AN ABSTRACT OF THE DISSERTATION OF

Michael Stephen Rappé for the degree of Doctor of Philosophy in Genetics presented on May 21, 1997. Title: Diversity of Bacterioplankton and Plastid SSU rRNA Genes from the Eastern and Western Continental Shelves of the United States.

Redacted for privacy

Abstract approved: _____

Stephen J. Giovannoni

The phylogenetic diversity of two continental shelf picoplankton communities was examined by analyzing SSU (16S) ribosomal RNA (rRNA) genes amplified from environmental DNA with bacterial-specific primers and the polymerase chain reaction (PCR). Picoplankton populations collected from the pycnocline (10 m) over the eastern continental shelf of the United States near Cape Hatteras, North Carolina, and surface seawater (10 m) from the western continental shelf of the United States 8 km west of Yaquina Head, Oregon, served as sources of bulk nucleic acids used in this study. A total of 285 SSU rRNA gene clones were analyzed in the two libraries, more than doubling the number previously available from seawater samples. In contrast to previous studies of bacterioplankton diversity from the open-ocean, a large proportion of the rDNA clones recovered in this study (38%) were related to plastid SSU rRNA genes, including plastids from bacillariophyte, prymnesiophyte, cryptophyte, chrysophyte, and prasinophyte algae, as well as a number of unique plastid rRNA gene clones for which no close phylogenetic relatives were discovered. A majority of the bacterial gene clones recovered

(72% of bacterial clones) were closely related to rRNA gene lineages discovered previously in clone libraries from open-ocean marine habitats, including the SAR86 cluster (γ Proteobacteria), SAR83, SAR11, and SAR116 clusters (all α Proteobacteria), the marine Gram-positive cluster (Actinomycetes), the marine group A/SAR406 cluster, and a cluster of environmental clones within the flexibacter-cytophaga-bacteroides phylum. A majority of the remaining bacterial clones were phylogenetically related to the γ and β subclasses of the Proteobacteria, including an rDNA lineage within the Type I methylotroph clade of the β subclass. The abundance of plastid rDNAs and the lack of cyanobacterial-related clones, as well as the presence of β Proteobacteria, are features of these coastal picoplankton gene clone libraries which distinguish them from similar studies of oligotrophic open-ocean sites. Overall, however, these data indicate that a limited number of as yet uncultured bacterioplankton lineages, related to those previously observed in the open-ocean, can account for the majority of cells in these coastal marine bacterioplankton assemblages.

Diversity of Bacterioplankton and Plastid SSU rRNA Genes from the Eastern and Western Continental Shelves of the United States

by

Michael Stephen Rappé

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented May 21, 1997

Commencement June 1998

Doctor of Philosophy dissertation of Michael Stephen Rappé presented on May 21, 1997

APPROVED:

Redacted for privacy

Major Professor, representing Genetics

Redacted for privacy

Director of Genetics Program

Redacted for privacy

Dean of Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for privacy

Michael Stephen Rappé, Author

ACKNOWLEDGEMENTS

First, I must thank Dr. Steve Giovannoni, my advisor for the last five years. His enthusiasm and support for my work is gratefully appreciated. I have benefited from the contributions of all past and present members of our laboratory, and thank them all. For several years now, Terah Wright, Doug Gordon, Dr. Ena Urbach, Marcelino Suzuki, Brian Lanoil, and Dr. Katharine Field have been a valuable resource to my research with many useful comments, suggestions, and help in all phases of this work. In addition, I thank Volker Huss and Linda Medlin for providing access to sequences prior to their publication, and analyses which contributed to the identification of the plastid SSU rRNA genes described here. Lee Kerkoff provided the nucleic acid sample used in the construction of the Ocean Marigins clone library, and Doug Gordon was responsible for the design of probe SAR83R. This work was supported by Department of Energy grant FG0693ER61697, a two-year contract from Dr. Paul Kemp at Brookhaven National Lab, and teaching assistantships from the Genetics Program and Microbiology Department at Oregon State University.

The people who contributed most to my work are the members of my family: my wife Jodi and daughter Kylie. Without their support and understanding over the last few years, none of this would have been possible. The same can be said for my brothers, Ed, Tom, and Ken, and my mother and father, who contributed their unconditional support.

CONTRIBUTION OF AUTHORS

Dr. Stephen J. Giovannoni was involved in all aspects of this work. Kevin Vergin assisted in the data collection for both clone libraries analyzed in this study, and Marcelino Suzuki performed the LH-PCR analyses in Chapters 4 and 5.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF TABLES

Diversity of Bacterioplankton and Plastid SSU rRNA Genes from the Eastern and Western Continental Shelves of the United States

CHAPTER 1
INTRODUCTION

1.1 The ecological role of bacterioplankton in the world's oceans

Over the past two decades, the role heterotrophic bacteria play in marine food webs and oceanic biogeochemical cycles has received considerable attention (e.g. see Azam and Cho 1987, Sherr and Sherr 1988, Fuhrman et al. 1989). In addition to their significant role as planktonic primary producers, bacteria in the water column appear to be the vehicle by which dissolved organic matter (DOM) is reconverted into a particulate form, allowing for the potential re-entry of DOM into the marine food web (Pomeroy 1974, Carlson et al. 1994). Recognition of this "microbial loop" in the marine food web has made it clear that bacterial turnover and metabolic activity are critical and often dominant processes intricately involved in global carbon and nitrogen budgets, as well as the production and consumption of greenhouse gasses (Azam et al. 1983, Ducklow et al. 1986, Williamson and Holligan 1990). Estimates of bacterial turnover and metabolic rates indicate that a large fraction of oceanic primary production cycles through bacterioplankton (Cole et al. 1988, Billen et al. 1990). Bacteria mediate much of the oxygen consumption and most of the carbon remineralization in marine waters (Ducklow and Carlson 1992).

Several recent studies indicate that a wide range of phenotypic and genotypic diversity exists within and between heterotrophic bacterioplankton

populations, and the bacterial fraction of planktonic ecosystems appears to be no less structurally dynamic than is commonly observed in the eukaryotic photoautotrophic and heterotrophic fractions of the marine planktonic community (Lee and Fuhrman 1991, Rehnstam et al. 1993, Fuhrman et al. 1994, Smith et al. 1995, Martinez et al. 1996). Methodological constraints, however, have led to a lack of consideration for individual species dynamics within heterotrophic bacterioplankton communities in systems level studies. The effect that spatial and temporal variation in bacterioplankton population structure has on the cycling of carbon and associated biogenic elements in the world's oceans is largely unknown (Azam et al. 1995, Kemp et al. 1995).

Currently in marine systems, bacterial processes are measured as community averages, with little if any characterization of the abundance or physiological activity of individual species. Unmeasured changes in community composition can invalidate predictions of the response of bacteria to a particular stimulus, unless the response is uniform across the community. Treating all species within the community as one "black box" masks genotype specific responses when all measurements are community averages. Lack of information on the structure of bacterioplankton communities makes it possible to measure average process rates but difficult to gain a mechanistic understanding of the factors which control bacterial processes.

1.2 The "great plate count anomaly"

In order to detect quantitative and qualitative changes in the structure of bacterioplankton populations, a knowledge of the types of organisms that can be potentially most abundant in a given seawater sample is necessary. As

with most ecosystems, the diversity of organisms within marine bacterioplankton populations has traditionally been assayed by culture-dependent methods (e.g. see ZoBell 1946, Baumann et al. 1972, Kita-Tsukamoto et al. 1993). It has long been known, however, that numbers of bacteria in seawater estimated by CFU counts are orders of magnitude lower than those estimated by epifluorescence direct counts (Jannasch and Jones 1959, Hobbie et al. 1977, Kogure et al. 1979, Ferguson et al. 1984). A significant proportion of cells visualized microscopically appear to be viable, but do not form visible colonies on plates (Staley and Konopka 1985), though this result has been debated (Zweifel and Hagström 1995, though see Choi et al. 1996). It is not currently known how much of the information derived from pure cultures of marine bacteria can be related to the activity of bacterioplankton *in situ*.

1.3 Molecular approach to assessing microbial diversity

In the early 1980s, Norman Pace and coworkers proposed that microbial diversity might be explored by analyses of nucleic acids isolated from natural samples without resorting to microbial cultivation (Olsen et al. 1986). Early studies focused on the 5S rRNA molecule, but it did not take long for the larger 16S rRNA to become the standard in these studies. The gradual refinement and application of these methods has profoundly changed the fields of microbial ecology and evolution. 16S rDNA sequences have been retrieved from such diverse habitats as soils (Liesack and Stackebrandt 1992, Stackebrandt et al. 1993), lake water (Bahr et al. 1996), thermal springs (Ward et al. 1990, Barns et al. 1994, Reysenbach et al. 1994), peat bogs (Hales et al. 1996), and the hindguts of termites (Ohkuma and Kudo, 1996, Paster et al.

