Temporal genetic diversity in phytoplankton in the Gulf of Naples

Thesis

How to cite:

For guidance on citations see FAQs.

© 2006 Sarah Margaret McDonald.

https://creativecommons.org/licenses/by-nc-nd/4.0/

Version: Version of Record

Link(s) to article on publisher’s website:
http://dx.doi.org/doi:10.21954/ou.ro.0000fe69

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.

oro.open.ac.uk
TEMPORAL GENETIC DIVERSITY IN PHYTOPLANKTON IN THE GULF OF NAPLES

SARAH MARGARET MCDONALD
BSc Marine Biology, University of Liverpool
MSc Advanced Methods in Taxonomy and Biodiversity, Imperial College London

Doctor of Philosophy

Sponsoring Establishment
Stazione Zoologica Anton Dohrn

December 2006
DONATION

X 589.40415
To my beloved and incredible grandmother,
Granny Yam,
who set the example,
and is sorely missed.

14th August 1913 - 25th August 2006

They sought it with thimbles,
they sought it with care,
They pursued it with forks and hope,
They threatened its life with a railway-share,
They charmed it with smiles and soap.

The Hunting of the Snark
Lewis Carroll 1876
Table of Contents

THESIS ABSTRACT ................................................................................................................................. 5
ACKNOWLEDGEMENTS .......................................................................................................................... 6
LIST OF FIGURES .................................................................................................................................... 7
LIST OF TABLES ....................................................................................................................................... 8
LIST OF ABBREVIATIONS USED .......................................................................................................... 9
CHAPTER ONE ......................................................................................................................................... 10
GENERAL INTRODUCTION
THE PHYTOPLANKTON .......................................................................................................................... 10
SPECIES CONCEPTS AND DIVERSITY IN THE MARINE ENVIRONMENT ........................................... 11
CRYPTIC, PSEUDO-CRYPTIC AND HIDDEN DIVERSITY ...................................................................... 16
TOOLS FOR ASSESSING GENETIC DIVERSITY AT SEA ...................................................................... 17
DIVERSITY IN PHYTOPLANKTON ON A TEMPORAL SCALE .............................................................. 20
GULF OF NAPLES AND THE MARECHIARA LONG TERM ECOLOGICAL PROGRAMME .................. 22
AIMS ..................................................................................................................................................... 25
CHAPTER TWO ......................................................................................................................................... 27
ULTRAPHYTOEUKARYOTE DIVERSITY IN THE GULF OF NAPLES OVER AN ANNUAL CYCLE
INTRODUCTION ....................................................................................................................................... 27
MATERIALS AND METHODS ................................................................................................................ 30
Characterisation of the MC sampling site .......................................................................................... 30
Sampling ................................................................................................................................................. 31
DNA extraction and amplification of environmental samples .......................................................... 31
Preparation and amplification of control DNA .................................................................................. 32
Dot blot hybridisation ......................................................................................................................... 33
Construction of clone libraries ........................................................................................................... 35
Magnesium Concentrations ................................................................................................................. 36
Sequences from cultured strains .......................................................................................................... 37
Phylogenetic Analyses ......................................................................................................................... 38
RESULTS ................................................................................................................................................. 39
MC Ecological Programme Data ....................................................................................................... 39
Dot Blots ................................................................................................................................................ 40
Clone Libraries ....................................................................................................................................... 42
Magnesium concentrations .................................................................................................................. 53
Comparison of Methods ...................................................................................................................... 53
DISCUSSION .......................................................................................................................................... 54
Method Biases – Dot Blots and Clone Libraries ................................................................................ 54
Where is Micromonas? ......................................................................................................................... 59
Dot Blot Hybridisations ...................................................................................................................... 61
Clone Libraries ....................................................................................................................................... 67
CHAPTER SUMMARY ............................................................................................................................ 71
CHAPTER THREE ...................................................................................................................................... 73
ULTRAPHYTOEUKARYOTE DIVERSITY IN DEPTH PROFILES DURING FOUR SEASONS
INTRODUCTION ....................................................................................................................................... 73
MATERIALS AND METHODS ................................................................................................................ 76
Environmental Data ............................................................................................................................. 76
Sampling ............................................................................................................................................... 76
DNA extraction and amplification of environmental samples .......................................................... 76
Preparation and amplification of control DNA .................................................................................. 76
Dot blot hybridisation ......................................................................................................................... 76
Flow cytometry ..................................................................................................................................... 76
RESULTS ............................................................................................................................................... 77
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environmental Data</td>
<td>77</td>
</tr>
<tr>
<td>Dot Blot Hybridisations</td>
<td>80</td>
</tr>
<tr>
<td>Flow cytometry</td>
<td>84</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>85</td>
</tr>
<tr>
<td>Dot Blot Hybridisation</td>
<td>85</td>
</tr>
<tr>
<td>CHAPTER SUMMARY</td>
<td>91</td>
</tr>
<tr>
<td><strong>CHAPTER FOUR</strong></td>
<td>93</td>
</tr>
<tr>
<td><strong>DIVERSITY IN THE DIATOM GENUS PSEUDO-NITZSCHIA OVER AN ANNUAL CYCLE</strong></td>
<td></td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>93</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>95</td>
</tr>
<tr>
<td>Sampling</td>
<td>95</td>
</tr>
<tr>
<td>DNA extraction and amplification</td>
<td>95</td>
</tr>
<tr>
<td>Primer development</td>
<td>97</td>
</tr>
<tr>
<td>SSCP</td>
<td>97</td>
</tr>
<tr>
<td>Clone libraries</td>
<td>99</td>
</tr>
<tr>
<td>Known sequences</td>
<td>99</td>
</tr>
<tr>
<td>Phylogenetic analyses</td>
<td>99</td>
</tr>
<tr>
<td>RESULTS</td>
<td>100</td>
</tr>
<tr>
<td>Primer development</td>
<td>100</td>
</tr>
<tr>
<td>SSCP</td>
<td>100</td>
</tr>
<tr>
<td>Seasonality - Counts and Clone Libraries</td>
<td>102</td>
</tr>
<tr>
<td>Diversity - Clone Libraries</td>
<td>104</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>109</td>
</tr>
<tr>
<td>Primers</td>
<td>109</td>
</tr>
<tr>
<td>SSCP</td>
<td>110</td>
</tr>
<tr>
<td>Diversity and Seasonality in Pseudo-nitzschia</td>
<td>111</td>
</tr>
<tr>
<td>CHAPTER SUMMARY</td>
<td>115</td>
</tr>
<tr>
<td><strong>CHAPTER FIVE</strong></td>
<td>117</td>
</tr>
<tr>
<td><strong>DIVERSITY WITHIN THE DIATOM PSEUDO-NITZSCHIA GALAXIAE</strong></td>
<td></td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>117</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>121</td>
</tr>
<tr>
<td>Isolation of strains</td>
<td>121</td>
</tr>
<tr>
<td>Extraction and amplification of DNA</td>
<td>122</td>
</tr>
<tr>
<td>Cloning</td>
<td>123</td>
</tr>
<tr>
<td>Sequence Analysis</td>
<td>123</td>
</tr>
<tr>
<td>RESULTS</td>
<td>124</td>
</tr>
<tr>
<td>Cloning</td>
<td>124</td>
</tr>
<tr>
<td>Sequence Analysis</td>
<td>124</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>128</td>
</tr>
<tr>
<td>Different markers, different sensitivity</td>
<td>128</td>
</tr>
<tr>
<td>Defining taxa in microalgae</td>
<td>132</td>
</tr>
<tr>
<td>The genetic composition of P. galaxiae blooms - Different bloom, different genotype</td>
<td>134</td>
</tr>
<tr>
<td>Different bloom, different size</td>
<td>136</td>
</tr>
<tr>
<td>Recurrence of genotypes in same period</td>
<td>138</td>
</tr>
<tr>
<td>How this pattern could arise</td>
<td>140</td>
</tr>
<tr>
<td>CHAPTER SUMMARY</td>
<td>143</td>
</tr>
<tr>
<td><strong>CHAPTER SIX</strong></td>
<td>145</td>
</tr>
<tr>
<td><strong>CONCLUSIONS AND FUTURE PROSPECTIVE</strong></td>
<td></td>
</tr>
<tr>
<td>Conclusions</td>
<td>145</td>
</tr>
<tr>
<td>Future Prospective</td>
<td>147</td>
</tr>
<tr>
<td><strong>CHAPTER SEVEN</strong></td>
<td>149</td>
</tr>
<tr>
<td>REFERENCES</td>
<td></td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>176</td>
</tr>
</tbody>
</table>
This thesis dealt with genetic diversity in groups of phytoplankton on a temporal scale. The aim of the study was to investigate diversity and the seasonal distribution of that diversity during the year.

Chapter Two examined the diversity and seasonal distribution of ultraplankton <5 μm in surface waters between 2003 and 2004 using 16S class level probes and clone libraries. Chapter Three examined the distribution of ultraplankton using the 16S probes over four depth profiles during 2005. The results showed the presence of Prymnesiophyceae and Cryptophyceae during the whole year and at most depths. The Chrysophyceae showed a marked preference for summer. All three classes showed highest signal just below the surface between 2 and 10 m. The Pelagophyceae were present in February in surface waters, but at 50 to 70 m in May and August. Diversity within these classes was high and other taxonomic classes were detected in the clone libraries. The seasonality within genotypes in the Cryptophyceae was visible and comparisons to known cultures could be made.

The diatom genus *Pseudo-nitzschia* was investigated in Chapter Four using newly developed genus specific primers. Thirteen genotypes were detected during the year, some of which showed seasonal preferences, for example *P. multistriata* in October, and others that were present most of the year, for example *P. galaxiae* LSU-clade 2 present from May until October. The three annual blooms of *P. galaxiae* were investigated in Chapter Five using LSU and ITS. The blooms were shown to be dominated by different ITS genotypes that recur in the same period of the year, every year. The results from *P. galaxiae* indicated a possible mechanism for speciation whereby populations may become separated in time rather than space.
Acknowledgements

The Stazione Zoologica ‘Anton Dohrn’ for the four year studentship that allowed me to come and experience Naples, fulfil my ambition of living abroad and learning another language - as well as do my PhD.

The Society for General Microbiology President’s Fund for a Short Research Visit and the Marine and Freshwater Microbial Biodiversity International Opportunities Fund for awards that enable me to spend time at the University of Warwick.

Adriana Zingone, my supervisor, for her invaluable enthusiasm, imagination and excitement for research, for wanting to tell stories and for supporting me. Dr Dave Scanlan, my second supervisor, for all his help, support, ideas and advice.

Carmen Minucci who filtered water for me. Diana Sarno for the data on phytoflagellates and the members of the Area Gestione Ambiente ed Ecologia Costiera delle Aree Temperate e Polari for the environmental data, especially Immacolata Santarpia for the HPLC data and Ferdinando Tramontano from the Laboratorio di Oceanographia. Wiebe Kooistra for his help and advice on molecular biology.

Everyone in Micro2 at the University of Warwick for making both my visits so useful and enjoyable, especially to Sophie Mazard, Ludwig Jardillier and Amy Kirkham for all their patience and help, and to Lee Humphreys for encouraging Friday afternoon meetings.

The crew of the Vettoria, especially Ciro and Augusto, who collected samples for me over three years and who put up with me being on the boat.

The staff of the Servizio di Biologia Molecolare, especially Elvira and Marco, for all their help and advice with sequencing and for putting up with me constantly bothering them.

Maria Pia Postiglioni for helping me with the Multiphor and teaching me the basics of SSCP, also to the members of the laboratory of Dr Paolo Macchia at the Policlinico di Napoli for their time and assistance with SSCP.

Nina Lundholm from the University of Copenhagen and Bente Edvardsen from the University of Oslo for supplying culture material and DNA of *Pseudo-nitzschia decipiens* and *Chrysochromulina*.

My friends: Prof. David Mann for encouraging me to apply to the SZN. Everyone in Botanica Marina, especially Cho, mio fratello Ugo and Domenico for all their friendship laughter and encouragement, but also Alberto, il mio compagno di stanza Raffaele, Isabella, Federica and Mariella. Alison, Elin, Enrique, Francesca, Mily, Nik, Pam, Paola, Vix and finally Nikos and Luigi (the Muppets). My parents for all their unconditional support and keeping me on an even keel. Frank Sinatra for being my sanity and Nennella for fried mozzarella.

E alla fine, la città di Napoli per lo sviluppo del mio istinto di sopravvivenza.
List of Figures

FIGURE 1.1: 16S rDNA TREE OF EUKARYOTE PHYLOGENY........................................... 14
FIGURE 1.2: MAP OF THE GULF OF NAPLES............................................................ 23
FIGURE 2.1: ENVIRONMENTAL DATA........................................................................... 39
FIGURE 2.2: DOT BLOT HYBRIDISATION FOR THE CHRY1037 CHRYSOPHYCEAE PROBE 40
FIGURE 2.3: PERCENTAGE RELATIVE HYBRIDISATION PLOTS..................................... 42
FIGURE 2.4: LM COUNTS WITH PIE CHARTS FOR THE SIX CLONE LIBRARIES........... 43
FIGURE 2.5: RAREFACTIO N CURVES FOR EACH OF THE SIX CLONE LIBRARIES........... 45
FIGURE 2.6: PRYMNESIOPHYCEAE NJ TREE (TOP HALF ONLY)................................. 46
FIGURE 2.7: PRYMNESIOPHYCEAE NJ TREE CONTINUED (BOTTOM HALF)................ 47
FIGURE 2.8: CRYPTOPHYCEAE NJ TREE .................................................................. 49
FIGURE 2.9: CHRYSOPHYCEAE NJ TREE .................................................................. 50
FIGURE 2.10: NJ TREE REMAINING SEQUENCES....................................................... 51
FIGURE 2.11: MAGNESIUM CONCENTRATIONS......................................................... 53
FIGURE 2.12: COMPARISON OF METHODS.................................................................. 54
FIGURE 3.1: CTD DATA FOR ALL PROFILES.............................................................. 78
FIGURE 3.2: NUTRIENT CONCENTRATIONS FOR ALL PROFILES.............................. 79
FIGURE 3.3: PERCENTAGE RELATIVE HYBRIDISATION FOR THE PRYMNESIOPHYCEAE 80
FIGURE 3.4: PERCENTAGE RELATIVE HYBRIDISATION FOR THE CRYPTOPHYCEAE PROBE 81
FIGURE 3.5: PERCENTAGE RELATIVE HYBRIDISATION FOR THE CHRYSOPHYCEAE PROBE 82
FIGURE 3.6: PERCENTAGE RELATIVE HYBRIDISATION FOR THE PELAGOPHYCEAE PROBE 83
FIGURE 3.7: TOTAL SIGNAL FROM THE FOUR PROBES OVER THE 2005 DEPTH PROFILES... 83
FIGURE 3.8: CYTOMETRY DATA.................................................................................. 84
FIGURE 3.9: PERCENTAGE OF CRYPTOPHYCEAE..................................................... 85
FIGURE 4.1: SSCP MDE GEL..................................................................................... 101
FIGURE 4.2: PSEUDO-NITZSCHIA SPECIES ABUNDANCES FROM 0 M ST. MC .......... 102
FIGURE 4.3: ABUNDANCES OF THE RARE PSEUDO-NITZSCHIA............................... 103
FIGURE 4.4: PIE CHARTS COMPARING COUNTS AND LIBRARIES............................ 104
FIGURE 4.5: RAREFACTIO N CURVES FOR THE SIX LIBRARIES............................... 106
FIGURE 4.6: NJ TREE FOR THE SEQUENCES FOUND IN THE SIX CLONE LIBRARIES 107
FIGURE 4.7: NJ TREE FOR ALL THE SEQUENCES SIMILAR TO P. GALAXIAE............ 108
FIGURE 5.1: THE THREE P. GALAXIAE SIZE CLASSES.............................................. 119
FIGURE 5.2: NJ TREE FOR LSU SEQUENCES FROM P. GALAXIAE........................... 125
FIGURE 5.3: NJ TREE FOR ITS SEQUENCES P. GALAXIAE........................................ 127
List of Tables

TABLE 2.1: PRIMERS USED IN THIS CHAPTER................................................................. 32
TABLE 2.2: STRAINS USED AS CONTROL DNAs....................................................... 33
TABLE 2.3: 16S rDNA OLIGONUCLEOTIDE PROBES.................................................. 34
TABLE 2.4: CULTURES SEQUENCED DURING THIS STUDY.......................................37
TABLE 2.5: CLONE LIBRARY DETAILS...........................................................................44
TABLE 2.6: DETAILS OF THE SEQUENCES IDENTIFIED FROM THE SIX CLONE LIBRARIES. ... 45
TABLE 4.1: PRIMERS USED AND DESIGNED IN THIS STUDY............................ 96
TABLE 4.2: DIATOM SPECIES USED IN THE PSEUDO-NITZSCHIA PRIMER DEVELOPMENT.... 97
TABLE 4.3: CULTURES USED AS CONTROLS IN THE OPTIMISATION OF SSCP............... 98
TABLE 4.4: CLONE LIBRARY DETAILS....................................................................105
TABLE 5.1: LSU AND ITS PRIMERS USED IN THIS STUDY.....................................122
TABLE 5.2: BASE PAIR DIFFERENCES IN THE ITS ALIGNMENT...............................126
TABLE 5.3: UNCORRECTED (‘p’) DISTANCES BETWEEN THE SEVEN ITS CLADES.........126
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>19'BF</td>
<td>19'-BUTANOYLOXYFUcoxanthin</td>
</tr>
<tr>
<td>19'HF</td>
<td>19'-HEXANOYLOXYFUcoxanthin</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>bp</td>
<td>Base Pair</td>
</tr>
<tr>
<td>CTAB</td>
<td>Hexadecyltrimethylammonium bromide</td>
</tr>
<tr>
<td>CTD</td>
<td>Conductivity Temperature Depth</td>
</tr>
<tr>
<td>ctDNA</td>
<td>Chloroplast Deoxyribose Nucleic Acid</td>
</tr>
<tr>
<td>DMS</td>
<td>Dimethylsulfide</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic Acid</td>
</tr>
<tr>
<td>EM</td>
<td>Electron Microscopy</td>
</tr>
<tr>
<td>HPLC</td>
<td>High Pressure Liquid Chromatography</td>
</tr>
<tr>
<td>ITS</td>
<td>Internal Transcribed Spacer ribosomal DNA</td>
</tr>
<tr>
<td>L:D</td>
<td>Light:Dark</td>
</tr>
<tr>
<td>LM</td>
<td>Light Microscopy</td>
</tr>
<tr>
<td>LSU</td>
<td>Large Subunit ribosomal DNA</td>
</tr>
<tr>
<td>MC</td>
<td>MareChiara Sampling Station</td>
</tr>
<tr>
<td>MP</td>
<td>Most Probable Number</td>
</tr>
<tr>
<td>NJ</td>
<td>Neighbour Joining</td>
</tr>
<tr>
<td>OTU</td>
<td>Operational Taxonomic Unit</td>
</tr>
<tr>
<td>PC</td>
<td>Polycarbonate</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PPE</td>
<td>Photosynthetic Picoeukaryote</td>
</tr>
<tr>
<td>PVP</td>
<td>Polyvinylpyrrolidone</td>
</tr>
<tr>
<td>rbcL</td>
<td>Ribulose-1, 5-Bisphosphate Carboxylase Large Subunit</td>
</tr>
<tr>
<td>rDNA</td>
<td>Ribosomal Deoxyribose Nucleic Acid</td>
</tr>
<tr>
<td>SDC</td>
<td>Serial Dilution Culture</td>
</tr>
<tr>
<td>SSCP</td>
<td>Single Strand Conformation Polymorphism</td>
</tr>
<tr>
<td>SZN</td>
<td>Stazione Zoologica di Napoli</td>
</tr>
<tr>
<td>TIN</td>
<td>Total Inorganic Nitrogen</td>
</tr>
<tr>
<td>tufA</td>
<td>A plastid gene encoding the elongation factor Tu</td>
</tr>
</tbody>
</table>
Chapter One

General Introduction

THE PHYTOPLANKTON

Victor Hensen (1835-1924) was the first not only to study plankton, but also to provide practical tools to measure its abundance (Mills 1989). Along with his colleagues, Karl Brandt and Hans Lohmann, he provided a coherent hypothesis for the factors controlling the annual cycle of plankton in temperate seas. Data from studies of annual cycles in the outer Kiel Fjord, Germany and the North Sea provided the basis for Hensen’s monograph *Über die Bestimmung des Planktons oder des im Meere treibenden Materials an Pflanzen und Thieren* (1887) and described the first quantitative work on plankton, as well as new equipment for measuring plankton. This was also the establishment of the word ‘plankton’ for the floating marine animals and plants, which Hensen later described as ‘this blood of the sea’ (dies Blut des Meeres). The word ‘plankton’ was coined, at Hensen’s request, by his colleague Richard Forster (1843-1922), professor of classical philology and elocution at Kiel, and comes from the Greek meaning ‘wanderer’ or ‘drifter’. It was intended to signify both living and dead organisms of the water column, as Hensen said, “I mean thereby everything that drifts in the water, whether shallow or deep, living or dead.” Curiously, Hensen never used the word in this broad sense, instead he used it to represent only the living organisms accessible to his nets. The definition has been stabilised to mean the living component and detritus is used to describe the dead particles (Zeizschel 1978).

The phytoplankton is the autotrophic component of the plankton (‘phyton’ Greek for plant). The phytoplankton in the sea is the fundamental source of energy upon which the marine environment is based. The phytoplankton makes up some 90 % of primary
production at sea and is important even in coastal areas, where phytoplanktonic production matches that of micro- and macro-benthonic algae (Charpy-Roubaud and Sournia 1990). Using solar energy they convert carbon dioxide and water into organic compounds by means of photosynthesis. The annual cycle of phytoplankton in response to temperature, light and nutrients drives the production of the oceans. Phytoplankton has been studied by systematists since the first descriptions of dinoflagellates in 1753 (Taylor 1987) and diatoms in 1786 (Werner 1977) and quantitatively by biological oceanographers since Hensen and yet the question of what exactly makes up the plankton has still to be answered.

**SPECIES CONCEPTS AND DIVERSITY IN THE MARINE ENVIRONMENT**

There are traditionally two species concepts that have been used to describe species, these being the typological species concept and the biological species concept (Mayr 1948). The typological species concept states that species are well-circumscribed classes and separated from each other by unbridgeable gaps. Species are traditionally recognised under this concept by clear discontinuities in morphology. This concept has been used for centuries and in many groups it works exceedingly well. The biological species concept arose from observations of sympatric sibling species that cannot be separated on the basis of traditional taxonomic characters. The biological species concept (Mayr 1948), states ‘species are groups of interbreeding natural populations that are reproductively isolated from other such groups’. Species are recognised under this concept by their ability to produce viable offspring.

It has proved difficult to delimit microalgal taxa under either concept. Morphology is notoriously poor, given that there are few characters to be gained from single cells. The ultrastructure of the diatom frustule yields more characters, but their investigation often
requires EM and laborious preparations in order to observe differences. Crossing experiments in microalgae in the laboratory are fraught with experimental errors due to false results (either positive or negative). The success or otherwise of a crossing experiment in the lab does not necessarily translate to the environment. The peculiarities of the diatom life cycle make crossing experiments difficult especially in heterothallic species, such as those within *Pseudo-nitzschia*, and can easily result in false negatives because it is difficult to recreate the required conditions in the laboratory. Sexual reproduction in diatoms is also not easily observed in natural samples as gametes are often indistinguishable from vegetative cells and fusing pairs may be confused with dividing cells (Montresor and Lewis 2005). These problems with phytoplankton species have led to the advocacy of pluralism when delimiting taxa, particularly in diatoms (Mann 1999), employing as many methods as possible to collect as many data as possible. This idea is best demonstrated by the case of the freshwater epipelic diatom *Sellaphora pupula* (Kützing) Mereschkowsky. The study of diversity within this morphospecies began in the early 1980s and five new sympatric species were described in 2004 (Mann et al. 2004) using data from morphometric analyses of valve shape, from 18S, ITS and \textit{rbcL} sequences and from mating experiments.

The issue of how many species are present at sea is not purely academic. The unknown diversity hides novel metabolisms, such as photoheterotrophy in the sea (Rivkin 1987) that force a re-evaluation of carbon and energy fluxes in the oceans. These processes must be understood so that, for example, precise models of global change can be constructed. Second, the unknown microorganisms are the largest potential reservoir of useful genes for medicine and biotechnology. Knowledge of microbial diversity is also essential to understand evolution (Pedrós-Alió 2006). Species in biology are comparable to elements in chemistry. To understand the structure and functioning of an
ecosystem it is essential to know the different elements of which it is composed, i.e. the
distribution of organisms in space and time (Zeisschel 1978). It is also important to
understand diversity on several taxonomic levels, from wide groups, such as the
cyanobacteria capable of fixing nitrogen, down to the species within genera that have a
particular function or property, such as those that contain toxic and non toxic species.

The problem with assessing diversity in the marine environment has been that the
oceans were assumed to contain little of the Earth’s biodiversity. Estimates of the
number of marine species have been put forward as low as 200,000 species in total
(Briggs 1994). The question posed by Griffiths (1989) is still a pertinent; ‘Do we
classify what we see or see what we classify?’ So perhaps the most important problem
in assessing diversity in the marine environment is one of perception. If high numbers
of species are not expected, they may not be looked for. The microbial species in the
plankton have large population sizes and ease of distribution due a lack of physical
barriers and are assumed to be present wherever their habitat requirements are met
(Fenchel 2005). These large populations sizes coupled with the lack of physical barriers
theoretically remove the possibility of allopatric speciation in planktonic species and so
have lead to the conclusion that species diversity in the plankton should be low.
However, step-wise advances in the description of planktonic diversity have been made
with the advent of new methods for identification. The first increase in species
descriptions coincided with the invention of the light microscope (references in Taylor
1987 and Werner 1977), a further increase with the electron microscope (e.g. Johnson
and Sieburth 1982 or Joint and Pipe 1984), and then more recently with the
development of genetic techniques for analysing DNA extracted from natural samples
(Delong et al. 1989; Diez et al. 2001a; Diez et al. 2001b; Fuller et al. 2006a; Guillou et
al. 2004; Moon-Van Der Staay et al. 2001; Moon-Van Der Staay et al. 2000; Rappé et
It has been noted that photosynthetic organisms in the terrestrial environment are characterised by few classes with many species whereas the marine realm is phylogenetically diverse, but species poor, particularly for the unicellular algae (Andersen 1992; Baldauf 2003; Briggs 1994; Vincent and Clarke 1995). The differences in photosynthetic apparatus between algal classes can be greater than the difference found between angiosperms, gymnosperms, ferns and mosses combined (Andersen 1992). There are roughly 18 classes with marine planktonic representatives, some of which contain only 1 or 2 species (e.g. Chlorarachniophycaceae and Rhodophyceae Sournia et al. 1991) as compared to Dicotyledones that contains around 220,000 species. The most important planktonic classes in terms of numbers of species...
are the diatoms (Bacillariophyceae) and the dinoflagellates (Dinophyceae) followed by the Prymnesiophyceae, which includes the coccolithophorids. These are followed by the Cryptophyceae, and the green lineage Prasinophyceae that is divided into several different clades in both 16S (Fuller et al. 2006a) and 18S rDNA (Guillou et al. 2004). The Chrysophyceae was formerly a large class including several groups that have since been divided into new classes and is generally considered to be more abundant in freshwater systems but has some representatives in the marine environment. After these are several recently erected classes such as Pelagophyceae and Bolidophyceae that so far contain only a few species. The phylogenetic diversity between these classes can be seen from the analysis of 16S rDNA (Figure 1.1). Many classes are distantly related, falling in the major clades following secondary (e.g. Bacillariophyceae) and tertiary (Dinophyceae) endosymbiosis events.

The disparity between the number of species on land and in the sea is due to a number of factors including problems of scale, sampling and identification. The problem of scale manifests itself in a lack of understanding of the factors affecting microorganisms, as Andersen (1992) points out ‘if humans could shrink to the height of 10 μm to explore the aquatic realm, then our knowledge of microorganisms would increase in quantum leaps.’ As most representatives of the photosynthetic classes are in the nanoplanckton size fraction, i.e. smaller than 20 μm (Sournia et al. 1991), this is particularly important when studying the phytoplankton. Studies have shown that the differences in abundance and species composition can change on the scale of centimetres (Owen 1989). These problems of fine scales affect sampling and identification. Plankton is difficult to sample due to its temporal and spatial fluidity. The identification of species of several phytoplankton groups is made almost impossible, especially using light microscopy, due to a paucity of morphological characters. Phytoplankton species have been said to
have a convergent morphology in that species from different phylogenetic lineages have very similar morphologies, essentially little balls differing only in the colour of their pigments (Potter et al. 1997). In many groups, the only way of distinguishing closely related taxa is using electron microscopy, which is time consuming and so unsuitable for monitoring programmes. The advent of molecular techniques and a more careful examination of morphology have lead to the discovery of hidden or cryptic diversity within phytoplankton groups that challenge the idea that marine microbial diversity is low.

**CRYPTIC, PSEUDO-CRYPTIC AND HIDDEN DIVERSITY**

Diversity can be measured and defined in a number of ways. It can be measured as the number of different species or as a function of the abundance ratios of different taxa. It can be thought of as the total number of species including rarely observed taxa or as only those that are abundant and make a regular or significant contribution to the energy flow of a particular ecosystem. The first can be called biodiversity and is difficult to measure because the rare taxa may be missed. Diversity can be defined as the components that are active and abundant at one particular time and place (Pedrós-Alió 2006). Studies of a particular site over time give more useful information about diversity, the words of a particular story, than snap shots along transects. Single site studies over time give the core species that are present every year with the occasional species that are rarely found. Core species are responsible for carbon and energy flow in the ecosystem whereas occasional species wander sporadically to areas at their limits of tolerance and eventually disappear (Pedrós-Alió 2006), but are useful in terms of allowing flexibility and adaptation to changes in conditions.

