCCAATGGCCGTCTCGCGGTTGAGCCGCGAGATATCACATCAGAC
TTAAGGAACCGCCTGCGCGCGCTTTACGCCCAATAATTCCGGACAACGCTTGCCACCTACGT

33 CCAGCAGAGTCGCCTTCGCCACTGGTGTTCCTCCACATCATCTGACGCTATTTCACCGCTACACTTGGAATTCCACT
CACCTCTCCTGCACTCAAGATCACCAGTTTCCAATGAACGTCCACGGTTGAGCCGTGAGATATCACATCAGACTTAC
TGAACCGCCTGCGCGCGCTTTACGCCCAATAATTCCGGACAACGCTTGCCACCTACGT

34 CGCTTCGCACTGGTGTTCTCCATCATCTCTGACGCTATTTCACCGCTACACGTGGAATTCCAGCTTCTCCTCTCCTGC
ACTCAAGTTCTCCAGTTTCCAATGACCGTCTCGCATTGAGCCGTGAGCATTTCACATCAGACTTAAGAGAACCGCCT
AGCGCGCGCTTTACGCCCAATAATTCCGGACAACGCTTGCCACCTACGT

35 CTTCGCACTGGTGTCTCCATCATCTCTGACGCATTTCACCGCTACACAGTGGATTCCAGCTCTCCTCTCCTGCACTC
AAGTTCTCCAGTTTCCAATGACCGTCTCGCGGTTGAGCCGTGAGCATTTCACATCAGACTTAAGAGAACCGCCTGCG
CGCGCTTTACGCCCAATAATTCCGGACAACGCTTGCCACCTACGT

36 ACCAGTGAGTCGCCTTCGCCACTGGTGTTCCTCCACATATCTACGCATTTCACCGCTACACTTGGAATTCCACTCTC
CTCTTCTGTACTCAAGTCTCCCAGTTTCCAATGACCCTCCACGGTTGAGCCGTGGGCTTTCACATCAGACTTAATAGA
CCGCCTGCGCGCGCTTTACGCCCAATAATTCCGGACAACGCTTGCCACCTACGT

38 AGTCTGATGTGAAGCGTGCTGTCTCAGCCATAGTACGCTTTGGTAACTGTCTAACTTGAGTGCAGAAGATGAGAGTG
GCTTCCATGTGTAGCGGTGAAATGCGTAGATATATGGAGGAACACCGGTGGCGAAGAGCGGCTCTCTGGCCTGGTA
ACTGACGCTGAAGGCCTCGA

39 TGAGTCTGATGTGAAGCCGCTGTCTCAGCGATAGGCGCTCTGTACTGTCTGCTGAGTGCAGTAGAGAGAGTGATCA
TGTGTAGCGCTGCATGCGTAGAGTATATGAGACACCGGTGGCGGAGAGGGCCTGCTCTGGCTTGTAACTGACGCTG
GAGGCTCGA

40 AGTCTGATGTGAGCCGCGTCTCAGCGTAGTACGCATGTACTGTTACTGAGTGCAGTAGAGAGAGTGCATCATGTGTA
GCGCTGATGCGTAGATATATGAGACACCGGTGGCGAAGCGCGGCTCTCTGGCCTGTAACTGACGCTGAGGCC

41 TAGTCTGATGTGCTGCCTGCGTCTCAGCGTAGTACGCTTGTAGCTGCTACTGAGTGCAGTAGAGAGAGTGCATCATG
TGTAGCGCTGCATGCGTAGATATATGAGTACACAGGTGGCGGAGGCGGCTCTCTGGTCTGTATCTGACGCTG

42 GCGTGCGCGAGCGTAGCTGTCTCGATAGTCTGATGTGAAGCTTGCTGTCTCAGCGATAGTACGCACTGTAACTGTCT
GCTGAGTGCAGAGAGAGAGTGATCATGTGTAGCGTGATGCGTAGATATATGAGACACAGTGGCGAAGAGCGCTCTC
TGGCTTGTAACTGACGCTGAGCTCGAAAAGCGTGGGGGAGCACCAGGATTTAGATCCCCTTG

50 TAGGGCTAGGACTACCGGGGTATCTAATCCCGTTCGCTCCCCTAGCTTTCGCGCCTCAGCGTCAGAAAAGATCCAG
CACACCGCTTTCGCCACCGGCGTTCCTTCCGATCTCTACGCATTTCACCGCTCCACCGGAAGTTCCGTGTGCCCCTA
TCTCTCTCAAGACAAATAGTTTCGAGGGCACTTCTCTAGTTAAGCTAGAGGCTTTCACCCTCGACTTATTCACCCGCC
TACGCGCCCTTTAAGCCCAGTGATTCCGAATAACGTTCGCACGGTTCGT

51 CAACGTTTAGGGCCAGGACTACCGGGGTATCTAATCCCGTTCGCTCCCCTGGCTTTCGTGCCTCAGCGTCAGTGAG
GGTCCAGCGTGCCGCCTTCGCCACCGGAGTTCCTAACGATATCAACGCATTTCACCGCTCCACCGTTAGTTCCGCA
CGCCCCTACCCCCCTCGAGACCGCCAGTATTCAAGGCACTTCCCCGGTTGAGCCGGAGGATTTCACCTCAAACTTG
GCAGCCCGCCTACGCACCCTTTAAGCCCAGTGATTCCGAACAACGTTCGCTCGGTTCGT

52 TAGTCCGTTTAGCTGTGTCACCTCACAAGTATACTTGCTGACGTCTGGCATTCATCGTTTACGGTGTGGACTACCAG
GGTATCTAATCCTGTTTGCTCCCCACACTTTCGCACCTCAGCGTCAGTATCGAGCCAGTAAGCCGCCTTCGCCACTG
GTGTTCCTCCGAATATCTACGAATTTCACCTCTACACTCGGAATTCCACTTACCTCTCTCGAACTCAAGACTAGCAGT
ATTAAAGGCAGTTCCAGGGTTGAGCCCTGGGATTTCACCTCTAACTTACTAATCCGCCTACGTGCGCTTTACGCCCA
GTAATTCCGAACAACGCTAACCCCCTCCGT

53 GGACTACCGGGGTATCTAATCCCGTTCGCTCCCCTGGCTTTCGTGCCTTAGCGTCAGAAAAGTCCCAGTAAGCCGC
TTTCGCCACCGGGTGTTCCTGATAATATCAACGCATTTCACCGCTCCACCATCAGTTCCGCTTACCTCTGACTCCCTC
AAGCAATATGGTTTCAAGCGCAGTTCCACGGTTGAGCCGTGGTATTTCACACCTGACCTATATCGCCGCCTACGCAC
CTTTTAAGCCCAGTGATTCCAAATAACGTTCGCCCAGTTCGT

56 TTACGCTTTCGCTCCGGGCGGAAGAGGCGAACCTCCTCCCCCCTAGTACCCATCGTTTACGGCCAGGACTACCGGG
GTATCTAATCCCGTTCGCTCCCCTGGCTTTCGTGCCTCAGCGTCAGTCGAGCCCCAGTCTGTCGCTTTCGCCACCG
GCGTTCCTTCCGATATCAACACATTTCACCGCTCCACCGGAAGTTCCACAGACCCCTAGCTGACTCAAGCCGGCCG
GTTTCGAACGCAATGCCTCAGTTGAGCTGAAGCCTTTCACATCCGACCTGGTCGGCCGCCTACGCACCCTGTAAGC
CCAGTGATTCCGAATAACGTTCGCCCAGTACGT

61 ATTATTGGGCGTAAGCGCGCGCAGGCGGTTCCTTAAGTCTGATGTGAAATCTCGCGGCTCAACCGCGAGCGGCCAT
TGGAAACTGGGGAACTTGAGTGCAGGAGAGGGGAGCGGAATTCCACGTGTAGCGGTGAAATGCGTAGAGATGTGG
AGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGTAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCGAACAG
GATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAAGTGTTAGAGGGTATCCACCCTTTAGTGCTGCAG
CAAACGCATTAAGCACTCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGACGGGG

62 ATTATTGGGCGTAAGCGCGCGCAGGCGGTTCCTTAAGTCTGATGTGAAATCCCGCGGCTCAACCGCGAACGGCCAT
TGGAAACTGGGGAACTTGAGTGCAGGAGAGGGGAGCGGAATTCCACGTGTAGCGGTGAAATGCGTAGAGATGTGG
AGGAACACCAGTGGCGAAGGCGGCTCTCTGGTCTGTAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCGAACAG
GATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAAGTGTTAG

63 GGTCCGTAGGCGGATTTGTAGTCAGAGGTGAATCCTGCGGCTCAACTGTGGATTGCCTTTGATACTGGTTGTCTTGA
GTTATTATGAAGTGGGTGGAATGCGTCGTGTAGCGGTGAAATGCATAGATATTACATAGAACACCAATTGCGAAGGC
AGATCACTAATGATTTACTGACGCTGATGGACGAAAGCGTGGGTAGCGAACAGGATTAGATACCCTGGTAGTCCACG
CCGTAAACGATG

64 CGTAGGCGGTATTATTAAGTCTAGAGGTGAATCCCACGGCTCAACTGTGGAACTGCCTTTGATACTGGTAGTCTTGA
GTTCGAGAGAGGTGAGTGGAATTCCTAGTGTAGCGGTGAAATGCATAGATATTACATGGAACACCAATTGCGAAGGC
AGATCACTGGCTCGTACTGACGCTGAGGTACGAAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACG
CCGTAAACGATG

65 CCGTAGGCGGATTTAGTCAGAGGTGAATCTTGCAGCTCTAACTGTGGATTGCCTTTGATACTGGATGTCTTGAGTTC
TTATGAAGTGGGTGGAATGTGTCGTGTAGCGGTGAAATGCATAGATATTACATAGAACACCAATTGCGAAGGCAGAT
CACTAATACTGTACTGACGCTGAGGGACGAAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCG
TAACGATGATTACTAGATGTTGG

66 TACTGGGCGTAAGCGCACGTAGGCGGATTAGTAAGTTAGAGGTGAAATCCCAGGGCTCAACCCTGGAACTGCCTTT

AATACTGCTAGTCTTGAGTTCGAGAGAGGTAAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGA
ACACCAGTGGCGAAGGCGGCTTACTGGCTCGATACTGACGCTGAGGTGCGAAAGTGTGGGGAGCAAACAGGATTA
GATACCCTGGTAGTCCACACCGTAAACGATGAATGCCAGACGTCAGCAAGTATACTTGTTGGTGTCACACCTAACGG
ATTAAGCATTCCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTTGACGGGG

67 TCCGTAGGCGGATCATTAAGTCTGTGGTGAAATCCCACGGCTCAACTGTGGAACTGCCTTTGATACTGCTGACCTTG
AGTACGAAAGAGGTAAGTGGAATTCCTAGTGTAGAGGTGAAATGCGTAGATATTACTCGGAACACCGATGGCGAAG
GCGGCTTACTGGCATGTATACTGACGCTGAGGTACGAAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTC
CACGCCGTAAACGAT

68 ATTATTGGGCGTAAGCGCGCGCAGGCGGTTCCTTAAGTCTGATGTGAAATCTCGCGGCTCAACCGCGAGCGGCCAT
TGGAAACTGGGGAACTTGAGTGCAGGAGAGGGGAGCGGAATTCCACGTGTAGCGGTGAAATGCGTAGAGATGTGG
AGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGTAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCGAACAG
GATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAAGTGTTAGAGGGTATCCACCCTTTAGTGCTGCAG
CAAACGCATTAAGCACTCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGATTGACGGGG

69 CTGCGCGTGCAGGAGACTGCCTATGGGTGTAACTGCTAAACAGGGAAGAAACCCCCCCACGAGTACAGCTTGCTGT
ACTGTAGAATACAGAAGGCTAACTTATGCCAGCACCGCGGTAATAAATA

70 GGAGAAGCTAGGTGAACTGCATTAAAGAGGGAAAAAAACGCCTACCTGTAAAGCTTGTTCTTTGGACTGTTCGAATA
AAGATCGGCTAACTCCATGCCGCAGCCGCGATATA

72 TGAGTGATGAAGCTAGGGTGTAAAGCTGCATAACAGAGGAGAAAATGACGGTACCTGGAAAAGAAACCCCTTGATAA
CTCCGAGCCAGCAGCCGCGATAATA

76 TGCGCGTGTGTGATGATGTCTAGGGTCGTAAGCACCGATCAGAGATGATAATGACAGTATCTGGTAAAGAAACCCCG
GCTAACTCCGTGCCAGCAGCCGCGATAAT

77 GAGTGAGAAGCTAGGCGTAAGCGCTAAAGAGATGATAATGACTGTACCTGGTAAGATTCCCTTTGATAACTCCAAGC
CAGCACCCGTAATTATATGCCACACCGCGATAATA

79 ATGCGCGTGCAGTGATGACTGCTATGGTGTAAGCTGCATAAAGAGAGAGAAACTTATCTACGTGTACTTCTTGACTG
TACTGTAAGAATAAGGATCGGCTAACTCCATGCCTGCAGCCGCGAT

80 CTGATCTAGTCATGTCGCGTGCAGTGATGACTGCCCTATGGCGTAAGCTCCTTTGATAGAGAGAACGTACTGTACGC
TGTAGAAGATATGACGGGACTGTACCAATACGGATCGGCTGACTCTATGCCATCAGCCGCGCATAATATAAT

81 CATGCGCGTGCAGTGATAGACTGCCCTATGGCGTAAGCTGCTTTGATAGAGGAAAGAAACGTCTGTACCTGTCCAG
AGTTGCACCGTACTGTACCAATAAGGATCGCCCTAACTCCATGCCAGCAGCCGCGCATAATA

82 TGCCGCGTGTGTGATGATTGCTTCAGGGTCGTAACGCTCTTTCGATAGGAGATGATAATGACAGGTATCTGGCATGA
GAAACCCCGGCTAACTCCGCTGCCATGCAGCCGCGATAATA

83 CATGCGCGTGCAGTGATGATGCTATGGGCGTAAGCTGCAATAAAAGGAGGAAACGTCTGTACCTGTGAAGATTGTCT
TGTATAATTCCAATAAGGACCGGGTAACTCTATGCCAGCAGCCGCGGTAATA

84 CGTGAGTGATGAGGCCTTAGGGTCGTAAAGCTCTCTCAACAGGGGGAGAAATGACTGTACCTGATAAGAAATTTCG
GAGTAACTCCGTGCCAGCAGCCGCGATTATA

85 GCGCGTGCAGTGAGACGCTATGGTGTAAGCTGCCTAAGAGAGGAGGAAATGTTTTTACCTGTGAAAAAATGCCCTG
GACAATAACAAGAAAGAACGGGCAACTATATGCCAGCGCCGCGGTAATA

86 TGCGCGTGCAGTGATGAATGCCTATGGTGTAAGCTCTTGAAGAGGAGAAATGACTGCACCTGTGAAAGACTGACCG
GACTATATCAATAAGGACCGGGTGACTCTATGCCAGCAGCCGCGATAATa

87 CGCGTGTGTGATGAAGGTCTTAGGGTTGTAAAGCTCTTTCGCTTGAAATGATAATGACTGTATCTGGTAAAGAAAGC
CCGGCTTACTTCGTGCCAGCAGCCGCGATAATAT

88 CTGTATGCTAGCCATGCCGCGTGTGTGATGAGGCCTTAGGGTTGTAAAGCTCTTTCGTCAGAGATGATAATGACTGT
ACCCGAGAAGAAAGTCCGGCTAACTTCGTGCCAGCAGCCGCGATAAT

89 CTGATCAGCATGCCGCGTGCAGTGAGACTGCCCTATGGGTGTAAACTGCTTCTATATAGGAAGAAACACCCCTACGT
GTAGGGGCTTGACGGTACTGTAAGAATAAGGATCGGCTAACTCCGTGCCAGCAGCCGCGATAATA

90 TGCGCGTGCAGTGATGACATGCCTATGGCTGTAAAGCTGCATTAACACAGGAAAGATAACTGCACTGTACCTGTAAA
AGCATTGCCCGGACTAATTCCAATAAAGCACCCCTAACTCTATGCCAGCAGCCGCGATAATA

91 TGCGCGTGCAGTGATGACTGCCTATGGCGAAAGCTGCTCAATAGAGAAGGAAACGACTGCACGTGTGAAAGAATGA
CGGGACAATAACAAGAAAGAACGGCTAACTCTATGCCAGCAGCCGCGATAAT

92 CTGATCTAGCCATGCCGCGTGTGTGATGATGCCCTAGGGTCGTAAGCTCTTTCGACTGAGAGATAATGACTGTACCT
GGAGAAGAAAGCCCGGCTAACTCCGTGCCAGCAGCCGCGATAA