1996). The results of these studies have often been startling: not only does it appear that the most abundant microbes in natural ecosystems are predominantly undescribed microbial species, but often these microbes cannot confidently be placed within any of the bacterial phyla described previously by Woese (Woese 1987).

How diverse are bacteria in natural systems? Liesack and Stackebrandt (Liesack and Stackebrandt 1992, Stackebrandt et al. 1993) analyzed the diversity of bacteria in an Australian soil sample by cloning 16S rRNA genes from environmental DNA. Sequence analysis of 30 clones and dot blot hybridization of 83 additional clones with taxon-specific oligonucleotide probes revealed three major bacterial groups. Representing 50% of the total, the most abundant clone type was closely related to nitrogen-fixing members of the α subclass of Proteobacteria, an unsurprising result given the source of the environmental sample. Six percent of the clones affiliated with the Planctomycetales, though they showed no close similarity to any described species. This was very surprising, since all Planctomycetales described to date come from aqueous habitats. The third major clone type, representing 20% of the total, appeared to form a unique line of descent within the *Bacteria*, possibly sharing a common ancestry with *Chlamydia* and the Planctomycetales.

A similar pattern of phylogenetic diversity was found in the intestinal microflora of the termite *Reticulitermes speratus* (Ohkuma and Kudo 1996). Of 55 rDNA clones, 20% were members of the Proteobacteria and showed close relationships with cultivated species of enteric and sulfate-reducing bacteria. The remaining clones showed no close relationships to known bacterial taxa, though members of the spirochetes (18%), flexibacter-cytophaga-bacteroides phylum (16%), and low G + C division of the Gram-positive bacteria (27%)

were recovered. 18% of the clones, among which 14% formed a single cluster, could not be placed within any of the major phyla of the *Bacteria*, and appeared to form unique lines of descent.

The termite and soil studies are used here as examples for the observation that the majority of microbial diversity found in natural habitats consists of uncultured, uncharacterized taxa. It is also evident that some clone types may be sufficiently divergent to form unique lines of descent with the *Bacteria*, and represent what are potentially undescribed phyla.

1.4 The diversity of bacterioplankton in the oligotrophic open-ocean

Among the first environments to receive significant attention in molecular studies were subtropical regions of the Atlantic and Pacific Oceans (Giovannoni et al. 1990a, Schmidt et al. 1991). Though the first studies analyzed surface seawater samples, samples from 80, 100, 250 and 500 meters, marine snow, and marine sediments have also been examined (Giovannoni et al. 1990a, Britschgi and Giovannoni 1991, Schmidt et al. 1991, DeLong 1992, Fuhrman et al. 1992, DeLong et al. 1993, Fuhrman et al. 1993, Mullins et al. 1995, Giovannoni et al. 1996b, Gordon and Giovannoni 1996, Gray and Herwig 1996, Field et al. 1997, Wright et al. 1997).

Important questions addressed by these studies included: (1) How diverse are bacterial rDNA lineages in a single sample? (2) Do environmental rDNA clones match rDNA sequences of known species? and (3) Are the same phylogenetic lineages found in different oceans?

Analyses of rDNA libraries cloned by different investigators using different methods, as well as supporting hybridization data, have shown that

16S rRNA genes isolated from natural samples of prokaryotic plankton (including both *Bacteria* and *Archaea*) are diverse, but predominantly fall into ten distinct phylogenetic lineages (Fig. 1.1) (Giovannoni et al. 1996a). Eighty-six percent of all bacterial 16S rDNAs cloned from marine picoplankton fall into one of the eight bacterial phylogenetic groups shown in Figure 1.1. Of this bacterial fraction, only 25% of the clones belong to the SAR83 clade of the α Proteobacteria or SAR6/SAR7 (Marine *Synechococcus/Prochlorococcus*) cyanobacterial cluster, members of which have been cultured. The remaining six groups are completely composed of uncultured bacterial groups.

Though the majority of cultured marine microbial species are Proteobacteria, only four of the ten major marine rRNA gene clusters are members of this class - the SAR11, SAR83, and SAR116 clusters of the α subclass, and the SAR86 cluster of the γ subclass. Interestingly, these four clusters are clearly distinct from the major genera of cultivated marine bacteria (mostly of the γ subclass) for which 16S rDNA sequences are available. The SAR83 cluster is, however, closely related to the aerobic photoheterotroph *Roseobacter denitrificans* (similarity = 0.96), as well as marine isolates from Oregon coast seawater (Suzuki et al. 1997). The SAR11 cluster is a rapidly evolving, deeply branching clade of the α subclass of Proteobacteria, and often appears as the deepest branch within this group (Giovannoni et al. 1990a). The marine *Synechococcus/Prochlorococcus* cluster is phylogenetically affiliated with the cyanobacteria, in particular to the marine *Synechococcus* group and *Prochlorococcus marinus*. It is not surprising that a cluster of cyanobacteria was recovered in clone libraries from pelagic marine samples since it was previously known that oxygenic phototrophs often constitute a significant proportion of the total bacteria in these environments. Although related to the

Figure 1.1. Composite phylogenetic tree displaying relationships among the most widespread bacterioplankton 16S rRNA gene clusters. This tree was inferred from comparisons of about 1000 nucleotide positions by the neighbor joining method. Strain OR412 is an undescribed heterotrophic bacterium. Of the eight bacterial clusters and two archaeal groups shown, only the marine *Synechococcus/Prochlorococcus* cluster and the SAR83 cluster have members that are in culture. Many other lineages found in marine aggregates or in a single clone library are not shown.