Three kinds of diversity are mentioned within this thesis, hidden, cryptic, and pseudo-
cryptic diversity. Hidden diversity is the unexpected diversity that has been revealed by the application of genetic methods. Cryptic diversity has been recognised in a number of groups and involves genetic differentiation to the point where sexual reproduction is not possible, but where morphological differences cannot be detected. Pseudo-cryptic diversity is where a subtle but consistent morphological difference can be found after genetic studies reveal the need to look for it. This is best illustrated by the case of the species within the diatom genus *Skeletonema* Greville (Kooistra et al. in press; Sarno et al. in press; Sarno et al. 2005; Zingone et al. 2005). *Skeletonema costatum* (Greville) Cleve emend. Zingone et Sarno was described on the basis of a set of morphological characters that genetic studies showed were masking several distinct taxa. Only once the genetic studies revealed these taxa could the morphology be re-examined and new more useful characters chosen to distinguish the various taxa. In the case of some of the species formerly classified within *S. costatum*, these morphological characters are only visible in electron microscopy and so the taxa are referred to as pseudo-cryptic to highlight the difficulty in distinguishing them.

**TOOLS FOR ASSESSING GENETIC DIVERSITY AT SEA**

In many cases the entities discovered by genetic methods can only be referred to as OTUs (Operational Taxonomic Units). Their identity cannot be determined to any taxonomic level due to a lack of sequences from known taxa that can be used for comparison. Roughly speaking genetic methods have been used in two broad ways, either investigating strains of known species for the purposes of taxonomy, phylogenetics or population genetics, or using community DNA to assess diversity within broad groups and determine the distribution of known taxa.

Amplified fragment length polymorphisms (AFLPs) were used to examine diversity
within the prymnesiophyte *Emiliania huxleyi* (Iglesias-Rodríguez et al. 2002) where two major taxonomic clades of *E. huxleyi* were identified. The first was associated with the Northern Hemisphere blooms, while the other was found in the Southern Hemisphere. However, no two clones shared the same genotypes indicating a high level of diversity within this cosmopolitan species.

Randomly amplified polymorphic DNA (RAPD) has been used to investigate diversity in the dinoflagellates *Prorocentrum micans* Ehrenberg (Shankle et al. 2004) and *Gymnodinium catenatum* Graham (Bolch et al. 1999). In *P. micans* blooms were sampled over a two year period and a significant difference between two groups was found. The first group made up about 40% of all samples and was the common member of the blooms during the sampled period. The second group contained isolates that were genetically diverse and changed over time. The work on *G. catenatum* was an attempt to identify the source population for the Tasmanian blooms, as it is hypothesized that *G. catenatum* was introduced into this area via ships ballast. Results showed that the Tasmanian samples were equally related to those from Spain/Portugal and from Japan, making the source of the introduced population equivocal. However, differences were detected within the Tasmanian populations showing the blooms were composed of localized estuary bound subpopulations with limited genetic exchange between them.

At a finer level, microsatellites have been used to examine population level differences in *Ditylum brightwellii* (T. West) Grunow in Van Heurck (Rynearson and Armbrust 2000; Rynearson and Armbrust 2004; Rynearson and Armbrust 2005; Rynearson et al. 2006), in *Pseudo-nitzschia multiseries* (Hasle) Hasle (Evans et al. 2004) and in *P. pungens* (Grunow ex Cleve) Hasle (Evans and Hayes 2004; Evans et al. 2005). The work on *D. brightwellii* revealed a structure to the population over a relatively small
geographic scale caused by hydrology. The population with Puget Sound was isolated from that outside the Sound by a narrow shallow sill at the mouth of the Sound. There was also a temporal structure, with the bloom between February and April being different from that in May and June. These two populations also showed differences in rDNA ITS (internal transcribed spacer) and average cell size. The work on *Pseudo-nitzschia* (Peragallo) Hasle showed that isolates of *P. multiseries* collected 15 km apart had more similarities than those isolated from more than 200 km apart, indicative of some biogeographic structure to the population (Evans et al. 2004). *P. pungens* in the North Sea was shown to be a panmictic population with a high level of diversity at the population level indicating frequent interbreeding (Evans and Hayes 2004; Evans et al. 2005).

Taxonomic studies investigating problematic species and complexes have used mainly the rDNA markers small subunit 18S and ITS to clarify species boundaries. For example, various studies revealed several taxa in previously described morphological species in the diatom genus *Pseudo-nitzschia* (Peragallo) Hasle (Lundholm et al. 2002a; Lundholm et al. 2002b; Lundholm and Moestrup 2002; Lundholm et al. 2003; Lundholm et al. 2006; Orsini et al. 2004; Orsini et al. 2002) and showed the biogeographical limits of genotypes within the prasinophyte *Micromonas pusilla* (Butcher) Manton et Parke (Šlapeta et al. 2006).

Community DNA, extracted from large volumes of filtered seawater, has been analysed using universal primers for a number of markers. 18S has been used to construct clone libraries (Diez et al. 2001b; Guillou et al. 2004; López-Gracia et al. 2001; Moon-Van Der Staay et al. 2001), with DGGE (denaturing gradient gel electrophoresis) to separate out fragments on the basis of differences in their sequence and therefore migration
through a gel (Diez et al. 2001a; van Hannen et al. 1998; Zeidner and Béjà 2004) and for probes targeting various groups (Moon-Van Der Staay et al. 2000). These studies revealed a high level of diversity both at the level of numbers of different OTUs and at the level of novel lineages. Fluorescence in situ hybridisation (FISH) probes were constructed against some of these novel lineages and used to give some information on the morphology of the cells belonging to each lineage (Massana et al. 2002). Analyses of the sequences generated by the clone libraries show examples of both cosmopolitan and biogeographically restricted OTUs (Massana et al. 2004). The photosynthetic marker *psbA*, encoding for protein D1 of photosystem II reaction centre, was used on BAC (Bacterial Artificial Chromosome) library samples from the Red Sea, the Mediterranean and the Hawaii Ocean Time-series (HOT) to target specifically the photosynthetic component. This revealed the presence of a high number of green algal sequences belonging to the Prasinophyceae, but also sequences from the Prymnesiophyceae and Cryptophyceae (Zeidner et al. 2003). The chloroplast rDNA marker 16S has been used to construct probes against a number of photosynthetic classes showing the importance of the Chrysophyceae along an Arabian Sea transect (Fuller et al. 2006b). A recent paper using the *rbcL* (ribulose-1, 5-bisphosphate carboxylase Large subunit) marker to construct clone libraries showed the diversity within Jiaozhou Bay, China and how diversity changes during the year (Yongjian et al. 2006).

### Diversity in Phytoplankton on a Temporal Scale

The studies of genetic diversity in the marine environment mentioned above have generally concentrated on spatial scales, often using transects (Fuller et al. 2006a, b) or single sampling dates (Diez et al. 2001a; Diez et al. 2001b). Temporal studies of phytoplankton have for the most part used bulk methods such as ocean colour or
chlorophyll concentrations to focus on primary production and phytoplankton biomass changes during the year (e.g. Dandonneau et al. 2004). Phytoplankton may be divided into broad groups such as diatoms, dinoflagellates and nanoflagellates. Temporal diversity has been assessed using flow cytometry (Bec et al. 2005; Calvo-Díaz et al. 2004; Grégori et al. 2001; Jacquet et al. 2002) or HPLC (high pressure liquid chromatography) to investigate changes in groups detected using accessory photosynthetic pigments as markers (Letelier et al. 1993; Marty and Chiaverini 2002; Marty et al. 2002; Suzuki et al. 2002). These studies give an idea of the main groups involved in seasonal dynamics of phytoplankton, but can be confusing due to difficulties in assigning one particular group to a set of pigments. Not all classes have an unequivocal marker pigment and not all members of a class may contain a particular pigment. There are also differences in pigment concentrations and complements depending on the physiological state of the cell and on the depth from which it was sampled (Andersen et al. 1996). Data on the seasonal distribution of different groups come from long-term sampling programmes such as in Helgoland (Wiltshire and Dürselen 2004), the Gulf of Naples (Ribera d'Alcalá et al. 2004; Zingone et al. 2002; Zingone and Sarno 2001; Zingone et al. 1999b) and the Continuous Plankton Recorder (Richardson et al. 2006) and from short term studies over the course of one or two years (Caroppo et al. 1999; Charles et al. 2005; Gailhard et al. 2002; Liu et al. 2002; Mercado et al. 2005; Rodríguez et al. 2003; Sackmann et al. 2004; Shankle et al. 2004). Data from the Helgoland Reede time series show the changes in the occurrence of certain species over a thirty year period, for example the slight shifts in the timing of the bloom of the diatom Guinardia delicatula (Cleve) Hasle over time (Wiltshire and Dürselen 2004). The BATS (Bermuda Atlantic Time Series), HOT (Hawaii Ocean Time-series) and DYFAMED (Dynamics of Atmospheric Fluxes in the Mediterranean Sea) time series all have a biogeochemical and physical focus (Karl and Michaels 1996; Marty
and so work on the phytoplankton communities has not been concentrated on diversity of organisms at these sites. Several studies over limited time periods during the course of the long-term programmes have shown the seasonal distribution of groups such as the diatoms (Scharek et al. 1999), 19'Hf containing nanoflagellates (Marty et al. 2002) and picoplankton (Campbell and Vaulot 1993) and measured the seasonal variations in primary productivity (Marty and Chiavérini 2002).

**Gulf of Naples and the MareChiara Long Term Ecological Programme**

The Gulf of Naples has been the site of marine biological study since the opening of the Stazione Zoologica di Napoli by the Polish Darwinist Anton Dohrn in 1873. The station was intended as a laboratory for visiting scientists to collect material, make observations and conduct experiments taking advantage of the institute’s facilities, while being free to pursue their own ideas and projects. Anton Dohrn funded the Stazione by introducing the ‘Bench’ system, where space could be rented by anyone for annual fee. A system that worked so well, it has been applied across the world and is still in use today.

The Gulf of Naples is characterised by two subsystems, one consisting of open waters influenced by the Tyrrhenian Sea and the other of coastal waters influenced by runoff from the city of Naples and the surrounding conurbation (Carrada et al. 1981; Carrada et al. 1980; Marino et al. 1984; Ribera d'Alcalà et al. 2004). Water enters the Gulf from the Tyrrhenian Sea from the south between the island of Capri and the Sorrento peninsular and exits to the north between Monte di Procida and the islands of Procida and Ischia. The Gulf also has inputs from the heavily polluted Sarno river, the city of Naples and the surrounding area. The MareChiara sampling station (MC) is situated on
the border between the two subsystems at about the 100m isobath 1 km from the coast (40°48.5'N, 14°15'E) (Ribera d'Alcalà et al. 2004). The inputs influencing this site change with the seasons, meteorological and hydrological conditions. The border between the two Gulf subsystems shifts with water movements, whereby the MC site may at different times of the year have the characteristics of either subsystem.

The MC station has been sampled fortnightly between 1984 and 1991 and then every week since 1995 giving a data set including CTD (conductivity temperature density), nutrient concentrations, pigment concentrations, phytoplankton and zooplankton counts.

Figure 1.2: Map of the Gulf of Naples showing the Marechiara (MC) sampling station. Modified from Ribera d'Alcalà et al. (2004).

This data set was used during this thesis for comparison and to put any results into context. The phytoplankton counts are carried out on samples fixed in 0.8-1.6 % neutralised formaldehyde. Cells are counted using an inverted microscope after
sedimentation of variable volumes of seawater depending on cell concentration. Cells are counted at 400X magnification from two transects representing about 1/30 of the whole of the bottom of the sedimentation chamber. Cells smaller than 2 μm are generally not detected unless very abundant (Ribera d'Alcalà et al. 2004). Diatoms and small flagellates of around 4 μm are the most abundant groups recorded at the MC station (Ribera d'Alcalà et al. 2004). In summer, diatoms may make up more than 70% of the sample (Zingone et al. 1990), while flagellates are the dominant group during the winter when cell numbers are low (Ribera d'Alcalà et al. 2004). Hence these two groups will be the focus of this thesis. One of the most important findings from the MC data set is the distinct seasonality of a number of species from diatoms to dinoflagellates to copepods (Ribera d'Alcalà et al. 2004; Zingone and Sarno 2001). The assumption from these data is that diversity is linked to seasonality by the ecological requirements of the different species and congeneric species will have a distinct seasonal distribution. This is an assumption that will also be tested during this thesis.

The small sized phytoflagellates have been studied in terms of single species (e.g. Cerino 2004; Cerino and Zingone in press; Throndsen and Zingone 1994; Zingone et al. 1999a; Zingone et al. 1999b) using cultures isolated using the serial dilution culture (SDC) method (Throndsen 1995), of which the following are a few examples. The seasonal occurrence of the prasinophyte *Micromonas pusilla* has been studied using serial dilution cultures and epiflorescence microscopy. This species shows a mainly winter distribution with a yearly maximum in March, which sharply declines towards the end of spring and this species is undetectable during the summer months (Zingone et al. 1999b). The prymnesiophyte genus *Phaeocystis* Lagerheim has been studied from a number of cultures isolated in the Gulf of Naples leading to the description of two new species, *P. jahnii* Zingone and *P. cordata* Zingone et Chrétiennot-Dinet (Zingone et al. 24
Species within the Cryptophyceae have been investigated genetically and morphological using 18S and EM and single strains isolated at different times of the year using SDC (Cerino 2004; Cerino and Zingone in press). The major groups contributing to this fraction have been investigated using SDC (Zingone et al. 1992) and HPLC but most of these data are unpublished.

Diatoms have been the subject of more intensive study in the Gulf of Naples leading to the description of new species (e.g. De Stefano and Marino 2001) and the reorganisation of the genus *Skeletonema* (Kooistra et al. in press; Sarno et al. in press; Sarno et al. 2005; Zingone et al. 2005) as well as studies of the dynamics of various species during the year (Ribera d'Alcalà et al. 2004; Zingone et al. 2002; Zingone et al. 2003; Zingone et al. 1990; Zingone and Sarno 2001). A number of toxic species are found in the Gulf, mainly within the genus *Pseudo-nitzschia* leading to a focus on this genus (Amato et al. in prep; Amato et al. submitted; Orsini et al. 2004; Orsini et al. 2002). Several newly described species within this genus are difficult to recognise in light microscopy (Lundholm et al. 2003; Lundholm et al. 2006). Members of the more easily recognised species are still difficult to identify if found as single cells. As such, this genus will also be studied during this thesis examining new ways of tracking taxa using genetic methods that do not suffer from the problems of identifying morphology in LM.

**AIMS**

The main aim of this thesis is to explore the most important taxa present during the year in the Gulf of Naples, the major players. The focus of this thesis is an assessment of genetic diversity at three taxonomic levels on a temporal scale. The aim is two fold, to examine the seasonality of different taxa during the year and to investigate the diversity within these taxa at different times of the year. The classes present within the <5 μm
fraction of the phytoplankton will be examined using class level 16S probes. The diatom genus *Pseudo-nitzschia* will be studied at the level of genus to investigate the number of taxa that can be recognised by genetic methods and the seasonal occurrence of these taxa. The species *Pseudo-nitzschia galaxiae* Moestrup et Lundholm will be studied at the intra-specific level to examine seasonal dynamics of different blooms of the same morphological taxon during the year.
Chapter Two

_Ultraphytoeukaryote Diversity in the Gulf of Naples over an Annual Cycle_

**INTRODUCTION**

Small cells between 1 and 10 μm have been observed in the plankton for more than a century (see Johnson and Sieburth 1982 and references therein). The study of these small cells could be said to have begun in the Mediterranean. Some of the earliest documentation on small flagellates comes from studies by Lohmann between 1900 and 1910 in the Straits of Messina in the Mediterranean and at Kiel in Germany (Mills 1989; Throndsen and Zingone 1994) using either filter paper or centrifugation rather than plankton nets. The morphology of several ‘classical’ species was described by examining cultures from the Mediterranean (Throndsen and Zingone 1994). The advent of electron microscopy enabled the examination of morphology at a finer scale (Joint and Pipe 1984) and led to the description of several new species, for example within the prymnesiophyte genus *Chrysochromulina* (Leadbeater 1972a; Leadbeater 1972b).

Primary production was measured first from cultures isolated using SDC (Throndsen 1976). Further studies led to estimates for small eukaryotes of up to 80% of primary production in oligotrophic waters (Li et al. 1983), although it was work by Lohmann that first highlighted these small cells as the base of the food chain (cited in Mills 1989).

The first paper listed in the ISI Web of Knowledge database to use the term picoplankton, defined as <2 μm, is Johnson and Sieburth (1982) describing eukaryotic phototrophs of bacterial size in the marine environment. With the development of flow cytometry in the 1970s, these small photosynthetic cells could be identified in natural samples in bulk (Li and Wood 1988), and HPLC enabled their accessory pigments to be assessed giving an approximate classification based on diagnostic marker pigments for
each algal class. Given the abundance and importance of small eukaryotes in the world’s oceans, detailed assessments of the taxonomy and diversity of this fraction are needed in order to understand the dynamics and ecology of the marine environment.

However, the assessment of the diversity within the smallest size fractions of the plankton is difficult for a number of reasons. Their morphology is simple, most are little ‘balls’ (Potter et al. 1997) and although their pigments can give some idea of taxonomic class, diagnosis using chemotaxonomy is often equivocal. The Cryptophyceae are the only class with an unequivocal marker accessory pigment, alloxanthin contained within all members of the class so far tested. Most other algal classes have combinations of shared pigments requiring algorithms, such as those employed by CHEMTAX (Mackey et al. 1998), to distinguish the different groups based on the proportion of different pigments to chlorophyll \( a \). However as mentioned in Chapter One, the use of pigments is problematic in that not all members of a class contain the same pigments in the same ratios. Enrichment and SDC can give a biased view of the diversity in that they capture only those organisms that grow well in artificial cultures (Johnson and Sieburth 1982). Genetic studies in the last decade or so have highlighted the taxonomic diversity of these small cells. 16S rDNA clone libraries targeting bacteria revealed the presence of six algal plastid sequences belonging to the chromophyte lineage including heterokonts, prymnesiophytes and cryptophytes (Rappé et al. 1995). Potter et al. (1997) used 18S rDNA to identify the lineages of nine undescribed algal strains finding that they belonged to three distinct eukaryotic phylogenetic lineages: the heterokont, prymnesiophyte and green algae. Further studies using universal 18S primers on community DNA revealed the importance of different classes within the small size fractions. For example, a study of samples from the Mediterranean showed the importance of the Prasinophyceae, but also the presence of sequences from the
Prymnesiophyceae, Dinophyceae, Pelagophyceae and Cryptophyceae (Diez et al. 2001a; Diez et al. 2001b). Molecular studies have revealed the presence of new taxonomic classes, for instance the Bolidophyceae (Guillou et al. 1999a; Guillou et al. 1999b) and novel lineages within the stramenopiles (Massana et al. 2004; Massana et al. 2002). However, despite these advances in genetic methods, even a recent issue of Vie et Milieu dedicated to picoplankton (2005, Volume 55, Issue 3/4) did not have a single article on the diversity of this fraction, but used the size fraction as a category. Another limitation is that few of these studies have concentrated on the temporal scale, opting for assessing diversity along transects giving snapshots of the presence of major groups.

Although the small size fraction in the Gulf of Naples is numerically dominated by the cyanobacteria, in terms of biomass small eukaryotes are more important (Casotti et al. 2001). The category of 'undetermined phytoflagellates <10 μm' has been recognised during the course of the MC Long-Term Ecological Programme as making a significant contribution in terms of numbers and biomass (Carrada et al. 1980; Marino et al. 1984; Ribera d'Alcalà et al. 2004). Small flagellates of around 4 μm, along with diatoms, are by far the dominant group throughout the year at the MC station (Ribera d'Alcalà et al. 2004). As far as biomass is concerned, these small flagellates make the third largest contribution after the diatoms and the dinoflagellates. Although small flagellates have been shown to be important in the Gulf of Naples for some time (Carrada et al. 1981), small coccoids belonging to unidentified algal groups have been more frequently observed at MC in the last few years (Ribera d'Alcalà et al. 2004). This increase in small cells suggests that a change is occurring in the trophic conditions of the Gulf of Naples with an increase in the relative importance of the microbial loop (Ribera d'Alcalà et al. 2004). It is therefore important to investigate the identity and temporal
occurrence of the major groups of small eukaryotes in the Gulf of Naples in order to understand this ecosystem (Carrada et al. 1981).

The definition and nomenclature for the various size fractions within the phytoplankton can be confusing. Pico-, ultra- and nano- are prefixes with varying definitions in terms of cell size. For this study, the size fraction $<$5 μm and the prefix ultra- were chosen for the following reasons. The average cell size of the undetermined flagellates observed in the Gulf of Naples is around 4 μm (D. Sarno pers. comm.). As a project investigating the pigment composition and primary production of the $<$5 μm fraction was already underway (Santarpia 2005), it was decided to investigate the same size fraction using available 16S probes in order to determine the major photosynthetic classes. This fraction will be referred to as the ultraplankton following the definition used by Li (Casotti et al. 2001; Li et al. 1993) in order to avoid confusion with the term picoplankton that is usually defined as $<$3 μm. As there is a lack of genetic studies on a temporal scale, the changes in diversity within the ultraplankton at a single sampling site were investigated over a period of fifteen months. These changes were assessed using the class level probes to investigate whether there was a seasonal pattern of taxonomic classes within the ultraplankton. Clone libraries were used to give an idea of the diversity present in the Gulf of Naples within those taxonomic classes and to investigate whether different OTUs were present at different times of the year.

MATERIALS AND METHODS

Characterisation of the MC sampling site

The MC sampling site is sampled every week by members of the Area Gestione Ambiente ed Ecologia Costiera delle Aree Temperate e Polari SZN as part of the Long Term MC Ecological Programme. Data were collected for use by colleagues at the
SZN, using a SBE 911plus CTD with salinity, temperature and oxygen sensors and a SCUFA fluorometer. The data were processed using SeaSave Data Processing software. The CTD was connected to the automatic Carousel sampler equipped with twelve 12 L Niskin bottles used to collect the water samples. Samples for nutrient analyses were collected in high density polyethylene vials directly from the Niskin bottles and immediately stored at -20 °C until analyses in the laboratory. LM counts of phytoplankton species found within samples, also collected using the Niskin bottles, from 0 m fixed using the Utermöhl method (Utermöhl 1958) are carried out by Diana Sarno of the Servizio Speciale Tassonomia e Identificazione del Fitoplankton Marino, SZN. These data are presented at the beginning of the results section to illustrate the conditions found at the MC station during the study period. The numbers of photoflagellates <10 μm observed in LM are displayed with the clone library data in the Clone Libraries section to illustrate the abundances of cells when the libraries were constructed.

**Sampling**

Samples were collected from 0 m at the MC station from July 2003 to September 2004 approximately once a week, using the twelve 12 L Niskin bottles mounted on the Carousel sampler. Five litres of seawater were pre-filtered through a 90 mm diameter polycarbonate (PC) with a 5 μm pore size filter using a vacuum pump at 200 mmHg. The filtrate was filtered through a 90 mm PC filter of 0.2 μm pore size for one hour. The 0.2 μm filter was cut into sections, placed in Eppendorf tubes, frozen immediately in liquid nitrogen and stored at -80 °C.

**DNA extraction and amplification of environmental samples**

DNA was extracted from the stored filter using CTAB extraction buffer (2 % CTAB, 200 mM Tris HCl pH 8.0, 50 mM EDTA, 1.4 M NaCl and 2.5 % PVP) modified from Doyle and Doyle (1987) for 45 min at 65 °C before protein removal using at least two
washes with chloroform-isoamyl alcohol (24:1). DNA was precipitated using isopropanol at -20 °C for one hour. The 16S rRNA region was amplified using the marine algal plastid biased primer PLA491F and the reverse primer OXY1313R (see Table 2.1) designed to be biased towards photosynthetic eukaryotes and exclude cyanobacteria.

Table 2.1: Primers used in this chapter.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequences 5'-3'</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLA491F</td>
<td>GAG GAA TAA GCA TCG GCT AA</td>
<td>Fuller et al. (2006a)</td>
</tr>
<tr>
<td>OXY359F</td>
<td>GGG GAA TYT TCC GCA ATG GG</td>
<td>West and Scanlan (1999)</td>
</tr>
<tr>
<td>OXY1313R</td>
<td>CTT CAY GYA GGC GAG TTG CAG C</td>
<td>West and Scanlan (1999)</td>
</tr>
</tbody>
</table>

Amplification was carried out in a total reaction volume of 50 µl containing 1 µl of environmental DNA (approximately 40 ng), 200 µM deoxynucleoside triphosphates, 1.1 mM MgCl₂, 1 mg ml⁻¹ BSA (bovine serum albumin, Sigma), 1 µM each primer and 2.5 U Taq polymerase in 1× enzyme buffer -Mg (Roche Diagnostics GmbH, Mannheim, Germany). The reaction conditions comprised 80 °C for 4 min followed by 30 cycles of 94 °C for 30 sec, 62.2 °C for 30 sec and 72 °C for 40 sec, with a final extension of 5 min at 72 °C. PCR products were purified using QIAquick PCR purification kit (Qiagen Ltd.) following the manufacturers instructions, quantified on an agarose gel and stored at -80 °C.

**Preparation and amplification of control DNA**

Control DNA for the dot blot hybridisations was prepared from PCR products from strains corresponding to each specific probe obtained from the Provasoli-Guillard National Center for Culture of Marine Phytoplankton (CCMP) and the Roscoff Culture Collection (RCC) (purified 16S PCR products kindly provided by Nick Fuller, for a list of strains used see Table 2.2) and amplified using the same primer pair and PCR conditions as for the environmental samples before purification in the same manner.
Dot blot hybridisation

PCR products were denatured using 2 M NaOH and blotted onto nylon membranes (Zetaprobe, Bio-Rad Laboratories Ltd, Hemel Hempstead, Herts, United Kingdom) and the DNA was then cross-linked to the membrane using a Stratalinker® 2400 UV Crosslinker (La Jolla, California). For the environmental samples, 30 ng of PCR product was blotted in triplicate. For the control DNAs, a dilution series ranging from 2.5 ng to 40 ng was blotted (2.5, 5, 10, 20, 30 and 40 ng). Blots were pre-hybridised in 5 ml Z-hyb buffer (1 mM EDTA, 0.5 M Na2HPO4 [pH 7.2], 7 % [wt/vol] SDS) for at least one hour at 37 °C.

Table 2.2: Strains used as Control DNAs and the corresponding probes, for details of these strains see Fuller et al. (2006a), the Prasinophyceae clades refer to different clades found within this class in the 18S phylogeny (Guillou et al. 2004).