93 CCCTAGGGGGTAGGAAAGCTCTTTCCCTTTGAGGAAGGAAAATGACCGTACCGAAAAAGAACCCCGGCCTTTACTC
CGTGCCGCCGCCCCGCGAATAATGATA

94 AGCATGCCGCGTGCAGTGATGACTGTCCTAGGGCGTAAGCTGCTTCTCATAGGAGGAAACGATTGTACGTGTACGG
GCTTGACGGGACTGTTCGAATAAGGAACGGCTAACTCTATGCCAGCAGCCGCGATAATA

95 TGCCGCGTGAGTGATGACTTGCCTCGGGTCGTAAAGCACTTTCGACTAGAGATGATAATGACAGTATCTGGAAAAGA
AACCCCGGCTAACTCCGTGCCAGCAGCCGCGATAAT

96 CTGATCTAGCCATGCCGCGTGTGAGATGAATGCCTAGGGTCGGAAAGCACTTTCGCCATGAGAAGATAATGACAGT
ATCTGGAAAAGAAACCCCGGCTAACTCCGTGCCAGCAGCCGCGATAATA

97 TGCCGCGTGCGTGAGACTGCCCCTATGGCTGAAAGCTGCTTCGATAGAAGGAAAGAAACATACCTACGAGTACGAG
CTTGACGGTACTGTAAGAATAAGGACCGGCTAACTCCATGCGGCAGCCGCGATCTA

98 TGCCGCGTGTGTGATGATGCCTAGGGTGTAAGCTCTTTCGAAGAGAAGATAATGACTGTACCTGGAAAAGAAAGCCC
GGCTAACTCCGTGCCAGCAGCCGCGATTATA

99 TGCCGCGTGTGTGATGAGTGTCCTAGGGTCGTAAAGCACTTTCGCATCAGAGATAGATAATGACTGTATCTGGTAAA
GAAACCCCGGCTAACTCTCGTGCCAGCAGCCGCGATAA

100 TGCCGCGTGAGTGATGAGTGTCCTCTAGGGTTCGTAAAGCATCTTTCAGAATCAGAGATAGATAATGACTGTATCTG
AATAAGAAGACGCTCCGGTTAACTCTCGTGCCAGCAGCCGCGATAA

101 TGCCGCGTGCAGTGAGACTGCCCTATGGCTGTAAACTGCTTCCTCTTGAGGAAGAAACATCCGTACGTGTACGGAG
ATGACGGTACTGTACGAATAAGGACCGGCTAACTCCATGCCAGCAGCCGCGATAATA

102 CCATGCCGCGTGCAGTGATGACTGCCCTATGGTTGTAAACTGCTTCTATATAGGAAGAAACGTCCTTACGTGTAAGG
GGATGACGGTACTGTAAGAATAAGGACCGGCTAACTCCATGCCAGCAGCCGCGATAATA

103 TGCGCGTGAGTGATGTAGGTCTTAGGGTCGTAAAGCTCTTTCCATGAGATGATAATGACAGTACCTGGTAAAGAAAC
CCCGGCTAACTCCGTGCCAGCAGCCGCGATAA

104 TGCCGCGTGTGTGATGATTGCTTCAGGGTCGTAACGCTCTTTCGATAGGAGATGATAATGACAGGTATCTGGCATGA
GAAACCCCGGCTAACTCCGCTGCCATGCAGCCGCGATAATA
105 TGTAGTGCGAAGTGCCTTCGGGTTGTAAACCTCTTCCGTAGAAGAAGATACTGACGTGACGGGACCTGCAAACAAC
GCGCCGGCTAGCTACGTGCCAGCAGCCGC
106 CTGATCAGCCATGCCGCGTGCAGGAGACTGCCCTATGGGTTGTAAACTGCTACATTACAGGAAGAAACCTCTGTAC
GTGTCCGGAGCTGACGTTACTGTACGAATAAGGATCGGCTAACTCCATGCCAGCAGCCGCGATAA
107 TGCCGCGTGAGTGATGATGCCTTAGGGTCGGAAAGCTCTTTCGCTAGGGGATGAAAATGACAGTACCTGGTAAAGA
AACCCCGGTTAACTCCGAGCCAGCAGCCGCGACTAA
108 CTGAGCTAGCACTGCGCGTGAGCGATGATGTCCTCGGAGCTAGAAGCACTGTCGAATAGGGAAGAACAATGACCGT
TCGAACAGGGCGGCACCTTGACGGTACGTGACGAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGATATTA
109 CGCGCGTGCGCGATGACTGCCTCGGGCTGTAAGCCGCGACCATCAAGGGAAGCGCAAGTGACCGAACCTGCAGGA
ACAATTGCCGGTTACTTACGTGCCAGCCGCCGCTATATTAGTGCccagcagccgcgataata
110 TGCCGCGTGCAGTGATGATGCCCTAGGGTGGAAAGCTCTTTCAATAGAGAGGAAAATGACTGTACCTGTACCAAGAA
TGACCGGATAATTCCATGACAGCTCGCGCGACTCTATTCCGTGCGCCGCGCCGCTATAATA
111 CCATGCCGCGTGCAGTGATGACTGCCCTATGGGTTGTAAACTGCTTCAATACAGGAAGAAACCTCTGTACGTGTACA
GAGCTGACGGTACTGTACGAATAAGGATCGGCTAACTCCATGCCAGCAGCCGCGATAATA
112 CATGCCGCGTGCAGTGATGACTGCCCTATGGCTGTAAACTGCTTCTATAGGGAAGAAACATCTCTACGTGTACAAGG
CTTGACGGTACTGTACGAATAAGGATCGGCTAACTCCATGCCAGCAGCCGCGATAATATAATa
113 TGCCGCGTGCAGTGATGACTGCCCTATGGTTGTAAACTGCTTCAATACAGGAAGAAACACACTCTACGTGTAAAGGC
TTGACGGTACTGTACGAATAAGGATCGGCTAACTCCATGCCAGCAGCCGCGATAATA
114 CCTCTGATCTAGCCATGCCGCGTGCAGTGATGACTGCCCTATGGGTTGTAAACTGCTCCATAACAGGAAGAAACACC
TCTACGTGTAGAGGCTTGACGGTACTGTAACAATAAGGATCGGCTAACTCCATGCCAGCAGCCGCGATAATA
115 TGCCGCGTGCAGTGATGAATGCTAGGGTGTAAGCTCTTTCGTCAGAGGTGAAATGACTGTACCTGTCAAAGAATGAC
CGGCTAACTCCGAGCCAGCTCCCGCGACTCTATGCCAGCAGCCGCGATAATA
116 TGCTGCGTGAGTGATGAGGCCTTTGGGCGTAAAGCTCTTTCGTCGGAGAGATAATGACTGTACCCGAATAAGAAGG
TCCGGCTAACTTCGCTGCCAGCAGCCGCGATAATATATA
117 CTGATCCAGCCATGCCGCGTGCAGTGATGACTGCCCTATGGGCTGTAAACTGCTCAATCACAGGAAGAAACACACT
CTACCTGTAGGGGCTTGACGGGACTGCAAGAATAAGGACCGGCTAACTCCATGCCAGCAGCCGCGATAATA
118 TGCCGCGTGCAGGATGACTGCCCTATGGGCTGTAAACTGCTTCAATACAGGAAGAAACACCTCTACGTGTAAAGGCT
TGACGGTACTGTAAGAATAAGGATCGGCTAACTCCATGCCAGCAGCCGCGATAATA
119 CGCGTGAGTGATGATGTCTTAGGGTCGTAAAGCTCTCAATAAGAGGATGATAATGACAGTACCTGGTAAAGAAACCC
CGCGCTAACTCCGCTGCCAGCAGCCGCGATAATATATA
120 TGCCGCGTGAGTGATGAAGGCCTTCAGGGTCGTAAAGCTCTTCTCACTCAGAGGATGATAATGACAGGTACCTGGT
AAAGAAACCCCGCGCTAACTCCGCTGCCAGCAGCCGCGACTAATATAATA
121 CATGCCGCGTGTGTGATGATGCCTAGGGTGTAAGCTCTTCCATCAGAGAAGATAATGACGGTACCTGGAAAAGAAA
GCCCGGCTAACTTCGCGCCAGCACCCGCGATAATATGCTA
122 ACGCCGCGTGCGCGGATGACGGCCTTCGGGCTGTAAACCGCGTTCATCGCGGACGAAGCGTACCTGACGGTACCT
GCAGAACATTGCCGGCTCTTACGTGCCAGCCGCCGCTATATTAGTGCCAGCAGCCGCGATAATA
123 CTGATGCAGCGATGCCGCTTGTAGTGAGAGGCCTTTGGGTTGTAAAGCTCTTTCGTCGGGGAGAAATGACTGTACC
CGAATAAGAAGGTCCGGCTAACTTCGTGCCAGCAGCCGCGATAATATA