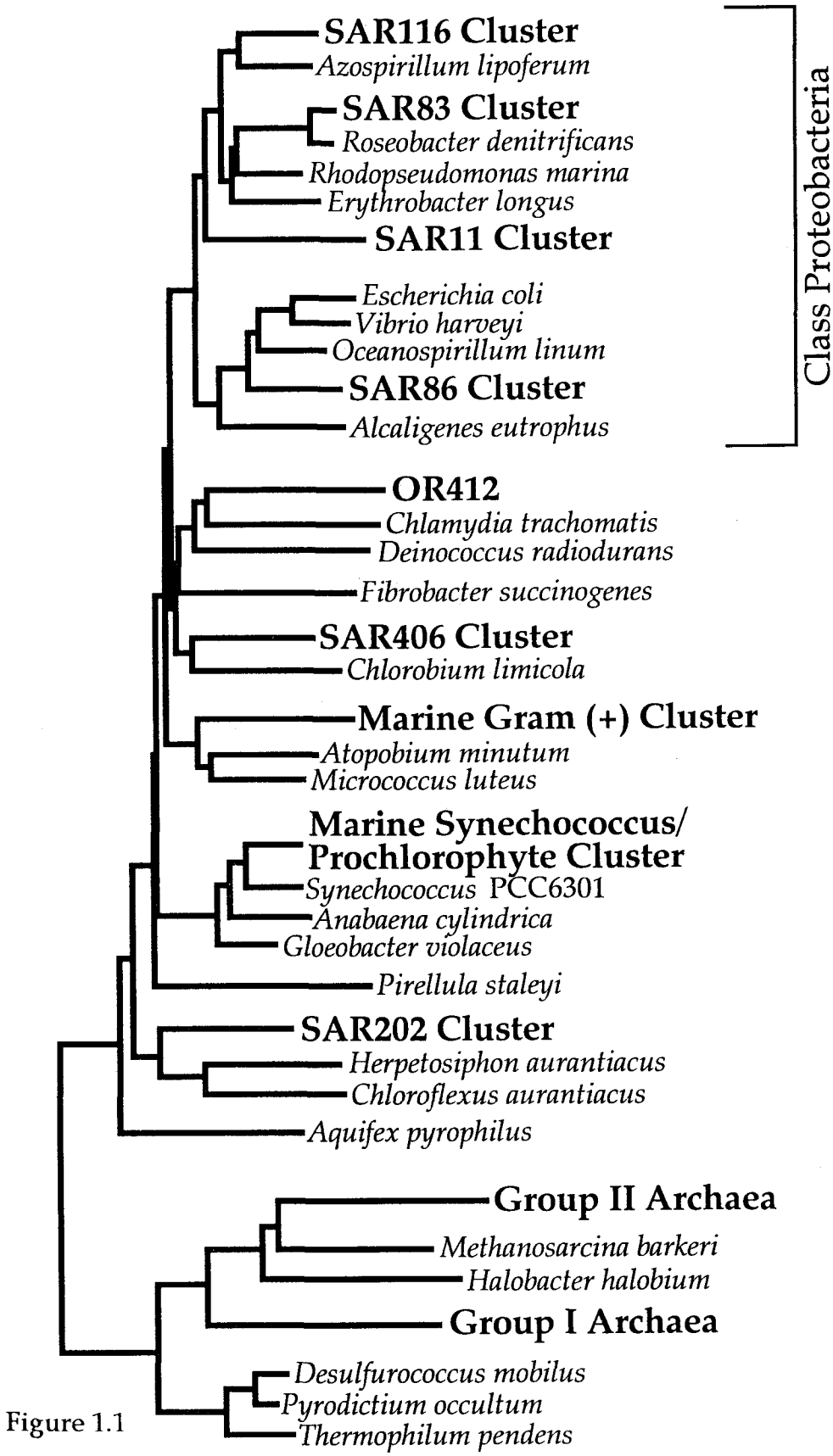

SAR116 Cluster
Azospirillum lipoferum
SAR83 Cluster
Roseobacter denitrificans
Rhodopseudomonas marina
Erythrobacter longus
SAR11 Cluster
Escherichia coli
Vibrio harveyi
Oceanospirillum linum
SAR86 Cluster
Alcaligenes eutrophus
Class Proteobacteria
OR412
Chlamydia trachomatis
Deinococcus radiodurans
Fibrobacter succinogenes
SAR406 Cluster
Chlorobium limicola
Marine Gram (+) Cluster
Atopobium minutum
Micrococcus luteus
Marine Synechococcus/
Prochlorophyte Cluster
Synechococcus PCC6301
Anabaena cylindrica
Gloeobacter violaceus
Pirellula staleyi
SAR202 Cluster
Herpetosiphon aurantiacus
Chloroflexus aurantiacus
Aquifex pyrophilus
Group II Archaea
Methanosarcina barkeri
Halobacter halobium
Group I Archaea
Desulfurococcus mobilus
Pyrodictium occultum
Thermophilum pendens

Figure 1.1

high G + C (Actinomycete) division of the Gram-positive bacteria, the marine Gram-positive cluster shows low similarity to all of the major high G + C Gram-positive SSU rRNA gene lineages, and instead forms a deep, unique branch in this phylum (Fuhrman et al. 1993). The phylogenetic positions of the SAR202 and SAR406 clusters, like those of the SAR11 and Marine Gram-positive clusters, are relatively deep with respect to the phyla in which they occur. The SAR202 cluster is the deepest branch of the *Chloroflexus/Herpetosiphon* phylum (the Green Non-Sulfur phylum) (Giovannoni et al. 1996b). The SAR406 cluster is a distant relative of the genus *Fibrobacter* and the green sulfur bacteria, which includes *Chlorobium*. It has no definitive phylum affiliation (Gordon and Giovannoni 1996). The remaining two groups are novel lineages within the Archaea - the Group I Archaea are peripherally related to hyperthermophilic Crenarchaeota, while the Group II Archaea appear to share a common ancestry with the Euryarchaeote *Thermoplasma acidophilum* (DeLong 1992, Fuhrman et al. 1992, DeLong et al. 1994).

The similarity between genes cloned directly from seawater and their closest relatives among cultured species represented in databases is shown graphically in Figure 1.2. Even though approximately 74% of marine bacterial species described in the systematic literature are represented in 16S rDNA sequence databases, the median similarity between marine bacterial genes retrieved from nature and genes from cultured marine microbial species is only ca. 0.87 (Fig. 1.2). The significance of minor sequence variation between environmental gene clones and sequences from cultivated species is not known. However, when an environmental clone is significantly dissimilar to any known gene sequence (similarity <0.96), it is unlikely to represent a

Figure 1.2. Chart summarizing the percent similarity to the nearest phylogenetic neighbor for environmental 16S rRNA gene clones from pelagic marine samples. This data set is composed of all published pelagic marine 16S rRNA gene clone libraries as well as unpublished data from 80 m and 250 m in the Sargasso Sea.

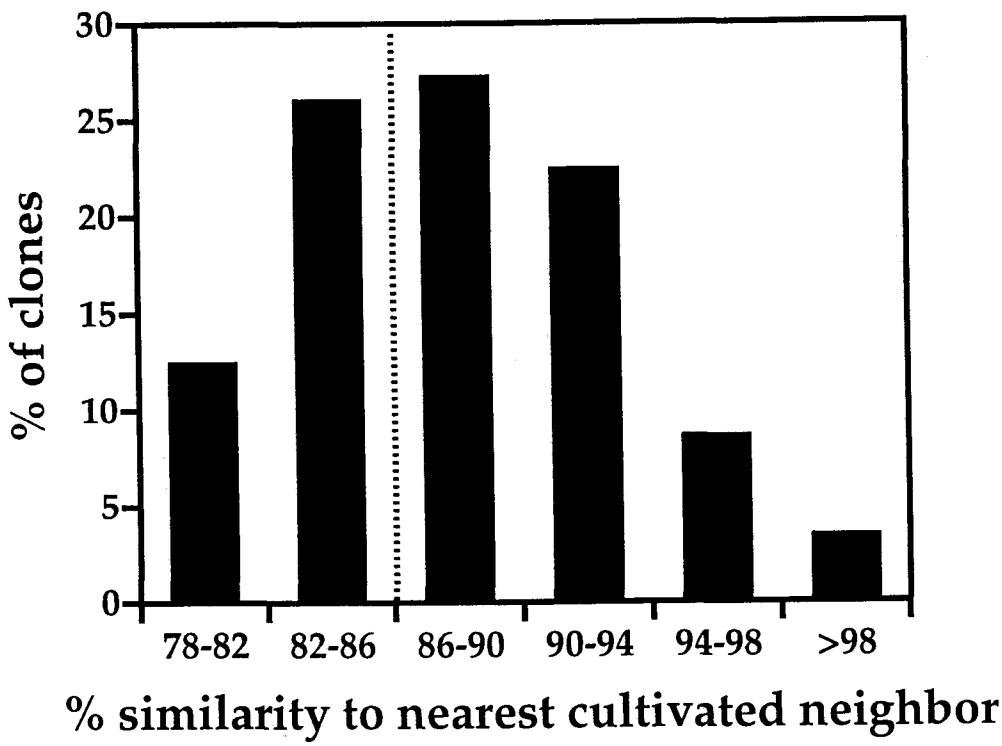

30
25
20
15
10
5
0
% of clones
78-82
82-86
86-90
90-94
94-98
>98
% similarity to nearest cultivated neighbor

Figure 1.2

cultured strain. To put these 16S rRNA sequence similarities in perspective, consider that bacterial 16S rDNAs diverge at an average rate of 1% per 50 million years (Moran et al. 1993). Even allowing for broad variation in rates of evolution, the data still support the conclusion that most marine bacteria do not share recent (i.e. in the Cambrian or later) common ancestors with cultured species.

In summary, the following general conclusions can be drawn from studies of bacterioplankton environmental clone libraries: First, the most abundant rDNA genes recovered in clone libraries do not correspond to cultured species. Second, many of the same novel gene lineages are found in clone libraries prepared from different oceans and by different methods. And third, these lineages (phylotypes) are often composed of multiple gene lineages (clusters) more related to each other than to any other gene lineage.

An unexpected aspect of bacterial diversity uncovered by these studies is the existence of 16S rRNA gene clusters: related sets of genes which are more similar to each other than to any rRNA gene sequences from a cultivated organism (Giovannoni et al. 1996b). From a phylogenetic perspective, gene clusters resemble gene clades from cultivated organisms, but because gene cluster sequences come from uncultivated organisms, it is not known whether or not they represent separate cellular lineages. An analysis of the 16S rRNA gene sequences from the seven *rrn* operons of *E. coli* strain PK3 (Cilia et al. 1996) revealed inter-operon similarities between 98.9%-100.0%, indicating that the divergence between *rrn* operons of the same organism may explain the shallowest branches in an environmental gene cluster, but is not sufficient to explain the high degree of sequence divergence characteristic of the gene clusters described here. The significance of these gene clusters in natural

populations is currently unknown, but many studies of natural microbial populations, from seawater and other habitats, have now observed them (Britschgi and Giovannoni 1991, Schmidt et al. 1991, DeLong et al. 1993, Fuhrman et al. 1993, Barns et al. 1994, DeLong et al. 1994, Hales et al. 1996, Ohkuma and Kudo 1996).

1.5 Objectives of this thesis

Only a small number of environmental SSU rRNA gene clones have been analyzed from coastal seawater samples. DeLong and coworkers compared SSU rRNA gene clones from marine aggregates to 18 SSU rRNA gene clones obtained from bacterioplankton in the surrounding water (DeLong et al. 1993). Of these 18 clones, 13 were members of the SAR11 cluster. Though these samples were too small to draw firm conclusions, they did indicate that perhaps more productive, nutrient-rich coastal habitats contained some of the same SSU rRNA gene lineages as are found in the open-ocean.