<table>
<thead>
<tr>
<th>Class</th>
<th>Strain</th>
<th>Probe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bolidophyceae</td>
<td>RCC238</td>
<td>CHLA768</td>
</tr>
<tr>
<td>Chlorarachniophyceae</td>
<td>CCMP621</td>
<td>CHLA768</td>
</tr>
<tr>
<td>Chrysophyceae</td>
<td>CCMP296</td>
<td>CHRY1037</td>
</tr>
<tr>
<td>Chrysophyceae</td>
<td>RCC21</td>
<td>CHRY1037</td>
</tr>
<tr>
<td>Cryptophyceae</td>
<td>CCMP1868</td>
<td>CRYP862</td>
</tr>
<tr>
<td>Bacillariophyceae</td>
<td><em>Thalassiosira weissflogii</em></td>
<td></td>
</tr>
<tr>
<td>Dictyochophyceae</td>
<td>RCC382</td>
<td></td>
</tr>
<tr>
<td>Euglenophyceae</td>
<td>CCMP390</td>
<td></td>
</tr>
<tr>
<td>Eustigmatophyceae</td>
<td>RCC92</td>
<td>EUST985</td>
</tr>
<tr>
<td>Eustigmatophyceae</td>
<td>RCC438</td>
<td>EUST985</td>
</tr>
<tr>
<td>Pavlovophyceae</td>
<td>CCMP609</td>
<td>PAVL665</td>
</tr>
<tr>
<td>Pelagophyceae</td>
<td>RCC100</td>
<td>PELA1035</td>
</tr>
<tr>
<td>Pelagophyceae</td>
<td>RCC286</td>
<td></td>
</tr>
<tr>
<td>Pinguiophyceae</td>
<td>RCC503</td>
<td>PING1024</td>
</tr>
<tr>
<td>Prasinophyceae clade II</td>
<td>RCC344</td>
<td>PRAS540</td>
</tr>
<tr>
<td>Prasinophyceae III</td>
<td>RCC618</td>
<td></td>
</tr>
<tr>
<td>Prasinophyceae IV</td>
<td>RCC500</td>
<td></td>
</tr>
<tr>
<td>Prasinophyceae V</td>
<td>RCC261</td>
<td></td>
</tr>
<tr>
<td>Prasinophyceae VI</td>
<td>RCC137</td>
<td>PRAS826</td>
</tr>
<tr>
<td>Prasinophyceae VII</td>
<td>RCC287</td>
<td></td>
</tr>
<tr>
<td>Prymnesiophyceae</td>
<td>RCC186</td>
<td>PRYM666</td>
</tr>
<tr>
<td>Trebouxiophyceae</td>
<td>CCMP243</td>
<td>TREB708</td>
</tr>
<tr>
<td>Trebouxiophyceae</td>
<td>CCMP508</td>
<td></td>
</tr>
<tr>
<td>Probe</td>
<td>$T_d$ (°C)</td>
<td>Sequence (5’-3’)</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>CHLA768</td>
<td>49</td>
<td>CCA TTC TCT CCC CTC GCT</td>
</tr>
<tr>
<td>CHRY1037</td>
<td>52</td>
<td>GCA CCA CCT GTG TAA GAG</td>
</tr>
<tr>
<td>CRYP862</td>
<td>42</td>
<td>CCG TCA ATT CCT TTG AGT TT</td>
</tr>
<tr>
<td>EUB908</td>
<td>35</td>
<td>CCG TCA ATT CCT TTG AGT TT</td>
</tr>
<tr>
<td>EUST985</td>
<td>49</td>
<td>CAC TTC TAG CAA ACC CTG</td>
</tr>
<tr>
<td>PAVL665</td>
<td>37</td>
<td>TAG AAA TCC CTC CTA CCC</td>
</tr>
<tr>
<td>PELA1035</td>
<td>52</td>
<td>ACC ACC TGT GTG TGT CTA</td>
</tr>
<tr>
<td>PING1024</td>
<td>47</td>
<td>ACG TAT TCC TTA CGG CAC</td>
</tr>
<tr>
<td>PRAS540</td>
<td>50</td>
<td>CGA ATA ACA CTT GCA TCC</td>
</tr>
<tr>
<td>PRAS826</td>
<td>58</td>
<td>GAT TCG CGT ATC CCC TAG</td>
</tr>
<tr>
<td>PRYM666</td>
<td>44</td>
<td>CTA GAA ATT CCC TCT ACC</td>
</tr>
<tr>
<td>TREB708</td>
<td>44</td>
<td>CCT TTG GTG TTC CTC CCG</td>
</tr>
<tr>
<td>SYN1258</td>
<td>47</td>
<td>TTG TCC TCG CGA ACT TGC</td>
</tr>
<tr>
<td>S1PRO634R</td>
<td>50</td>
<td>GCC GAT CAG TTT CCA CTG</td>
</tr>
<tr>
<td>S2PRO640R</td>
<td>41</td>
<td>ATA CTC AAG CCT TTG AGT TT</td>
</tr>
</tbody>
</table>

The membranes were hybridised at 30 °C overnight in fresh buffer containing a $^{32}$P radioactively labelled oligonucleotide (see Table 2.3 for a list of probes used). Three 15 min washes at 30 °C were carried out using wash buffer (0.2 x SSPE, 0.1 % [wt/vol] SDS). A 10 min stringency wash was carried out using pre-warmed wash buffer at the temperature determined by the wash curves (Fuller et al. 2006b). The dissociation
temperatures (T_d) used for each probe are displayed in Table 2.3. The membranes labelled with the class specific probes were exposed overnight in phosphorimager cassettes and those labelled with the eubacterial probe were only exposed for five hours. The amount of hybridisation was quantified using a Fujifilm FLA-5000 phosphorimager and Total Lab software (Nonlinear Dynamics Ltd.). The probes were stripped from membranes to enable reuse using two 20 min washes with buffer solution (0.1x SSPE, 0.5 % SDS) at 95 °C. The relative hybridisation of the picoeukaryote class specific probes to the total oxygenic phototroph 16S rDNA was calculated using the following equation:

\[
\% \text{ relative hybridisation} = \left( \frac{S_{\text{env}}}{E_{\text{env}}} \right) \cdot \left( \frac{S_{\text{con}}}{E_{\text{con}}} \right)^{-1} \times 100
\]

where \( S_{\text{env}} \) and \( E_{\text{env}} \) represent hybridisation to environmental DNA of the specific and eubacterial probes, respectively, and \( S_{\text{con}} \) and \( E_{\text{con}} \) are the slopes of the specific and eubacterial probe-binding curves, respectively, calculated by the dilution series of homogenous control DNA. Percentage relative hybridisation was calculated for each environmental replicate and averaged to produce the mean % relative hybridisation for each sample. Values of below 2 % were considered to be background noise.

**Construction of clone libraries**

Six clone libraries were constructed using a TOPO TA Cloning® kit (Invitrogen™ Life Technologies, Carlsbad, California) for 28th October (library MC597) and 1st December 2003 (MC601), 9th February (MAMA64), 25th March (250304), 17th May (MC615) and 6th July (MC622) 2004 (see Results Table 2.5 for numbers of clones examined) from PCR products obtained from the PLA491F-OXY1313R primer pair. Plasmid purification was obtained in high throughput using the Millipore Montage Plasmid Miniprep Kit (Millipore Corporate 290 Concord Rd. Billerica, MA 0182 - USA) and a robotic station, Beckman Coulter’s Biomek® FX Laboratory Automation Workstation,
equipped with ORCA® robotic arm (Beckman Coulter, Fullerton, CA) using one 96 well plate per library. Sequence reactions were obtained with the BigDye Terminator Cycle Sequencing technology (Applied Biosystems, Foster City, CA), purified in automation using the Millipore Montage SEQ96 Sequencing Reaction Cleanup Kit (Millipore Corporate 290 Concord Rd. Billerica, MA 0182 - USA) and the Biomek FX. Products were analysed on an Automated Capillary Electrophoresis Sequencer 3730 DNA Analyzer (Applied Biosystems). The MC615 May library corresponds to the period of maximum total phytoplankton species abundance for the MC site (Ribera d'Alcalà et al. 2004), the rest of the libraries correspond to dates when the dot blot specific probe signals were the lowest in comparison to the eubacterial probe. Coverage of each clone library (Romari and Vaulot 2004) was calculated using $1-(N_c/N)$ where $N_c$ is the number of cumulative unique sequences and $N$ is the number of useful sequences (excluding primers dimers, chimeras, those clones for which there was no reliable sequence and prokaryote sequences). Rarefaction curves (Romari and Vaulot 2004) were plotted to illustrate how accurately each library sampled the diversity of each sample.

**Magnesium Concentrations**

In order to test the conditions of the PCR and ensure that the primers were biased towards eukaryotes, two libraries were constructed using PCR products from a reaction mix containing 3 mM MgCl₂. These libraries were from MC601 1st December 2003 and MAMA64 9th February 2004 and the libraries were constructed as above using one 96 well plate per library. Both these libraries were extended using one 96 well plate containing PCR products from a PCR mix containing 1.1 mM MgCl₂ (48 clones per sampling date).
Sequences from cultured strains

rDNA 16S from taxa belonging to the Cryptophyceae, Prasinophyceae Clade II and Prymnesiophyceae (Table 2.4) was sequenced using OXY359F and OXY1313R primer pair (Table 2.1) because this gives a slightly longer fragment than PLA491F and allowed the region of the PLA491F primer to be checked for mismatches.

Table 2.4: Cultures sequenced during this study. DNA for the Cryptophyceae cultures was courtesy of Federica Cerino. Bente Edvardsen supplied the DNA for the Prymnesiophyceae indicated with an asterisk. The other cultures were held within the SZN culture collection. DNA was extracted using the CTAB method outlined in the section 'Species sequenced as part of this study'.

<table>
<thead>
<tr>
<th>Species</th>
<th>Class</th>
<th>Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cryptochloris sp.</td>
<td>Cryptophyceae</td>
<td>C94</td>
</tr>
<tr>
<td>Hemiselmis sp.</td>
<td>Cryptophyceae</td>
<td>C15</td>
</tr>
<tr>
<td>Plagoselmis prolonga</td>
<td>Cryptophyceae</td>
<td>C27</td>
</tr>
<tr>
<td>Proteomonas sulcata</td>
<td>Cryptophyceae</td>
<td>C28</td>
</tr>
<tr>
<td>Rhinomonas paucana</td>
<td>Cryptophyceae</td>
<td>C5</td>
</tr>
<tr>
<td>Rhizomonas sp. 3</td>
<td>Cryptophyceae</td>
<td>C42</td>
</tr>
<tr>
<td>Micromonas pusilla</td>
<td>Prasinophyceae II</td>
<td>PART</td>
</tr>
<tr>
<td>M. pusilla</td>
<td>Prasinophyceae II</td>
<td>MP1</td>
</tr>
<tr>
<td>M. pusilla</td>
<td>Prasinophyceae II</td>
<td>MP2</td>
</tr>
<tr>
<td>M. pusilla</td>
<td>Prasinophyceae II</td>
<td>MPCO</td>
</tr>
<tr>
<td>M. pusilla</td>
<td>Prasinophyceae II</td>
<td>P7/1</td>
</tr>
<tr>
<td>M. pusilla</td>
<td>Prasinophyceae II</td>
<td>Norbal</td>
</tr>
<tr>
<td>Chrysochromulina acanthi</td>
<td>Prymnesiophyceae*</td>
<td>UIO T20</td>
</tr>
<tr>
<td>C. aff. Polylepis</td>
<td>Prymnesiophyceae*</td>
<td>PLY200</td>
</tr>
<tr>
<td>C. campanulifera</td>
<td>Prymnesiophyceae*</td>
<td>UIO J10</td>
</tr>
<tr>
<td>C. cymbium</td>
<td>Prymnesiophyceae*</td>
<td>UIO R18</td>
</tr>
<tr>
<td>C. herdleins</td>
<td>Prymnesiophyceae*</td>
<td>CCMP284</td>
</tr>
<tr>
<td>C. hirta</td>
<td>Prymnesiophyceae*</td>
<td>hi-1</td>
</tr>
<tr>
<td>C. leadbeateri</td>
<td>Prymnesiophyceae*</td>
<td>UIO THE</td>
</tr>
<tr>
<td>C. rotalis</td>
<td>Prymnesiophyceae*</td>
<td>UIO TH2</td>
</tr>
<tr>
<td>C. simplex</td>
<td>Prymnesiophyceae*</td>
<td>UIO JomfB</td>
</tr>
<tr>
<td>C. throndsenii</td>
<td>Prymnesiophyceae*</td>
<td>UIO K11</td>
</tr>
<tr>
<td>Dicrateria inornata</td>
<td>Prymnesiophyceae*</td>
<td>CCMP355</td>
</tr>
<tr>
<td>Emiliana huxleyi</td>
<td>Prymnesiophyceae*</td>
<td>UIO BOF</td>
</tr>
<tr>
<td>Imantonia rotundata</td>
<td>Prymnesiophyceae*</td>
<td>UIO 101</td>
</tr>
<tr>
<td>Isochrysis cf. galbana</td>
<td>Prymnesiophyceae*</td>
<td>UIO Ryfylke</td>
</tr>
<tr>
<td>Phaeocystis cordata</td>
<td>Prymnesiophyceae</td>
<td>Phaeonap 1</td>
</tr>
<tr>
<td>P. globosa</td>
<td>Prymnesiophyceae</td>
<td>Napoli</td>
</tr>
<tr>
<td>P. jahnii</td>
<td>Prymnesiophyceae</td>
<td>B5</td>
</tr>
<tr>
<td>Phaeocystis sp.</td>
<td>Prymnesiophyceae</td>
<td>MC679 20m-5 IV</td>
</tr>
</tbody>
</table>

The DNA of the Cryptophyceae taxa was kindly provided by Federica Cerino, SZN and was from cultures isolated as part of a study of this class in the Gulf of Naples (Cerino...
2004; Cerino and Zingone in press). The Prasinophyceae Clade II material came from cultures isolated from the Gulf of Naples and held at the SZN. Approximately 30 ml of each culture was filtered using a syringe onto a 47 mm diameter PC filter of a 0.2 μm pore size from which DNA was extracted using the same CTAB method as for the environmental samples. Some of the Prymnesiophyceae sequences came from DNA kindly supplied by Bente Edvardsen and others from cultures held within the SZN, from which DNA was extracted using the same method as for the Prasinophyceae Clade II cultures. The PCR for all these sequences used a mix consisting of a total reaction volume of 50 μl containing 1 μl genomic DNA (approximately 40 ng), 200 μM deoxynucleoside triphosphates, 1 μM each primer and 2.5U Taq polymerase in 1x enzyme buffer containing 1.5 mM MgCl₂ (Roche Diagnostics GmbH, Mannheim, Germany). The reaction cycle was 80 °C for 5 min, followed by 30 cycles of 95 °C 1 min, 55 °C 1 min and 72 °C 1 min and an extension step of 72 °C for 5 min. PCR products were run on 1 % agarose gels, the bands excised and purified using a Gel Purification Kit from Qiagen (Qiagen Ltd.). Sequence reactions were obtained with the BigDye Terminator Cycle Sequencing technology (Applied Biosystems, Foster City, CA), purified in automation using the Millipore Montage SEQ96 Sequencing Reaction Cleanup Kit (Millipore Corporate 290 Concord Rd. Billerica, MA 0182 - USA) and a robotic station Biomek FX (Beckman Coulter, Fullerton, CA). Products were analysed on an Automated Capillary Electrophoresis Sequencer 3730 DNA Analyzer (Applied Biosystems).

**Phylogenetic Analyses**

Contigs from the forward and reverse library sequences were constructed using Seqman II 3.61 (DNASTAR Inc.) and aligned using Clustal W in BioEdit 7.0.1 (Hall 1999). Alignments of sequences from each clone library, including the sequences from the control DNA, were analysed using preliminary Neighbour Joining trees with
uncorrected ‘p’ distances, in order to identify the class of each clone library sequence and the proportion of each class found in the libraries. At this point the more obvious chimeric sequences, detected by examining the alignment by eye, were removed from the analysis. All those sequences belonging to the Chrysophyceae, Cryptophyceae and Prymnesiophyceae were then analysed separately using additional sequences from members of those classes from both the GenBank and the ARB (Ludwig et al. 2004) databases and from species sequenced as part of this study in an attempt to identify the sequences to species. The respective alignments were analysed in PAUP* (Phylogenetic Analysis Using Parsimony and other methods, Swofford 2002) using the modelblockPAUPb10 from Modeltest 3.7 (Posada and Crandall 1998) in order to assess the best-fit substitution model. The likelihood settings generated by Modeltest for the best-fit model were used to construct neighbour joining trees using PAUP*. Trees were rooted at the midpoint. Bootstrap analyses were performed using 1000 replicates and the same likelihood settings. Long branches were investigated using BLAST (Altschul et al. 1990) to check for chimeric sequences.

RESULTS

MC Ecological Programme Data

Figure 2.1: Environmental data from nutrient analyses and CTD measurements showing the variation in TIN SiO$_2$~ and PO$_4$~$^2$~ (left panel) and temperature and salinity (right panel) during the study period courtesy of the members of the Area Gestione Ambiente ed Ecologia Costiera delle Aree Temperate e Polari.
Nutrient, temperature and salinity data are presented in Figure 2.1 for samples from 0 m during the study period. The MC site is characterised by temperatures of between 14 °C and 25 °C with a slightly lowering of salinity during spring and early summer. Nutrient levels are generally higher in winter, but there are peaks of input during the summer months.

**Dot Blots**

Figure 2.2 shows the raw data from the phosphorimager for one membrane hybridised with the CHRY1037 Chrysophyceae probe and the specific blot for the control DNA from the chrysophyte culture. The environmental blots show a weaker signal during the winter. The probe is specific to the Chrysophyceae control DNA and does not cross hybridise with DNA from the other classes. Excel files containing the outputs from TotalLab are available on the accompanying CD.

![Dot Blot Image](image)

**Figure 2.2:** Dot blot hybridisation for the CHRY1037 Chrysophyceae probe showing a membrane with the triplicate blots for each environmental samples (left), samples were blotted in triplicate per date in four columns starting from the sample from 22nd July 2003, and the membrane with the concentration gradient of the Chrysophyceae culture DNA for the control (right).

Following the membrane above, the calculation for the % relative hybridisation of the 22nd July sample in the top left hand corner of the left panel would be as follows. The calculation of the first term uses the signal from the specific probe in the environmental sample, S_{env}, which in this case equals 44.80057507. The signal from the eubacterial
probe for the same spot equals 19854144.38. Therefore the first term is 
44.80057507/19854144.38 = 0.702203487. The second term comes from the graphs 
drawn from the concentration gradients blotted for the control DNA, the blots in the 
right panel of Figure 2.2, for both the specific probe and the eubacterial probe. The 
CHRY1037 probe figures are $S_{con} = 2577337.48$ and $E_{con} = 1644341.04$ so the second 
term is $2577337.48/1644341.04 = 1.567398379$. The % relative hybridisation for the 
first blot from the 22nd July sample equals $0.702203487/1.567398379*100 = 
44.80057507 \%$.

Four of the specific probes used gave mean % relative hybridisation values of above 2 
%. They were PRYM666 (Prymnesiophyceae), CRYP862 (Cryptophyceae), 
CHRY1037 (Chrysophyceae) and PELA1035 (Pelagophyceae) presented in Figure 2.3.

The PRYM666 Prymnesiophyceae probe gave a signal in all samples throughout the 
sampling period (Figure 2.3) between a low of 1.7 %, which is not above background 
oise, on 28th October 2003 and a high of 61 % on 16th February 2004 with an average 
of 35 % over the year.

The Cryptophyceae probe (Figure 2.3) indicates that this class was also present in most 
samples. There was a high of 45 % on 4th March 2004 and a general low during the 
summer and early autumn of 2003 that was not repeated the following year. There were 
peaks during November, March, May, July and the end of August.

The Chrysophyceae probe registered during the summer months of 2003 and 2004 with 
a high of 64 % during September 2003 and a general low during the winter 2003/4 with 
signals that generally do not rise above background (Figure 2.3).
Figure 2.3: Percentage Relative Hybridisation plots for the four probes that gave signals above 2 %. Bars indicate the standard deviation between the three spots per sampling date. The dotted lines indicate the dates when the clone libraries were constructed.

PELA1035, the Pelagophyceae probe, gave signals above background on six dates between December 2003 and late March 2004 (Figure 2.3) with a peak of 6.75 % on 2nd February 2004.

Clone Libraries

Figure 2.4 shows the percentage of each taxonomic class found in the six clone libraries plotted on the variations in abundance of phytoflagellates <10 μm (courtesy of Diana Sarno) at the MC sampling station between 22nd July 2003 and 6th September 2004. The abundance data show a low during the winter with generally higher abundances in this size class during the spring and summer.
The pie charts indicate a higher level of class level taxonomic diversity during the winter when abundance is lower and the reverse during the summer with dominance by only three classes, Chrysophyceae, Cryptophyceae and Prymnesiophyceae. Chrysophyceae sequences are not present in the two winter libraries (MC601 and MAMA64) and Prymnesiophyceae sequences were found in all the libraries. Pelagophyceae sequences were only found in the February library that is from the period in which the probe gave the highest signal. Sequences from classes for which probes are not currently available were also found, for example for Prasinophyceae Clades I (Pyramimonadales) and II (Mamiellales) and Dictyochophyceae.
Table 2.5: Clone library details, including the number of clones sequenced per library, the number of readable non chimeric sequences per library (useful sequences), the number of genotypes identified and the coverage value for each library indicating how well the diversity within each sample was estimated by the library.

<table>
<thead>
<tr>
<th>Library</th>
<th>Date</th>
<th>Number of clones sequenced</th>
<th>Number of useful sequences</th>
<th>Number of distinct genotypes</th>
<th>Coverage Value 1-(Nc/N)%</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC597</td>
<td>28th October 2003</td>
<td>96</td>
<td>86</td>
<td>30</td>
<td>65</td>
</tr>
<tr>
<td>MC601</td>
<td>1st December 2003</td>
<td>141 (94 at 3 mM MgCl2, 47 at 1.1)</td>
<td>79</td>
<td>25</td>
<td>68</td>
</tr>
<tr>
<td>MAMA64</td>
<td>9th February 2004</td>
<td>141 (94 at 3 mM MgCl2, 47 at 1.1)</td>
<td>69</td>
<td>36</td>
<td>48</td>
</tr>
<tr>
<td>250304</td>
<td>25th March 2004</td>
<td>96</td>
<td>79</td>
<td>55</td>
<td>30</td>
</tr>
<tr>
<td>MC615</td>
<td>17th May 2004</td>
<td>104</td>
<td>92</td>
<td>46</td>
<td>50</td>
</tr>
<tr>
<td>MC622</td>
<td>6th July 2004</td>
<td>96</td>
<td>82</td>
<td>42</td>
<td>49</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>674</strong></td>
<td><strong>487</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Rarefaction curves give an idea of how well the diversity within a sample has been assessed by the number of clones examined. As is evident from Figure 2.5, none of the libraries have reached saturation meaning that the diversity within each sample has not been fully sampled. However, the curves for the winter libraries are generally nearer to an asymptote than the rest. The curve for the March library (250304) is the steepest and corresponds to the period of the spring bloom. This is further confirmed by the coverage values shown in Table 2.5, which indicate that the March library is the least well covered as the number of genotypes recovered is high in relation to the number of clones examined.
Table 2.6: Details of the sequences identified from the six clone libraries. The number of sequences belonging to each taxonomic class is indicated (plain text) with the number of OTUs (bold) in that class found in that library and the number of OTUs only found in that particular library (italics). The totals indicate the total number of sequences (plain), the number of OTUs found across the six libraries (bold) and the number of OTUs restricted to a single library (italics).

<table>
<thead>
<tr>
<th>Class</th>
<th>MC597</th>
<th>MC601</th>
<th>MAMA64</th>
<th>Library</th>
<th>MC615</th>
<th>MC622</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacillariophyceae</td>
<td>5/2/1</td>
<td>2/2/0</td>
<td>2/2/2</td>
<td>1/1/1</td>
<td>1/1/1</td>
<td>1/1/1</td>
<td>12/9/6</td>
</tr>
<tr>
<td>Chrysophyceae</td>
<td>10/7/5</td>
<td>-</td>
<td>-</td>
<td>5/3/1</td>
<td>25/14/9</td>
<td>21/6/4</td>
<td>61/26</td>
</tr>
<tr>
<td>Cryptophyceae</td>
<td>49/9/4</td>
<td>61/11/3</td>
<td>15/7/1</td>
<td>10/8/1</td>
<td>34/4/1</td>
<td>10/3/0</td>
<td>179/19/1</td>
</tr>
<tr>
<td>Dictyochophyceae</td>
<td>-</td>
<td>-</td>
<td>1/1/1</td>
<td>-</td>
<td>1/1/1</td>
<td>-</td>
<td>2/2/2</td>
</tr>
<tr>
<td>OM270</td>
<td>-</td>
<td>-</td>
<td>1/1/1</td>
<td>-</td>
<td>-</td>
<td>1/1/1</td>
<td>2/2/2</td>
</tr>
<tr>
<td>Pelagophyceae</td>
<td>-</td>
<td>-</td>
<td>2/1/1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2/1/1</td>
</tr>
<tr>
<td>Prymnesiophyceae</td>
<td>15/10/6</td>
<td>11/9/3</td>
<td>35/21</td>
<td>63/43</td>
<td>31/26</td>
<td>49/31</td>
<td>204/121</td>
</tr>
</tbody>
</table>

Figure 2.5: Rarefaction curves for each of the six clone libraries calculated using the cumulative number of unique genotypes found against the number of clones examined, in order to assess the saturation of each library.
Figure 2.6: Prymnesiophyceae NJ tree (top half only, continues from green circle overleaf) with Maximum Likelihood parameter settings calculated by Modeltest (best fit model TIM+1+G A = 0.2985, C = 0.1657, G = 0.2654 and T = 0.2704, A ≪ G = 8.0139, C ≪ T = 13.3944, A ≪ C = 1 and C ≪ G = 1.4867, the positional rate variation across the alignment y = 0.385 and the proportion of invariable sites I = 0.549) showing the diversity within the Prymnesiophyceae genotypes found. Blue sequences are those from the six Gulf of Naples libraries, the numbers in brackets refer to the total number of sequences corresponding to that genotype, red are those from other libraries, black are culture sequences and those in green are the cultures sequenced as part of this study (see Table 2.4 for details). The boxes indicate where library sequences and those from known cultures group together. Bootstrap values are from 1000 replicates.
The class Prymnesiophyceae shows the greatest diversity in the clone libraries with 121 distinct genotypes in 204 sequences (alignments for the trees constructed and the chromatogram files for each clone are available on the accompanying CD). Figure 2.6 and Figure 2.7 show a tree of the distinct genotypes and known sequences from GenBank and sequences from *Chrysochromulina*.

Figure 2.7: Prymnesiophyceae NJ tree continued (bottom half) showing the association between the clone library sequences and those from *Phaeocystis*, *Imantonia*, *Isochrysis*, *Emiliania* and *Chrysochromulina* (orange boxes).
obtained during this study. It is not possible from this tree to identify many of the genotypes. However, sequences belonging to *Emiliania huxleyi*, *Imantonia rotunda*, *Isochrysis* cf. *galbana*, *Phaeocystis cordata*, *P. jahnii*, *P. globosa* and the genus *Chrysochromulina* were found (orange boxes). The tree also gives some indication of seasonal distribution of the genotypes. For example, sequences belonging to *Phaeocystis cordata*, *P. jahnii* and *P. globosa* were not found in the two summer libraries.

Figure 2.8 shows that, of the classes found in this study, the Cryptophyceae clone library sequences have the closest match to known cultures. There is a correspondence between sequences from the cultures identified by Federica Cerino (Cerino 2004; Cerino and Zingone in press) and sequences found in the clone libraries, indicated by the orange boxes. For example, sequences matching *Cryptochloris* sp. strain C94, *Plagioselmis prolunga* strain C27, *Hemiselmis* sp. strain C15, *Rhodomonas* sp. 3 strain C42 and *Proteomonas sulcata* strain C28 were all found in the various libraries.

Unfortunately marine Chrysophyceae are little studied and the lack of 16S gene sequences of marine algal plastids makes identifying most of the genotypes found difficult if not impossible. However, Figure 2.9 shows that some of the genotypes do match sequences found in other clone libraries. For example a sequence from the May library is identical to the uncultured eukaryotic clone HE17 from Helgoland (indicated in Figure 2.9 by a purple box). The orange boxes on Figure 2.9 correspond to groups that contain Gulf of Naples sequences and sequences from the ‘AMBITION’ cruise in the Arabian Sea (Fuller et al. 2006a). The May library is the most diverse with 14 of the 24 genotypes being found in this library. It also contains 9 genotypes that are not found in the other libraries.
Figure 2.8: Cryptophyceae NJ tree constructed using ML settings (best fit model K81uf+I+G A = 0.2735, C = 0.2126, G = 0.2821 and T = 0.2319, A ≈ G = 5.8410, C ≈ T = 5.8410, A ≈ T = 0.3832, A ≈ C = 1 and C ≈ G = 0.3832, the positional rate variation across the alignment γ = 0.9072 and the proportion of invariable sites I = 0.802) using the unique genotypes from the six clone libraries (blue), the numbers in brackets refer to the total number of sequences corresponding to that genotype, with sequences from GenBank corresponding to cultures (black) and to other clone libraries (red), sequences from cultures isolated during Cerino & Zingone (in press) (green) and from Dinophysis sequences in GenBank (purple). The bootstrap values are from 1000 replicates. The orange boxes indicate where there is a correspondence between the Cerino & Zingone cultures and the sequences from the clone libraries.
Figure 2.9: Chrysophyceae NJ tree constructed using ML settings (best fit model TrN+1+G A = 0.2896, C = 0.1789, G = 0.2739 and T = 0.2576, A \& G = 3.9412, C \& T = 5.7361, A \& C = 1 and C \& G = 1, the positional rate variation across the alignment \( \gamma = 0.6965 \) and the proportion of invariable sites \( I = 0.6104 \)) of 16S sequences from the six Gulf of Naples clone libraries (blue) including sequences from GenBank corresponding to cultures (black) and clone libraries (red). The numbers in brackets refer to the total number of sequences corresponding to that genotype. The purple box at the top of the tree shows the match between a Gulf of Naples sequence and a sequence from a clone library from Helgoland. The orange boxes indicate the groups containing sequences from the Gulf of Naples and Red Sea clone libraries.
Figure 2.10: NJ tree constructed using ML settings (best fit model GTR+1+G A = 0.286, C = 0.1934, G = 0.2685 and T = 0.2521, A ° G = 5.3583, C ° T = 7.4588, A ° T = 1.7974, A ° C = 0.952 and C ° G = 0.5347, the positional rate variation across the alignment γ = 0.608 and the proportion of invariable sites I = 0.3933) showing the remaining sequences from the six clone libraries in blue, the numbers in brackets refer to the total number of sequences corresponding to that genotype, with representatives from cultures (black) from GenBank and ARB, from other clone libraries (red) and from newly sequenced cultures from the Gulf of Naples (green). Bootstrap values are from 1000 replicates.