Table 5.3

1 TGCCGCGTGTGTGATGATGCCTTTGGGCTGTAAAGCTCTTTCGTAAGGGAAGAAAATGACTGTACCCGAATAAGAAG
GTCCGGCTAACTTCGTGCCAGCAGCCGCGATAATATA
2 TGCCGCGTGAGTGATGAATGGCCTTTGGGTGTAAAGCTCCTCCGTACAGAGAGGAAGAAAATGACTGTACCCGAAT
AAGAAGGTCCGGTTAACTTCGTGCCAGCAGCCGCGATAATATA
3 GCCGCGTGAGTGACGAAGGCCTCTGGGCTGTAAACCTCTTTCCTCAAGGAAGATGATGACTGTACACTTGAAGAATA
GGCCGCCTATATTTCCGCGCCCGCCCCCGAGAATATA
4 ATGCCGCGTGAGTGATGATGGCCTTTGGGCTGTAAGCTCTTCCGTCGGGGAAGAAAATGACTGTACCCGAATAAGA
AGGTCCGGTTAACTTCGTGCCAGCAGCCGCGATAATATa
5 CGATGCCGCGTGAGTGACGATGGCCTTTGGGCTGTAAAGCTCTTCCGTCGGGGAAGAAAATGACTGTACCCGAATA
AGAAGGTCCGGTTAACTTCGTGCCAGCAGCCGCGATAATATA
6 ATGCAGCGATGCCGCGTGAGTGATGAAGGCCTTTGGGCTGTAAAGCTCTTCTCGTCGAGGGAAGAGATGACTGTAC
CCGAATAAGAAGGTCCGGCTAACTTCGTGCCAGCAGCCGCGATAATATA
7 CTGATGCAGCGATGCCGCGTGAGTGATGAAGGCCTTTGGGTTGTAAGCTCTTCTCGTCGGGAAGAGATGACTGTAC
CCGAATAAGAAGGTCCGGCTAACTTCGTGCCAGCAGCCGCGATAATATA
8 ACTCTGATGCAGCGATGCCGCGTGAGTGAGATGCTTTGGCTGTAAGCTCTTTCGTCGGGAGAGATGACTGTACCGA
ATAAGATGTCCGGCTAACTTCGTGCCAGCAGCCGCGATAATATA
9 TGCCGCGTGAGTGATGAGGCTTTGGGCTGTAAGCTCTCCGTACGGGAGAGATGACTGTACACGTGTAAGATGGTCC
GGCTAACTTCGTGCCAGCAGCCGCGATAATATA
11 ATGCCGCGTGAGTGATGAAGGCCTTTGGGCTGTAAAGCTCTTTCCTCGGGAGATAATGACTGTACCCGAATAAGAAG
GTCCGGCTAACTTCGTGCCAGCAGCCGCGATAATATA
13 GTGATGATGATCTAGGGTGTAACGCATCTCAGTCAGGAGAGAAATGACGTGTACCTCGCACTAAGAGCCTCCGGCT
AATTTCTTGCCACCAGCCGCGATATAATA
15 CTGATGCAGCTATGCCGCGTGAGTGATGATGGCCTTAGGGCTGTAAAGCTCTTCTCGTCGGAGAGATAATGACTGTA
CCCGAATAAGAAGGTCCGGTTAACTTCGTGCCAGCAGCCGCGATAATATA

Table 5.4

1 TGATCAGCATGCGCGTGTGTGATGAGGCTTAGGTGTAAAGCACTTTCACTGTGAAGATGATGACGGTAACCAGAGAA
GAGCCCCGGCTAACTTCGTGCCAGCAGCCGCGATAAT
2 TGCATGCAGCTATGCGCGTGTGTGATGCAGCTCAGACTGTAAAGCTACTCTCACTGGTGAGATGATGACTGTAACCG
AGAGAGAGCTCCGGCTAACTTCGTGCCAGCAGCCGCGAT
3 TGATCTAGCGATGCGCGTGTGTGATGAGCTAGACTGTAAGACTCTGTCAGTCAGGAGATATGACTGTACCTGAAGAG
AGATCCGGCTAACTCCGTGCCAGCAGCCGCGATAA

4 ACTGATCAGCGATGCGCGTGAGTGATGAGGCTTAGGGTGTAAACTCTTCGTCAGGGAAGATAATGACGGTACCTGA
AGAAGAGATCCGGCTAACTCCGTGCCAGCAGCCGCGATAATA

5 ATGAGGCCTTAGGGTTGTAAAACTCTTTCGTCAGGGAAGATAATGACGGTACCTGAAGAAGAAGATCCGGCTAACTC
CGTGCCAGCAGCCGTGATAAT

6 AGTTGATGCAGCTATGCTGCGTGTGTGATGAGCTAGGCTGTAAGCTACTCTCATCTAGGTGAGATGATGACTGTACC
TGCAGAGAGCTCCGGCTAACTTCGTGCCAGCAGCCGCGATAT

7 ACCTGATGCAGCACTGCCGCGTGAGTGATGAGCTTTAGGCCTGTAACGCTCTTTCATCTGGTGAAGATAATGACTGT
ACACTGGAGGAGTAGCCCCGGCTAACTTCGTGCCAGCAGCCGCGATAATA

8 CTGATGCAGCTATGCTGCGTGTGTGATGATGACTTAGATGTAAAGCTCTCTCATACTGTGAGTATGACGTGTAGCTG
TATAGAGCTCTGTCTAACTCGTGCAGCAGCGCGAT

9 CTGCTGCAGCTATGCGCGTGTGTGATGAGCTTAGGCTGTAAGCTCTCTCATCGGTGAGATGATGACTGTACCCGGA
GAGAGCTCCGGCTAACTCTCGTGCCAGCAGCCGCGATAA

10 ACTCTGCATGCTAGCGATGCTGCGTGTGTGATGCAGCTAGGACTGTAAAGCTCTCTCTCAGTGAGATGATGACTGTA
CCTGTAGAGAGCTCCGGCTAACTCGTGCCAGCAGCCGCGATA

11 CTGCATGCAGCTATGCTGCGTGTGTGATGCAGACTCAGGCTGTAAGCTACTCTCATCTAGGTGAGATGATGACTGTA
CTGAGAGAGAGCTCCGGCTAACTACGTGCCAGCAGCCGCGAT

12 TCTGATGCAGCTATGCCGCGTGTGTGATGAGGCATAGGCTGTAAGCTCTTTCATCTGGTGAGATGATGACTGTACCT
GGAGAGAGCTCCGGCTAACTCTCGTGCCAGCAGCCGCGATA

13 CTGATGCTAGCTATGCTGCGTGTGTGATGAGACATAGGATGTAAGCTCTCTCGTCAGTGAGATAATGACTGTACCTG
GAGAGAGCTCCGGCTAACTCCGTGCCAGCAGCCGCGATAA

14 TGCATCCAGCTATGCCGCGTGTGTGATGAAGGCCTTCAGGGTTGTAAAGCACTTTCACTGGTGAAGATGATGACGGT
AACCAGAGAAGAAGCCCCGGCTAACTCTCGTGCCAGCAGCCGCGATAAT

15 TGATGCAGCACTGCAGCGTGAGTGATGACCTCTAGCCTGTAACGCTCTCTCGTCAGGAGAGAATGACTGTACTTGA
GAGTAGTCAGGCTAACTTCGTGCCAGCCGGCGCGGTA

16 ACTGATGCAGCATGCTGCGTGTGTGATGAGCTAGACTGTAAGCTACTTTCATCTGTGACGATGATGACTGTAGCTGC
AGAGAGCTCCGGCTAACTCGTGCCAGCAGCCGCGAT

Table 5.5

1 ATTACTGGGCGTAAAGCGTGCCGCAGGGCGGTTGTGAAGACAGATGTGAAATCCCTGGGCTTAACCTAGGAATTGC
ATTTGTGACTACATAACTAGAGTGTGTCAGAGGGAGGTAGAATTCCACGTGTAGCAGTGAAATGCGTA

2 ATACTGGGCGTAAGCGTGCGCAGGCGGTTGTGTAAGACAGATGTGAATCCCTGGGCTTAACCTAGGAACTGCATTT
GTGACTGCATGACTAGAGTGTGTCAGAGGGAGGTAGAATTCCACGTGTAGCAGTGAAATGCGTA