Chapters two and three of the work described here were executed as part of the U.S. Department of Energy's Ocean Margins Program. This program is focused on the eastern continental shelf of the United States near Cape Hatteras, North Carolina; its goal is to study the storage, flux, and transformations of carbon in ocean margin systems. Although continental shelves account for only a small percentage of the total surface area of the worlds oceans, they may provide a significant sink for atmospheric carbon dioxide and a major pathway for the export of organic matter into deep ocean basins (Walsh et al. 1985, Rowe et al. 1986, Walsh 1991). As a first step in analyzing the role of bacterioplankton community structure on organic carbon

transformations in this region, we investigated the phylogenetic diversity of this continental shelf picoplankton community by an analysis of PCR-amplified SSU rRNA genes. Our immediate objectives were (1) to determine whether the bacterioplankton in this region of the continental shelf were similar to those which had been found previously in open-ocean samples, and (2) to develop probes for the detection of key species in ecological studies planned for the future.

Chapters four and five of this study were sparked by a personal interest in the complex hydrography and high productivity common to the coast of Oregon, U.S.A. The Oregon coast represents an unusual area of high productivity with wind-driven coastal upwelling processes seasonally dominating the region's oceanography (Huyer 1977, Huyer 1983). Similar hydrographical events occur in both the northern and southern hemispheres along eastern ocean boundaries (Barber and Smith 1981). The high nutrient availability in these regions results in high productivity, such that they may be responsible for half the world's catch of fish even though they represent only 0.1% of the world's surface area (Ryther 1969). High phytoplankton abundance and elevated levels of primary production are characteristic of the Oregon coast during upwelling events (Small and Menzies 1985, Hood et al. 1992). Two ecologically significant questions were addressed by the construction of a SSU rRNA gene clone library from surface seawater collected over the western continental shelf of the Pacific Ocean off Oregon, USA. First, we were able to compare the SSU rRNA gene clones recovered in the Oregon coast clone library to the SSU rRNA genes obtained from bacterial isolates cultured from the same seawater sample. This comparison revealed that there was very little overlap between the SSU rRNA genes recovered from the

cultures and gene clones (Suzuki et al. 1997). In addition, I was able to compare the picoplankton SSU rRNA gene clones recovered off the Oregon coast to those recovered off Cape Hatteras, North Carolina, as well as to published clone libraries previously constructed from oligotrophic regions of the open-ocean, resulting in significant insights into the diversity of SSU rRNA types in coastal environments.

CHAPTER 2

CHROMOPHYTE PLASTID 16S RIBOSOMAL RNA GENES FOUND IN A CLONE LIBRARY FROM ATLANTIC OCEAN SEAWATER

Michael S. Rappé, Paul F. Kemp, and Stephen J. Giovannoni

Published in the *Journal of Phycology*,
Phycological Society of America, Lawrence, Kansas,
December, 1995, volume 31, number 6, pages 979-988

2.1 Abstract

In recent years, the cloning of ribosomal RNA genes from natural plankton communities has provided insight into the biodiversity of marine bacterioplankton. Small eukaryotic phytoplankton, like bacterioplankton, can be difficult to cultivate or identify routinely by morphological characteristics. We used bacterial-specific 16S rDNA primers to amplify genes from picoplankton samples collected on 0.2 μm filters by filtration from a depth of 10 m in the pelagic region over the continental shelf off Cape Hatteras, North Carolina. Nucleic acid sequencing and probe hybridization revealed that chromophyte plastid genes comprised 25% of the genes in a library of 170 clones. The plastid genes belonged to two groups within the Chromophyta: the Prymnesiophyceae and the Bacillariophyceae. Comparisons revealed substantial diversity among the bacillariophyte gene sequences, but the species from which the genes originated could not be identified because few sequences from cultured bacillariophytes are available. The prymnesiophyte genes also could not be identified, although they were most similar (similarity=0.94) to plastid genes from the coccolithophorid *Emiliania huxleyi* (Lohmann) Hay and Mohler strain PML92D. These results provide evidence of abundant chromophyte plastid 16S rRNA genes in the water over the continental shelf off Cape Hatteras. The results also suggest that plastid 16S rRNA genes may provide suitable genetic markers for studying phytoplankton biodiversity and biogeography.

2.2 Introduction

Several studies in recent years have demonstrated the usefulness of sequencing environmental 16S rRNA gene clones as a method for understanding the diversity of microorganisms in natural communities (Britschgi and Giovannoni 1991, DeLong et al. 1993, Fuhrman et al. 1993, Reysenbach et al. 1994, Mullins et al. 1995). This approach relies on the cloning of genes directly from mixed assemblages of cells collected from the environment, and thus circumvents the biases inherent in microbial cultivation. Cloning methods also involve biases, different from those inherent in cultivation, but have lead to the discovery of many novel bacterioplankton and Archaea which appear to be distributed widely in seawater. Recently, rDNA probes have been designed from environmental sequences for applications to ecological questions, including the determination of the spatial and temporal distributions of bacterial species (Giovannoni et al. 1990a, DeLong 1992).

Like bacterioplankton, small eukaryotic phytoplankton present problems for ecologists interested in biodiversity because they are often difficult to cultivate and present few morphological characters which can be used to routinely discriminate between biological species (Brand 1986, Campbell et al. 1994). Unidentified "green cells" constitute a major fraction of phytoplankton assemblages at our study site, a region on the continental shelf near Cape Hatteras selected for intensive studies of the carbon cycle by the Department of Energy Ocean Margins Program. Marshall and Ranasinghe (1989) reported that unidentified picoplankton (1-3 μm) consisting of cyanobacteria, Chlorophyta and "other unidentified green cells" accounted for 50.6% of the annual average cell concentration in water samples taken over a

ten year period (ending in 1985) from the eastern continental shelf of the United States. Molecular techniques employing plastid 16S rRNA genes, including *in situ* hybridization of taxon-specific oligonucleotide probes and the cloning and sequencing of plastid rRNA genes, could provide methods to circumvent the problem of identifying these cells (DeLong et al. 1989, Simon et al. 1995).

Ribosomal RNA gene cloning and sequencing has been used for investigations of chloroplast phylogeny, thereby providing databases of gene sequences which might be used to identify unknown plastid genes cloned from natural samples (Giovannoni et al. 1988b, Turner et al. 1989, Urbach et al. 1992). Plastid evolution is particularly interesting because of the diversity of chloroplasts and their apparent monophyletic origin within the cyanobacterial phylum of the bacteria (Fox et al. 1980). Chloroplast rRNA genes are so closely related to those of cyanobacteria that the same bacterial-specific polymerase chain reaction amplification primers work for both. However, phenotypically distinct plastid-containing groups within the Chlorophyta, the Chromophyta, and the Rhodophyta can be distinguished on the basis of plastid 16S rDNA sequence signatures (Delaney and Cattolico 1989, Douglas and Turner 1991, Markowicz and Loiseaux-de Goër 1991). Thus, comparisons of plastid rDNAs to sequence databases for known phytoplankton species can be a valuable tool for the identification and taxonomy of eukaryotic phytoplankton (Cattolico and Loiseaux-de Goër 1989).

A limitation of this approach is that the 16S rRNA plastid genes sequenced to date do not adequately represent the diversity of cultivated eukaryotic phytoplankton. Only eleven plastid 16S rDNA sequences from members of the Chromophyta are available through Genbank (accession

numbers in parentheses) and the Ribosomal Database Project (RDP; Maidak et al. 1994): the cryptophytes *Cryptomonas* ϕ (X56806) and *Pyrenomonas salina* (X55015), the phaeophyte *Pylaiella littoralis* (X14873, X14803), the prymnesiophytes *Ochrosphaera neapolitana* (X80390), *Ochrosphaera* sp. (X65101), *Isochrysis* sp. (X75518), and *Emiliania huxleyi* (X82156), the bacillariophytes *Skeletonema costatum* (X82154) and *Skeletonema pseudocostatum* (X82155), the raphidophyte *Heterosigma akashiwo*, formerly *Olisthodiscus luteus* (M82860, M34370; Hara and Chihara 1987), and the partially sequenced chrysophyte *Ochromonas danica.* (X53183). No representative plastid 16S rRNA gene sequences have been published for such chromophyte algae groups as the Eustigmatophyceae, Synurophyceae, and Xanthophyceae.

The unique plastid rRNA genes described here were encountered unexpectedly by the random sequencing of environmental 16S rDNA clones from picoplankton. The initial purpose of this study had been to identify bacterioplankton rRNA genes common to the eastern continental shelf of the United States off Cape Hatteras, North Carolina, so that oligonucleotide probes for ecological investigations of bacterioplankton might be assembled. The discovery of the plastid genes presented an opportunity to obtain genetic data which might eventually lead to an alternative method of identifying phytoplankton at division, species and population genetic levels. Such an identification scheme might include organisms which had eluded identification by other methods because they had not been cultivated and/or lacked clearly defined morphological characters. In addition to assembling a database of genetic sequence information, we also sought to identify unique sequence domains which might be used as sites for the annealing of diagnostic oligonucleotide probes.