Figure 2.10 shows the rest of the genotypes found displayed in a NJ tree in order to illustrate their closest relatives. Diatoms belonging to Skeletonema were found in the winter libraries (MC597 and MC601). The rest of the diatom sequences do not identify...
exactly with any known sequences but the March sequence 250304-76 groups with a clone library sequence from the Southern Ocean and the rest form a group with the library sequence OM20 from Cape Hatteras, North Carolina. The Dictyochophyceae are polyphyletic in this tree, but this topology is not supported by bootstrap. The MC615-64 May sequence lies in a clade including the Dictyochophyceae sequences found in clone libraries from the AMT15 cruise in the mid Atlantic Ocean and the sequence for the RCC382 strain of *Mesopedinella arctica*. The other two Dictyochophyceae culture sequences, *Dictyochophyte* sp. RCC332 and *Rhizochromulina* sp. CCMP1253, form a separate clade near the Pelagophyceae. The Dictyochophyceae MAMA64-21 February library sequence forms a single branch between the OM270 clade and the Prasinophyceae Clade II Mamiellales. In the February library there are two Pelagophyceae sequences that are identical to a *Pelagomonas calceolate* sequence from the strain RCC100 isolated in the North Pacific. A number of Prasinophyceae Clade I sequences were found that fall into two genotypes, one of which does not seem to be closely related to any known sequences and the other which differs at only one position from the library sequence OCS162 from the Oregon Coast Study. Two sequences group closely with the Cape Hatteras sequence OM270 (MAMA64-51 February and MC622-32 July) and so form a separate clade. The Prasinophyceae Clade II sequences fall into three genoptyes, two *Micromonas* and one *Ostreococcus*. The two *Micromonas* genotypes group with the sequences from cultures collected in the Gulf of Naples and Partenope (all Tyrrhenian Sea) and a culture sequence from the Spanish coast of the Mediterranean (RCC434). The *Ostreococcus* sequence is close to a culture sequence (RCC393) isolated in the Tyrrhenian Sea from the PROSOPE cruise Station 9, which is located between the Italian coast and Corsica.
Magnesium concentrations

The two libraries using the 3 mM MgCl₂ PCR mix are displayed in Figure 2.11 with the single 96 well plate library (47 clones each from MC601 and MAMA64) using the 1.1 mM MgCl₂ mix.

The figure shows that, in the libraries formed using 3 mM MgCl₂, the primers are not specific to eukaryotes and so also amplify bacteria and cyanobacteria and the reverse primer OXY1313R forms a primer dimer. The charts on the left show that when the primers are used under more stringent conditions, they also seem to exclude some eukaryote classes. For example Prasinophyceae Clade II sequences are not found in either library under conditions with a lower concentration of MgCl₂ and Dictyochophyceae were not found in the MAMA64 lower MgCl₂ library.

Figure 2.11: Charts indicating the identity of the sequences found in the MC601 and MAMA64 libraries constructed using PCR products from reaction mixes with different concentrations of magnesium chloride. The numbers next to each segment show the number of sequences found in that group.

Comparison of Methods

Data from Cerino (2004) relating to the abundance of cryptophytes between October and December 2003 provided a means of comparing the dot blot hybridisation with the methods used for estimating abundance as shown in Figure 2.12. Although the %
relative hybridisation does not relate to an absolute number of cells, there is a correspondence with the other methods in terms of the trend observed. There is a peak in all methods on 28th October and at the end of November. The different methods for counting cells give different estimates of cells numbers, with the Most Probable number (MP) calculated from SDC being the lowest. Flow cytometry generally gives the highest abundance, except on 28th October 2003.

![Graph](image.png)

Figure 2.12: Comparison of dot blot % relative hybridisation for the CRYP862 Cryptophyceae probe with data from Cerino (2004) on the abundance of cryptophytes between October and December 2003. Graph adapted from Figure 4.4 in Cerino (2004) to include % relative hybridisation.

**DISCUSSION**

**Method Biases – Dot Blots and Clone Libraries**

Before discussing the results of the two methods employed during this chapter, an evaluation of these methods and possible areas of error and bias is needed. Filtering of sea water samples to extract size fractionated community DNA has inherent biases. As both methods rely on PCR, there are a number of biases connected with the dynamics of the reaction that are relevant to this and the following chapters and so will be discussed.
in full in this chapter. The involvement in both methods of a ribosomal multi-copy marker is another possible source of problems and again relates to this thesis as a whole. The probes have weaknesses that should be discussed, as does the construction of clone libraries. However, the use of such methods even with the accompanying problems has many advantages as tools for working towards increasing the knowledge of so far understudied groups within the uliplankton.

The filtering of samples undoubtedly excludes some cells smaller than 5 μm, particularly those with long flagella or processes, such as some Prymnesiophyceae and diatoms, and those forming mucilaginous colonies, such as some *Phaeocystis* species, which can clog the filter and also exclude other flagellates. However the fact that none of the sequences recovered in the clone libraries belonged to classes lacking uliplanktonic representatives suggest that at least the filtering excluded organisms larger than 5 μm. This also rules out the problem of cells breaking and chloroplasts of larger organisms passing through the filter.

There are a number of biases involved in PCR using community DNA. Sequences with high G+C content are less well amplified than lower G+C content templates, because higher G+C content sequences dissociate with lower efficiency (von Wintzingerode et al. 1997), as demonstrated by the preferential amplification of yeast DNA over DNA from thermophilic archaia in a mixed template (Reysenbach et al. 1992). When degenerate primers are used, templates with G or C in the degenerate position are more likely to be amplified than A or T (Polz and Cavanaugh 1998). DNA in the regions flanking the template can also inhibit amplification (Hansen et al. 1998). Since single-stranded molecules must react with free primer to initiate extension reactions, the rate of formation of primer-template hybrids will be influenced by the proportion of template
molecules in a single-stranded state (Suzuki and Giovannoni 1996) and this can lead to
different proportions of amplicons in relation to the original proportions of target
copies. In general, PCR is not fully quantitative and all target genes may not be
amplified with the same efficiency (Moon-Van Der Staay et al. 2000).

rDNA copy number may bias the results and is the main reason for not drawing
conclusions about abundance from the number of sequences found in a clone library or
the % hybridisation of a specific probe. The number of copies of 16S per cell is not only
a question of the number of copies in the ctDNA ribosomal operon, but also of the
number of ctDNA molecules per chloroplast and the number of chloroplasts per cell.
Some planktonic plastids contain up to 650 copies of small circular genomes per plastid,
with each usually containing two ribosomal operons per genome (Ersland et al. 1981).
This study was done on the raphidophyte *Olisthodiscus luteus*, which is 12-19 μm in
size and like all raphidophytes has many chloroplasts. This species was shown to have a
variable number of copies of the ctDNA per organelle depending on the age of the
culture. It was also shown that the absolute amount of ctDNA found within a
*Olisthodiscus luteus* cell remains constant but that the amount of ctDNA per plastid is
inversely proportional to the number of chloroplasts to which that DNA must be
distributed (Cattolico 1978). In general ultraplanktonic classes have one or two
chloroplasts so the copy number is likely to be considerable less than in the
Raphidophyceae, but there is still the possibility that the pattern of more copies of
ctDNA per organelle when fewer chloroplasts are present is followed. *Ochromonas
minima*, a chrysophyte with individuals small enough to be in the below 5 μm fraction,
has a single chloroplast and *Hemiselmis* species (Cryptophyceae) all have a single
chloroplast as does *Micromonas* (Prasinophyceae II), while the Prymnesiophyceae
generally have two, but there is no information on the number of ctDNA molecules per
chloroplast. rDNA copy number is generally considered to increase in proportion to cell length (see Figure 2 Zhu et al. 2005) especially in the case of 18S. The Chloroplast Genome Database project lists four microalgal species for which the chloroplast genome has been sequenced, all of which contain one or two copies of 16S (*Nephroselmis olivacea* Prasinophyceae 2 copies, *Guillardia theta* Cryptophyceae 2, *Emiliania huxleyi* Prymnesiophyceae 2, *Chlorella vulgaris* Trebouxiophyceae 1 copy (Cui et al. 2006). The cyanobacteria so far sequenced contain between 1 and 5 copies of this gene per genome (Klappenbach et al. 2001). Of the marine cyanobacteria sequenced, *Prochlorococcus* has between 1 and 2 copies (three strains sequenced) and the *Synchecoccus* strain WH8102 has one copy. Given the small cell size in the <5 µm fraction, it seems unlikely that ultraplankton species will contain a high number of copies, the range is likely to be between 1 and 10 as for 18S (Zhu et al. 2005).

One of the main problems with dot blot hybridisation lies in the accuracy of the probes not just in terms of non-specific reactivity but also in terms of how well the probes hybridise to the target. This is perhaps well illustrated by the results from the clone libraries. The samples for the libraries were chosen because they correspond to low total hybridisation by the four probes giving positive results. Yet none of these libraries was dominated by a class not covered by the probes. The low hybridisation may be due to problems with the quality of the community DNA or the PCR products from these samples that affected the hybridisation of the probe rather than the fact these classes were not present. It is also possible that the stripping of the set of membranes before a second round of hybridisation may have affected the quality of the DNA on the membranes.
There is a general difficulty in developing 16S probes for certain groups. For example, the Bacillariophyceae and Bolidophyceae do not form separate clades and any probe will therefore not be able to distinguish the two classes. Other classes are so poorly represented that developing probes is difficult as they are likely to underestimate diversity within the class. The Dictyochophyceae is a good example of such a class, of which there are very few 16S sequences in the ARB and GenBank databases.

Intra-individual and intra-specific variability in the 16S operon have not been studied in microalgal species. It is therefore difficult to give a taxonomic rank to the various sequences found in the clone libraries. What degree of difference makes a different species or genus? Do all the sequences found within, say the Prymnesiophyceae, correspond to new species or are some of them intra-specific or -individual variation? These questions could be answered by studying cultures but a study on one species would not be applicable to all species within that class. However, culturing these taxa may not be easy in the first place, given that none of these entities have been cultured so far. Plastid 16S and 23S rDNA gene sequences are highly conserved, thus are less likely to show intra-individual polymorphisms. A case of intra-individual variation in plastid rRNA genes was reported for a perennial holoparasitic (nonphotosynthetic) angiosperm, *Cynomorium coccineum* (Gracia et al. 2004), but has not been reported in the plastid of a microalga.

Chimeric sequences can be a problem in clone libraries as they can be difficult to detect when known sequences from cultures are not available for comparison. Chimeras form during PCR when two different DNA molecules with high similarity compete with the primers during the annealing step (von Wintzingerode et al. 1997). As the 16S sequences recovered in the Gulf of Naples clone libraries are generally more than 80%
similar, with some regions that are highly conserved across diverse lineages, the likelihood of chimeras forming is reasonably high. At the 30 PCR cycles used in this study around 15% of the sequences could be chimeras (von Wintzingerode et al. 1997). Chimeric sequences can be recognised due to the similarity of different regions to different organisms. The chimeric sequences found in the Gulf of Naples clone libraries were in most cases easily recognised in that the beginning and the end of the sequences belonged to different classes and this could be seen in the alignment. Other sequences had smaller regions that BLAST to highly diverse organisms (for example a cryptophyte and *Synechococcus*). The percentage of chimeric sequences found in each library ranged between 0 and 9.5%. The concentration of magnesium chloride seems to have some effect, in that the percentage of chimeric sequences in the products from the PCR mix with 3 mM MgCl₂ was roughly double that for the 1.1 mM MgCl₂ mix in both libraries (see Figure 2.11). The difficulty came in recognising chimeras when both halves of the sequences aligned to another clone library sequence, as in the case of MAMA64-51 February and MC622-32 July that are similar to the Cape Hatteras clone library sequence OM270. It is not sure whether all three sequences are chimeras as they form a separate clade outside the known classes. Clone OM270 was unique among the sequences investigated in the OM library and appeared as an outgroup to prymnesiophytes and cryptophytes (Rappé et al. 1998). The secondary structure of this clone was analysed in order to investigate whether it was chimeric and failed to uncover any artifacts (Rappé et al. 1998).

Statistical analysis have not been carried out on any of the data presented in this thesis due to the unsuitability of the data to such analysis. True replicates have not been used and the use of PCA to link certain classes to environmental conditions as used in Fuller et al. (2006b) does not take into account that taxonomic classes are not comparable in
terms of numbers of species. Such comparisons must be carried out at the level of species in order to be biologically meaningful.

Where is *Micromonas*?

The dominance of members of the prasinophyte order Mamiellales (Clade II) at other coastal sites (Bec et al. 2005; Not et al. 2004; Not et al. 2005) highlights the obvious absence of this group in the results of this study at MC in the Gulf of Naples. *Micromonas pusilla*, one of the most important and abundant species in the Mamiellales, has been shown to be abundant in the Gulf of Naples by studies using SDC (Throndsen and Zingone 1994; Zingone et al. 1999a) and epifluorescence microscopy (Zingone et al. 1999a). These two methods gave comparable results (Zingone et al. 1999a) showing a peak in cells numbers during March. This peak is further confirmed by pigment analyses that show a peak in the prasinophyte pigment Prasinoxanthin during March (Santarpia 2005) in surface waters at MC. The omission of this group in the results from the clone libraries may be due to a number of reasons. However, the presence of Prasinophyceae Clade II sequences in the libraries constructed using the higher concentration of MgCl₂ indicates that the problem probably lies with the PCR. The Prasinophyceae Clade II sequence OM5 has As in the degenerate positions of the OXY1313R primer and 3 other mismatches towards the centre. OCS182 has a G and an A in the degenerate positions and also the same 3 other mismatches as OM5. As mentioned above, G/C in a degenerate position is more likely to be amplified and the mismatches make annealing even more unlikely. The 'absence' of this group is likely to be due to these mismatches in the primers. Even if these problems could be resolved by designing a new primer, there are problems developing a probe for Prasinophyceae as it has been shown to be polyphyletic in 18S (Guillou et al. 2004), a pattern that is repeated in 16S (Fuller et al. 2006a). The PRAS540 probe does
not hybridised well to control DNA and did not show any signal in the natural samples, so a new probe for the Prasinophyceae Clade II will have to be designed.

There is evidence that other green algae are also important between May and July in surface water as shown by higher concentrations of chl \( b \) and violaxanthin (Santarpia 2005), two pigments that are found in all groups of Prasinophyceae and Chlorophyceae (Sym and Pienaar 1993). The fact that Prasinophyceae sequences were only detected in the winter libraries (October and December) does not help clarify what these green algae are. The question therefore remains as to what algae make up this bloom between May and July. The methods employed in the present study have not shed any light on this, as there is a lack of probes for these groups and this will have to be addressed in the future.

**Dot Blot Hybridisations**

The nine probes targeting photosynthetic eukaryotes in the <5 \( \mu \)m fraction show that members of the Prymnesiophyceae, Cryptophyceae and Chrysophyceae dominate this size class at the MC station in the Gulf of Naples with a small bloom of Pelagophyceae during February. These classes show some seasonal variation in occurrence that will be discussed below in relation to the data from the MC Long-Term Ecological Programme and from other sites in the Mediterranean.

The data from the MC Long-Term Ecological Programme show generally higher abundances in the summer months (Ribera d'Alcalà et al. 2004). This is in contrast to the study conducted in the NW Mediterranean French coastal site in the Bay of Banyuls-sur-Mer, where picoeukaryotes (<3 \( \mu \)m) were abundant in winter, peaking during January with 21,000 cells ml\(^{-1} \) at the surface (Charles et al. 2005). Picoeukaryote numbers showed a correlation to refreshment of the study area by nutrient rich waters.
A study carried out in the Thau Lagoon, situated around 160 km north along the French coast from Banyuls-sur-Mer, found high abundances of picoeukaryotes in the summer months (Bee et al. 2005) showing a positive correlation to temperature. The study at Banyuls-sur-Mer concluded that the efficient assimilation of nutrients by picoeukaryotes, due to their small size, enables the exploitation of nutrient pulses (Charles et al. 2005) and that picoeukaryote abundance was mainly controlled by nutrient availability. The conclusion from the Thau Lagoon was that temperature and irradiance were more important in contributing to picoeukaryote growth, but that grazing pressure also played a part. In the Gulf of Naples region, nutrients never become limiting and the total phytoplankton bloom lasts from March to October (Ribera d'Alcalà et al. 2004). If picoeukaryote abundance is controlled by nutrients and/or light and temperature then the presence of high numbers in summer in the Gulf of Naples is logical. This is also confirmed by the phytoplankton counts from the MC Long-Term Ecological Programme showing a high percentage of the plankton is made up of phytoflagellates smaller than 10 µm during the summer months (Zingone et al. 1990). The probes do not give any idea of abundance but do help towards understanding which are the major groups within the smallest size fraction.

The probes show that the Prymnesiophyceae are present in the Gulf of Naples all year round and do not show any preference for a particular season. This class has been shown to be important in the Gulf of Naples (Ribera d'Alcalà et al. 2004; Zingone et al. 1999a; Zingone et al. 1990) and is the most species rich class within the Gulf after the diatoms and dinoflagellates (A. Zingone and D. Sarno, pers. comm.). It has also been shown to be important in other studies investigating the smaller size fractions of the plankton (Fuller et al. 2006b; Moon-Van Der Staay et al. 2000; Thomsen et al. 1994). The Prymnesiophyceae are a mainly marine class of algae including the
coccolithophorids, such as *Emiliania huxleyi*, that form calcified plates. Most species are photosynthetic, but there are examples of non-photosynthetic species in both marine and freshwater environments. The unique features of this class are the haptonema, a filiform organelle that may be many times longer than the cell body, and unmineralised scales that, in their simplest form, consist of a two layered plate with each layer being composed of radially arranged fibrils (Green and Jordan 1994). However, these characters generally require EM studies on cultures in order to identify cells to class level. So in ecological studies at sea, this class is generally identified using pigments. 19’HF is used as the diagnostic pigment for this class, but there is considerable variation in content within the class. In a survey by Jeffrey and Wright (1994) only 60% of the strains tested were found to contain 19’HF, which is also found within some diatoms and pelagophytes (Santarpia 2005). The dinoflagellate *Gyrodinium aureolum* has been shown to contain 19’HF as its main carotenoid pigment (Tangen and Björnland 1981). *Dinophysis* is also known to steal plastids from pyrrhmesiophytes, (Koike et al. 2005; Takishita et al. 2002), which further complicates the picture. Despite these problems, it is important to identify members of this class for a number of reasons. There are a number of toxic species that form blooms affecting marine life, particularly fish species. These species are mainly within the genera *Chrysochromulina* and *Prymnesium*, for example *Prymnesium parvum*, *Prymnesium patelliferum*, *Chrysochromulina polylepis*, and *Chrysochromulina leadbeateri* (Meldahl et al. 1995). Coccolithophorids such as *Emiliana huxleyi* are important in that they form dense blooms that are highly reflective due to the nature of the coccoliths and may alter the earth’s albedo having a cooling influence on the climate. Coccoliths are also a major contributor to ocean floor limestone accumulation and represent the world’s largest long-term inorganic carbon sink (Graham and Wilcox 2000). *E. huxleyi* along with *Phaeocystis pouchetii* (another bloom forming pyrrnnesiophyte) are also producers of
DMS, which when released into the atmosphere acts as cloud condensation nuclei leading to increased cloud cover that again has a cooling effect on the climate.

The importance of the Cryptophyceae at sea has been demonstrated (Novarino 2003; Novarino 2005) and the results from the dot blot highlight the seasonal differences in occurrence of this class. The dot blot hybridisations show low % hybridisation between July 2003 and mid October 2003 and then a general high signal of around 30 % for the rest of the sampling period, with odd peaks and troughs. Data from the MC counts show a prevalence of cryptophytes during the year with cells being present in most samples (Cerino and Zingone in press). There are blooms of several different species at different times of the year. For example, there was a bloom of *Plagioselmis prolonga* at the end of April 2003, then of a species of *Hemiselmis* at the end of July that was followed by a slight bloom of a species of *Cryptochloris* between October and December (Cerino and Zingone in press). The dot blots match the counts in that they show the presence of Cryptophyceae throughout the year. However the counts do not show the importance of this group during the winter months (see Figure 47 Cerino and Zingone in press) that shows very low numbers during the winters of 2002 and 2003). This could be due to poor preservation of winter species in the fixed samples used for the counts. Another possibility is that there is a generally low abundance of cells in the winter so the % hybridization of the Cryptophyceae specific probe appears higher due to a lower amount of total 16S DNA in the samples. However, this would still mean that the Cryptophyceae form a significant proportion of the winter ultraplanktonic community in the Gulf of Naples. The comparison between fixed counts, cytometry and dot blot is positive in that the same trend is recovered by all methods at least for the Cryptophyceae. Dot blot hybridisation is not quantitative as mentioned in the section on errors and biases. However the fact that the same trend is recovered is encouraging.
This method is therefore extremely useful for identifying the major players in different samples.

The Cryptophyceae are rightly named the ‘hidden single cells’ as they are among the most cryptic algae. Despite forming large populations and being present in most environments, they are generally small (3-50 μm) and poorly preserved (Graham and Wilcox 2000). They have been much studied as a group from the evolutionary point of view as they are formed by a secondary endosymbiosis of a red alga with an as yet unknown eukaryote host and retain a number of features from the red algae including the remnants of its nucleus, the nucleomorph (Novarino 2003). However, there is evidence that this group may also be important as the basis of trophic webs given their high food value for herbivorous zooplankton (Cerino and Zingone in prep; Klaveness 1988). Cryptophytes were among the most grazed plankton by calanoid copepods in a lake study (Klaveness 1988) and it is possible that this may also apply to the marine environment. This class has a number of unique characters, including their bean-like asymmetrical ellipsoidal shape, the colour of the chloroplasts, the furrow/gullet complex and the ejectosomes that when discharged make the cell jump in the opposite direction (Cerino 2004; Novarino 2003). However the taxonomy of this group is still poor for a number of reasons. The cells are not well preserved by fixatives and there has been a reliance on characters that have since been shown to have little taxonomic value (Cerino and Zingone in prep). The situation is further complicated by the heteromorphic life cycle that has lead to different life stages being described as different genera (Cerino and Zingone in prep). Given their trophic importance as food items, it is important to understand their seasonal occurrence in order to understand ecosystem energy flows.
The Chrysophyceae show some seasonality, with the probe showing greater signal in summer. This class is generally regarded as being more important in freshwater environments (Thomsen 1986) and it has been largely overlooked at sea. In fact, in a 1995 treatment of the Chrysophyceae (Sandgren et al. 1995), the only marine 'chrysophyte' discussed as being of any importance at sea was *Aureococcus anophagefferens* (Nicholls 1995), which has been subsequently transferred to the Pelagophyceae (DeYoe et al. 1997). It is possible that this oversight is due to the fact that the pigment signature for the Chrysophyceae is the same as the diatoms and Prymnesiophyceae, all of which contain fucoxanthin and chlorophyll *c* making it difficult to distinguish these classes in oceanographic studies that rely on pigments. A recent study using dot blots on the AMBITION Arabian Sea transect revealed a dominance of chrysophytes along the whole length of the transect, but especially in more coastal areas towards the Gulf of Oman and the Straits of Hormuz (Fuller et al. 2006b). It is possible that, with more extensive use of the Chrysophyceae probe, this group will be shown to be more widespread in the marine environment than previously thought. The class Chrysophyceae was first recognised in the 18th Century (Andersen et al. 1999) and has at different times included taxa that have now been assigned to the Prymnesiophyceae, Pelagophyceae, Dictyochophyceae and Synurophyceae (Tomas 1997). The class is named for the golden brown colour of the chloroplasts (‘*chrysos*’ is Greek for gold) and includes photo-, mixo- and hetero-trophic species. There are a number of colonial species with cellulose or chitin lorica. The class is identified on the basis of the formation of a silica walled cyst called a stomatocyst and a combination of pigments including chlorophyll *a*, *c*₁ and *c*₂ and fucoxanthin (Graham and Wilcox 2000). The stomatocysts provide a good fossil record of past chrysophytes (Graham and Wilcox 2000) and freshwater genera have been extensively studied. The few genera with marine representatives include *Ochromonas*, *Dinobryon* and *Chromulina* (Tomas
1997). *Ochromonas minima* Throndsen is small enough to be present in the ultraplankton (Tomas 1997) as are some members of *Picophagus*, *Tetraparma* and *Triparma* (Vaulot et al. 2004). There is a significant bloom of cells attributed to *Ochromonas minima* in the Gulf of Naples in June and July (A. Zingone unpublished data). *Dinobryon coalescens* Schiller and *D. faculiferum* (Willén) Willén are both present in the Gulf of Naples. *D. coalescens* has a bloom in March and April, but forms colonies with lorica and is too large to appear in the ultraplankton fraction. *D. faculiferum* is present between April and September, but is also too large to be found in the <5 μm fraction.

**Clone Libraries**

The clone libraries show a high level of diversity within the classes found as well as backing up the data from the dot blots by demonstrating the seasonality of some genotypes. The six clone libraries show a large amount of diversity both at the class level and within classes. The results will be discussed according to class.

The diversity shown within the Prymnesiophyceae is in keeping with the few other genetic studies. Moon-van der Staay et al (2001) found high levels of diversity within 18S clone libraries from the Pacific. The Oregon Coastal Study clone library was 61% Prymnesiophyceae sequences in a library containing 109 plastid clones (Rappe et al. 1998). The class contains 11 genera with 80 species that do not form coccoliths and 40 genera containing a total of 200 species that do (Graham and Wilcox 2000). The diversity seen in the clone libraries could also be a reflection of the lack of 16S sequences for known species. The total number of Prymnesiophyceae species seen in the Gulf of Naples is around 100 (A. Zingone unpublished data) of which only *Phaeocystis* has been studied in any detail (Zingone et al. 1999a). There are no 16S sequences for *Prymnesium*, and only around 5% of the heterococcolithophorid species
have been established in culture (Thomsen et al. 1994). The actual number of Chrysochromulina species may be more than 100 (Thomsen et al. 1994), of which C. apheles, C. minor, C. elegans and C. pyramidosa are all picoplanktonic (Thomsen et al. 1994), but no 16S sequences are available for these species. In fact there is only one pico prymnesiophyte culture, for C. minor, available in the culture collections. Given the paucity of available sequences and cultures, it is possible that some of the diversity found does not correspond to unknown entities, but to known species that have not yet been sequenced. The diversity of this class in the Gulf of Naples is of importance ecologically and should be investigated further in order to characterize this important group, especially as the lack of repetitive OTUs in the libraries makes tracking the seasonality of the different taxa difficult.

The recent study using newly isolated Cryptophyceae cultures from the Gulf of Naples (Cerino 2004) has allowed some comparisons to be made between the Gulf of Naples clone libraries and well-studied cultures. The ‘Plagioselmis prolonga C27-like’ sequences were the most numerous and are found in all libraries, but with almost equal numbers of sequences being found in the MC597 October library and the MC615 May library. Plagioselmis prolonga C27 cells were most abundant during May 2003, but this was the most frequently recorded species during the year (Cerino and Zingone in press). The ‘Cryptochloris sp. C94-like’ sequences were all found in the winter libraries. Cerino & Zingone (in press) show a peak of Cryptochloris sp. C94 in December 2003. The sequences matching Hemiselmis sp. C15 were only found in the MC622 July library, a taxon that was only seen by Cerino & Zingone (in press) during July and August. A sequence matching Rhodomonas sp.3, a late autumn species, was recovered in the MC597 October library. The dinoflagellate Dinophysis sequences are identical to Teleaulax amphioxeia AY453067. This group also contains the MC601-127 December
genotype, which was only found in the winter libraries, with the highest number of sequences being found in the December library.

*Dinophysis* is a genus of dinoflagellates that has been shown to undergo kleptoplastidy, in that they 'steal' plastids by ingesting cryptophytes (Janson 2004) or prymnesiophytes (Koike et al. 2005), which may explain why *Dinophysis* species have proved so difficult to maintain in culture. It is hypothesised that *Dinophysis* species ingest the most abundant cells in the water column (Janson 2004). In Figure 2.8, there are clone library sequences that show identity to *Teleaulax amphioxeia* and various *Dinophysis* species. This does not necessarily indicate presence of *Dinophysis* after filtration, as even the smallest *Dinophysis* cells are 36 μm wide. These sequences are more likely to be *Teleaulax amphioxeia*, a cryptophyte with cells of 4-6 μm in width, narrow enough to pass through a 5 μm filter. This species has been shown to be the one of the most common Cryptophyceae species along the Swedish coast and generally has a colder water distribution including the Gulf of Finland, western Baltic, Skagerrak, Atlantic coast of Belgium, east coast of America and Japan (website), which may explain why sequences identical to this species were all found in the December library. A species of *Teleaulax* has been recorded in the Gulf of Naples, generally in the winter and spring months (A. Zingone, unpublished data and Cerino and Zingone accepted), but it was not possible to preserve it long enough in culture to identify the species (Cerino and Zingone in press). Given the conservative nature of 16S as a marker, the taxon found in the Gulf of Naples is not necessarily the same as *Teleaulax amphioxeia* found in the North Sea, but it seems reasonable to assume that it is not *Dinophysis*.

Some genetic studies have identified sequences similar to members of the Chrysophyceae. For example, six 18S clones found in libraries from the Mediterranean,
Antarctica and the North Sea that were closely related to *Paraphysomonas foraminifera*. As this is a heterotrophic species, these sequences are not likely to be from phototrophic organisms (Diez et al. 2001b). Clone libraries from selected stations along the AMBITION Arabian Sea transect showed high numbers of novel chrysophyte sequences (Fuller et al. 2006a). The results from the Gulf of Naples libraries also show the importance of Chrysophyceae. The clone libraries show a high diversity with no affiliation to known sequences. As the known ultraplanktonic members of the Chrysophyceae have not been sequenced for 16S, it is not possible to say whether this diversity reflects a totally unknown diversity. The fact that several Gulf of Naples sequences show similarity to clone library sequences is interesting, particularly in that they are from such different regions. MC615-92 May shows 100 % identity to a sequence from Helgoland. The Mediterranean and the North Seas are very different with different physical and biological conditions. It is interesting to find the same sequence in both places, but this may not reflect species level identity, as 16S is a slowly evolving conserved gene. It could represent the same species complex or genus. Other Gulf of Naples sequences show some similarity to the Arabian Sea sequences, but not exact identity. The discovery of these sequences highlights the need for a more in depth investigation of this class at sea. The finding of eight genotypes unique to the May library shows that there is a seasonal pattern to the diversity within the Chrysophyceae.