3 ATACTGGGCGTAAGCGTGCGCAGGCGTTGTGTAAGACAGATGTGAATCCCTGGGCTTAACCTGGGAACTGCATTTG
TGACTGCATGGCTAGAGTGTGTCAGAGGGAGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAC

4 ATACTGGGCGTAAGCGTGCCGCACGCGGTTTTGTAAGTCTGATGTGAATCCCCGGGTCTCAACCTGGGAATTGCAT
TTGTAGACTGCATGGCTAGAGTCTGGCAGAGGGGGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAC

5 ATACTGGGCGTAAAGCGTGCGCAGGCGGTTGTGTAAGACAGATGTGAATCCCTCGGTCTTAACCTAGGAATTGCATT
TGTGACTGCATAACTAGAGTGTGTCAGAGGGAGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAC

6 ATACTGTGGCGTAAGCGCGCGCAGGCTGGATTGTTAAGTCAGCATGTGAAGCCCCGAGTTCAACCTGGGCACTGCA
TCCGATACTGTCTGACTAGAGTGTGGTAGAGGGAGGTAGAATTCCACGTGTAGCAGTGAAAAGCGTA

7 AGTCTGCATGTGAAGCTCGAGCTCACCTGTCACTGCATCCGATACTGTCAGACTAGAGTGCGTAGAGGCAGGTAGA
TTCCACGTGTAGCAGTGAATAGCGTA

8 ATACTGTGCGTAACGCGCGCGCAGGCGGTCTGATAGTCAGCATGTGAAGCTCCGGTCTCACCTGGACTGCATCCGA
TACTGTCAGACTAGAGTGTGGTAGAGGGAGGTAGATTCCACGTGTAGCAGTGAGATGCGTA

9 TATCGGATCTGGGCGTAACGCGCGCGAGGGTCTGTTAAGTCGGATGTGAAAGCCCCGGTGTTCAGAGCCTGGGAA
CTGCATCCGATACTGGCAGACTAGAGTGTGGTAAGAGGAGGTAGAATTCCACGTGTAGCAGTGAAATGCGTA

10 ATACTGGGCGTAAGCGCGCGTAGGTGGTTGTTAAGTCAGCATGTGAAGCCCGTCTCAACCTGGGAACTGCATCCGA
TACTGTTTGACTAGAGTGTGGTGGAGGGAGGTAGAATTCCACGTGTAGCAGTGAAATGCGTA

11 ATACTGTGCGTAAGCGCGCGCAGGCTGGTCTGTTAGTCAGGATGTGAAGCCCCGGGTTCAACCTGGGACTGCATCC
GATACTGGCAGACTAGAGTGCGGGAGAGGGAGGTAGAATTCCACGTGTAGCAGTGAAATGCGTA

12 ATACTGGGCGTAAAGCGTGCGCAGGGCGGTTGTGAAGACAGATGTGAAATCCCTGGGCTTAACCTAGGAATTGCAT
TTGTGACTACATAACTAGAGTGTGTCAGAGGGAGGTAGAATTCCACGTGTAGCAGTGAAATGCGTA

Table 5.6

3 TATCAAGTTAGATGTGAATCTGAGGCTCAACCTCTAGCTGCATCTAATACTGATAAGCTAGAGTACTAGAGAGGTAGT
AGAATTCTTAGTGTAGCGTGGAATGCGTAGATATTAAGAGGCATACCAATGGCGAAGCGAACTTTCTGGATAGATAC
TGACACTGAGGTGCGAAAGCTGGGGAGCAAAACAGGATTAGATACCCCTGGTAGA

4 CTAAGTTAGATGTGAATCACGGTCTTAACTTCGAACTGCATCTAATACTGGTTGCTAGAGTACTAGAGAGGTAAGTAG
ACTCTTAGTGTAGCGTGGAATGCGTAGATATTAAGAGGCATACCAATGGCGAAGGAACTTCTCTGGATAGATACTGA
CACTGAGGTGCGAAAGCTGGGGGAGCAAACAGGATTAGATACCCCTGGTAGA

5 TAGTCAGATGTGAATCTCCAGCTCAGCTCGACTGCATCTGATACTGCTGACTAGAGTACTATAGAGGTAGTAGATCT
CTAGTGTAGCGTGATGCGTAGATATGTGACGATACCAGTTGCGAAGCGGCTTCTGGACATATACTGACGCTGAGGT
GCGAACAGCGTGGGGAGCGACCGGGATTAGATACCCCCTGGAAGAA

7 TGTGAATGCCCTGGCTTAAGCTGAGGAACTGCATCCAAACTGTCATCACTAGAGTACGGGAGAGGAGAGTAGAATTC
ATGGTGTAGCGGTGCAATGCGTAGATATCATGAGGAATACCCAGTGGCGAAGCGAGGCTCTCTGGATCTGTACTGA
CGCTGAGGTGCGAAAGCTGGGGTAGCAAAACAGGATTAGATACCCCTGATAGA

8 GTTAGTTAGGATGTGAAGCCTGGCTCAACCTGAGTACTGCATCCAGACTGTCTCACTAGAGTACGATAGAGGGAGTA
GAATCACTAGTGTAGCGTGAATGCGTAGATATCATGAGAATACCAGTGGCGAAGGGGCTCTCTGGATCTGTACTGAC
ACTGAGGTGCGAAAGTGGGGTAGCGAACAGGATTAGATACCCCTGGTAGA

9 GGATGTGAAGCCTGGCTCAACCTGGAACTGCATCCAGAACTGTCATCACTAGAGTACGATAGAGGGAAGTAGAATTC
ATAGTGTAGCGTGAATGCGTAGATATTATGAAGAATACCAGTGGCGAAGGGGCTCCCTGGATCTGTACTGACACTGA
GGTGCGAAAGCTGGGGTAGCAAAACAGGATTAGATACCCCTGGTAGA

10 AACCTGGAACTGCATCAGTACTGTTTGACTAGAGTACGGGAGAGGAGAGTAGAATTCATGGTGTAGCGGTGAAATG
CGTAGATATCATGAGGAATACCAATGGCGAAGGAACTCTCTGGACCGTTACTGACGCTGAGGTGCGAAAGTGGGTA

GCAAACAGGATTAGATACCCCTGGTAGA

11 TAGCTAGCATGTGAAGCCTGGTTCAACCTGGAACTGCATCCAGAACTGTCTGACTAGAGTACGATAGAGGGAAGTAG
AATTCACAGTGTAGCGTGAATGCGTAGATATCATGAGAATACCAGTGGCGAAGCGGACCTCCTGGACATAATACTGA
CGCTGAGGTGCGAAAGTGGGGTAGCGAACAGGATTAGATACCCCTGATAGA

12 AACCTGGAACTGCATCCAGAACTGTCTGACTAGAGTACGGTAGAGCGAGGTAGAATTCACTAGTGTAGCGTGTAATG
CGTAGATATCATGAGAATACCAGTGGCGAAGCGGGCCTCCTGGACCTATACTGACGCTGAGGTGCGAAAGTGTGGG
TAGCGAACAGGATTAGATACCCCTGATAGA

13 TAGCTAGATGTGAAGCCTGGCTCAACCTGGAACTGCATCTAGAACTGTCTGACTAGAGTATGGTAGAGGGAAGTAGA
ATTCACTGGTGTAGCGTGAATGCGTAGATATCATGAGGAATACTCAGTGGCGAAGCGGGCTCTCTGGATCTATACTG
ACGCTGAGGTGCGAAAGTTGGGGTAGCGAACAGGATTAGATACCCCTGATAGA

14 AACCTGGAACTGCATCCAGAACTGTCATGACTAGAGTACGATAGAGGGAAGTAGAATTCACTAGTGTAGCGTGAATG
CGTAGATATCATGAGAATACCAGTGGCGAAGCGGACTCTCTGGATCTATACTGACGCTGAGGTGCGAAAGTGTGGG
TAGCGAACAGGATTAGATACCCCTGATAGA

15 TAGGATGTGAAGCCTGGCGCAAGCCTGAGAACTGCATCCAGAACTGTCATCACTAGAGTCACTGGATAGAGGGAAG
TAGAATTCACTGGTGTAGCGTGCAATGCGTAGATATCATGAAGAATCACTCAGTGGCGAAGCGGACTCCCTGGATCT
GATACTGACGCTGAGGTGCGAAAGTGTGGGTAGCGAACAGGATTAGATACCCCTGATAGA