2.3 Materials and methods

On 13 May 1993, a picoplankton sample was collected on 0.2 μm filters by filtration from a depth of 10 m at station 18 (35°59'N, 75°08'W), approximately 80 km northeast of Cape Hatteras, North Carolina, on R.V. Gyre cruise number 93-06. Total cellular nucleic acids were isolated from the picoplankton sample by lysis with proteinase K and SDS and phenol/chloroform extraction as previously described (Giovannoni et al. 1990b).

rRNA gene amplification. Bacterial and plastid 16S rRNA genes were amplified from the environmental sample of genomic DNAs using two general bacterial 16S rRNA primers (27F, 5'-AGAGTTTGATCMTGGCTCAG-3'; 1522R, 5'-AAGGAGGTGATCCANCCRCA-3'; Giovannoni 1991) and the polymerase chain reaction (Saiki et al. 1988). PCR reactions were performed in a MJ Research thermal cycler (MJ Research, Watertown, Massachusetts). The amplification reaction conditions were 5.0% acetamide, 0.2 μM of each amplification primer, 200 μM of each dNTP, *Pfu* DNA polymerase reaction buffer (10 mM KCl, 20 mM Tris·HCl pH 8.2, 6 mM $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, 0.1% TritonX-100), 2.5 units of *Pfu* DNA polymerase (Stratagene, La Jolla California), and 10 ng of template DNA in a total volume of 100 μL. Prior to the addition of *Pfu* DNA polymerase to the reaction mixture, a modified version of "hot start" PCR was employed. In a microcentrifuge tube on ice, all PCR reagents were added to the reaction mixture except for *Pfu* DNA polymerase and layered with mineral oil. The reaction mixture was then placed in the thermal cycler and heated to 95°C for 5 min. After cooling to 85°C for 1 min, enzyme was added to the reaction through the mineral oil

without removing the reaction tubes from the thermal cycler. The reactions were then cooled to the annealing temperature of 55°C to begin normal cycling. The amplification conditions were 95°C for 1 min, 55°C for 1 min, and 72°C for 1 min extended 5 sec per cycle for 35 cycles. Following the final cycle, the reaction was extended at 72°C for 10 min.

Clone library construction. The blunt-ended rDNA amplification product was cleaned using a QIAquick-spin PCR purification column (Qiagen, Chatsworth, California) following the manufacturer's instructions. The amplification product was visualized by electrophoresis through a 1.0% agarose gel in 1X TAE (40 mM Tris-acetate, 1 mM EDTA) containing 0.75 $\mu g \cdot mL^{-1}$ ethidium bromide. A single band of amplification product was observed, 1.5 Kb in size. After quantification with a Shimadzu UV160U spectrophotometer, the amplified rDNA was inserted into the *Sma*I restriction site of the phagemid vector pBluescript KSII- (Stratagene) by blunt-end ligation as previously described (Giovannoni et al. 1990b) with the following modifications: the ligation reaction mixture, consisting of 260 ng of insert, 50 ng of vector, and 1 unit of T4 DNA ligase in ligation buffer (10 mM Tris-HCl pH 8.3, 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.4 mM ATP), was incubated at 16°C for 18 h. The ligation product was diluted 2:3 with TE buffer (10 mM Tris-HCl, 1 mM EDTA pH 8.0) and used to transform competent *E. coli* XL1-blue cells (clone numbers 1-105; Stratagene) or *E. coli* MAX Efficiency DH10B cells (clone numbers 106-303; Gibco BRL, Bethesda, Maryland). Transformants were subsequently plated onto LB agar plates (1% tryptone, 0.5% yeast extract, 0.5% NaCl, 1.5% bacto-agar) containing the antibiotic ampicillin (200 $\mu g \cdot mL^{-1}$) and spread evenly with 40 µL of X-Gal (40 $mg \cdot mL^{-1}$) and 100 µL of IPTG (100 mM). The 170 clones which constitute the library described here are those that

either hybridized to a taxon-specific oligonucleotide probe or contained a full size (1.5 Kb) insert. They are numbered discontinously from 1-303. Some of the clones with lengths of less than 1.5 Kb apparently contained fragments of rRNA genes, which is observed commonly in clone libraries of this type. For long-term storage, stab cultures of all were made in LB top agar (1% tryptone, 0.5% yeast extract, 0.5% NaCl, 0.7% bacto-agar) and kept at 4°C. Clones were also stored in 7.0% DMSO at -80°C.

rRNA gene sequencing and phylogenetic analysis. Template plasmid DNAs were prepared from overnight cultures of positive transformants by alkaline lysis using a Prep-A-Gene plasmid purification kit (Bio-Rad, Richmond, California) and quantified spectrophotometrically. Plasmid DNAs were visualized by electrophoresis through a 1.0% agarose gel containing 0.5 $\mu g \cdot mL^{-1}$ ethidium bromide in 1X TAE. A pBluescript plasmid known to contain the correct sized insert (1.5 Kb) was used as a positive size standard while one lacking an insert, isolated from a blue colony, served as a negative control. Plasmid DNAs from clones containing full-length inserts were sequenced on both strands using an ABI model 373A automated sequencer (Applied Biosystems, Foster City, California) and dye-terminator chemistry. DNA sequence data from cloned 16S rDNA genes were manually aligned to bacterial and plastid sequences obtained from the Ribosomal Database Project (Maidak et al. 1994) using the Genetic Data Environment (GDE) v2.0 sequence analysis software package supplied by Steve Smith. Regions of ambiguous alignment were excluded from the analysis. The Phylogeny Inference Package (PHYLIP v3.5; J. Felsenstein) was used for all phylogenetic analyses except for maximum likelihood, which used the program fastDNAml distributed by the RDP (Maidak et al. 1994). Evolutionary distances were calculated using the

program DNADIST and the Kimura 2-parameter model for nucleotide change, with a transition/ transversion ratio of 2.0 (Kimura 1980). The program NEIGHBOR was used to reconstruct phylogenetic trees from evolutionary distance data by the neighbor-joining method (Saitou and Nei 1987). Parsimony trees were reconstructed using the program DNAPARS. A total of 100 bootstrapped replicate resampling data sets for both DNADIST and DNAPARS were generated using SEQBOOT with random sequence addition and global rearrangement (Felsenstein 1985). All phylogenetic trees were edited using the program Treetool provided by Mike Maciukenas for the RDP. The program gRNAid v1.4, supplied by Shannon Whitmore (Mentor Graphics, Wilsonville, OR), was used to construct secondary structural models for the cloned 16S rDNAs. Secondary structural models were manipulated manually for optimal agreement with published 16S rRNA secondary structures (Gutell 1993). All gene sequences were submitted to the RDP program CHECK_CHIMERA to aid in the detection of chimeric gene artifacts (Maidak et al. 1994).

Accession numbers. Nucleotide sequences were filed in Genbank under the accession numbers: OM13, U32667; OM19, U32668; OM20, U32670; OM21, U32671; OM22, U32669.

Colony blot construction. Colony blots (pinblots) were prepared using an adaptation of the methods of Jeff Stein (RBI, LaJolla, CA). 213 transformants were included in the pinblots, of which 43 were later determined to be not of full length and thus not included in the library analysis. Transformants were grown in microtiter plate wells containing 100 µL of LB broth (with ampicillin, 200 $\mu g \cdot mL^{-1}$) overnight at 37°C. Microtiter plate-sized Zeta-probe membranes (Bio-Rad) were placed on 150 x 15 mm culture plates containing approximately

75 mL LB agar (with ampicillin, 200 $\mu g \cdot mL^{-1}$). A Kleinschmidt 96A-PR microtiter plate replicator (M. Kleinschmidt, Palm Springs, California) was used to inoculate the membranes from the microtiter plates. Inoculated membranes were incubated at 37°C for 6 h on the agar plates (colony size of approximately 1.5 mm in diameter). The membranes were then placed in a square baking dish on Whatman 3 mm chromatography paper saturated with 5% SDS/2X SSC (150 mM NaCl, 15 mM Na_3citrate·$2H_2O$ pH 7.0), allowed to stand 2 min, and microwaved for 2.5 min on high power. The membranes were submerged for 20 min at 37°C in 50 mM Tris-HCl pH 8.0, 50 mM EDTA, 100 mM NaCl, 1% sarcosyl, and 250 $\mu g \cdot mL^{-1}$ proteinase K. The membranes were baked for 15 min at 80°C and UV-crosslinked (1200 Joules) in a Ultra-Lum UVC-515 ultraviolet multilinker to fix the DNA to the membrane.