The other classes found within the clone libraries shed light on the high taxonomic diversity within the ultraplanktonic fraction. Some of the members of the diatom genus *Skeletonema* when not in colonies are small enough to pass through a 5 µm filter, but that said the sequences could belong to another genus, as the 16S does not have a very
high resolution in the diatoms. The other classes are little studied within the Gulf of Naples but the indication of their presence is a stimulus for further study.

The results of this chapter indicate a clear seasonality at the class and species level with both blots and clones showing that classes and species are not present all year. This attributes an ecological value to the large amount of hidden diversity within the smaller fraction of the plankton. Although it is possible that this diversity corresponds to already described species for which 16S sequences are not available, the addition of sequences from cultures outside the Gulf of Naples did not ‘fill in the gaps’. For instance the addition of the Norwegian and North Sea *Chrysochromulina* sequences in the Prymnesiophyceae tree did not help identify any of the Gulf of Naples library sequences. However the addition of sequences from cultures isolated in the Gulf of Naples gave more information, for example in the case of the Cryptophyceae and *Phaeocystis*, possibly providing an indication of endemism or at least highlighting the need for cultures from the study area when interpreting clone library data. The likelihood is that the majority of the sequences in these libraries are new taxa (Pedrós-Alió 2006). If Pedrós-Alió (2006) is right and the diversity of a site can be retrieved by PCR then the diversity of Gulf of Naples is phenomenal.

**Chapter Summary**

- Dot blot hybridization provides a large scale means of assessing class level diversity over time using community DNA and bypassing the biases associated with isolation and culturing of single strains.
- Clone libraries give a display of the genetic diversity within a sample and a means of highlighting new genotypes as well as contributing to the seasonal data on known species.
• The data from the dot blots illustrate the seasonal patterns of occurrence of different classes within the ultraplankton, highlighting the major players at different times of the year.

• The Prymnesiophyceae are shown to be important components during most of the year, as are the Cryptophyceae.

• The Chrysophyceae are shown to be abundant in the summer months, an interesting results given that this class is largely overlooked at sea.

• The Pelagophyceae are shown to make a small contribution in February.

• The clone libraries highlight the high diversity within the three major classes found in the Gulf of Naples.

• The number of genotypes found possibly reflects already described species that have not been cultivated and characterized genetically, or entirely new genotypes.
Chapter Three

_Ultraphytoeukaryote Diversity in Depth Profiles during Four Seasons_

**INTRODUCTION**

The conditions in the marine environment vary with season and depth. Changes in temperature during the year affect the density of the water and induce changes in the structure and stability of the water column. Salinity also changes during the year due to inputs from different water masses (Carrada et al. 1980; Ribera d’Alcalà et al. 2004), for instance in the Gulf of Naples where water of higher salinity generally enters the Gulf at depth during the winter from the Tyrrenhenian Sea and lower salinity waters come from terrestrial runoff and may accumulate at the surface during the summer months due to reduced advection (Ribera d’Alcalà et al. 2004). Nutrient concentrations change during the year and with depth according to the different inputs and to biological activity. As with salinity, nutrients are determined according to the two different water masses in the Gulf, the open Tyrrenhenian nutrient rich waters and nutrient laden terrestrial runoff.

At any time of the year, there are differences in light with depth as a consequence of exponential attenuation. Light is absorbed in different ways down the water column so there is a decrease in the amount of light and in its quality (Falkowski and Raven 1997). Different wavelengths are absorbed at different depths creating the need for different pigments to maximise the efficiency of light capture and photosynthesis. All these differences create a number of niches throughout the year and at different depths not present in the terrestrial environment. The phylogenetic diversity seen within the marine photosynthetic community is therefore much higher than on land (Andersen 1992;
Baldauf 2003; Briggs 1994). Green plants dominate terrestrial environments whereas at sea major groups including the Heterokontophyta (diatoms, Prymnesiophyceae, Chrysophyceae), Cryptophyta and Cyanophyta employ a wide range of different pigments (Falkowski et al. 2004). The changing conditions down the water column allow the segregation of different populations at different depths and may even promote speciation as can be seen in different ecotypes within the prasinophyte Ostreococcus (Rodriguez et al. 2005) and Prochlorococcus (Rocap et al. 2003; West and Scanlan 1999; West et al. 2001) adapted to different depths and light regimes.

A flow cytometric study of ultraplankton along several depth profiles in the Atlantic showed a prevalence of smaller eukaryotic cells at depth with the average cell size decreasing from 3-4 μm at 20 and 65 m to 2-3 μm at 70 and 80 m (Li and Wood 1988). The eukaryotes showed a peak in abundance at 70 m before which there was only a slight decrease in abundance with depth and after which there was a more rapid decrease. It was concluded that the change in mean chl a fluorescence per cell with depth represented a physiological response to changes in light levels rather than the presence of two different populations, such as a ‘bright’ and ‘dim’ population sometimes seen in cyanobacteria. An investigation of several depth profiles in the eastern Mediterranean basin showed a decrease in eukaryotic ultraplankton below around 100 m (Li et al. 1993). However these profiles were in an area with low total chl a and such areas are generally considered to have the chl a accounted for by the <10 μm fraction (Li et al. 1993).

Investigations of differences in phytoplankton abundance and species composition with depth at MC carried out at the beginning of the MC Long-Term Ecological Programme demonstrated a seasonal cycle and the differences between upper and lower layers in the
water column (Ribera d'Alcalà et al. 2004; Zingone et al. 1990; Zingone et al. 1995). Under conditions of mixing, Prymnesiophyceae and Cryptophyceae, identified using HPLC, were shown alternate with the diatoms as the dominant group along the whole water column. When the water column became stratified, diatoms tended to dominate in surface waters down to 5-10 m with flagellates and cyanobacteria being present along the length of the water column (Ribera d'Alcalà et al. 2004). Prymnesiophyceae and other undetermined nanoflagellates were shown to be important in deeper layers and these occurrences were unaffected by the peaks in nutrients that are characteristic of the MC station during summer stratification (Ribera d'Alcalà et al. 2004). However both HPLC and the LM counts lump classes together. Therefore, the aim of this chapter is to determine the vertical distribution of the major classes within the ultraplankton using 16S class level probes that can distinguish more accurately between the major classes. The main question is whether the small flagellates, which dominate phytoplankton communities in subsurface waters for a large part of the year, exhibit specific patterns of vertical distribution at the class level and whether this pattern varies with the seasonal differences in the vertical structure of the water column. As station MC has variable conditions at depth depending on the time of year, the extremes of mixing and temperature were sampled to give an idea of how vertical distribution of ultraplankton may be influenced by depth and hydrological conditions. The results of this chapter are part of a collaborative project on the vertical distribution of ultraeukaryotes, involving the study of pigments and primary production (Immacolata Santarpia, Area Gestione Ambiente ed Ecologia Costiera delle Aree Temperate e Polari, SZN) and flow cytometry (Raffaella Casotti, Ecophysiology Laboratory, SZN).
**MATERIALS AND METHODS**

**Environmental Data**

CTD data were collected by the members of the Area Gestione Ambiente ed Ecologia Costiera delle Aree Temperate e Polari SZN for the four profiles using the same apparatus as outlined in Chapter 2. Nutrient concentrations were also collected in the same manner as for Chapter 2.

**Sampling**

Samples were collected from the MC station from 0 m, 2 m, 5 m, 10 m, 20 m, 30 m, 40 m, 50 m, 60 m and 70 m on four dates during 2005, 8^{th} March, 24^{th} May, 30^{th} August and 12^{th} December representing coldest mixed, coldest stratified, warmest stratified and warmest mixed conditions respectively. 70 m was not sampled on 24^{th} May and 30^{th} August due to a problem with the Niskin bottle at that depth. Water samples were collected using an automatic Carousel sampler from SeaBird Electronics, equipped with twelve 12 L Niskin bottles. Five litres of seawater were filtered using the method outlined in Chapter 2. 1 ml from each depth before and after filtration through 5 μm was fixed using 100 μl of 10 % paraformaldehyde and 10 μl of 8 % gluteraldehyde and stored at -80 °C for analysis using flow cytometry.

**DNA extraction and amplification of environmental samples**

For method see Chapter 2.

**Preparation and amplification of control DNA**

For method see Chapter 2.

**Dot blot hybridisation**

For method see Chapter 2.

**Flow cytometry**

Flow cytometry analyses of the four depth profiles were carried out in collaboration with Raffaella Casotti at the SZN using the protocol and settings previously described.
(Casotti et al. 2005; Marie et al. 1999) on a FACScalibur flow cytometer (Becton Dickinson, USA). This flow cytometer is not equipped to detect blue green cryptophytes (those that contain Cr-Phycocyanin) so it does not record species such as *Hemiselmis*. Samples were run and analysed by Raffaella Casotti. The data were compiled by Sarah McDonald.

**RESULTS**

**Environmental Data**

The CTD data are presented for the four depth profiles in Figure 3.1. These data show that the profile in March is mixed, except for a slight stratification in salinity with less saline waters between 0 and around 15 m. The fluorescence data for this profile were measured using a different fluorometer than the other profiles and are therefore not comparable with the other profiles, but still show homogeneity along the water column. Stratification is stronger in the May profile with salinity showing a stronger difference between 0 and about 5 m with fluorescence and oxygen following the same pattern. Temperature and density show a shallower decline with depth, most of the change being in the first 15 m. The August profile is more strongly stratified. There is a thermocline, with temperatures being around 25 °C between 0 and 20 m and then declining gradually reaching 15 °C at 70 m. Salinity shows a sinusoidal pattern with lower salinity waters between 0 and 7 m followed by higher salinity between 7 and 20 m when salinity begins to decline reaching 37.8 psu at around 30 m, which continues until 70 m. Density and oxygen also show changes at 7 and 20 m, with density increasing between 0 and 7 m reaching a plateau between 7 and 20 m before increasing gradually towards 70 m. Oxygen shows a more variable pattern increasing between 0 and 7 m, being constant between 7 and 20 m where it begins to increase gradually between 20 and 30 m, finally there is shallow dip between 47 and 67 m and a sharp decline between 67 and 70 m.
The December profile is completely mixed with constant values for all variables down to 70 m.

Figure 3.1: CTD data for the four profiles showing the changes in the variables during four different seasons. March shows slight stratification in salinity but mixing in the other variables. May shows more stratification with all variables showing differences with depth. There is stratification of the water column during August and the constant nature of all the variables in December indicates mixing of the whole water column.

Nutrient data are presented in Figure 3.2 showing the changes in TIN, PO$_4^{3-}$, and SiO$_4^{4-}$ over the four depth profiles. Nutrients show roughly constant values between 0 and 10 m in the March profile, followed by rapid changes between 10 and 20 m, the values change little towards 70 m. May shows a marked difference in SiO$_4^{4-}$ and TIN concentrations between 0 m and 10 m dropping from around 10 µmol l$^{-1}$ to 2 µmol l$^{-1}$. PO$_4^{3-}$ concentrations show a less dramatic decline, but do decrease gradually after a
slight increase between 0 and 2 m. August show much higher concentrations of PO₄³⁻, again with an increase between 0 and 2 m followed by a sharp decline from 0.125 µmol l⁻¹ at 2 m and 0.044 µmol l⁻¹ at 10 m and a further slight increase between 50 and 70 m. SiO₄⁴⁻ and TIN are less concentrated than the May profile, but again show sharp declines between 0 and 10 m and both follow PO₄³⁻ in showing slight increases between

Figure 3.2: Nutrient concentrations for the four depth profiles showing the increase in SiO₄⁴⁻ and TIN between March and May followed by a decrease towards December. PO₄³⁻ shows an irregular pattern being high in March lower in May, at its highest concentration in August before being less concentrated in December. All nutrients are more highly concentrated in surface waters.
50 and 70 m. Nutrients are low along the December profile, with both SiO$_4$$^{+}$ and TIN being around 2 µmol l$^{-1}$ for the whole profile. PO$_4$$^{3-}$ shows a more variable pattern declining between 0 and 2 m before increasing towards 20 m and declining again towards 70 m.

**Dot Blot Hybridisations**

Four of the probes produced signals above 2 %, PRYM666 for the Prymnesiophyceae, CRYP862 for the Cryptophyceae, CHRY1037 for the Chrysophyceae and PELA1035 for the Pelagophyceae (Excel files containing the outputs from TotalLab are available on the accompanying CD). The PRYM666 Prymnesiophyceae probe (Figure 3.3) almost always gives the highest signal at all depths and in all profiles, but shows some variation with depth. The MC654 March profile shows a general increasing trend with depth from 76 % at 2 m to 89 % at 70 m, however there are spikes at 0 and 5 m where the signal drops below 70 %. MC665 May shows an increase in signal between 0 m and 10 m followed by a decrease from 95.7 % to 22.8 % between 10 and 60 m. The MC679 August profile shows marked variations in signal with depth in a very irregular pattern,
with peaks at 0, 10 and 40 m and a decrease in signal from 85 % to 18 % between 50 and 60 m. The MC693 December profile shows an increase between 0 and 10 m from 36 % to 84 % followed by a decrease to 60 % at 40 m before returning to 80 % at 70 m.

The CRYP862 Cryptophyceae probe (Figure 3.4) gives signals above background in all profiles, with generally higher values in the upper layers. The MC654 March profile shows a peak at 10 m followed by a brisk decline and rather homogeneous values down to 70 m. In the MC665 May samples, 0 m is the highest signal and there is a smaller peak at 30-40 m. The stratified MC679 August profile shows a peak at 2-5 m followed by a sharp decline with a very small increase at 50 m. The MC693 December profile shows some variation in signal between 0 and 20 m followed by a gradual decline below 30 m.

The CHRY1037 probe (Figure 3.5) shows the highest hybridisation values in August, with a peak at 5 m followed by a gradual decline in signal towards 60 m, and the lowest
values, almost below detection levels, at all depths in March. The MC665 May and MC693 December profiles show intermediate values and very similar patterns, with subsurface peaks in signal at around 2 to 5 m and then a deeper but smaller peak at 30 to 40 m.

The Pelagophyceae PELA1035 probe (Figure 3.6) gives signals above background at depth in the MC654 March, MC665 May and MC679 August profiles increasing towards 60 m. The MC693 December profile shows a small increase in signal at 10 m but the signals are only slightly above background.

Figure 3.7 shows the total signal by the four probes, indicating the periods and depths where coverage by these four probes is not complete. The MC654 March profile is the best covered, with the two summer profiles being less well covered at depth. The MC693 December profile has low coverage in surface waters.
Figure 3.6: Percentage relative hybridisation for the PELA1035 Pelagophyceae probe over the four depth profiles.

Figure 3.7: Total signal from the four probes over the 2005 depth profiles.
Flow cytometry

The data shown in Figure 3.8 show the abundance of eukaryote cells <5 μm over the four depth profiles. All profiles show a marked decrease in cell numbers with depth and a general decrease in abundance from March to December. The MC654 March profile shows the highest abundance of cells reaching just above 8000 cells ml\(^{-1}\) at 2 m. The MC654 March and MC665 May profiles show the most dramatic decreases with depth. In the MC654 March profile the abundance falls to just above 2000 cells ml\(^{-1}\) at 70 m and in MC665 May from 5500 cells ml\(^{-1}\) at 0 m to around 500 at 60 m. The MC693 profile has the lowest abundance of cells with most of the profile having less than 1000 cells ml\(^{-1}\).

![Figure 3.8: Cytometry data giving the cell numbers for the ultraplankton (<5 μm) fraction.](image)

The percentage of cryptophytes detected by the flow cytometer relative to total numbers of eukaryotes is displayed in Figure 3.9. This is the only class that can be detected by the flow cytometer, which, as mentioned, cannot detect all cryptophytes as those containing Cr-Phycocyanin are missed. The MC654 March profile shows that this class
is not a large percentage of the eukaryotes present with a very slight increase towards 70 m. The other two profiles match the % relative hybridization data, with the MC679 May profile showing peaks in percentage at 0 m, between 20 and 30 m and at 70 m. There are two peaks in the MC679 August profile at 2 and 50 m, the latter being more pronounced than the one shown by the probe (Figure 3.4). Members of this class could not be detected in the samples from the December MC693 profile.

![Graphs showing Cryptophyceae cytometry data for different profiles](image)

Figure 3.9: Percentage of Cryptophyceae in each cytometer sample for each depth profile, calculated using the number of cryptophyte cells over the total number of eukaryotes.

**DISCUSSION**

**Dot Blot Hybridisation**

The results from the probes show that the classes Chrysophyceae, Cryptophyceae and Prymnesiophyceae are not only important at the surface during the year (Chapter Two), but are also present in significant proportions at depth. All classes show variations in signal with depth and season. There are often subsurface peaks in signal between 2 and 10 m for all the classes found. Another common feature of all the profiles is sudden

85
peaks along the profile possibly indicating the presence of different species with distinct depth preferences. The depth analyses also highlight the importance of the Pelagophyceae at depths of between 60 and 70 m.

The Prymnesiophyceae are shown to be present all year round at most depths with spikes in signal that do not match the data on pigments from HPLC (Santarpia 2005 and unpublished data). 19'HF is generally used as the marker pigment for the Prymnesiophyceae, but it is difficult to match its concentration with the signal from the probe in any of the profiles. In the MC665 May profile, the pigment is at its highest concentration for the four profiles, with almost equally high values at 0 m and 20 m (0.35 µg l⁻¹ and 0.21 µg l⁻¹ respectively) (Santarpia 2005), whereas the probe shows its highest signal at 10 m where the concentration of the pigment is much lower (0.09 µg l⁻¹). The MC679 August profile shows a similar lack of correspondence, with the concentration of 19'HF decreasing steadily with depth (from 0.06 to 0.04 µg l⁻¹) and the probe showing erratic changes in signal with depth. In the MC693 December profile, the strong ‘inverse’ stratification is not repeated by pigment data that do not show any differences in 19'HF with depth (I. Santarpia unpublished data). These differences between pigment and probe signal could be due to the fact that the pigment was only sampled at 0 m, 10 m, 20 m and 40 m and so could be less subtle in picking up fluctuations in the class with depth. The fact that not all Prymnesiophyceae contain 19’HF could also explain the differences.

The cryptophyte probe shows the importance of this group in the samples of the December profile and to a lesser extent in the March profile. During stratification this class seems to show some preference for waters just below the surface having the highest signal from the probe at 5 m in August. The surface waters of the Gulf of
Naples during the summer months can be eutrophic caused by runoff from the city of Naples and SDC data show that, under these conditions, 0 m is dominated by a few species of small diatoms, which may explain the exclusion of other groups from surface waters. In the May and August profiles, there is a close match between the % of cryptophyte cells seen in the cytometer and the signal from the probe. In the May profile, the cytometer data show peaks at 0 m, between 20 and 30 m and at 60 m, the probe has a similar pattern with peaks at 0 m and between 30 and 40 m. The August profile has peaks both in the % of cryptophyte cells and in % relative hybridisation at around 2-5 m and 50 m. The only disparity between the dot blots, flow cytometry and HPLC is in the December profile. The abundance of total eukaryote cells is lowest in December and cryptophyte cells were not detected in these samples in the December profile. Alloxanthin was also not detected in samples from this profile. Yet the Cryptophyceae probe gives the highest signal in these samples. This is matched by data from Fuller et al (2006b) in the Arabian Sea where cryptophytes were detected in samples where the flow cytometer did not detect any cells. Given the parity between the methods in the other profiles, it seems unlikely that the probe would be wrong in the December profile. It is possible that it is due to PCR being more sensitive (Fuller et al. 2006b) or to the fact that for the dot blots 5 litres of water were filtered whereas cytometry relies on a sample of 1 ml and HPLC on 3 litres. It may be that when eukaryote cell numbers are low, there is a lower amount of 16S detected by the eubacterial probe, therefore the signal from the Cryptophyceae seems higher. It is still possible to say that this class makes up a significant proportion of the winter population of ultraeukaryotes in the Gulf of Naples. Overall the different methods show similar trends, even if the probe is not quantitative, giving an indication of the reliability of the dot blot method for tracking trends in the occurrence of different classes during the year.
The Chrysophyceae show a similar pattern to the results of Chapter 2 in that they are mainly present in summer, present at low levels in December and are almost below detection in the March profile. The May and August profiles show that the signal is greatest just below the surface, in May at 2 m and at 5 m in August. This is again likely to be due to the dominance in surface waters during the summer of a few diatom species. The May profile shows an interesting increase in signal at 30 m that could possibly be due to low light adapted species within this class. The August profile shows a decrease in signal with depth. As there is no unequivocal marker pigment for this class, it is not possible to verify the probe signals using cytometry or pigment data.

The Pelagophyceae are shown by the PELA1035 probe to be present at depth between March and August, with the highest signal being in August. This is in contrast to the peak seen in the surface water data from Chapter 2 where the Pelagophyceae were present in February at the surface. This class has been previously shown to be important at depth at the BATS Hydrostation S in the Atlantic and the HOT ALOHA station in the Pacific by pigments analyses using 19'BF as a marker pigment and epifluorescence microscopy (Andersen et al. 1996). Several of the pelagophyte strains held in the world’s culture collections have been isolated from the deep chlorophyll maximum (Guillou et al. 1999b) and are maintained at low light levels (Not et al. 2002), indicating a physiological preference for the lower light levels in deeper waters. This class has also been shown to be present in mesotrophic waters (Fuller et al. 2006b), which may explain their presence at depth during the summer when the surface waters tend to be more eutrophic in the Gulf of Naples. The marker pigment used by Andersen et al. (1996), 19’BF, does not show a clear match to the probe apart from in the August profile when its concentration increases considerably between 20 and 40 m. The
inconsistency between the pigment and the probe in the other profiles could be due to
the Pelagophyceae being detected below the final HPLC sample, i.e. below 40 m, or the
fact that this pigment is not a reliable indicator for the Pelagophyceae.

From the graph showing the total coverage by the four probes, it is difficult to assess
whether the depths and periods with less than 100 % correspond to periods when other
groups without a 16S probe were present. The data from Chapter 2 indicate that even
when the probes do not approach 100 %, the libraries do not show dominance by a class
lacking a 16S probe. The fact that the four probes do not add up to 100 % could be due
to the PCR product not amplifying the most abundant class due to reannealing dynamics
(Suzuki and Giovannoni 1996). It could also be due to the probe not binding well to all
members of a particular class in those samples, perhaps due to mismatches in the target
region. The lowest levels of total signal are in the MC665 May profile towards 60 m
and between 0 and 5 m in the MC693 December profile. All pigments, from zeaxanthin
to 19'HF, show a drop in concentration between 20 and 40 m in the May profile,
leading to the conclusion that the drop in signal is not due to a lack of coverage by the
probes, but is more probably a real drop in all groups with depth. The only pigment
showing some stratification in the MC693 December profile is chl c2. Chl c2 has a
slightly higher concentration at 10 m than the rest of the profile (0 m 0.018 μg l\(^{-1}\)
10 m 0.022 μg l\(^{-1}\) 20 m 0.019 μg l\(^{-1}\) and 40 m 0.017 μg l\(^{-1}\)). This pigment is characteristic of the
golden brown algae, the Heterokonta (diatoms, Bolidophyceae, Prymnesiophyceae,
Chrysophyceae and Pelagophyceae). The Utermohl counts show a high proportion of
undetermined phytoflagellates in December at MC693 0 m, despite there being
relatively few cells (622 cells ml\(^{-1}\)) 71 % of them were phytoflagellates <10 μm. The
low total signal between 0 and 5 m could be due to the probes not binding well, or to the
presence of another class not covered by the probes, but with a similar pigment profile, essentially any other golden brown algal class.

The flow cytometry data show that the Gulf of Naples does not follow the pattern seen in the Bay of Banyuls (Charles et al. 2005) or the Thau lagoon (Bec et al. 2005) in that small eukaryote cells are most abundant in March rather than January or during the summer months. This is likely to be due to the different hydrographical conditions creating the two subsystems within the Gulf of Naples, between which the station MC is situated. The winter occurrence of eukaryotes in the Bay of Banyuls was linked to influxes of nutrients (Charles et al. 2005) and the presence of eukaryotes in summer in the Thau lagoon was correlated with high temperature and light (Bec et al. 2005).

March in the Gulf of Naples is the beginning of the annual increase in cell numbers and the high abundance of small eukaryotes may reflect their ability to efficiently assimilate nutrients due to their small size allowing them to increase in numbers before the main phytoplankton bloom in May. Their dominance in winter may also reflect their ability to utilise the lower concentrations of nutrients during these months, seen in the nutrient data in Figure 3.4.

Any differences in vertical distribution may indicate that class specific sets of pigments have a depth related functional role. This may account for the presence of the Pelagophyceae at depth and for the spikes in signal in the other classes that may be due to depth-adapted taxa within those classes. Fuller et al (2006a) showed marked differences in abundance of cells smaller than 3 μm below depths of between 30 and 50 m at the open water sites Stn 2 and Stn 10 in the Arabian Sea. Dot blot hybridisations revealed the dominance by chrysophytes with prymnesiophytes being the second most abundant class (Fuller et al. 2006b). Cryptophytes were much less abundant and were
restricted to waters between 20 and 30 m depth at the higher-nutrient and coastal Stations 9 to 11. The Trebouxiophyceae were also only found at depth, but much deeper than the cryptophytes, at around 200 m. These data indicate a depth adaptation by certain members of each class. The data also imply that these classes are made up of different members in the Arabian Sea and the Gulf of Naples given the differences in depth distribution, for example the presence of pelagophytes at depth in Naples and in surface waters in the Arabian Sea.

**CHAPTER SUMMARY**

- As with the results of Chapter Two, the main classes detected using the probes are Prymnesiophyceae, Cryptophyceae, Chrysophyceae and Pelagophyceae.
- The Prymnesiophyceae are present in all profiles showing spikes in signal with depth in the March, August and December profiles.
- Cryptophyceae are also present in all profiles and dot blot data support the idea that this class is an important proportion of the winter population despite low abundance of eukaryotes.
- The Chrysophyceae again show a summer distribution with a decrease with depth during stratification in May and August.
- The Pelagophyceae are present at depth during March, May and August and are below detection in the December profile, possibly indicating an adaptation or preference for deeper waters.
- All classes show a preference for depths around 5-10 m rather than surface waters, possibly due to the dominance of surface waters by a few species small diatoms.
• It is not always easy to match the results from the dot blots with the results from pigment analyses, perhaps highlighting the difficulty in using pigments as unequivocal markers for different classes.

• Spikes in signal down the profile may indicate the presence of different species within each class adapted to distinct depths, indicating the need for finer level probes to identify species groups within the ultraplankton.
Chapter Four

Diversity in the Diatom Genus Pseudo-nitzschia over an Annual Cycle

Introduction

*Pseudo-nitzschia* is a diatom genus containing several potentially toxic species. Although the morphology of the frustule, as with all diatoms, yields a number of useful taxonomic characters, the identification of some species within this genus is problematic. The genus contains a number of species complexes that in recent years have been split into new species that are not easily distinguishable in LM (Lundholm et al. 2002b; Lundholm et al. 2003). These include *P. pseudodelicatissima, P. calliantha*, *P. caciantha* formerly included within *P. pseudodelicatissima*, and *P. delicatissima, P. dolorosa* and *P. decipiens* that were previously described as *P. delicatissima*. A number of different studies conducted at SZN using cultured strains of *Pseudo-nitzschia* isolated from the Gulf of Naples have identified other new species. For example, in the isolation of approximately 100 strains from cells that resembled *P. delicatissima* or *P. pseudodelicatissima*, eight taxa were identified including two new genotypes, *P. calliantha* clade 2 and *P. delicatissima* clade 2 (Amato et al. in prep). All these taxa are referred to as pseudo-cryptic because of the difficulty in detecting morphological differences between genetically distinct entities. However, it is possible that the isolation of single cells may grossly under-sample the number of distinct taxa present in any given area. The strains isolated by Amato et al. were isolated over several months (between January and April) from samples containing thousands of cells per ml. A method that does not require the isolation of strains is needed to investigate whether further taxa exist within *Pseudo-nitzschia*.
A consequence of this higher level of diversity is that it must be possible to track these entities in the environment in order to understand ecosystem dynamics and energy flows. If the pseudo-cryptic taxa are lumped together in their complexes ecological meaningful groups are not being analysed. This is especially important when considering the groups within *Pseudo-nitzschia* as some members of a complex may be toxic and others not, some strains of a particular morphological taxon have been shown to be toxic and others not. A clearer understanding of the taxonomy of toxic taxa is essential for the study of toxicity in *Pseudo-nitzschia*, particularly whether the environmental influences the toxicity of a particular taxon. It is important from the monitoring point of view to be able to track the temporal occurrence of these different species. Information gathered during the MC Long-term Ecological Programme on phytoplankton abundance shows the regular occurrence of some species whereas other ‘species’ (e.g. ‘*P. delicatissima*’ and ‘*P. pseudodelicatissima*’) seem to be present most of the year (Zingone et al. 2002; Zingone et al. 2003). Within this ecological programme, the pseudo-cryptic species are lumped together in their species-complex names, and a catchall category of *Pseudo-nitzschia* spp. includes single cells as *Pseudo-nitzschia* species are difficult to identify unless the overlap between adjacent cells can be observed in colonies.