16 TGTGAAGCCTGGCTCAACCTGATAACTGCATCCAGAACTGTCTGACTAGAGTATTGTAGAGGGAAGTAGAATTCATG
GTGTAGCGTGTGTATGCGTAGATATCATGAGGAATACCAGTGGCGAAGCGGACCTCCTGGACATGATACTGACGCT
CAGGTGCGAACAGTGTGGGTAGCGAACAGGATTAGATACCCCTGGTAGA

17 TTAAGCTAGATGTGAAGCCCGGGTCCACCTGAGATGCTGCATTTAGAACTGGCAGACTAGAGTCATTGGAGAGGGG
AGTGGAATTCCAGGTGTAGCGGTGAAATGCGTAGATATCTGGAGGAACATCAGTGGCGAAGGGGACTCTCTGGCCA
AAGACTGACGCTCATGTGCGAAAGTGTGGGTAGCGAACAGGATTAGATACCCCTGGTAGA

18 AGCTAGATGTGAAGCCCGTGTCAACCTGGCATGCTCATTTAGAACTGGCAGACTAGAGTCTTGGAGAGGGAGTGGA
ATTCCAGGTGTAGCGGTGAATGCGTAGATATCTGGAGTAACATCAGTGGCGAAGCGACTCCCTGGCCAAAGACTGA
CGCTCATGTGCGAAAGTGTGGGTAGCGAACAGGATTAGATACCCCTGATAGA

19 TGGGATGCTCATTTAGAACTGGCAGACTAGCAGTCTTGGAGAGGGAGTGGAATTCCAGGTGTAGCGGTGAAATGCG
TAGATATCTGGAGGAACATCAGTGGCGAAGGGACTCCCTGGCCAAAGACTGACGCTCATGTGCGAAAGTGTGGGTA
GCGAACAGGATTAGATACCCTGGTAGA

20 CGTGTGGCGTTGAAATGCTTAGATATGTGGAGGACACCAGTGTCGAAGCGGCCTCCCTGGCTCGACACTGACGCTG
AAGGTGCGAAAGTGGGGGGGAGCAAACAGGATTAGATACCCCTGGTAGAA

21 AACCTGGACTGCATTTAATACTGTATACTAGAGTACGAGAGAGCGAGTAGGAATTCTTGGGGTAGCGGTGAAATGCG
TAGATATTATGAGGAATACCAGTGGCCGAAGCGAGCTCTCTGGCTCGATACTGACGCTGAGGTGCGAAAGGTGGGG
GGAGCAAACCAGGATTAGATCCCCTGGTAGGA

22 AATTAAGTCAGAGGTGAATCTTGCAGCTCAACTGTAACATTGCCTTTGATACTGGTTATCTTGAGTCATTATGAAGTA
GTTAGAATATGTAGTGTAGCGTTGAAATGCATAGATATTACATAGCAATACCAATTGCGAAGGAGCACTACTAATAAT
GTACTGACACTGATGGACGAAAGCGTGGGTAGCGAACAGGATTAGATACCCTGATAG

24 AGATAGATGTGAATCTCTGGTCTCAACCTGGCACTGCATCTGATACTGTCTATCTAGAGCTGAGAGAGGTAGCAGCC
TCACAGTGTAGAGTGCATGCGTAGATATTGTCAATCATACCAGAGGCGAAGGGAGACTTCTCTGGACATTTACTGAC
GCTGAGATGCGAAAGCGTGGGGGAGCAAAACAGGGATTAGATACCCCTTGGTAGA

Table 5.7

1 GATACTGGCGTACGCGAGCGTACGCTCTCTGTAGTCTGATGTGAAGCTGCAGTCTCAGCTCTGCACTGCACTGCGA
GCTATTATCTAGAGTATAGTAGAGCAGTGATCATGTGTAGAGTGAATTCGTAGATATTAGAGACACCAGTGGCGAAA
GCAGACTCTCTGGCCTCATATACTGACGCTGAGGTACGAAAGCGTGGGTAGCAAAACAGGAATAGATACCCCTGGT
AGAA

2 ATATCGGCGTACGTACGTAGCTGTCATAGTCTGATGTGCAGCTCACGTCTCAGCGTCAGTCATGACTGCAGACTGAG
TACAGAGAGAAGTGATCACGTGTAGCGTGATGCGTAGAGATGTGAGACTCAGTGGCGAAGGGACTCTCTGGCCTGT
AACTGACGCTGAGCCCGAAG

4 GCATACTGGCGTACGCGCCGTACGCTATATCATCAGTCTGATGTGAATCTCGCGTCTCAGCGTCGCACTGCTCTAGC
TGCTGACTGAGCTCGAGAGAGTGAGTGACTCACGTGTAGCAGCTGAATGTCGTAGATATTCGAAGACACAGTGGCG
AAAGCGGCTCTCTGGGCCTCGTATACTGACGCTGAGGCTGCGAAAGTGTGG

5 GGACTACTGGGCGCAAGCGCGCGTAGGCGTGTCTGTAGTCTGGAGTGAATCTCCGATCTCAACTCTGCCACTGCCT
TCAAAACTACATTTCTAGAGTTCGGTAGAGGAGAGTGGAATTCCTAGTGTAGAGGTGAAATTCGTAGATATTAGGAG
GAACACCAGTGGCGAAGCGCGACTCTCTGGGCCGAATACTGACGCTGAGGCTCGAAGCCATGGGGTAGCAACACA
GGATTAGATACCCTGATGAGA

7 GCATACTGGCGTACGCGAGTGTACGCTCTCTGCTAGTCTGATGTGAGCCGCAGTCTCAGCTGCTAGCACTGCATGC
GACTGACTCATCTAGAGTACTAGTAGAGCAGTGATCATAGTGTAGCAGTGAATGCGTAGATATCAGAGACACAGTGG
CGAAAGCGATCTTCTGGCTGATATACTGACGCTGAAGCTCCGAAAGGCGTGGGGAAGCAAACAGGAATTAGGATAA
CCCTGGTAGA

8 CGGCGTACTGTGCGTAACGCGAGCGTACGACGGATCTGCTAGTCTGGAGGTGAAGCTCACGAGCTCAACTCTAGAA
CTGCCTTCTAAACTACATTTCTAGAGTCTGGGAGAGGAGAGTGGAATTCCTAGTGTAGAGGTGAAATTCGTAGATATT
AGGGGGAACACCAGTGGCGAAGGAGATTCTCTGGGCCAATACTGACGCTAAGGCTCGAAAGCATGGGAAGCGAAC
AGGATTAAATACCCTGGT

10 ATACTGGCGTACGCGCGCGTACGCGGATCTGCTAGTCTGATGTGAGTCCATCAGTCTCAGCTGCAGCACTGCATGC
GAAACTGTCGATCTAGAGTATCGTAGAGCAGAGTGATCATAGTGTAGCAGTGACTGCGTAGATATCAGAGACACAGT
GGCGAAGCGACTCTCTGGCTCGATACTGACGCTGAGCTGCGAAAGCGTGGGGAGCAAACAGGAATAGATACCCTT
GATAGA

11 GACTACTGCAGTACGCAGCGTACGCGGTCTGTAGTCTGATGTGAGCCTCGTCTCACGAGCACGTGCATGACTGTCG
ACTAGAGTGCGTAGAGAGAGTGCATCATAGTGTAGCAGTGACTGCGTAGATATGTGAGACACCAGTGGCGAAAGGG
ACTCTCTGGTCTCGTAACTGACGCTGAGGTGCGAAGGCGTGGGAAAGAAAACAGGATTTAGATCCCTTGGTAGGA

12 ATACTGGCGTACGCGCGCGTAGACGTGATCTGCTAGTCTGATGTGCAGTCCGATCAGTCTCACGCTGCACGTGCAC
TGAACTGTCATCTAGAGTATCGGAGAGGTGAGTGCATCATAGTGTAGCAGTGAATGCGTAGATATCAGAGACACCAG
TGGCGAAGCGACTCTCTGGCTCGTATCTGACGCTGAGGCTACGAAAGCGTGGGGAGCAAACAGGATTAGATACCCC
TGGATAGGA

14 AGTCTGATGTGCAGCTCCGTCTCAGCGTAGTCATGTACTGCAGACTGAGTACAGAGAGAGAGTGATCACGTGTAGC
GTGCACTGCGTAGAGATGTGAGACACCAGTGGCGAGGGGACTCTCTGGCCTGTAATTGACGCTGAGCGCGAAACG
CTGGGGGCCACCCGGGATTAGTTCCCCTGGAAGAA