Oligonucleotide probes and colony hybridization. To aid in screening the clone library, two oligonucleotide probes were constructed using an Applied Biosystems automated DNA synthesizer (Table 2.1). The probe CHRYS1 (5'-ACT AAT TCT ACC TTA GGC-3') was designed to complement the domain of clones OM12, OM13, OM19, OM20, and OM22 which corresponded to *E. coli* 16S rRNA positions 1464-1481 (18 bases). The probe HAP1 (5'-ACT AGC CCT ACC TTA GAC-3') was designed to complement clone OM21 in the same domain as CHRYS1. The two probes were screened for possible cross-hybridization to unrelated clones by submitting the probes to the program CHECK_PROBE, available through the RDP, for comparison to sequences in the RDP (Maidak et al. 1994). The bacterial probe 338R was hybridized to the pinblots to check for the presence of a 16S rRNA gene clone in the inserted DNA of the recombinant plasmid. Each oligonucleotide probe

Table 2.1. Alignment of the CHRYS1 and HAP1 probe target sites with the plastid 16S rRNA target site regions (*E. coli* numbering) from representative chromophyte algae.

Table 2.1

CLASS	SPECIES	PLASTID 16S rRNA SEQUENCE (5'-3')
		1464--GCCUAAGGUAGAAUUAGU--1481 (CHRYS1 target site)
environmental clones	OM12	1444--cuaaccu-uuu-ggaggagggc--GCCUAAGGUAGAAUUAGU--gacuggggugaagucguaac--1501
	OM13	--cuaaccg-uuu-ggaggaggac--GCCUAAGGUAGAAUUAGU--gacuagggugaagucguaac--
	OM19	--cuaaccg-uuu-ggaggaggau--GCCUAAGGUAGAAUUAGU--gacuagggugaagucguaac--
	OM20	--cuaaccu-uuu-ggaggagggc--GCCUAAGGUAGAAUUAGU--gacuggggugaagucguaac--
	OM22	--cuaaccg-uuu-ggaggaggau--GCCUAAGGUAGAAUUAGU--gacuggggugaagucguaac--
Bacillariophyceae	*Skeletonema costatum*	--cuaacca-uuu-gggggagggu--GCCUAAGGUAGAAUUAGU--gacuggggugaagucguaac--
	Skeletonema pseudocostatum	--cuaacca-uuu-gggggagggu--GCCUAAGGUAGAAUUAGU--gacuggggugaagucguaac--
	Corethron criophilum	--uuaaccg-uuu-ggaggaggac--GCCUAAGGUAGAAUUAGU--gacuagggugaagucguaac--
Raphidophyceae	*Heterosigma akashiwo*	--cuaaccg-uuu-ggaggaaggc--GCCUAAGGUAG**G**AUU**G**GU--gacuggggugaagucguaac--
Phaeophyceae	*Pylaiella littoralis*[b]	--uuaaccu-uuuuggagagaggc--GCCUAAGGUA**AGGC**UAGU--gacuggggugaagucguaac--
Euglenophyceae	*Euglena gracilis*	--uugccug-aaa-agagggaaau--**A**CCUAAGG**CCUGGC**U**G**GU--gacuggggugaagucguaac--
		1464--GUCUAAGGUAGGGCUAGU--1481 (HAP1 target site)
environmental clone	OM21	--uuaaccg-uuu-ggagaacgac--GUCUAAGGUAGGGCUAGU--gacuggggugaagucguaac--
Prymnesiophyceae	*Emiliania huxleyi*	--cuaacca-uucgggagggcggc--G**C**U**U**AAGGUAGGGCUAGU--gacuggggugaagucguaac--
	Ochrosphaera neapolitana	--cuaacca-uuu-ggaggacgac--GUCUAAGGUAGGGCUAGU--gacuggggugaagucguaac--
	Ochrosphaera sp.	--cuaacca-uuu-ggaggacggc--GUCUAAGGUAGGGCUAGU--gacuggggugaagucguaac--
Cryptophyceae	*Cryptomonas* ϕ[c]	--uuaacca-uuc-ggaggggggc--G**C**CUAAGGUAGGG**U**UAGU--gacuggggugaagucguaac--
	Pyrenomonas salina[c]	--uuaacca-uuu-ggaggggggc--G**C**CUAAGGUAGGG**U**UAGU--gacuggggugaagucguaac--
Rhodophyceae	*Cyanidium caldarium*[c]	--cuccauaacaaugguaguaagc--G**C**CUAAGG**CAAAAU**UAGU--gacuggggugaagucguaac--
	Galdieria sulphuraria	--cuaaccu-uaauggaggaggau--G**C**CUAAGG**C**AGGGCUAGU--gacuggggugaagucguaac--
	Griffithsia pacifica	--u----------gc--**A**CCUAAGGUAGGGCUAGU--gannggggugaagucguaac--
	Antithamnion sp.	--cuauccu-uuuuagaggguuc--**A**CCUAAGGU**U**GGGCUAGU--gacuggggugaagucguaac--
	Palmaria palmata	--cuaaccu-uuu-ggaggggggc--G**C**CUAAGG**C**AGGGCUAGU--gacuggggugaagucguaac--

[a] Underlined letters in bold face indicate mismatches with the CHRYS1 or HAP1 target site.

[b] *P. littoralis* contains two base mismatches with the HAP1 target site.

[c] *Cryptomonas* ϕ, *P. salina*, and *C. caldarium* contain two base mismatches with the CHRYS1 target site.

was [32]P-labeled on its 5' terminus with T4 polynucleotide kinase as previously described (Giovannoni et al. 1988a).

Pinblots were prehybridized in 15 mL Z-hyb buffer (1.0 mM EDTA, 0.50 mM NaH_2PO_4 pH 7.2, 7.0% SDS) for 45-60 min at room temperature. After decanting the pre-hybridization buffer, the blots were hybridized in 6.0 mL Z-hyb buffer containing 200 µL of [32]P-labeled oligonucleotide (25-50 ng of oligonucleotide) for 8-16 h at room temperature. The blots were washed three times for 15 min each in 25 mL wash buffer [0.2X SSPE (150 mM NaCl, 1.0 mM EDTA, 10 mM $NaH_2PO_4.H_2O$ pH 7.4), 0.1% SDS] at room temperature, and one time for 30 min at the probe hybridization stringency temperature (T_H). The T_H for each probe was determined empirically by washing at successively higher temperatures. T_H was 40°C for probe CHRYS1, and T_H was 45°C for probe HAP1 (data not shown). Blots were stripped of probe by washing three times in 25-50 mL wash buffer at 70°C for 10 min each wash. All blots were visualized using a Molecular Dynamics PhosphorImager SI (Molecular Dynamics, Sunnyvale, California) as well as autoradiographically with x-ray film.

2.4 Results and discussion

In the course of sequencing the 5' and 3' domains of clones selected at random from a clone library of 16 rRNA genes, six plastid 16S rRNA gene sequences were fortuitously discovered. Although plastid rDNA sequence databases are fairly extensive, with over 50 sequences currently available (Maidak et al. 1994), they fall far short of fully representing the diversity of cultured phytoplankton. Nonetheless, these databases are currently sufficient

for the identification of plastid 16S rRNA gene sequences to at least division level.

Phylogenetic analyses revealed that the rDNA clones isolated in this study belonged to two groups within the Chromophyta (Fig. 2.1). The first group, with a complete clone sequence from one representative (OM21), was most closely related to the plastid 16S rRNA genes of *Ochrosphaera neapolitana* strain CCMP593 (similarity=0.93) and *Emiliania huxleyi* (Lohmann) Hay and Mohler strain PML92D (similarity=0.94), two prymnesiophytes. The topology of the phylogenetic tree in Figure 2.1 was determined by the neighbor-joining method, but analyses using Wagner parsimony and the maximum likelihood method yielded the same result with respect to the position of OM21 and the Prymnesiophyceae. This relationship was supported by 100% of the bootstrap replicates in both the neighbor-joining and Wagner parsimony trees.

The second group of plastid clones, represented by one complete (OM20) and four partial rDNA clone sequences (OM12, OM13, OM19, and OM22), were most closely related to the Bacillariophyceae (Fig. 2.1, Fig. 2.2). Both the neighbor-joining and Wagner parsimony analyses supported the close association of OM20 with the bacillariophytes *Skeletonema costatum, Skeletonema pseudocostatum* and *Corethron criophilum* (L. Medlin, unpubl. sequence) in 100% of the bootstrap replicates. In all analyses, OM20 also remained closely associated with the 16S rDNA clone AGG56 (DeLong et al. 1993). AGG56 was isolated by PCR amplification and cloning of bacterial 16S rRNA genes from macroscopic phytodetrital material collected in the Santa Barbara Channel. A majority (55%) of the rDNA clones collected from these aggregates belonged to a restriction fragment length polymorphism group represented by the clone AGG56 (DeLong et al. 1993).