The aim of this study was to develop and use a genus specific primer to track different genotypes within *Pseudo-nitzschia* directly from community DNA to overcome the problem of under-sampling the diversity present by using cultures isolated from single cells. LSU rDNA was chosen as the marker region because there are a high number of sequences available for both *Pseudo-nitzschia* and other diatoms making the design of primers more accurate. The initial idea was to use a polyacrylamide gel to separate different DNA fragments in a PCR product from a natural sample based on slight
differences in their sequences, and subsequently identify the species present in each sample by the position of the bands on the gel. The method chosen for this purpose was SSCP (Single Strand Conformation Polymorphism), where single stranded DNA is denatured and migrated through the gel where differences in the secondary structure of the fragments, caused by differences in their sequence, result in variation in the migration of the fragment in the gel. Clone libraries from different dates were planned to construct a database of *Pseudo-nitzschia* sequences from selected periods in the Gulf of Naples to be compared to the results from the SSCP and to sequences from strains held in the culture collection of the SZN. Sampling was carried out over 15 months to give the seasonal distribution of the various taxa within *Pseudo-nitzschia* in an attempt to clarify whether some taxa really do occur throughout the year.

**Materials and Methods**

**Sampling**

Samples of seawater were collected from 0 m at the MC sampling site between July 2003 and September 2004. Five litres were filtered onto a 3 μm filter at 200 mmHg using a vacuum pump. The filters were cut in sections, immediately frozen in liquid nitrogen and stored in Eppendorf tubes at -80°C until the extraction of DNA.

**DNA extraction and amplification**

DNA was extracted from sections of the filter using the method outlined in Chapter 2. A nested PCR amplification using three sets of primers was performed (see Table 4.1 for all primers used). The universal LSU rDNA primers D3Ca and DIR were first used to amplify a large fragment (c. 800 bp) from the community DNA extracted from the filters. The D3Ca/DIR fragment was amplified in a total reaction volume of 50 μl containing 1 μl environmental DNA (approximately 40 ng), 200 μM deoxynucleoside triphosphates, 1 μM each primer and 2.5 U Taq polymerase in 1× enzyme buffer
containing 1.5 mM MgCl₂ (Roche Diagnostics GmbH, Mannheim, Germany). The PCR conditions for D3Ca/DIR were 94 °C for 4 min followed by 35 cycles of 94 °C for 1 min, annealing at 55 °C for 1 min 30 sec and elongation at 72 °C for 1 min, followed by further elongation at 72 °C for 5 min. This PCR product was purified by excising the band from a 1 % (w/v) agarose gel and using a Gel Purification Kit from Qiagen (Qiagen Ltd.). This product was then used as a template in a PCR using newly developed primers specific to *Pseudo-nitzschia* amplifying a 350bp fragment in the D1 region, D1-186F and D1-548R. The PCR mix consisted of 1 μl of purified PCR product (approximately 40 ng), 200 μM deoxynucleoside triphosphates, 2.5 mM MgCl₂, 1 μM each primer and 2.5 U Taq polymerase in 1× enzyme buffer -Mg (Roche Diagnostics GmbH, Mannheim, Germany) in a total volume of 50 μl. The PCR conditions were 94 °C for 4 min followed by 35 cycles of 94 °C for 1 min, annealing at 62 °C for 35 sec and elongation at 72 °C for 1 min 20 sec, followed by further elongation at 72 °C for 5 min. This fragment proved too long to distinguish between closely related *Pseudo-nitzschia* species in the SSCP analyses, hence a new primer was developed to give an approximately 91bp fragment in conjunction with D1-186F, D1-258R. Amplification used a mix consisting of 1 μl of purified PCR product (approximately 40 ng), 200 μM deoxynucleoside triphosphates, 1.3 mM MgCl₂, 0.2 μM each primer and 2.5 U Taq polymerase in 1× enzyme buffer -Mg (Roche Diagnostics GmbH, Mannheim, Germany) in a total volume of 50 μl.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence 3'-5'</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3Ca</td>
<td>ACG AAC GAT TTG CAC GTC AG</td>
<td>Scholin et al. (1994a)</td>
</tr>
<tr>
<td>DIR</td>
<td>ACC CGC TGA ATT TAA GCA TA</td>
<td>Lenaers et al. (1989)</td>
</tr>
<tr>
<td>D1-186F</td>
<td>GTT CCT TGG AAA AGG ACA GCT GA</td>
<td>This study</td>
</tr>
<tr>
<td>D1-548R</td>
<td>AGA CAT CAA CTC TGA CTG</td>
<td>This study</td>
</tr>
<tr>
<td>D1-258R</td>
<td>GCA ATC CCA AAC AAC TCG ACT C</td>
<td>This study</td>
</tr>
</tbody>
</table>
The PCR conditions were 80 °C for 4 min followed by 35 cycles of 94 °C for 30 sec, annealing at 62.4 °C for 15 sec and elongation at 72 °C for 20 sec, followed by further elongation at 72 °C for 1 min. Both reverse primers were phosphorylated to enable digestion by exonuclease (see below).

**Primer development**

The primers developed in this study were found by eye using an alignment of known *Pseudo-nitzschia* LSU sequences held in GenBank and in the SZN database (courtesy of A. Amato). The 350 bp D1-186F/D1-548R fragment contains two loop regions with a conserved stem region in between. The shorter D1-186F/D1-258R contains just one loop region and part of the stem. The primer specificity was tested using BLAST (Altschul et al. 1990) and PCR using DNA extracted from cultures of other common diatom species in the Gulf of Naples using the CTAB extraction method outlined in Chapter 2 (culture material kindly provided by Wiebe Kooistra, see Table 4.2 for details).

<table>
<thead>
<tr>
<th>Species</th>
<th>Isolating Scientist</th>
<th>Isolation Date</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Odontella mobiliensis</em></td>
<td>Wiebe Kooistra</td>
<td>26th January 2004</td>
<td>GON MC</td>
</tr>
<tr>
<td><em>Ditylum brightwellii</em></td>
<td>Wiebe Kooistra</td>
<td>26th January 2004</td>
<td>GON MC</td>
</tr>
<tr>
<td><em>Leptocylindrus danicus</em></td>
<td>Wiebe Kooistra</td>
<td>26th January 2004</td>
<td>GON MC</td>
</tr>
<tr>
<td><em>Chaetoceros curvisetus</em></td>
<td>Wiebe Kooistra</td>
<td>26th January 2004</td>
<td>GON MC</td>
</tr>
</tbody>
</table>

**SSCP**

The protocols used were taken from the Multiphor II system handbook and Dohrmann and Tebbe (2004). The digestion of the reverse phosphorylated strand was carried out using Lambda Exonuclease (New England BioLabs) with 1x reaction buffer and 2.5 U of enzyme added to 30 µl of PCR product to make a final reaction volume of 40 µl. Digestion was carried out at 37 °C for 45 min. Immediately following digestion the products were purified using a QIAquick PCR purification kit (Qiagen Ltd.). The PCR
products used in the optimisation of SSCP were run both digested and undigested and fragments from the D1-186F/D1-548R and D1-186F/D1-258R primer pairs were tested. The gels tested were: pre-casted ExcelGel™ SDS Homogenous 12.5 gels (Amersham Biosciences) run with buffer strips and 0.5x MDE Gel solution gels (Cambrex) with and without 10 % glycerol that were run with continuous 0.6x TBE buffer. The MDE gels were cast using a SDS and Native PAGE IEF Kit (Amersham Biosciences) that includes glass plates and clamps.

Table 4.3: Cultures used as controls in the optimisation of SSCP. GON is Gulf of Naples. All strains except \textit{P. delicatissima} 2 were isolated from net samples; \textit{P. delicatissima} 2 was isolated from a water sample collected from the beach.

<table>
<thead>
<tr>
<th>Species</th>
<th>Strain</th>
<th>Isolating Scientist</th>
<th>Date</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{P. delicatissima}</td>
<td>AL90</td>
<td>Alberto Amato</td>
<td>21\textsuperscript{st} April 2004</td>
<td>GON MC</td>
</tr>
<tr>
<td>\textit{P. galaxiae}</td>
<td>SM26</td>
<td>Sarah McDonald</td>
<td>9\textsuperscript{th} June 2004</td>
<td>GON MC</td>
</tr>
<tr>
<td>\textit{P. fraudulenta}</td>
<td>AL75</td>
<td>Alberto Amato</td>
<td>25\textsuperscript{th} February 2004</td>
<td>GON MC</td>
</tr>
<tr>
<td>\textit{P. dolorosa}</td>
<td>AL25</td>
<td>Alberto Amato</td>
<td>26\textsuperscript{th} January 2004</td>
<td>GON MC</td>
</tr>
<tr>
<td>\textit{P. caciantha}</td>
<td>AL37</td>
<td>Alberto Amato</td>
<td>21\textsuperscript{st} January 2004</td>
<td>GON MC</td>
</tr>
<tr>
<td>\textit{P. multistriata}</td>
<td>1</td>
<td>Alec Lüdeking</td>
<td>1\textsuperscript{st} February 2005</td>
<td>GON MC</td>
</tr>
<tr>
<td>\textit{P. pseudodelicatissima}</td>
<td>AL31</td>
<td>Alberto Amato</td>
<td>26\textsuperscript{th} January 2004</td>
<td>GON MC</td>
</tr>
<tr>
<td>\textit{P. delicatissima} 2</td>
<td>AL95</td>
<td>Alberto Amato</td>
<td>18\textsuperscript{th} June 2004</td>
<td>Gulf of Gaeta</td>
</tr>
<tr>
<td>\textit{P. calliantha}</td>
<td>AL76</td>
<td>Alberto Amato</td>
<td>25\textsuperscript{th} February 2004</td>
<td>GON MC</td>
</tr>
</tbody>
</table>

All gels were run on a Multiphor II Electrophoresis Unit with a MultiTemp III Thermostatic Circulator and an electrophoresis power supply EPS 3501 (Amersham Biosciences). Gels were run at 4, 10, 15 and 20 °C. 6 μl of sample was loaded that included 3 μl of PCR product and 3 μl of loading buffer (10 mM NaOH, 0.25 % xylene cyanole [wt/vol], 0.25 % bromophenol blue [wt/vol], 95 % formamide [vol/vol] Dohrmann and Tebbe 2004). These samples were denatured at 94 °C for 5 mins and held on ice before loading on the gel. The cultures used in the optimisation are listed in Table 4.3.
Clone libraries

Clone libraries were constructed for six dates in 2004, chosen using the results of the phytoplankton counts (courtesy of Diana Sarno) to sample those dates where the highest diversity in *Pseudo-nitzschia* was observed attempting to avoid large peaks where single species dominated the sample (see Figure 4.2). Those dates were 5th April, 26th May, 22nd June, 20th July, 31st August and 26th October 2004. The libraries were constructed for the longer D1-186F/D1-548R fragment, as there is more variation in this fragment. The products from two 50 µl PCR reactions were pooled and purified using a QIAquick PCR purification kit (Qiagen Ltd.) and libraries constructed using TOPO TA Cloning® kit (Invitrogen™ Life Technologies, Carlsbad, California) and the automated Miniprep and sequencing as outlined in Chapter 2. Coverage of each clone library and rarefaction curves were also calculated as in Chapter 2.

Known sequences

For the purposes of phylogenetic analysis of the clone libraries, sequences from known *Pseudo-nitzschia* species were downloaded from GenBank. A number of *Pseudo-nitzschia* LSU sequences from Gulf of Naples isolates were kindly provided by Alberto Amato. *P. galaxiae* sequences from the Chapter 5 study were also added. In addition DNA was extracted using the CTAB method outlined in Chapter Two from material for the recently described *P. decipens* Mex12 strain, kindly provided by Nina Lundholm, and was amplified and sequenced for LSU.

Phylogenetic analyses

Contigs and alignments were constructed using Seqman and BioEdit as in Chapter Two. LSU sequences for *Pseudo-nitzschia* species from GenBank and from the control cultures (see Table 4.3) were added to the alignment to help the identification of the genotypes within the libraries. All libraries were analysed together. Two alignments were analysed, one containing genotypes that had base changes in either of the two loop
regions and for which there were at least two library sequences, and the other containing genotypes that differed from a known taxon by only one base change situated in the stem region and for which there was only one representative. Phylogenetic analyses were carried out in PAUP* using Neighbour Joining tree building method with the Maximum Likelihood settings generated by Modeltest.

RESULTS

Primer development

The D1-186F, D1-258R and D1-548R primers were shown to be specific to *Pseudo-nitzschia* by BLAST searches. Both primer pairs gave products of the correct length. The specificity was further proved by successful amplifications and sequencing from genomic DNA of known cultures of *Pseudo-nitzschia* isolated from the Gulf of Naples (Table 4.3). Further confirmation came from PCRs using genomic DNA from other diatom species found in the Gulf of Naples (Table 4.2) that either gave no product or a fragment considerably longer than expected. The D1-186F/D1-548R fragment contained enough variation within the two loop regions to separate all known genotypes within *Pseudo-nitzschia*, but the single loop region within the D1-186F/D1-258R fragment did not contain enough polymorphism to distinguish *P. australis* from *P. multistriata*. As *P. australis* is not present in the Gulf of Naples, this primer set is sufficiently discriminating for this study. The D1-186F/D1-548R pair clone libraries yielded only *Pseudo-nitzschia* sequences indicating that the primers are specific enough not to amplify other organisms.

SSCP

SSCP was carried out using PCR products from both cultures and natural samples. The cultures were amplified directly from genomic DNA using the specific primer pairs and
the natural samples were amplified using the nested PCR with the universal and specific LSU primers.

An example of a SSCP gel run using the longer fragment is shown in Figure 4.1. The separation between closely related *Pseudo-nitzschia* by either primer pair under any of the conditions and gel types tested was not sufficient to distinguish between the ten *Pseudo-nitzschia* cultures. This can be seen in the above figure, particularly between the bands for *P. delicatissima*, *P. dolorosa* and *P. delicatissima* clade 2. The mixed sample (containing PCR products in equal concentrations from the ten cultures) in the centre of the gel appears as a fuzzy band and it is not possible to distinguish the ten species within this sample. Digestion did not make it any more possible to distinguish between closely related species as can be seen by the two lanes per culture, one containing digested PCR products that should appear as one band and the other containing undigested products that should contain two bands.
The seasonality of the seven taxa recognised in light microscopy during the study period (courtesy of Diana Sarno) can been seen from Figure 4.2. Figure 4.3 shows the same data with *P. delicatissima* and *P. galaxiae* removed to highlight the trends of the less abundant species.
Figure 4.3: Same data as Figure 4.2 but with *P. delicatissima* and *P. galaxiae* removed to demonstrate more clearly the abundances of the less numerous species.

Figure 4.4 shows the percentage composition of the *Pseudo-nitzschia* populations as recorded with the two methods, counts of the seven morphospecies (including the catchall *Pseudo-nitzschia* spp. category) in Utermohl fixed samples and the thirteen LSU genotypes in clone libraries. The clone libraries detected a higher number of taxa at all sampling dates, apart from the April library where there were less. There is little match between the counts and the libraries. Some genotypes showed restriction to one particular library, for example, *P. calliantha* 2 October, *P. delicatissima* 2 mostly April and *P. delicatissima* also mostly found in the April library. Whereas other genotypes were present in almost all libraries, for example *P. galaxiae* LSU clade II present in five of the six libraries from May to October.
Diversity - Clone Libraries

The nested PCR, first using universal LSU primers followed by the *Pseudo-nitzschia* specific primers, was necessary because it was not possible to directly amplify *Pseudo-nitzschia* from the natural samples using the specific primers. Occasionally, samples from the summer months amplified directly giving a very weak product.
Table 4.4: Details of the number of sequences sequenced per library, the number of genotypes found and the coverage of each library.

<table>
<thead>
<tr>
<th>Library</th>
<th>Date</th>
<th>Number of clones sequenced</th>
<th>Number of useful sequences</th>
<th>Number of unique genotypes</th>
<th>Coverage Value %</th>
</tr>
</thead>
<tbody>
<tr>
<td>050404</td>
<td>5th April 2004</td>
<td>48</td>
<td>31</td>
<td>2</td>
<td>93.5</td>
</tr>
<tr>
<td>260504</td>
<td>26th May 2004</td>
<td>48</td>
<td>21</td>
<td>4</td>
<td>80.9</td>
</tr>
<tr>
<td>220604</td>
<td>22nd June 2004</td>
<td>48</td>
<td>34</td>
<td>5</td>
<td>85.3</td>
</tr>
<tr>
<td>200704</td>
<td>20th July 2004</td>
<td>48</td>
<td>30</td>
<td>4</td>
<td>86.6</td>
</tr>
<tr>
<td>310804</td>
<td>31st August 2004</td>
<td>48</td>
<td>32</td>
<td>5</td>
<td>84.4</td>
</tr>
<tr>
<td>261004</td>
<td>26th October 2004</td>
<td>48</td>
<td>25</td>
<td>6</td>
<td>76.0</td>
</tr>
</tbody>
</table>

Table 4.4 shows the number of sequences from each library that proved to be useful and the number of genotypes found per library. Sequences were eliminated from this analysis if the contig was not readable, if it was suspected to be a chimeric sequence or if the polymorphism was only seen in one sequence and was found in the central stem region of the fragment. The detection of chimeras was done by eye using the alignment with sequences from known cultures, as these sequences were only found once they were assumed not to be new genotypes and removed from the analysis. Sequences differing from the known strains were only considered to be new genotypes if found more than once in the six libraries. The alignments and chromatogram files are available on the accompanying CD.

The coverage for each library is high despite the low number of clones sequenced and the high percentage of sequences removed from the analysis. Figure 4.5 shows the rarefaction curve for each library giving another measure of how well the diversity in each sample was covered by the number of clones sequenced. The April library is at an asymptote indicating that the diversity within this sample was well sampled by the number of clones examined. The rest of the libraries do not seem to be as near to an asymptote.
Figure 4.5: Rarefaction curves for the six libraries.

Figure 4.6 shows the NJ tree from the clone library sequences and includes some GenBank *Pseudo-nitzschia* sequences (black), some sequences from the SZN database (courtesy of Alberto Amato and shown in green) and three *P. galaxiae* sequences and one *P. decipiens* sequence obtained from cultures used in the study outlined in Chapter 5 (red). A total of thirteen genotypes were found. Ten of which had already been detected, including seven genotypes identical to those obtained from cultures of six described *Pseudo-nitzschia* species (two *P. galaxiae* LSU genotypes visible from the two sequences used in the original description Lundholm and Moestrup 2002). Three further genotypes recently identified from cultures from the Gulf of Naples were also found, *P. calliantha* clade 2 and *P. delicatissima* clade 2, which probably correspond to new species (A. Amato, pers. comm.) and one from cultures morphologically identical to *P. galaxiae* (*P. galaxiae* LSU clade III). Three completely new lineages were found, of which two were found within *P. galaxiae*, LSU clades IV and V in the figure, and one as a possible sister group to *P. delicatissima*, *Pseudo-nitzschia* new genotype.
Figure 4.6: NJ tree constructed using ML settings calculated using Modeltest (best fit model GTR+I+G A = 0.2691, C = 0.171, G = 0.2901 and T = 0.2698, A = 0.0986, A = 0.4148, A = 0.0986, C = 4.2536 and C = 0.5661, the positional rate variation across the alignment γ = 0.6885 and the proportion of invariable sites I = 0.41) for the sequences found in the six clone libraries. Sequences that did not exactly match known sequences were only included in this analysis if they were found more than once and the base changes were found in the loop regions, with the exception of P. calliantha, for which an exact match was not found.
Figure 4.7: NJ tree for all the sequences found that were similar to *P. galaxiae* constructed using ML settings calculated using Modeltest (best fit model K81uf+1+G A = 0.2442, C = 0.16, G = 0.3108 and T = 0.285, A G = 3.9877, A T = 0.793, A C = 1, C T = 3.9877 and C G = 0.793, the positional rate variation across the alignment γ = 0.5208 and the proportion of invariable sites I = 0.6791).

Figure 4.7 shows the tree resulting from the alignment including the genotypes that differed from a known taxon by a single base pair and were only represented by a single clone sequence. Only those found within *P. galaxiae* are shown as it is within this taxon.
that most of these sequences were found. A number of sequences that differed by a single base from types II and V were found. These base pair differences were each in a different position and found in the stem region of the D1-186F/D1-548R fragment.

**DISCUSSION**

The genus-level primers developed for *Pseudo-nitzschia* revealed a higher level of diversity within *Pseudo-nitzschia* in the Gulf of Naples than could be seen in LM. The results from the clone libraries gave some indication as to the seasonality of these taxa during the year, however it was not possible to use the shorter D1-186F/D1-258R primer pair to track the seasonality of these taxa more precisely using SSCP.

**Primers**

The two primer sets were shown to be specific to *Pseudo-nitzschia*. However it was not possible to amplify directly from community DNA and the primers had to be used as part of a nested PCR using universal LSU primers, which is a limitation. As it was necessary to perform two reactions, there was double the opportunity for errors to occur. Any errors that do occur, especially if they are in the first reaction, will be amplified in the second and will be more likely to be picked up in either of the two methods used in this study to detect taxa. It is possible that the primers did not amplify a product directly from the community DNA because of the presence of too many other organisms. Unlike the 16S primers used in the first two chapters, the *Pseudo-nitzschia* primers are targeting a small component of the phytoplankton that at its most abundant makes up around 20% of the phytoplankton in a sample, but only around 3% for most of the year (D. Sarno unpublished data). It is therefore less likely that the primers will find their target.
The long branches seen in a number of ‘species’, for example *P. galaxiae* and *P. calliantha*, are single base differences in the middle of the fragment. These base changes are outside the two loop regions, in positions that are conserved in all other sequences. The changes are at random, in that they are not found in the same position in more than one sequence. The chromatograms generally show a small second peak with the ‘correct’ base in either the forward or reverse sequence. It is probable that these sequences do not represent different genotypes. Rather they are Taq or sequencing errors, or possibly intra-individual differences in LSU. The clades indicated by bars in Figure 4.6 correspond to consistent base changes in the two loop regions that are seen in more than one sequence and are therefore more believable as new genotypes.

Neither primer pair proved useful for separating the *Pseudo-nitzschia* taxa from the Gulf of Naples using SSCP, for possible reasons that will be discussed below.

**SSCP**

The SSCP method was chosen for this study on the basis that it is a more reliable method than DGGE because there is no need for specialist equipment or GC clamps on the primers and the bands should be easier to interpret given the digestion of the anti-sense strand (Medlin et al. 2002). The LSU marker was chosen for the high number of available sequences enabling the accurate design of genus specific primers. SSCP is sensitive to both the length and the G/C content of the fragment. The shorter the fragment the more sensitive the method is to low numbers of base changes between taxa. The higher the G/C content the more easily the fragments are separated. The D1-186F/D1-258R fragment obtained in this study has a length of around 90 bp and a G/C content of around 46 % with all the known sequences of *Pseudo-nitzschia* having around the same G/C content. Although the length should be ideal, the G/C content is on the low side and may be the reason why the different taxa could not be clearly
separated. This low G/C content, in combination with the low diversity within the fragment, makes separation difficult. An alternative marker with a higher level of polymorphism, and preferably a higher G/C content, is needed to develop SSCP as a useful tool for tracking *Pseudo-nitzschia* taxa during the year. Possible candidates could be either within the ITS or *rbcL* markers, for which the number of sequences from diatoms is growing rapidly. Both have a higher level of resolution, particularly in *Pseudo-nitzschia* (see Chapter 5, Amato et al. in prep and Orsini et al. 2004), and may prove to be more useful markers for developing genus specific primers giving fragments with enough polymorphism.

**Diversity and Seasonality in *Pseudo-nitzschia***

The clone libraries not only detected several of the newly described species within *Pseudo-nitzschia*, but also showed the existence of new taxa and the occurrence of the different taxa during the year.

The differences between the number of taxa seen in the LM counts and the number of genotypes seen in the clone libraries raise two issues. The first is that the clone libraries detect more taxa, best illustrated by the difference seen on the 26th May where the counts detected one taxon and the libraries four. The libraries detect the pseudo-cryptic species such as *P. delicatissima* LSU-clade 2, as seen in the April library where the counts only detect *P. delicatissima* and the libraries split this taxon onto two, and ‘*P. pseuodelicatissima*’ in the October counts is split into the pseudo-cryptic species *P. calliantha* and *P. calliantha* LSU-clade 2 in the library. The other point is that the libraries detect taxa that are rare. When the counts contain a single species, the other taxa must be below 6000 cells per litre yet the PCR is able to amplify them. Several factors may contribute to one template being preferentially amplified over another. G/C content is one factor (Suzuki and Giovannoni 1996), but as all the *Pseudo-nitzschia*
fragments have very similar G/C content, this seems unlikely to cause the increased amplification of rare templates. Sukuzi & Giovannoni (1996) observed that, when using mixed templates, genes in the final product tended towards a 1:1 ratio regardless of the initial proportions of the different templates. This was explained by the fact that the higher the concentration of the template, the faster the reannealing reaction between two homologous single stranded molecules. If the template reanneals, it is not free to react with the primer, therefore the annealing step of the PCR becomes inhibited. This inhibition occurs more quickly for the more concentrated template, leaving the rarer templates free to be amplified. This seems likely in the case of the Pseudo-nitzschia libraries as one particular template, for instance the *P. galaxiae* template in the May library, is generally several orders of magnitude greater in abundance than the others as can be seen from the LM counts in Figure 4.2. This, however, does not explain the lack of amplification of *P. delicatissima* in the June, July and August libraries, when this taxon formed a significant proportion of the counts on those dates. It is possible that '*P. delicatissima*' seen in the counts was in fact one of the pseudo-cryptic taxa, for instance the new genotype that seems to be the sister group to *P. delicatissima*.

It is interesting that *P. delicatissima* 2 and *P. delicatissima* were found in the same library, and a similar pattern was found for *P. calliantha* and *P. calliantha* 2. *P. delicatissima* 2 is the same morphospecies as *P. delicatissima*, in that in LM it is not possible to distinguish them and the same goes for *P. calliantha* and *P. calliantha* 2. In the case of *P. delicatissima* 2 and *P. calliantha* 2, small morphological differences have been found after the genetic differences prompted a closer examination of the frustule (Amato et al. in prep). These differences are however difficult to observe in LM. The difference between *P. delicatissima* and *P. delicatissima* 2 is the width of the cell, but this is not clear in LM as the difference is small and the cells are difficult to measure as
the width will change depending on the orientation of the cell. *P. calliantha* and *P. calliantha* 2 share the following characters: the lanceolate valve shape, uniseriate striae, poroid density, and poroids that are split into several sectors. In order to distinguish the two, one must measure the percentage of poroids with a central sector, which is lower in *P. calliantha* 2 and examine the cell width (1.7-2.6 vs 1.4-1.8 μm) and ranges of fibula and stria density, being both wider in *P. calliantha* 2 than in *P. calliantha* (Amato et al. in prep).

The finding of these pairs of species in the same library indicates that they have similar ecological requirements making them not only pseudo-cryptic, but also allochronous species. As such they are interesting candidates for the study of pseudo-cryptic sympatric species at sea and how these species may arise. Cryptic species have been defined as having identical morphology, but being reproductively isolated (Mayr 2001). The assessment of species boundaries using genetic markers is primarily based on the concept that coexisting sexual organisms will share a genetic pool, which generally is not available to organisms of other species. Consequently, there is expected to be far fewer genetic differences within a species than between individuals from different species (Sáez et al. 2003). Therefore genetic differences between individuals have been understood to indicate the presence of cryptic species. These results raise a number of questions about the nature of pseudo-cryptic species. Truly cryptic species, those with no differences in morphology, have been judged to reflect neutral mutations in some genes that do not necessarily reflect functional diversity (Fenchel 2005). Fenchel (2005) concludes that the value of the species concept is that it provides information on the organism in terms of phenotypic properties. In the case of the genotypes within *Pseudo-nitzschia* the question is whether these differences reflect real functional differences or that the slight genetic and morphological differences are simply due to variations within
the species. The assumption of the previous work on the data from MC Long-Term Ecological Programme is that congeneric species have different ecologies reflected in their different timings during the year, the different conditions during the year being ideal for different species (Ribera d'Alcalà et al. 2004; Zingone et al. 2002; Zingone et al. 2003). At first glance the results from the clone libraries seem to support Fenchel's ideas, in that the different genotypes within *P. galaxiae*, *P. delicatissima* and *P. calliantha* are all found within the same library indicating a lack of functional meaning to the different genotypes, as they seem to have the same ecological requirements enabling them to occur at the same time. However, the clone libraries do not sample the entire year, nor are they closely enough spaced to conclusively say that these taxa are co-occurring. For example, the finding of *P. delicatissima* and *P. delicatissima* 2 in the same library (April) could reflect the end of the bloom of one genotype and the start of the bloom of the other. A finer sampling strategy perhaps using probes for the various genotypes could reveal a difference in occurrence that suggests a difference in ecological requirements. Physiological studies on cultures in the laboratory may also reveal functional differences in terms of the physiological limits of each taxon.

The situation seen within *P. galaxiae* is even more complicated. This species forms several blooms during the year, with those in March being separated in the LM counts from the May and August blooms on account of the cells being much smaller. The pattern of genotypes within LSU does not present a clear picture as to whether the distinction on the basis of size has a genetic meaning. Five genotypes were recovered in the libraries, but their distribution does not match the blooms. Two of these genotypes are known from the original description and strains have been isolated from the Gulf of Naples that match one or the other. Strains matching the AY081137 sequence from the Sydney 4 strain have all been of the small size category. Three further genotypes were
detected, one of which matched a small number of strains of medium sized cells isolated from the Gulf of Naples during the course of the study detailed in Chapter Five. The seasonal distribution of these genotypes does not show whether any of the blooms are formed by different genotypes as the most repeated genotype, *P. galaxiae* LSU-clade 2, is found in all libraries except April. The *P. galaxiae* morphospecies is the subject of the study in Chapter 5 where this diversity will be explored in more detail.