Table 5.8

1 TGGGCGTAAGCGTCCGCAGGCGGCTTTTCAGTCTGCTGTTAAATCGTGGAGCTTAACTCCATCATGGCATTGGATAC
TGTTGGGCTTGAGTGTGGTAGGGGCAGAGGGAATTCCCGGTGTAGCGGTGAAATGCGTAGATATCGGGAAGAACA
CCAGTGGCGAAGGCGCTCTGCTGGGCCATTACTGACGCTCATGGACGAAAGCCTGGGGAGCGAAAGGGATTAGAT
ACCCCGGTAGTCCTGGCCGTAAACGATGAACACTAGGTGTCGGGGGAATCGACCCCTTCGGTGTCGTAGCTAACGC
GTTAAGTGTTCCGCCTGGGGAGTACGCACGCAAGTGTGAAACTC

2 GCTTAAGGGTGCGTAGGCGGGTTTAAAGGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAATTGCGTTGGAAACC
ATAAGCCTTGAGGGAGTCAGAGGTAAGCGGAACTGATGGTGGAGCGGTGAAATGCGTTGATATCATCAGGAACACC
GGTGGCGAAAGCGGCTTACTGGGACTCTTCTGACGCTGAGGCACGAAAGCTAGGGGAGCGAACGGGATTAGATAC
CCCGGTAGTCCTAGCTGTAAACGATCAGTACTAGTCTGTGGGGACTTCCACATCCTCTCGGACGTAGCGAAAGTGTT
AAGTACTGCGCCTGGGGAGTATGGTCGCAAGGCTGAAACTCAAAGGA

3 AGTCTGATGTGAATCCCGCGGCTCAACTGTGGATTGCGTTGGATACTGTTAGACTTGAGTAAGTGAGAGGTAAATGG
AATTCCTGGTGTAGCGGTGAAATGCGTAGATATCAGGAAGAACACCGGTGGCGAAGGCGGCTTACTGGGACTCTAC
TGACGCTGAGGCACGAAAGCTTGGGGAGCGAACGGGATTAGATACCCTGGTAGTCCTCGCCGTAAACGATGAGTAC
TA

4 GGCTTAAGGCGTGCGTAGGCGGGTTTACAGGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAATTGCGTTGGAAA
CTGTAAGCCTTGAGGGAGTCAGAGGTAAGCGGAACTGATGGTGGAGCGGTGAAATGCGTTGATATCATCAGGAACA
CCGGTGGCGAAAGCGGCTTACTGGGACTCTTCTGACGCTGAGGCACGAAAGCTAGGGGAGCGAACGGGATTAGAT
ACCCCGGTAGTCCTAGCTGTAAACGATCAGTACTAGTCTGTGGGGACTTCCACATCCTCTCGGACGTAGCGAAAGT
GTTAAGTACTGCGCCTGGGGAGTATGGTCGCAAGGCTGAAACTCAAAGGAATTGAC

5 TGATTATTGGGCGTAAGCGCGCGCAGGCGGTTCCTTAAGTCTGATGTGAAATCTCGCGGCTCAACCGCGAGCGGCC
ATTGGAAACTGGGGAACTTGAGTGCAGGAGAGGGGAGCGGAATTCCACGTGTAGCGGTGAAATGCGTAGAGATGT
GGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGTAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCGAAC
AGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAAGTGTTAGAGGGTATCCACCCTTTAGTGCTGC
AGCAAACGCATTAAGCACTCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAGTTTGACGGGG

6 CCGGATGTGCATTGCAAACTGGGCATCTTTGATTGTGGGAGAGGGAAGAGGAATTCTCCGTGTATCGGTGACATGC
GTACATATCGGGAGGAACACCAGTGTCGAAGGCGGCTTTCTGGCCTGACTCTGACACTGAGGTGCGAAAGCGTGTC
GGAGCGAACAGGATTAGATTCCTGGTAATCCACGCCGTAAACGATGT

7 CGTGGAACTGCTTTGCATACTGATTGGCTAGAGTACGGGAGAGGAGAGCGGAACTCTTGGTGGAGCGGTGAAATGC
GTAGATATCAAGAGGAACACCAGAGGCGAAGGCGGCTCTCTGGCCCGAAACTGACGCTGAGGTGCGAAAGCCAGG
GTAGCGAACGGGATTAGATACCCCGGTAGTCCTGGCCGTAAACGATGGGCACTAGCTGGTGGTCGTACCTGTGCGA
TCGCCGGCGAAGCTAATGTGATAAGTGCCCCGCCTGGGGAGTATGGTCGCAAGGCTGAAACTCAAAGGAATTGACG
GGG

8 AACTCCATCATGGCAGTGGAACTGACGGCTAGAGTATGGTAGGGGCAGAGGGAATTCCCGGTGTAGCGGTGAAATG
CGTAGATATCGGGAAGAACACCAGTGGCGAAGGCGCTCTGCTGGGCCATTACTGACGCTCATGGACGAAAGCCGG
GGGAGCGAAAAGGGATTAGATACCCCTGTAGTCCCTGGCCGTAAACGATGAACGAGCTAACGCGTTAAGTGTCCCG
CCTGGGGAGTACGCCCGCAAGTTTAAAACTCAAGGAATGGACGGGGGA

9 TGGGCAGAAGCGCACGTAGGCGGCTTATCTAGTGCGAGTGTGAAGCCACGCGCTCACCCGTGGACTGCTTTGCATA
CTGATTGGCTAGAGTACGGGAGAGGAGAGCGGAACTCTTGGTGTAGCGGTGAATGCGTAGATATCAAGAGGAACAC
CAGAGGCGAAGGCGGCTCTCTGGCCCGAAACTGACGCTGAGGTGCGAAAGCCAGGGGTAGCGAACGGGATTAGAT
ACCCCGGTAGTCCTGGCCGTAAACGATGTGCACTAGCTGGTGGTCGTACCTGTGCGATCGCCGGCGAAGCTAATGT
GATAAGTGCCCCGCCTGGGGAGTATGGTCGCAAGGCTGAAACTCAAAGGAATTGACGGGGA

10 GGCGTAAGCGTGCGCAGGCGGTTTTTTAGTCTGATGTGAATCTCGCGGCTCAACCGCGAATTGCATTGGAGACTGG
GGGCTTGAGTGCGGCAGAGGGGGGTGGAATTCCAGGTGTAGCGGTGAAATGCGTAGATATGTGGAGGAACACCGG
TGGCGAAGGCGGCTCTCTGGGCTGCTACTGACGCTGATGCACGAAAGCGTGGGGAGCGAACGGGATTAGATACCC
TGGTAGTCCTGGCCGTAAACGATG

11 GGTTAAGGCGTGCGTAGGCGGGTTTAAAGGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAATTGCGTTGGAAAC
CATAAGCCTTGAGGGAGTCAGAGGTAAGCGGAACTGATGGTGGAGCGGTGAAATGCGTTGATATCATCAGGAACAC
CGGTGGCGAAAGCGGCTTACTGGGACTCTTCTGACGCTGAGGCACGAAAGCTAGGGGAGCGAACGGGATTAGATA
CCCCGGTAGTCCTAGCTGTAAACGATCAGTACTAGTCTGTGGGGACTTCCACATCCTCTCGGACGTAGCGAAAGTGT
TAAGTACTGCGCCTGGGGAGTATGGTCGCAAGGCTGAAACTCAAAGGAATTGACGG

12 GGTGTAGCGGTGAAATGCGTAGATATCTGGAGGAACACCGGTGGCGAAGGCGGCTTACTGGGCTGCTACTGACGC
TGAGGCACGAAAGCTTGGGGAGCGAACGGGATTAGATACCCCGGTAGTCCTCGCTGTAAACGAT

Table 5.9

1 TATTGGGCGTAAGCGTCCGCAGGCGGCTTTTCAAGTCTGCTGTTAAAACGTGGAGCTTAACTCCATCATGGCAGTGG
AAACTGATGGGCTTGAGTATGGTAGGGGCAGAGGGAATTCCCGGTGTAGCGGTGAAATGCGTAGATATCGGGAAGA
ACACCAGTGGCGAAGGCGCTCTGCTGGGCCATTACTGACGCTCATGGACGAAAGCCAGGGGAGCGAAAGGGATTA
GATACCCCAGTAGTCA