Figure 2.1. Phylogenetic tree showing evolutionary relationships of OM20 and OM21 to representative plastid and cyanobacterial 16S rRNA genes. The phylogenetic tree was inferred from evolutionary distance data by the neighbor-joining method. The number of bootstrap replicates which supported the branching order are shown above (neighbor joining) and below (Wagner parsimony) the relevant nodes, out of a total of 100 replicate samplings. Values below 50 are not shown. A total of 1221 bases were used in the analysis. 16S rDNA sequences for all reference plastids were obtained from the Ribosomal Database Project, except *Corethron criophilum* (L. Medlin, unpubl. sequence). The gene sequence of *Bacillus subtilis* was used to root the tree. AGG56 = unidentified plastid rRNA gene cloned from marine snow, SAR6 = uncultured marine prochlorophyte rRNA gene cloned from the Sargasso Sea, pt = plastid.

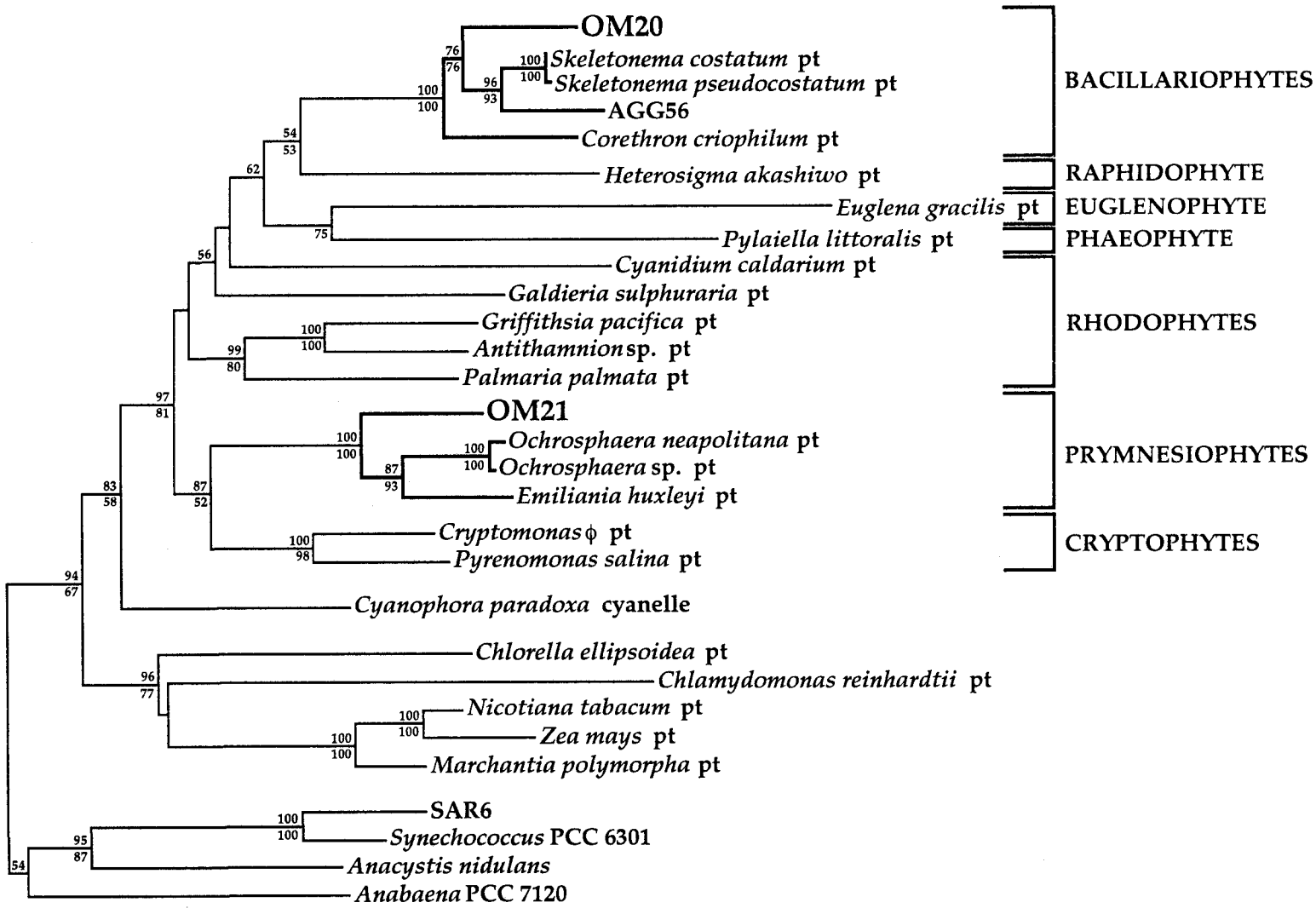

OM20
Skeletonema costatum pt
Skeletonema pseudocostatum pt
AGG56
Corethron criophilum pt
Heterosigma akashiwo pt
Euglena gracilis pt
Pylaiella littoralis pt
Cyanidium caldarium pt
Galdieria sulphuraria pt
Griffithsia pacifica pt
Antithamnion sp. pt
Palmaria palmata pt
OM21
Ochrosphaera neapolitana pt
Ochrosphaera sp. pt
Emiliania huxleyi pt
Cryptomonas ϕ pt
Pyrenomonas salina pt
Cyanophora paradoxa cyanelle
Chlorella ellipsoidea pt
Chlamydomonas reinhardtii pt
Nicotiana tabacum pt
Zea mays pt
Marchantia polymorpha pt
SAR6
Synechococcus PCC 6301
Anacystis nidulans
Anabaena PCC 7120
BACILLARIOPHYTES
RAPHIDOPHYTE
EUGLENOPHYTE
PHAEOPHYTE
RHODOPHYTES
PRYMNESIOPHYTES
CRYPTOPHYTES
0.10

Figure 2.1

Figure 2.2. Phylogenetic tree showing the relationships among five plastid 16S rDNA clones related to the Bacillariophyceae. The tree was inferred by the maximum likelihood method from a total of 477 nucleotide positions. The 16S rRNA gene sequences of the cyanobacteria *Synechococcus* PCC 6301 and SAR6 were used to root the tree.

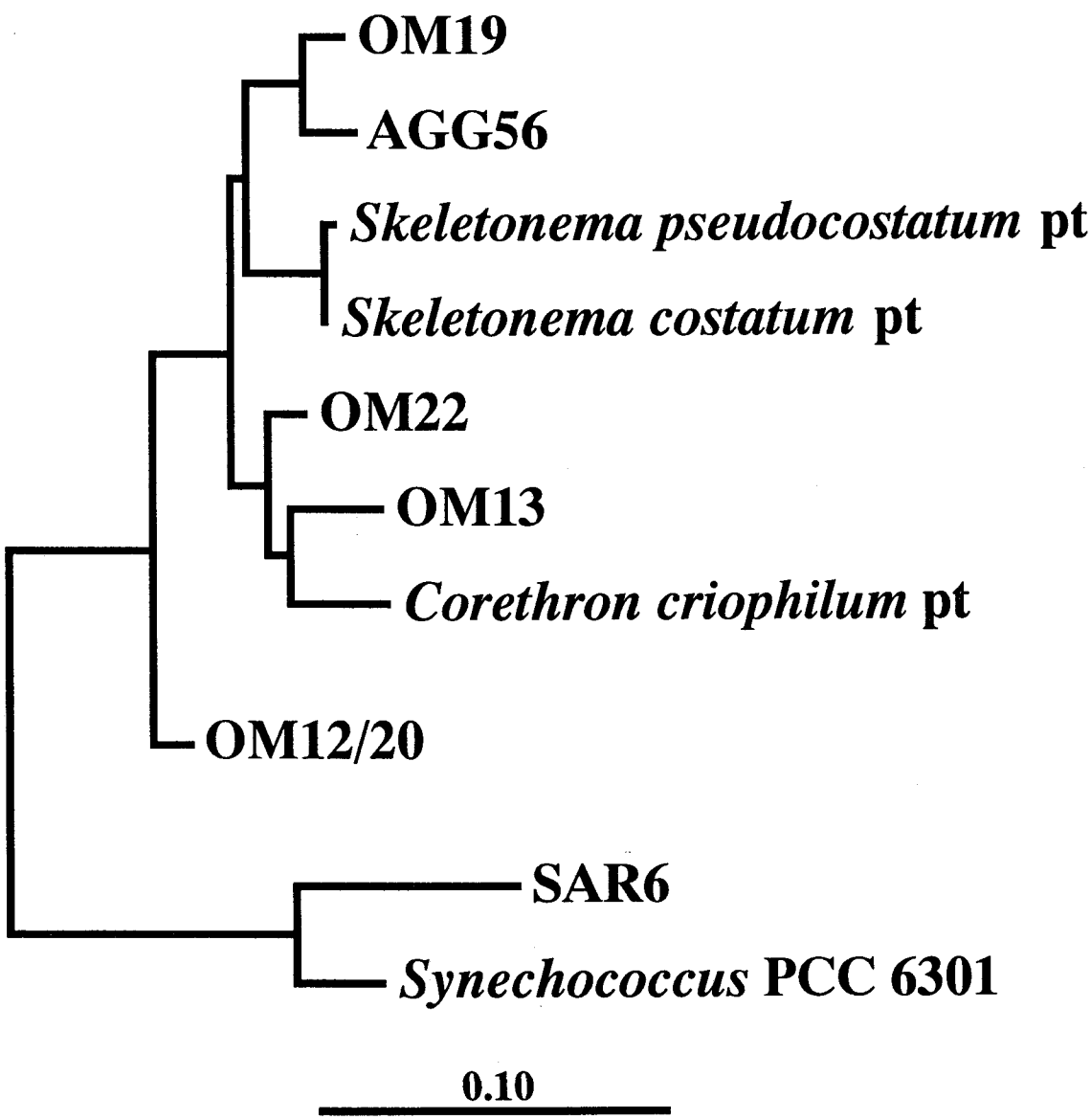

OM19
AGG56
Skeletonema pseudocostatum pt
Skeletonema costatum pt
OM22
OM13
Corethron criophilum pt
OM12/20
SAR6
Synechococcus PCC 6301
0.10