*P. fraudulenta* and *P. subfraudulenta* are not distinguishable in LM and there has always been some doubt as to which species was present in the Gulf of Naples. Only one sequence was found from this group, belonging to *P. fraudulenta*. This does not rule out the presence of *P. subfraudulenta*, but does work towards clarifying exactly which species are present. *P. multistriata* sequences are present in the August and October libraries. This is consistent with the counts that show a regular bloom of this species between late August and October.

The results from this chapter, as with those from Chapter Two, highlight the hidden diversity within morphologically homogeneous taxa and the need for more rigorous assessments of those taxa at sea. The ecology and functioning of an ecosystem cannot be understood unless the elements contributing to that ecosystems are properly recognised.

**CHAPTER SUMMARY**

- Two pairs of genus level primers specific to *Pseudo-nitzschia* were developed giving fragments of 350 bp and 90 bp.
- Using either fragment, it was not possible to separate the various taxa within *Pseudo-nitzschia* using SSCP.
Clone libraries constructed from environmental DNA gathered on six distinct sampling dates using the longer 350 bp fragment revealed thirteen genotypes, 3 of which had not been previously found. There is a better resolution using PCR in that more taxa are found, but the abundance for these taxa in the original sample cannot be determined due to the dynamics of the PCR.

The environmental DNA data give some indications on the different seasonality of distinct genotypes, including an indication of the temporal co-occurrence of pairs of pseudo-cryptic species within *Pseudo-nitzschia* such as *P. delicatissima/P. delicatissima* and *P. calliantha/P. calliantha*. 
Chapter Five

Diversity within the diatom Pseudo-nitzschia galaxiae

INTRODUCTION

Mayr (2001) defined cryptic species as ‘sympatric natural populations that did not interbreed even though they showed no evident taxonomic differences’. Cryptic diversity has been identified in many groups from crickets to *Paramecium* to foraminiferans and the advent of molecular techniques has revealed the magnitude of the phenomenon. One of the first molecular examinations of diversity within a morphological phytoplankton species was the work of Gallagher in the early 1980s. Gallagher identified a winter and summer population in the diatom *Skeletonema costatum* using allozyme banding patterns from five enzyme loci (Gallagher 1980). Physiological differences between these two populations were also observed, with the summer population having a higher growth rate, a lower amount of chlorophyll per cell and a higher carbon assimilation rate (Gallagher 1982). As the clones used are no longer available it is not possible to qualify these results in the light of the recent reorganisation of *Skeletonema costatum* into several new species (Sarno et al. in press; Sarno et al. 2005; Zingone et al. 2005). However it is likely that the winter population is *Skeletonema japonicum* Zingone et Sarno and the summer *Skeletonema grethae* Zingone et Sarno (M. Saggiomo, pers. comm.). In the case of the *Skeletonema costatum* complex, some of the new species are pseudo-cryptic and other members are cryptic as far as current methods of detecting morphological differences go.

This situation is repeated for a wide range of marine microalgal species and complexes examined, of which the following are a few examples. A study of various nuclear, mitochondrial and plastid markers (Ślapeta et al. 2006) for 17 strains of *Micromonas*
*pusilla* (Prasinophyceae) collected from throughout the world’s oceans revealed several lineages for which there are presently no morphological markers. The heterotrophic dinoflagellate *Oxyrrhis marina* Dujardin (Lowe et al. 2005) has been shown to contain several different ITS types and two different physiological responses to salinity that matched the habitat from which the strains were collected rather than their ITS type. In the dinoflagellate *Scrippsiella trochoidea* (Stein) Loeblich (Montresor et al. 2003), analyses of the ITS show genetic distances between various strains that were comparable to those between other morphospecies within *Scrippsiella*, supporting the hypothesis that there were cryptic species with similar morphological characteristics. As mentioned in Chapter Four, several *Pseudo-nitzschia* morphospecies have been shown to contain more than one genetic entity, which has led to a more careful examination of the morphology, and revealed subtle differences between them (Lundholm et al. 2002b; Lundholm et al. 2003; Lundholm et al. 2006; Orsini et al. 2004; Skov et al. 1997). Sáez et al. (2003) elevated the subspecies within the coccolithophorids *Calcidiscus leptoporus* (Murray et Blackman) Loeblich et Tappan and *Coccolithus pelagicus* (Wallich) Schiller to species level based on genetic data from ITS and tufA.

Assessments of temporal diversity within a phytoplanktonic species are few, being limited to those of Gallagher in the 1980s and more recent studies using RAPDs (Shankle et al. 2004) and microsatellites (Rynearson et al. 2006). The work by Shankle et al. (2004) found only one significant difference among blooms of the dinoflagellate *Prorocentrum micans* sampled from a single station in La Jolla, California. However, allozymes and RAPDs both have problems with underestimating genetic differences and RAPDs have particular problems with reproducibility as they rely on PCR kinetics (Medlin et al. 2000). As discussed in Chapter One, the two blooms of *Ditylum brightwellii* (Bacillariophyceae) within Puget Sound (Rynearson et al. 2006) are
different at the level of ITS and microsatellites. Microsatellites can be very sensitive markers and are generally used at the level of populations. It is therefore necessary to determine first whether the taxon to be examined is comprised of coarser genetic groups, for example using more slowly evolving markers.

Figure 5.1: Representatives of the three *P. galaxiae* size classes, A small, B medium and C large, in all cases the scale bar equals 5 μm (modified from Cerino et al. 2005).

*Pseudo-nitzschia galaxiae* (Lundholm and Moestrup 2002), a potential toxic pennate planktonic diatom, shows a pattern of multiple blooms during the year, to which genetic analysis can be applied to investigate temporal diversity at sea. *P. galaxiae* has been observed in the Gulf of Naples since the beginning of the phytoplankton sampling programme in 1983 where it has been separated into three size classes that occur at different times of the year. In early samples, *P. galaxiae* forms were recognised before its formal description. The small *P. galaxiae* form was identified as cf. *Phaeodactylum tricornutum* Bohlin, the doubt being that it had two chloroplasts instead of one. This morphotype occurs in March and April and has an approximate size range of 6-20 μm. The other two size classes were identified as *Pseudo-nitzschia* cf. *prolongatoides* (Hasle) Hasle due to their thin shape with a central swelling (Cerino et al. 2005). These classes occur in May (c. 20-50 μm) and August (c. 50-80 μm) (Figure 5.1). These size classes overlap considerably, for example it is possible to find cells larger than 20 μm in
March.

*P. galaxiae* was described in 2002 when it was found as a contaminant of a diatom culture from Mexico and was also directly isolated from a water sample from Australia (Lundholm and Moestrup 2002). Gulf of Naples strains from May and July 2001 were examined in EM and found to have the same ultrastructure as the original description, but *P. galaxiae* in the Gulf of Naples has a much wider size range (Cerino et al. 2005). It was concluded that the difference in occurrence related to either different size cohorts arising from different sexual events or that *P. galaxiae* was several cryptic species. A low level of polymorphism was detected in the LSU fragment, with the Gulf of Naples May strains being identical to the AY081136 sequence from Mexico and the second sequence from the original description, AY081137 from Sydney, having a single base difference.

The results from Chapter Four do not clarify whether the three blooms are genetically different as one genotype *P. galaxiae* LSU-clade2 is present from May until October. Chapter Four also showed there is a lack of a suitable marker for bulk tracking of genotypes. Therefore it was necessary to return to single cell isolations for want of a better method to investigate the differences between the blooms. The aim of this chapter was to investigate the genetic composition of each bloom over three years in order to determine whether the pattern of occurrence of *P. galaxiae* in the Gulf of Naples has a genetic basis or is simply a consequence of the diatom life cycle. LSU was sequenced for each culture to determine whether any of the rarer genotypes recovered in the clone libraries from Chapter Four made up any of the blooms in the Gulf of Naples. ITS was also used because it is a more quickly evolving region in the rDNA between the SSU, 5.8S and LSU subunits of the ribosome and as such is more suited to studies at lower
taxonomic levels (Medlin and Simon 1998). It is transcribed but is not present in the final ribosome and therefore able to evolve more quickly (Suh et al. 1993). Both regions have the advantage of being present in multiple copies in each cell making amplification easier. Intra-strain variability of ITS was also investigated in order to test the usefulness of ITS as a marker in *Pseudo-nitzschia*. As ITS is present in multiple copies in the cell, unless it is under the constraints of concerted evolution (Arnheim 1983), the various copies can be different, as observed for example within individuals of *Caulerpa racemosa* (Forsskal) J. Agardh (Chlorophyta) (Famà et al. 2000). This intra-individual variation could give a less reliable indication of differences between genotypes.

**MATERIALS AND METHODS**

**Isolation of strains**

DNA was available from five strains isolated by Luisa Orsini during 2001 from the MC station. Within the culture collection at the SZN were three strains collected in 2003 from MC. During August 2003, Alberto Amato isolated ten strains from net samples collected from MC. A further 77 strains of *P. galaxiae* were isolated from seawater collected at MC between August 2003 and August 2005. During the March and May blooms, it was necessary to first grow SDCs (Throndsen 1995) in order to have enough cells to isolate as the net samples collected every week from MC use a 40μm net through which most *P. galaxiae* cells passed. It was possible to isolate the larger cells found during August from the net sample in addition to isolating from SDCs. The strains were isolated using a glass micropipette into 24 well plates with F_{10} medium and grown for one week at 18°C 14:10 L:D at a light intensity of approximately 60 μE m^{-2} s^{-1}. Wells containing *P. galaxiae* were transferred into Petri dishes containing F_{2} medium.
and grown under the same conditions for approximately ten days (see Appendix A for collection details of all strains used).

**Extraction and amplification of DNA**

Extraction of genomic DNA was carried out using the CTAB method outlined in Chapter 2. DNA from 98 strains was amplified using primers targeting LSU and ITS, outlined in Table 5.1. LSU was amplified in a mix containing 1 μl genomic DNA (approximately 40 ng), 200 μM deoxynucleoside triphosphates, 1 μM each primer and 2.5 U Taq polymerase in 1× enzyme buffer containing 1.5 mM MgCl₂ (Roche Diagnostics GmbH, Mannheim, Germany). The PCR conditions for LSU were 94 °C for 4 min followed by 35 cycles of 94 °C for 1 min, annealing at 55 °C for 1 min 30 sec and elongation at 72 °C for 1 min, followed by further elongation at 72 °C for 5 min.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequences 5'-3'</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3Ca</td>
<td>ACG AAC GAT TTG CAC GTC AG</td>
<td>Scholin et al. (1994a)</td>
</tr>
<tr>
<td>D1R</td>
<td>ACC CGC TGA ATT TAA GCA TA</td>
<td>Lenaers et al. (1989)</td>
</tr>
<tr>
<td>ITS1</td>
<td>TCC GTA GGT GAA CCT GCG G</td>
<td>White et al. (1990)</td>
</tr>
<tr>
<td>ITS4</td>
<td>TCC TCC GCT TAT TGA TAT GC</td>
<td>White et al. (1990)</td>
</tr>
</tbody>
</table>

ITS used a mix containing 1 μl genomic DNA (approximately 40 ng), 200 μM deoxynucleoside triphosphates, 2mM MgCl₂, 1 μM each primer and 2.5 U Taq polymerase in 1× enzyme buffer -MgCl₂ (Roche Diagnostics GmbH, Mannheim, Germany). The conditions for ITS were 94 °C for 2 min followed by 35 cycles at 94 °C for 30 sec, annealing at 46 °C for 20 sec and 72 °C for 1 min with an extension at 72 °C for 5 min. Sequence reactions were obtained with the BigDye Terminator Cycle Sequencing technology (Applied Biosystems, Foster City CA) and purified in automation using a robotic station "Biomek FX" (Beckman Coulter, Fullerton CA).
Products were analyzed on an Automated Capillary Electrophoresis Sequencer "3730 DNA Analyzer" (Applied Biosystems). The sequences of the two strains used in the original description of *P. galaxiae* (Lundholm and Moestrup 2002) and 23 LSU sequences and 28 ITS sequences from GenBank were also included in the analysis. The LSU sequence from *P. decipiens* from Chapter Four was also included.

**Cloning**

Cloning of three strains from different years was carried out using pCR®2.1-TOPO® TA Cloning® kit (Invitrogen™ Life Technologies, Carlsbad, California) according to the manufacturers instructions. About 20 colonies per strain were picked, grown in LB with ampicillin added and purified using a QIAprep® Spin Miniprep kit (QIAGEN S.p.A., Milan) following the manufacturers instructions. Sequence reactions were obtained as above using the vector primers T7 and M13rev (Invitrogen™ Life Technologies, Carlsbad, California). Contigs from the forward and reverse sequences were formed and checked using Seqman II 3.61 and aligned using ClusterW in BioEdit 7.0.1. Intra-strain differences were observed in BioEdit.

**Sequence Analysis**

The contigs for the LSU and ITS sequences obtained were analysed as above using the same software. The unique LSU sequences for each sampling year were aligned with representatives of the closest relatives of *P. galaxiae* and a tree constructed using PAUP* and Neighbour Joining with Maximum Likelihood settings calculated from the modelblockPAUPb10 from Modeltest 3.7. A bootstrap analysis of 1000 replicates was also carried out for the NJ tree.

The ITS sequences were analysed in two ways. Two alignments were constructed, one including all positions and another including only those regions that could be aligned with reasonable confidence corresponding to the 5.8S and the stem regions in the two
spacers. Modeltest was preformed on both alignments using Modeltest 3.7. Trees were constructed for both alignments using Maximum Likelihood (only unique sequences were included to reduce computational time) and Neighbour Joining using the ML settings calculated by Modeltest (using unique sequences from each sampling date). Bootstrap analyses were performed for the two different tree construction methods (100 replicates for ML and 1000 replicates for NJ) to assess the strength of the clades found.

**RESULTS**

**Cloning**

25 colonies were sequenced for the *P. galaxiae* Gulf of Naples strain B157 collected on 24<sup>th</sup> January 2003. Eighteen of the sequences were identical and the remaining 7 showed single point mutations, each in a different position. Similar picture emerged from the colonies from the B58 strain isolated on 18<sup>th</sup> July 2001 and the SM1 strain isolated on 5<sup>th</sup> April 2004. The alignments of these sequences are found on the accompanying CD.

**Sequence Analysis**

The LSU sequences show a low level of diversity in *P. galaxiae* with differences at only two positions in the entire hyper-variable region (Figure 5.1). Based on these differences, three groups within the Gulf of Naples could be distinguished: the March sequences, the May and August sequences and a third group made up of strains collected between January and May (i.e., strains AL31, SM10, SM54 and SM77). Of the sequences from the original material used by Lundholm & Moestrup (2002), the AY081137 Sydney 4 sequence groups with the March Gulf of Naples sequences and the AY081136 Mex 23 sequence groups with the May and August sequences. All *P. galaxiae* sequences form a monophyletic group within *Pseudo-nitzschia* in this analysis.
Figure 5.2: NJ tree constructed using ML settings calculated using Modeltest (best fit model TVM+I+G
A = 0.2413, C = 0.1929, G = 0.3004 and T = 0.2654, A ≻ G = 2.8495, C ≻ T = 2.8495 A ≻ T = 1.4245
A ≻ C = 0.3299 and C ≻ G = 0.4426, the positional rate variation across the alignment γ = 0.763 and the
proportion of invariable sites I = 0.7997) for LSU sequences from
P. galaxiae, representing the unique
sequences from each sampling date, and other representatives from Pseudo-nitzschia
downloaded from GenBank or sequenced for this study using material kindly provided by Nina Lundholm (marked with an
asterisk). The numbers in brackets refer to the number of strains matching that genotype on that date. Bootstrap values are from 1000 replicates.
Using the total ITS region including 5.8S, seven different genotypes were identified within *P. galaxiae*. Table 5.2 shows pairwise base pair differences in a 662bp ITS alignment between strains corresponding to the seven ITS clades found. The level of difference between the different clades ranges from 4 bp to 44 bp, which corresponds to an uncorrected distance of between 0.605 and 6.693 % (see Table 5.3).

**Table 5.2:** Base pair differences in the 662 bp ITS alignment between 7 unique strains calculated using the pairwise base pair differences feature in PAUP*.

<table>
<thead>
<tr>
<th></th>
<th>SM65 March</th>
<th>SM52 May</th>
<th>AY257850 Mex23</th>
<th>Sydney4</th>
<th>SM55</th>
<th>SM99 August</th>
</tr>
</thead>
<tbody>
<tr>
<td>SM65</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SM52</td>
<td>7</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AY257850</td>
<td>12</td>
<td>11</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sydney4</td>
<td>31</td>
<td>32</td>
<td>32</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SM55</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>6</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SM99</td>
<td>42</td>
<td>40</td>
<td>44</td>
<td>42</td>
<td>38</td>
<td>-</td>
</tr>
<tr>
<td>SM54</td>
<td>41</td>
<td>39</td>
<td>43</td>
<td>39</td>
<td>35</td>
<td>4</td>
</tr>
</tbody>
</table>

**Table 5.3:** Uncorrected ('p') distances, expressed as a percentage, between 7 strains representing the seven ITS clades within *P. galaxiae* calculated using PAUP*.

<table>
<thead>
<tr>
<th></th>
<th>SM65 March</th>
<th>SM52 May</th>
<th>AY257850 Mex23</th>
<th>Sydney4</th>
<th>SM55</th>
<th>SM99 August</th>
</tr>
</thead>
<tbody>
<tr>
<td>SM65</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SM52</td>
<td>1.053</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AY257850</td>
<td>1.807</td>
<td>1.657</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sydney4</td>
<td>4.724</td>
<td>4.884</td>
<td>4.854</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SM55</td>
<td>3.813</td>
<td>3.822</td>
<td>3.805</td>
<td>0.913</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SM99</td>
<td>6.396</td>
<td>6.092</td>
<td>6.693</td>
<td>6.439</td>
<td>5.843</td>
<td>-</td>
</tr>
<tr>
<td>SM54</td>
<td>6.242</td>
<td>5.935</td>
<td>6.543</td>
<td>5.976</td>
<td>5.375</td>
<td>0.605</td>
</tr>
</tbody>
</table>

The NJ tree from the alignment excluding the loop regions for the unique *P. galaxiae* ITS sequences from each sampling date is displayed in Figure 2 with bootstrap values. *P. galaxiae* was monophyletic in all topologies, except the NJ tree from the alignment that included the loop regions. In this topology, *P. galaxiae* is not monophyletic as it includes the two *P. decipiens* sequences between Clades VI and VII. However, in the ML analyses the inclusion or exclusion of these regions had no effect on topology.
Figure 5.3: NJ tree constructed using ML settings calculated using Modeltest (best fit model GTR+1+G
A = 0.2246, C = 0.2277, G = 0.2297 and T = 0.318, A G = 2.1193, C T = 2.9092 A = 1.5141
A C = 0.8258 and C G = 0.7733, the positional rate variation across the alignment γ = 1.193 and the
proportion of invariant sites I = 0.3002) for ITS sequences from P. galaxiae, representing the unique
sequences from each sampling date, and other representatives of Pseudo-nitzschia downloaded from
GenBank. The numbers in brackets refer to the number of strains matching that genotype on that date.
Bootstrap values are from 1000 replicates.
Figure 5.2 shows seven clades well supported by bootstrap, three of which consist of sequences from the three annual blooms (March, May and August), indicating that each bloom is characterised by a different ribotype that recurs year after year. There are two rarer genotypes that do not seem to dominate during the bloom periods. These include SM55 that groups with the Sydney strain and AL31, SM10, SM54 and SM77 that, forming a separate clade, group with the August clade. The March and May clades group closely with the Mexico strain. The alignments for both markers and chromatogram files for all strains are available on the accompanying CD.

**DISCUSSION**

**Different markers, different sensitivity**

It is necessary to assess the differences in resolution between the two markers used. In general three markers have been used in molecular investigations in diatoms. Diatom molecular phylogenies have been constructed using 18S (Kooistra et al. 2003) for class level analyses of deep branches between the major groups in the centric and pennate diatoms. In fact within *Pseudo-nitzschia* this region does not resolve the relationships between the various species as it is not variable enough. LSU rDNA has generally been used within genera to resolve relationships between species, for example within *Skeletonema* (Kooistra et al. in press) and *Pseudo-nitzschia* (Lundholm et al. 2002a). Probes for various *Pseudo-nitzschia* species have also been developed for the D1-D3 LSU region (Miller and Scholin 1996; Miller and Scholin 1998; Scholin et al. 1996; Scholin et al. 1999; Scholin et al. 1994b), which proved successful in identifying some harmful species such as *P. australis*, *P. pungens* and *P. multiseries* in natural samples from California. However it has been difficult to apply these probes to strains from other regions (Orsini et al. 2002), possibly due to an undiscovered diversity in LSU within *Pseudo-nitzschia* morphospecies. ITS has been used for resolving relationships
between closely related species and within morphospecies such as *P. delicatissima* (Lundholm et al. 2006) and the *P. pseudodelicatissima/P. cuspidata* complex (Lundholm et al. 2003). This marker has also been used to investigate bloom composition in the Gulf of Naples (Orsini et al. 2004), showing that a single bloom was dominated by one ribotype.

The large subunit in *Pseudo-nitzschia galaxiae* distinguishes three groups, but this is on the basis of very little variation and two of the blooms have identical LSU types. The internal transcribed spacer reveals seven generally well-supported clades with the three blooms being comprised of separate ITS types. A subsequent study investigating chloroplast inheritance in *Pseudo-nitzschia* using the *rbcL* chloroplast marker included a limited number of the strains used in this chapter and found that they group into seven clades matching those found in ITS (J. H. Levialdi Ghiron, pers. comm.). The differences found between the resolution of LSU and ITS reflect already well known differences in the rate of evolution of these two markers and highlight the importance of choosing the right markers to answer the question (Medlin et al. 2002; Medlin and Simon 1998). It is also important to use more than one marker. For example, SSU resolves the relationship between the morphospecies in *Skeletonema* but not those within *Pseudo-nitzschia*. The fact that one marker is useful in one group does not mean it will give the same resolution in another. LSU has been used to investigate phylogenetic relationships within *Pseudo-nitzschia* (Lundholm et al. 2002a; Orsini et al. 2002) where it gives well-supported branches between species groups (Orsini et al. 2002). However, the results from the present study indicate that further diversity may have been overlooked by the lower resolution of the LSU marker. ITS has been used in various algal taxa to investigate intra-specific diversity (Lundholm et al. 2003; Lundholm et al. 2006; Montresor et al. 2003; Rynearson et al. 2006). Within *Pseudo-
nitzschia, ITS sequences revealed differences within *P. delicatissima* and *P. pseudodelicatissima* that stimulated a closer morphological investigation using EM, and several new species were described (Lundholm et al. 2003; Lundholm et al. 2006). The level of difference between the clades within *P. galaxiae* (0.605-6.693 % uncorrected pairwise distance) is low compared to the differences between for example *P. pseudodelicatissima* and *P. calliantha* that were originally described as the same species (25%). The difference between *P. delicatissima* strain Laesoe5 and *P. micropora*, which would appear to be its closest relative, is 5.8%, but the distance between *P. delicatissima* and *P. dolorosa*, again described as the same morphospecies, is 17.8%.

The rare genotypes in both LSU, discovered in the clone libraries, and ITS, found during this study, do form blooms in the Gulf of Naples. It is likely that these types represent part of the total diversity within *P. galaxiae*, but not within the Gulf of Naples. These types are evidence of rare taxa at the edge of their physiological or ecological tolerance that are not able for whatever reason to dominate any of the blooms in the Gulf. These types contribute to the flexibility of the Gulf ecosystem giving the potential for the system to adapt to changes in conditions. It may be that these types will be able to form blooms in the Gulf at some point in the future.

The changes in topology created by including the loop regions in the ITS NJ analysis (data not presented) were due in part to the limitations of the NJ analysis and in part to a large insertion of around 60 bp in *P. decipiens* at the beginning of ITS1. This insertion is not present in full in any other *Pseudo-nitzschia* species, but has some corresponding shorter inserts in the more distant relatives of *P. galaxiae*. In a distance NJ analysis, this pushes *P. decipiens* and these more distant relatives apart, artificially forcing *P. decipiens* between the *P. galaxiae* clades. ML analyses including and excluding the
loop regions of both spacers have the same topology because this analysis does not rely on distances. ML assess the probability that a given substitution model and tree will give rise to the observed data (Hillis et al. 1996). This method is frequently the least affected by sampling error and is generally more robust (Hillis et al. 1996). For these reasons, the inserts do not have an effect on the topology. With the present level of taxon sampling, the ribotypes within *P. galaxiae* form a monophyletic group.

As ITS has the problem of not always being homogeneous in individuals due to a lack of concerted evolution (Arnheim 1983), it was necessary to investigate intra-strain diversity in order to assess the usefulness of ITS as a marker. Concerted evolution is the evolution of two identical genes together so that they are both maintained in the original functional form and do not markedly diverge (Lawrence 1989). This process allows homogenisation between multiple copies within the genome of a single individual through DNA recombination mechanisms, such as gene conversion and unequal crossing over (Ruggiero and Procaccini 2004). If there were differences in ITS within strains of *P. galaxiae*, it would not be possible to use this marker to trace intraspecific diversity, as it would be impossible to distinguish the different kinds of diversity. Comparing the sequences obtained it could be seen that at least in *P. galaxiae*, the ITS copies in a unialgal culture are identical making this marker useful for assessments of intraspecific diversity.

The questions arising from the small but consistent differences found in LSU and ITS are: how important are these differences and do the various clades constitute new species?
Defining taxa in microalgae

The seven ITS clades within *P. galaxiae* could be described as new species based solely on genetic differences. This would fit within the phylogenetic species concept (Cracraft 1989) as the various clades form recognisable monophyletic groups. This distinction would be supported by ecological data showing that certainly three of the seven have different requirements as they bloom under different environmental conditions. There is also the morphological character of size. This is generally a weak character in diatoms and in the case of *P. galaxiae* does not involve discrete categories. When coupled with the other genetic and ecological data, it does become more significant, but still fails to distinguish five of the clades (clades I, III, IV, V and VII) as they all fall into the medium size class.

The genetic differences seen are slight when compared to the differences seen in other complexes within *Pseudo-nitzschia* (see section above). An investigation of sexual reproduction and genetic diversity in strains of *Pseudo-nitzschia* from the Gulf of Naples revealed a relationship between differences in secondary structure of ITS2 and the ability to produce viable offspring (Amato et al. submitted). Strains isolated from the Gulf of Naples identified as *P. pseudodelicatissima* and *P. cuspidata* have identical LSU sequences and six single sided compensatory base changes in the stem regions of ITS2, the loop regions are considerably more variable. Cross experiments between these two clades failed to produce gametes. Within Gulf of Naples *P. calliantha* strains, two ITS clades were found with three single sided compensatory base changes in ITS2, but again with very variable loop regions. Reproduction between strains belonging to these two clades proceeded as far as producing auxospores, which did not mature and initial cells were not formed. *P. multistriata* has two ITS genotypes, differing by two short inserts and two point mutations distributed between ITS1 and ITS2. These changes do
not seem to alter the secondary structure, as the inserts are elongations of loops and the point mutations also occur in loops. Successful sexual reproduction has been observed between these two clades (D. D'Alelio, pers. comm.). Similar data are required for the genotypes within *P. galaxiae*, which demonstrated far lower overall diversity in ITS than the strains tested both by Amato *et al.* and by D'Alelio. For example, a preliminary examination of the secondary structures of sequences from Clades VI and VII shows that three of the four different base pairs in the entire ITS sequence are located in ITS2, but all in loops rather than as single sided compensatory base changes. It is probable that sexual reproduction is possible between these two clades (A. Amato, pers. comm.). Hybrids between the seven ITS *P. galaxiae* types were not found during this investigation, but this could be due to only examining the genotype of 100 unialgal cultures.

Given that *P. galaxiae* is a potentially toxic species, a more comprehensive study of the toxicity of each clade would be useful in light of their different occurrences and may shed light on whether the clades should be described as different species. Two strains (SZN-B54 and SZN-B56, both belonging to Clade III May) out of seven tested were found to contain extremely low toxin concentrations (Cerino *et al.* 2005). The concentration of domoic acid was $3.6 \times 10^{-4}$ and $7.8 \times 10^{-7}$ pg per cell in SZN-B54 and SZN-B56, respectively (Cerino *et al.* 2005). This is lower than that found for *P. multistriata*, which was already one of the lowest levels reported. The seven strains tested belong to two clades, Clade III May and Clade VI August. Strains from Clade II March have not been tested for domoic acid concentration. This information would be especially important for the monitoring of harmful algal species given the occurrence pattern of *P. galaxiae* during the year.
Defining species is not esoteric as it has relevance for many fields particularly in the case of a medical important genus such as *Pseudo-nitzschia* (Mann 1999). In order to be of relevance to the many disciplines that use taxonomy, a species description must use many approaches and characters. These characters should also mean that the species is easily recognisable. In the case of *P. galaxiae* defining seven species based solely on ITS sequences would not be practical, as it would require the effort of isolation and cultivation of strains for genetic analysis before identifications could be made. Neither is the phylogenetic species concept a sound concept, rather it is a way of delimiting taxa and so is a variation of the typological species concept (Mayr 2001). As such the delimitation of species under this concept changes depending on which marker is used to construct the phylogeny and several doubts have been raised on the wisdom of using sequence divergence to delimit taxa (Fergurson 2002). Without further investigations, there is also no guarantee that ITS gives the complete picture of diversity in *P. galaxiae*. The results on diversity within *P. galaxiae* will need to be confirmed by mating data and a deeper morphological examination of the strains in order to determine whether this diversity should be defined at the species level. A separation of the seven clades into new species would be premature, but this does not mean that the pattern of blooms, slight genetic differences and slight morphological variations are not worth discussing as they do raise interesting questions as to the nature of the occurrence of diatom blooms.