2 TATTGGGCGTAAGCGTCCGCAGGCGGCTTTTCAGTCTGCTGTTAAAGCGTGGAGCTTAACTCCATCATGGCAGTGG
ACACTGATAGGCTTGAGTATGGTAGGGGCAGAGGGAATTCCCGGTGTAGCGGTGAAATGCGTAGATATCGGGAAGA
ACACCAGTGGCGAAGGCGCTCTGCTGGGCCATTACTGACGCTCATGGACGAAAGCCAGGGGAGCGAAAGGGATTA
GATACCCCAGTAGTCA

3 ATTATTGGGCGTAAGCGTCCGCAGGCGGCCTTTCAGTCTGCTGTTAAACGTGGAGCTTAACTCCATCATGGCAGTG
GAGACTGTATGGGCTTGAGTGTGGTAGGGGCAGAGGGAATTCCCGGTGTAGCGGTGAAATGCGTAGATATCGGGA
AGAACACCAGTGGCGAAGGCGCTCTGCTGGGCCATTACTGACGCTCATGGACGACAGCCAGGGGAGCGAAAGGGA
TTAGATACCCCAGTAGTCA

4 TATTGGGCGTAAGCGTCCGCAGGCGGCCCTTCAGTCTGCTGTTAAAAGTGGAGCTTAACTCCATCATGGCAGTGGA
ACTGTTGGGCTTGAGTGTGGTAGGGGCAGAGGGAATTCCCGGTGTAGCGGTGAGATGCGTAGATATCGGGAAGAA
CACCAGTGGCGAAGGCGCTCTGCTGGGCCATCACTGACGCTCATGGACGAAGCCAGGGGAGCGAAAGGGATTAGA
TACGCCAGTAGTCA

5 ATACTGGGCGTAAGCGCCCGTAGGCGGTATTGTCAAGTCTGCTGTTAAGCCTAGGGCTTAACTCCATCATGGCATTG
GATACTGATAGGCTGTGAGTATGGTAGGGAAGAGGGAGTTCCCGGTGTAGCGGTGAAATGCGTAGATATCAGAGAA
GACACCAGTGGCGAGGGCGACTTCTTGCTGGAGCTCGATATCTGACGCTGCAGTAGTGACGAAAGCTCTGGGTGA
GCAGAACAGGATTAGATACCCCAGTAGTCA

6 GGCGTAAGCGCGCGTAGCGGCTTATCAGTCTGATGTGTAAGCGCGGAGCTCACTCCAGCACTGCAGTGGATACTGA
TAGACTATGAGTATGAGTAGGGTAGATGGATTCCTAGGTGTAGCGGTGAATGCGTAGATATCATAGAGATCACCAGT
GGCGAAGGCGCTCTGCTGGATCATTACTGACGCTCAGAGACGAA

7 GCGTAAGCGCGCGTAGGCGTGTTTAACAGTCTGCTGTGAATCCCGGGGCTCACTCCAGCATGTGCAGTGGATACTG
ATAGACTAGAGTATGGTAGGGAAGATGGATTCCTGGTGTAGCGGTGAATGCGTAGATATCATAGAGATACCAGTGGC
GAAGGCGCTCTGCTGGATCGATACTGACGCTGATAGACGAAAGCGTGGGTAGCAAACAGGATTAGATACCCCAGTA
GTCA

8 GCGTAAGCGCGCGTAGGCGTGCTTAACAGTCTGATGTGTAATCCCGGGGCTCACCCCGGCATGTGCAGTGGATACT
GATAGACTAGAGTATGGTAGAGAAGATGGATTCCTGGTGTAGCGGTGAATGCGTAGATATCAGGAGATACCAGTGG
CGAAGGCGCTCTGCTGGATCATTACTGACGCTCATAGACGAAAGCGTGGGTAGCGAACAGGGATTAGATACCCCAG
TAGTCA

9 TTGGGCGTAAGCGCTCTGCAGGTGTGTTTATTCAAGTCTGCTGTTAAGCACTGGGAGCTTAACCCCATCAATGAGCA
ATGGATACTAGCATAGAGCTATGAGTATGGCTAGGGGTAGAGGGAATTTCTAGTGTAGCGGTGATATGCGTAGATAT
TAGGAGAGACACCGAGTGGCGAAGGCGCTCTAGCTGGGAGCCATTACTGACAGCTCAGTAGGACGAAAGCTCAGG
GTGAGCAGAAGGGGATTAGATACCCCCAGTAGTCA

10 TTGGGCGTAAGCGCTCTGCAGGTGTGTTTATCAAGTCTGCTGTTAAGCACTGCGAGCTTAACTCTATCAATGAGCAA
TGGATACTAGCATAGAGCTATGAGTATGGCTAGGAGTAGAGGGAATTTCTAGTGTAGCGGTGATATGCGTAGATATC
AGGAGAGACACCGAGTGGCGAAGGCGCTCTAGCTGGACCATTACTGACAGCTCAGTAGGACGAAAGCTCAGGGTG
AGCAGAAGGGGATTAGATACCCCATGTAGTCA

11 TTGGGCGTAAGCGTCTGCAGGTGTGTTTATCAAGTCTGCTGTTAAGACTGGAGCTTAACTCCATCAATGAGCAATGG
ATACTGCATAGACTATGAGTATGGCTAGGAGTCAGAGGGAATTTCTCAGTGTAGCGGTGAATGCGTAGATATTCAGG
AGAGACACCGAGTGGCGAAGGCGCTCTAGCTGGAGCATCACTGACAGCTCAGTAGGGACGAAAGCTCAGGGTGAG
CAGAAGGGGATTAGATACCCCCATGTAGTCA

12 TGGGCGTAAGCGCGCGTAGGCGGCTTGAAAGTCGGATGTGAATCCCTGGGCTCAACCTGGGATCTGCATTCGATAC
TGATAGACTAGAGTATGGTAGAGGAAAGTGGACTCCTGGTGTAGCGGTGAATGCGTAGATATCAGGAGATACCAGT
GGCGAAGGCGCTCTCCTGGATCGATACTGACGCTGATGTGCGAAAGCGTGGGTAGCAAACAGGATTAGATACCCCA
GTAGTCA

13 TACTGGGCGTAAGCGCGCGTAGGCGGTATGCTAAGTTGGGTGTGAATCCCCGGGCTTAACCTGGGAACTGCATTCA
AAACTTGCATACTGGAGTACGGAAGAGGCGAGTAGAATTCATGGTGTAGCGGTGAAATGCGTAGATATCATGAGGAA
TACCAATGGCGAAGGCAACTCGCTGGTCCGTAACTGACGCTGAGGTGCGAAAGCGTGGGTAGCAAACAGGATTAGA
TACCCCAGTAGTCA

14 AGTATTGGGCGTAAGCGTCCGCAGGCGGCTTTTCAAGTCTGCTGTTAAGCGTGGAGCTTAACTCCAGTCATGGCAG
TGGAACTGATAGGCTTGAGTATGGTAGGGGCAGAGGGAATTCCCGGTGTAGCGGTGAGATGCGTAGATATCGGGAA
GAACACCAGTGGCGAAGGCGCTCTGCTGGGCCATTACTGACGCTCATGGACGAAAGCCAGGGGAGCGAAAGGGAT
TAGATACCCCAGTAGTCA

15 TACTGAGCGTAGCGTCGTAGCTAGATGCTAGTCAGATGTGAATCTCTGAGCTCACTTAGTACTGCATCGATACTACT
GACTAGAGTATGATAGAGTAGTAGATCGTAGTGTAGCGATGATGCGTAGATATCGATGAGATACAGTGGCGAAGCGA
GCTCTCTGGATCTATACTGACGCTGAGAGACGAAAGCGTGGGTAGCAAACAGGATTAGATACGCCCAGTAGTCA

16 ATTGGGCGTAAGCGTCTGCAGGCGTGTTTTTCAAGTCTGCTGTTAAGACTGGAGCTTAACTCCATCATGGCAGTGGA
ACTGCATAGACTATGAGTATGGTAGGAGTAGAGGGAATTCCTAGTGTAGCGGTGAATGCGTAGATATCAGAGAAGAC
ACCGAGTGGCGAAGGCGCTCTAGCTGGAGCCATTACTGACAGCTCAGTAGGACGAAAGCTCAGGGTGAGCAGAAG
GGGATTAGATACCCC

Appendix II

Improved group-specific PCR primers for denaturing gradient gel electrophoresis analysis of the genetic diversity of complex microbial communities

Martin Mühling, John Woolven-Allen, J. Colin Murrell, and Ian Joint

The ISME Journal 2: 379-392 (2008)