**The genetic composition of *P. galaxiae* blooms - Different bloom, different genotype**

As in the case of *Pseudo-nitzschia delicatissima* (Orsini et al. 2004) and *Prorocentrum micans* (Shankle et al. 2004), the blooms of *Pseudo-nitzschia galaxiae* seem to be dominated by a single ribotype. This needs to be confirmed using a more
comprehensive method than the isolation of single strains as only a few strains were sampled for each bloom. A preliminary investigation using cultures, of the type undertaken during this investigation, are necessary as a basis. However the data need to be expanded preferably using methods such as probes that do not rely on isolating strains and on the vagaries of PCR. Given that a bloom may be millions of cells per litre, isolating only 100 is an obvious underestimate of the total diversity. However, in order to develop probes one must first have an idea of the diversity such as given by the present study.

Multiple blooms by the same species during the year are common and generally attract little attention (e.g. *Guinaria delicatula* in Helgoland Wiltshire and Dürselen 2004). The presence of multiple blooms is assumed to be evidence for the opportunistic nature and physiological plasticity of diatoms allowing them to bloom whenever environmental conditions are favourable. Physiological studies have demonstrated the plasticity of many diatom species, showing them to be capable of growth at a wide variety of temperatures and light intensities. Instead the blooms of *P. galaxiae* are made up of different ITS genotypes.

The implication of these results for the monitoring of phytoplankton species is that the morphospecies is not necessarily a good indicator of biodiversity. During monitoring programmes the typological species concept using morphology visible in light microscopy has been applied to assess diversity within natural samples, even to entities that bloom repeatedly during the year. The results from *P. galaxiae* add to the evidence that this may be underestimating the diversity present in cases where the genetic composition of the different blooms is not understood.
Different bloom, different size

The blooms of *P. galaxiae* are characterised by cells of different sizes progressing from small at the beginning of the year to large in August. The most obvious explanation for this pattern would be that they correspond to different cohorts resulting from sexual reproduction events. As diatoms reproduce asexually, the average cell size decreases. At first glance, the pattern of *P. galaxiae* could be explained by the small March cells being the result of a sexual reproduction event a year or perhaps two years previously. This was clearly seen in *Nitzschia sigmoidea* (Nitzsch) W. Smith (Mann 1988) where sexual reproduction events could be seen by the appearance of maximum sized cells. These populations could then be followed for several years as the average cell size decreased.

Given that the differences in sizes are accompanied by genetic differences the question then becomes why are different sizes found at different times of the year. The fact that the three size classes occur under different environmental conditions suggests that their ecological requirements may be different or that, in some way, different sized cells have different advantages under different conditions. March in the Gulf of Naples is generally characterised by a mixed layer of around 70 m, salinity of 37.85 psu and a temperature of 14 °C, May is at the beginning of stratification with a mixed layer of 10 m, salinity at yearly low of 37.5 psu and temperature of 18 °C and August has a mixed layer of 5 m, salinity of 37.9 psu and temperature of 25 °C (Ribera d'Alcalà et al. 2004).

The pattern of succession by different species in the plankton has been explained by a combination of turbulence and nutrients illustrated in the classic Margalef mandala (Margalef 1978). This predicts that large colonial diatoms will be favoured by high turbulence and high nutrients in spring whereas dinoflagellates will dominate the
summer conditions of low turbulence and low nutrients. According to the mandala, \textit{Pseudo-nitzschia} should all occur around April at medium levels of turbulence and nutrients, whereas in reality there are blooms of \textit{Pseudo-nitzschia} species at different times from February to October. The large \textit{P. galaxiae} size class (Clade VI) should be present in March, with the smallest cells being favoured by the low turbulence and low nutrients of August. The fact that the mandala cannot be applied to the Gulf of Naples has been previously noted (Zingone et al. 1990). Generally the Gulf of Naples is characterised by small diatoms in summer rather than dinoflagellates as predicted by the mandala. The hypothesised explanation for this is that as the Gulf of Naples never becomes nutrient depleted (Ribera d'Alcalà et al. 2004), the diatoms can continue to dominate during the year and there is little advantage to the dinoflagellates ability to remain in surfaces waters during the summer (Zingone et al. 1990; Zingone and Sarno 2001).

There could be other reasons as to why different sized cells could be favoured at different periods during the year. In Mann (1988) small \textit{Nitzschia sigmoidea} (Nitzsch) W. Smith cells were selected for during a year when there was an attack by the oomycete parasite \textit{Aphanomycopsis} Scherffel. This parasite attacks along the full length of the cell so larger cells are more vulnerable to infection. The effect of parasites on the \textit{P. galaxiae} populations at different periods in the year could explain the distribution of the size classes. Predation could also be a factor. Grazer numbers vary considerably during the year (Ribera d'Alcalà et al. 2004) and either different concentrations of copepods at different times of the year or size selection by copepod species could select for the different size classes.
Recurrence of genotypes in same period

The fact that the same ribotype occurs in the same period of the year in the three years sampled is further confirmation that the blooms are not size cohorts originating from different sexual reproduction events. The recurrence of species in particular periods of the year has been observed for a number of taxonomic groups during the course of the MareChiara sampling programme (Ribera d'Alcalà et al. 2004; Zingone and Sarno 2001). When the data from the time series are viewed at the community level, in order to try and assess whether a particular time of year is characterised by a particular species, there seems to be no pattern as no single species is always the dominant. However, if the situation is viewed species by species, each has a particular time of the year in which it is present in greatest abundance. All that changes from year to year is the amplitude of this abundance (Ribera d'Alcalà et al. 2004). This is seen across phylogenetic diverse groups for instance in cryptophyte species (Cerino and Zingone in press), many diatom species including all the *Pseudo-nitzschia* present in the Gulf (Zingone et al. 2002) and in the prasinophyte *Micromonas pusilla* (Zingone et al. 1999b).

This precise timing could be linked to two groups of factors, environmental cues (exogenous), an intrinsic clock (endogenous) or a combination of the two. In the Gulf of Naples, there does not seem to be any environmental factor that correlates exactly with the timing of blooms of species. Temperature and nutrients are more variable than the blooms and do not seem to be the triggers for the blooms.

It is possible that there is some intrinsic clock in phytoplankton that triggers the bloom, perhaps as a means of increasing numbers to ensure survival until the next year or as a means of facilitating sexual reproduction. An internal clock has advantages for small
organisms with short life spans as it means they do not rely on stochastic cues that may be easily missed (Lewis 1984). Lewis (1984) hypothesised that vegetative size reduction in diatoms could act as a clock between sexual reproduction events. It was thought that diatoms reproduced sexually very rarely, perhaps every 20 years and so the reduction in cell size gave a way of timing the last event without using some physiological mechanism that would be too costly for a unicellular organism. Diatoms have a specific size below which sex is possible, shown to be around 40% of the maximum size for pennates (Lewis 1983). However, the work on microsatellites indicates that sexual reproduction is much more important than previously thought and must be frequent in order to maintain the high diversity within populations observed by Evans et al. (2002) and Rynearson and Armbrust (2006). However, it is still possible that size reduction could play some part in timing of blooms if the time of the bloom were also the time of sexual reproduction. The size threshold below which sex is possible in *Pseudo-nitzschia* has been shown to be considerably wider than in other diatoms, around 70% of the maximum cell size (Amato et al. 2005). This would mean that they could reproduce after relatively few divisions, possibly every year rather than once every 20 years as hypothesised.

A third possibility would be a combination of exogenous and endogenous factors as seen in flowering plants. The only clearly consistent environmental factor in the Gulf of Naples is day-length. It has been suggested that photoperiod may be a factor in controlling the beginning of blooms by spore forming diatom species by initiating germination (Eilertsen et al. 1995). Putative cryptochromes, which are known to be responsible for photoperiod perception and flowering time in plants, have been found in the genome of *Thalassiosira pseudonana* Hasle et Heindal (C. Bowler and A. Falciatore, pers. com.). It is possible that other diatoms may have the ability to perceive
and respond to changes in photoperiod. Certainly the pattern of occurrence of phytoplankton species observed in the Gulf of Naples would not be noteworthy if they were species of plants. Consistencies in flowering time in plants are well documented, but comparisons between phytoplankton blooms and flowering of plants have always been discouraged as one is an increase in growth and the other is a sexual reproduction event. However, in terms of their precise timing, there are similarities that are worth studying in more depth. Most flowering plants have a specific flowering period controlled by endogenous perception of photoperiod using cryptochrome but modified by environmental factors such as temperature. The phenology of flowering plants can be anticipated or delayed by the particular conditions of a year. Phytoplankton show a similar pattern in that blooms are always within a particular period but the exact week varies slightly from year to year.

**How this pattern could arise**

The recurrent pattern of genetic and morphological differentiation between blooms of *P. galaxiae* raises interesting questions about how diversity is generated in the marine environment as well as providing the indication of a possible mechanism. Speciation is one of the least understood aspects of evolution (Schluter 2001) and there have been particular problems in understanding how speciation could occur at sea. It has even led to the ‘marine speciation paradox’ (Bierne et al. 2003), where species are differentiated despite large contact zones and a lack of geographic barriers. Allopatric speciation is particularly problematic for planktonic species given the large population sizes and ease of dispersal that make geographic isolation unlikely (Palumbi 1994). Darwin suggested that sympatric speciation would occur as species became ecologically divergent. Different populations would occupy different niches and would diverge into two species after many generations, without the need for geographic separation. Mayr (1948) showed that evidence of this kind of sympatric speciation could not be confirmed in the
cases from mammals, birds, butterflies or beetles examined and concluded that geographic isolation was the only possible mechanism of speciation. However he later points out that the absence of sympatric speciation in mammals and birds does not mean it is not possible for other groups (Mayr 2001). In the plankton, the theoretical absence of geographic barriers makes sympatric speciation the more satisfactory explanation for the diversity seen at sea (Norris 2000).

Three mechanisms have been identified by which sympatric speciation could occur, (1) disruptive selection acts against intermediates in a bimodal gene distribution pulling two populations apart, (2) allochronic speciation, owing to a drifting apart of the breeding season and (3) colonization of a new host, the case of host-specific species (Mayr 1982). The first two seem possible for the case of *P. galaxiae*.

The first mechanism is also referred to as ecological speciation (Coyne and Orr 1998; Schluter 2001) and may occur both in allopatry and sympatry (Schluter 2001). In the case of *P. galaxiae*, it is necessary to imagine a single population that blooms between March and August. Disruptive selection would negatively select cells blooming in April and July causing the separation into three blooms. The disadvantages to blooming in April or July would have to remain constant year after year in order to maintain the pattern of occurrence and allow genetic differentiation to occur through mutation and genetic drift. In this case selection could either act on cell size or on the timing of the bloom. Different sized cells could have advantages at different times of the year or there could be some particular disadvantage to blooming in April or July.

The second case would be possible if the three blooms were not just occurring at different times but if the blooms were also the occasion for sexual reproduction. Sexual
reproduction in *Pseudo-nitzschia* requires the fusion of gametes from opposite mating types (heterothallism) in order to restore maximum cell size (Davidovich and Bates 1998). In unialgal cultures of *P. galaxiae* enlargement has never been observed. Cell size falls over time and eventually, generally after less than one year, the culture dies indicating that *P. galaxiae* is a sexual heterothallic species. It is logical to assume that sex is more likely to occur during the bloom when cell concentrations are higher and there is a greater encounter rate between opposite mating types. Cells blooming in March would then be more likely to mate with other cells blooming at the same time. Marine sibling species with allochronous breeding seasons are common. Knowlton (1993) lists 16 examples of sympatric sibling marine invertebrate species with planktonic life stages that have asynchronous reproductive cycles. This situation is also seen in flowering plants and is referred to as assortative mating, where genetic variations in flowering time create temporal population structure (Weis and Kossler 2004). When examining the genetic structure of such populations, samples group by time of reproduction rather than the year in which they were collected (Hendry and Day 2005). This also applies to *P. galaxiae*, if this species were panmictic, genetic differences between strains would occur at random due to mutation and genetic drift. Strains would not group by period of occurrence or show greater diversity between the different seasonal blooms than within them. *P. galaxiae* has a temporal population structure that could make it more responsive to selection (Weis and Kossler 2004). Even if sexual reproduction is possible between the clades within *P. galaxiae*, if it mainly occurs within the bloom periods, differences between the blooms will accumulate. Hendry and Day (2005) propose the adoption of ‘isolation by time’, for populations that are isolated due to differences in reproductive time rather than isolated by distance, and this seems to apply to *P. galaxiae*. Whether selection acts on physiological characteristics or on the possible endogenous timing, separation between populations
could occur as a result of the tight coupling between environmental change and plankton dynamics (Hays et al. 2005).

The example of *P. galaxiae* illustrates the complexity of the phytoplankton and its similarities to higher plants and animals. Small is not necessarily simple. The evidence is growing that microbes follow similar patterns to more morphologically complex organisms. *P. galaxiae* demonstrates that it is right to compare microorganisms with plants and animals because the patterns do seem to be the same for all domains of life regardless of size (Whitfield 2005).

**CHAPTER SUMMARY**

- The three blooms of *P. galaxiae* during the year have been separated on the basis of size. This separation has a genetic basis as each bloom is dominated by a different ITS type. The LSU marker gives a lower level of resolution separating the March bloom from the other two.

- Rare genotypes are present in both LSU and ITS and seem to be an indication of further diversity within *P. galaxiae*. These genotypes demonstrate the flexibility of marine ecosystems where wanderers at the limits of their physiological tolerance could potentially form blooms were conditions to change in the future.

- These ITS types occur year after year in the same period of the year, indicating that the differences seen between the blooms are not due to cohorts from different sexual reproduction events.

- The pattern of morphological and genetic differences suggests a mechanism for speciation, that of separation by time rather than in space. The complexity of the patterns seen within *P. galaxiae* suggest that microbes should be thought of as
simple but may subject to processes as intricate as those controlling higher plants and animals.
Chapter Six

Conclusions and Future Prospective

Conclusions

The seasonality of several different planktonic groups and the diversity within these groups have been investigated during the year at the MC station in the Gulf of Naples. The main themes are that there is a higher level of diversity seen with the genetic methods than can be seen using morphological characters. This diversity has a seasonal pattern in many cases, with different genotypes being present at different times of the year.

The <10 μm phytoflagellate fraction seen in the phytoplankton abundance counts of the MC Long-Term Ecological Programme was shown to contain three main taxonomic classes, Prymnesiophyceae, Cryptophyceae and Chrysophyceae. These classes showed some seasonality in that the Chrysophyceae were present in summer and the Pelagophyceae were detected between December and March with a peak in February. The Prymnesiophyceae and Cryptophyceae were present all year round.

There is an indication from the clone libraries of the high diversity within these classes and how that diversity changes during the year. Some OTUs are present at particular times of the year, as demonstrated by the Cryptophyceae that show interesting seasonality reflected in the finding of different OTUs dominating different libraries. The results from the dot blots and clone libraries give an indication of the most important ultraplanktonic groups in the Gulf of Naples during the year and serve as a starting point for more in depth studies into their diversity and the role of the various species within these classes during the year.
The assessment of changes in taxonomic composition of the ultraplankton with depth showed spikes in signal possibly indicating the presence of species adapted to specific depths. The Pelagophyceae showed the most dramatic change with depth, being present between 50 and 70 m during May and August. The other classes showed some preference for subsurface waters between 2 and 10 m, possible due to exclusion from surface waters by high numbers of a few diatom species.

The results from the diatoms show similar seasonality and diversity with the finding of new genotypes and the demonstration of the seasonal occurrence of some already known taxa. *Pseudo-nitzschia* was shown to contain more taxa than can be detected in the phytoplankton counts based on morphology visible in LM. There is some indication of the distribution of these genotypes during the year from the libraries, but a more comprehensive sampling of different periods is needed to conclusively determine the timing of each taxon.

The use of two different markers to construct clone libraries highlights some methodological considerations. LSU in diatoms, and particularly in *Pseudo-nitzschia*, has been sequenced for a large number of strains of many different species. This enabled the detection of chimeras because representative sequences were available for comparison. 16S is less well studied in most eukaryotic ultraplanktonic groups and this made the detection of chimeras more difficult. It is possible that a greater percentage of the 16S library sequences are chimeric and this could not be detected because of a lack of sequences from known species for comparison. It is interesting that only three new genotypes were discovered within *Pseudo-nitzschia* whereas over 100 unknown
prymnesiophyte OTUs were detected. This is likely to be due to the more intensive study of diatoms over prymnesiphytes.

The genetic basis for multiple blooms by a single morphological species was investigated for the diatom *P. galaxiae* that blooms three times during the year in the Gulf of Naples. These three blooms were shown to be dominated by different ITS types, whereas LSU only showed a difference between the March bloom and those of May and August. For both markers, rarer genotypes were found that did not dominate a particular bloom period during any of the years sampled. It is likely that these types correspond to genotypes at the limits of their ecological tolerance that are not able to form a recognisable bloom. The pattern of occurrence of *P. galaxiae* with genetic and slightly morphological differences indicates a possible mechanism for speciation at sea, that of separation by time, where populations become separated on a temporal rather than spatial scale. The populations within *P. galaxiae* could have become separated either by selection favouring different sizes at different times of the year or by changes in the timing of the bloom, either of which would bring about non random mating if the majority of sexual reproduction takes place during the bloom.

**Future Prospective**

The development of genetic techniques using probes targeting the OTUs or clades within the major classes in the Gulf of Naples may prove useless in more accurately identifying the period of the year in which these taxa are present enabling a more focused sampling and culturing effort. FISH has already been used to identify cells belonging to the new lineages within the Stramenopiles (Massana et al. 2002) and may prove equally useful in identifying cells belonging to the lineages detected during this thesis. An increased sequencing effort in also needed in order to identify the sequences
found in the clone libraries, as many of the known taxa within less studied classes have not been sequenced.

In order to determine and describe a species it must be in culture (Chretiennot-Dinet and Courties 1997; Pedros-Alio 2006). The irony of this thesis is that a return to culturing is needed in order to understand the diversity seen with the ultraplanktonic classes and the new genotypes seen within *Pseudo-nitzschia*. This may require improvements in culturing techniques, given that most of these taxa probably have not been cultured and may be difficult to maintain under laboratory conditions. However, successful culturing of the SAR11 bacterial lineage (Rappe et al. 2002) and the number of new taxa recovered by SDC targeting the Cryptophyceae (Cerino 2004) indicate that the possibility of culturing many of the OTUs found during this thesis is not remote.

The significance of the diversity found is that the ecology of the Gulf of Naples cannot be fully understood until these taxa are well studied. It should be stressed that the molecular tools described in Chapter One and those used during this thesis should be used, and are only useful if used, as part of studies using other methods including cultures and morphological investigations. The use of a single genetic marker without information from ecology, physiology and morphology can simply confuse the picture. The results from this thesis highlight the need for more study into diversity and seasonality and show that it is a mistake to dismiss the smaller plankton size fraction and assume that small equals simple. This fraction is proving to be too complex and diverse to be overlooked.
Chapter Seven

References


149


DE BRUIN, A., B. W. IBELENGS, M. RIJKEBOER, M. BREHM, and E. VAN DONK. 2004. Genetic variation in Asterionella formosa (Bacillariophyceae): is it linked to


GUILLOU, L., W. EIKREM, M.-J. CHRÉTIENNOT-DINET, F. LE GALL, R. MASSANA, K. ROMARI, C. PEDRÓS-ALIÓ, and D. VAULOT. 2004. Diversity of picoplanktonic prasinophytes assessed by direct nuclear SSU rDNA sequencing of
environmental samples and novel isolates retrieved from oceanic and coastal marine ecosystems. Protist **155**: 193-214.


**Hansen, M. C., T. Tolkner-Nielsen, M. Givskov,** and **S. Molin.** 1998. Biased 16S rDNA PCR amplification caused by interference from DNA flanking the template region. FEMS Microbiology Ecology **26**: 141-149.


species complex (Bacillariophyceae): morphometric analysis, ultrastructure and mating data provide evidence for five new species. *Phycologia* **43**: 459-482.


ORSINI, L., G. PROCACCINI, D. SARNO, and M. MONTRESOR. 2004. Multiple rDNA ITS-types within the diatom Pseudo-nitzschia delicatissima (Bacillariophyceae) and their relative abundances across a spring bloom in the Gulf of Naples. Marine Ecology Progress Series 271: 87-98.


communities in a pluriannual time series at a coastal Mediterranean site (Gulf of Naples): an attempt to discern recurrences and trends. Scientia Marina 68 (Suppl. 1): 65-83.


costatum with the description of S. grevillei sp. nov. Journal of Phycology 41: 140-151.


## Appendix A

Appendix A: Details of all *P. galaxiae* strains used including collection dates and locations

<table>
<thead>
<tr>
<th>Strain</th>
<th>Collection Date</th>
<th>Sample Details</th>
<th>Isolating Scientist</th>
</tr>
</thead>
<tbody>
<tr>
<td>B54</td>
<td>29th May 2001</td>
<td>SDC 4II</td>
<td>Luisa Orsini</td>
</tr>
<tr>
<td>B55</td>
<td>29th May 2001</td>
<td>SDC 4II</td>
<td>Luisa Orsini</td>
</tr>
<tr>
<td>B58</td>
<td>18th July 2001</td>
<td>Net MC 0m 20μm</td>
<td>Luisa Orsini</td>
</tr>
<tr>
<td>P1</td>
<td>29th May 2001</td>
<td>SDC 4V</td>
<td>Luisa Orsini</td>
</tr>
<tr>
<td>P5</td>
<td>29th May 2001</td>
<td>SDC 4V</td>
<td>Luisa Orsini</td>
</tr>
<tr>
<td>B157</td>
<td>24th January 2003</td>
<td>NORBAL L20 150-96m</td>
<td>Federica Cerino</td>
</tr>
<tr>
<td>B158</td>
<td>24th January 2003</td>
<td>NORBAL L20 150-96m</td>
<td>Federica Cerino</td>
</tr>
<tr>
<td>B161</td>
<td>21st February 2003</td>
<td>Net MC 0m 20μm</td>
<td>Federica Cerino</td>
</tr>
<tr>
<td>AL 8</td>
<td>19th August 2003</td>
<td>Net MC588 0m 20μm</td>
<td>Alberto Amato</td>
</tr>
<tr>
<td>AL 10</td>
<td>19th August 2003</td>
<td>Net MC588 0m 20μm</td>
<td>Alberto Amato</td>
</tr>
<tr>
<td>AL 11</td>
<td>19th August 2003</td>
<td>Net MC588 0m 20μm</td>
<td>Alberto Amato</td>
</tr>
<tr>
<td>AL 15</td>
<td>19th August 2003</td>
<td>Net MC588 0m 20μm</td>
<td>Alberto Amato</td>
</tr>
<tr>
<td>AL 17</td>
<td>19th August 2003</td>
<td>Net MC588 0m 20μm</td>
<td>Alberto Amato</td>
</tr>
<tr>
<td>AL 24</td>
<td>19th August 2003</td>
<td>Net MC588 0m 20μm</td>
<td>Alberto Amato</td>
</tr>
<tr>
<td>AL 31</td>
<td>21st January 2004</td>
<td>Net MC 0m 5μm</td>
<td>Alberto Amato</td>
</tr>
<tr>
<td>AL 32</td>
<td>21st January 2004</td>
<td>Net MC 0m 5μm</td>
<td>Alberto Amato</td>
</tr>
<tr>
<td>AL 33</td>
<td>19th March 2004</td>
<td>Net MC 0m 5μm</td>
<td>Alberto Amato</td>
</tr>
<tr>
<td>AL 34</td>
<td>19th March 2004</td>
<td>Net MC 0m 5μm</td>
<td>Alberto Amato</td>
</tr>
<tr>
<td>SM 1</td>
<td>5th April 2004</td>
<td>SDC-I2</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 3</td>
<td>5th April 2004</td>
<td>SDC-I2</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 4</td>
<td>5th April 2004</td>
<td>SDC-I2</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 6</td>
<td>5th April 2004</td>
<td>SDC-I2</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 7</td>
<td>5th April 2004</td>
<td>SDC-I2</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 10</td>
<td>11th May 2004</td>
<td>SDC-I13</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 14</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 15</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 16</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 17</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 19</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 20</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 21</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 24</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 25</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 26</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 27</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 28</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 29</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 30</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 31</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 32</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 33</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 34</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 35</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 36</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 37</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 38</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 40</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 41</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 42</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 43</td>
<td>9th June 2004</td>
<td>SDC-I1</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>Strain</td>
<td>Collection Date</td>
<td>Sample Details</td>
<td>Isolating Scientist</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------</td>
<td>----------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>SM 44</td>
<td>25\textsuperscript{th} August 2004</td>
<td>Net MC629 0m 20\textmu{}m</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 45</td>
<td>25\textsuperscript{th} August 2004</td>
<td>Net MC629 0m 20\textmu{}m</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 46</td>
<td>25\textsuperscript{th} August 2004</td>
<td>Net MC629 0m 20\textmu{}m</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 47</td>
<td>25\textsuperscript{th} August 2004</td>
<td>Net MC629 0m 20\textmu{}m</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 48</td>
<td>25\textsuperscript{th} August 2004</td>
<td>Net MC629 0m 20\textmu{}m</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 49</td>
<td>25\textsuperscript{th} August 2004</td>
<td>Net MC629 0m 20\textmu{}m</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 50</td>
<td>25\textsuperscript{th} August 2004</td>
<td>Net MC629 0m 20\textmu{}m</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 52</td>
<td>8\textsuperscript{th} March 2005</td>
<td>SDC-I2 MC650</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 53</td>
<td>15\textsuperscript{th} February 2005</td>
<td>SDC-II MC651</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 54</td>
<td>15\textsuperscript{th} February 2005</td>
<td>SDC-I3 MC651</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 55</td>
<td>2\textsuperscript{nd} March 2005</td>
<td>SDC-III1 MC653</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 57</td>
<td>15\textsuperscript{th} March 2005</td>
<td>SDC-III11 MC655</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 58</td>
<td>15\textsuperscript{th} March 2005</td>
<td>SDC-IV3 MC655</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 59</td>
<td>15\textsuperscript{th} March 2005</td>
<td>SDC-IV3 MC655</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 60</td>
<td>15\textsuperscript{th} March 2005</td>
<td>SDC-IV3 MC655</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 61</td>
<td>15\textsuperscript{th} March 2005</td>
<td>SDC-V3 MC655</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 62</td>
<td>15\textsuperscript{th} March 2005</td>
<td>SDC-V3 MC655</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 63</td>
<td>15\textsuperscript{th} March 2005</td>
<td>SDC-V3 MC655</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 64</td>
<td>15\textsuperscript{th} March 2005</td>
<td>SDC-V3 MC655</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 65</td>
<td>8\textsuperscript{th} March 2005</td>
<td>SDC-IV3 0m MC654</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 67</td>
<td>17\textsuperscript{th} May 2005</td>
<td>SDC-II4 MC664</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 68</td>
<td>17\textsuperscript{th} May 2005</td>
<td>SDC-II3 MC664</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 69</td>
<td>17\textsuperscript{th} May 2005</td>
<td>SDC-III3 MC664</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 70</td>
<td>17\textsuperscript{th} May 2005</td>
<td>SDC-II4 MC664</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 71</td>
<td>17\textsuperscript{th} May 2005</td>
<td>SDC-II4 MC664</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 75</td>
<td>17\textsuperscript{th} May 2005</td>
<td>SDC-II4 MC664</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 77</td>
<td>17\textsuperscript{th} May 2005</td>
<td>SDC-III3 MC664</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 78</td>
<td>17\textsuperscript{th} May 2005</td>
<td>SDC-III3 MC664</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 79</td>
<td>17\textsuperscript{th} May 2005</td>
<td>SDC-III3 MC664</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 80</td>
<td>17\textsuperscript{th} May 2005</td>
<td>SDC-III3 MC664</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 81</td>
<td>24\textsuperscript{th} May 2005</td>
<td>SDC-II1 MC665</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 89</td>
<td>23\textsuperscript{rd} August 2005</td>
<td>MC678 net</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 97</td>
<td>23\textsuperscript{rd} August 2005</td>
<td>MC678 net</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 98</td>
<td>23\textsuperscript{rd} August 2005</td>
<td>MC678 net</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 99</td>
<td>23\textsuperscript{rd} August 2005</td>
<td>MC678 net</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 100</td>
<td>30\textsuperscript{th} August 2005</td>
<td>MC679 net</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 102</td>
<td>30\textsuperscript{th} August 2005</td>
<td>SDC-III1 0m MC697</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 103</td>
<td>30\textsuperscript{th} August 2005</td>
<td>SDC-III1 0m MC697</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 104</td>
<td>30\textsuperscript{th} August 2005</td>
<td>SDC-III1 0m MC697</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 105</td>
<td>30\textsuperscript{th} August 2005</td>
<td>SDC-III1 0m MC697</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 106</td>
<td>30\textsuperscript{th} August 2005</td>
<td>SDC-III1 0m MC697</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 107</td>
<td>30\textsuperscript{th} August 2005</td>
<td>SDC-III2 10m MC697</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 108</td>
<td>30\textsuperscript{th} August 2005</td>
<td>SDC-III2 10m MC697</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 109</td>
<td>30\textsuperscript{th} August 2005</td>
<td>SDC-III4 10m MC697</td>
<td>Sarah McDonald</td>
</tr>
<tr>
<td>SM 110</td>
<td>30\textsuperscript{th} August 2005</td>
<td>SDC-IV2 10m MC697</td>
<td>Sarah McDonald</td>
</tr>
</tbody>
</table>