Figure 2.2

A phylogenetic tree was based on partial sequences from the five cloned rDNAs related to the Bacillariophyceae, as well as AGG56 (Fig. 2.2). Two of the six clones were identical (OM20 and OM12), but the rest showed considerable variation. Sequence variation can be artifactual due to sequencing errors, misincorporation of bases in the polymerase chain reaction, and chimeric gene products. These possibilities were investigated with the program CHECK_CHIMERA, and by constructing secondary structural models for each of the partially sequenced rDNA clones.

The heterogeneity between clone types in a variable region of the 16S rRNA molecule (positions 141-222, using the *E. coli* 16S rRNA gene has a reference) was compared (Fig. 2.3). The variation is not random; it is concentrated in a limited number of nucleotide positions and, where expected, is nearly always accompanied by compensatory base changes across helices. It is improbable that this variation is artifactual. This region of the 16S rRNA gene is often used as a target for hybridization of oligonucleotide probes because of its relatively high variability in comparison to other areas of the gene. Probes designed to hybridize to this region of the molecule can be highly specific for the target gene, while failing to hybridize to closely related though not identical rDNAs. Similar sequence variation has been observed in prokaryotic picoplankton rRNA genes (Giovannoni et al. 1990a), but this is the first example of such clustering in plastid rRNA genes. The relationship of the sequence variation (Fig. 2.3) to the divergence of species and population genetic variation is unknown. For the present, the simplest hypothesis is that the variation represents an assemblage of diatom species. Resolution of this question will probably come about through future phylogenetic studies of cultured diatoms.

Figure 2.3. Comparison of the stem and loop structure defined by *Escherichia coli* 16S rRNA nucleotide base positions 141-222 for six rRNA gene clones related to the Bacillariophyceae plastid lineage. Bold nucleotides indicate regions of variability between the secondary structural elements. The nucleotide indicated by the arrow is a "U" residue in OM19 and a "G" residue in AGG56.

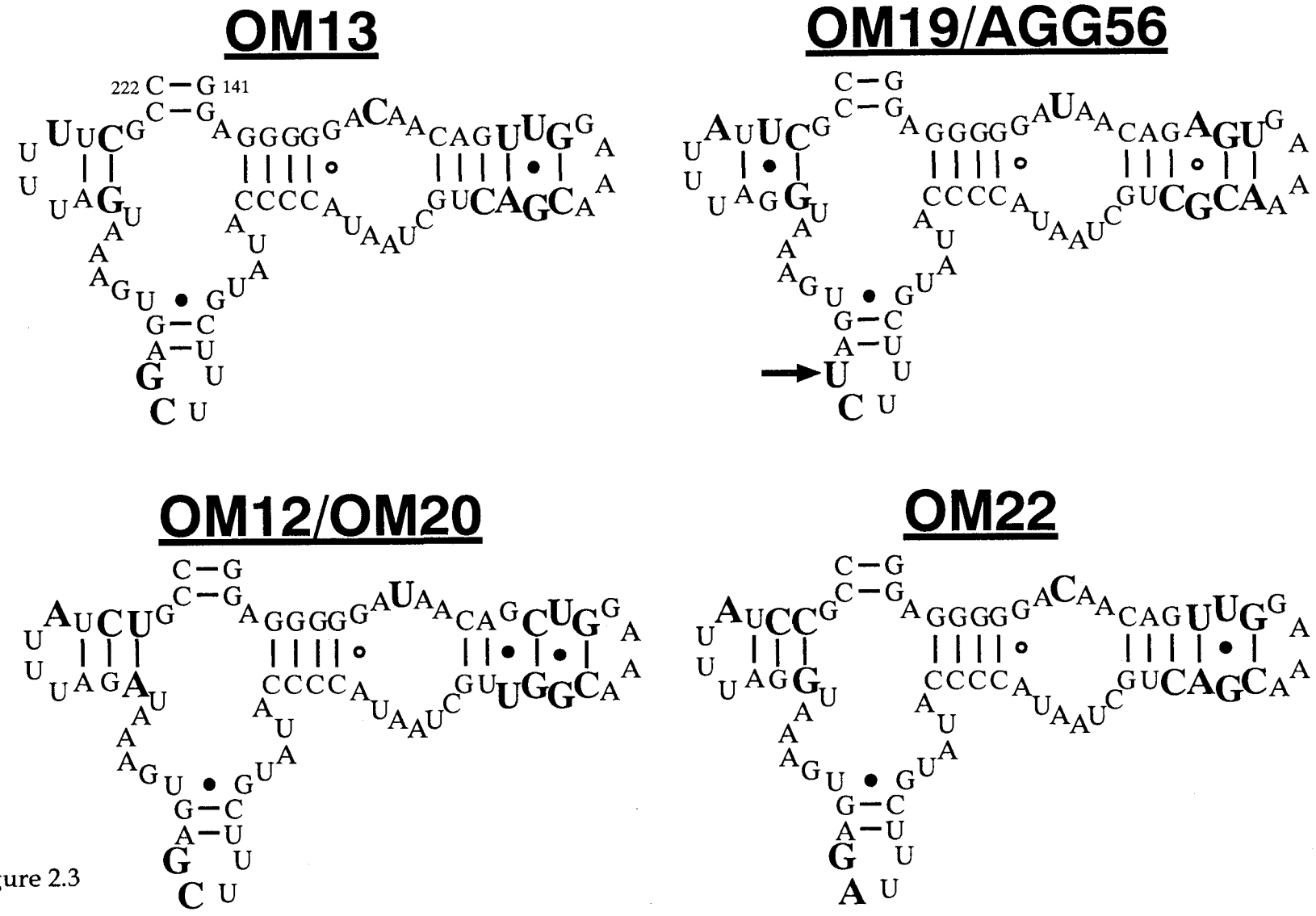


Figure 2.3

To confirm that the predicted 16S rRNA gene products could be folded into consensus secondary structures consistent with the general topological features of other established plastid and bacterial 16S rRNAs, secondary structural models were generated for the two complete plastid clone sequences, OM20 and OM21 (Fig. 2.4; Gutell 1993). No unusual base pairing in conserved helices was observed, which is a possibility if different portions of a gene sequence are derived from different organisms (Kopczynski et al. 1994). The program CHECK_CHIMERA, supplied by the RDP (Maidak et al. 1994), was also used to check for the presence of chimeric gene artifacts. Again, no evidence of chimeric gene sequences was found. Three single-base sequence signatures defined the cyanobacteria and chloroplasts (Fig. 2.4; Woese 1987). The presence of a "U" residue in OM21 and an "A" residue in OM20 at position 799 (*E. coli* 16S rRNA gene numbering), a "C" residue at position 1207, and a "G" residue at position 1233 support the phylogenetic position of these two clones within the cyanobacterial phylum of the bacteria.

Oligonucleotide probes were designed to selectively hybridize to unique regions of the two major chromophyte plastid types found in the clone library (Table 2.1). The probes were used to screen the remaining clones in the library. The probe CHRYS1 was designed to selectively hybridize to clones OM12, OM13, OM19, OM20, and OM22 and the diatoms *Skeletonema costatum, Skeletonema pseudocostatum* and *Corethron criophilum*. HAP1 was designed to selectively hybridize to clone OM21 and the prymnesiophytes *Ochrosphaera neapolitana* and *Ochrosphaera* sp. As cross-hybridization to non-target plastid or bacterial rRNA genes would give misleading results, probe specificity was tested by comparing the probe sequences in the RDP collection utilizing the program CHECK_PROBE (Maidak et al. 1994). Radioactive CHRYS1 and

Figure 2.4. Proposed secondary structural models of clones OM20 and OM21 16S rRNA. Arrows indicate signature nucleotides of the cyanobacterial and choroplast 16S rRNA gene lineage. Numbers refer to nucleotide positions in the *E. coli* 16S rRNA gene. Target sites for the probes CHRYS1, HAP1, and the primers 27F and 1518R are indicated.

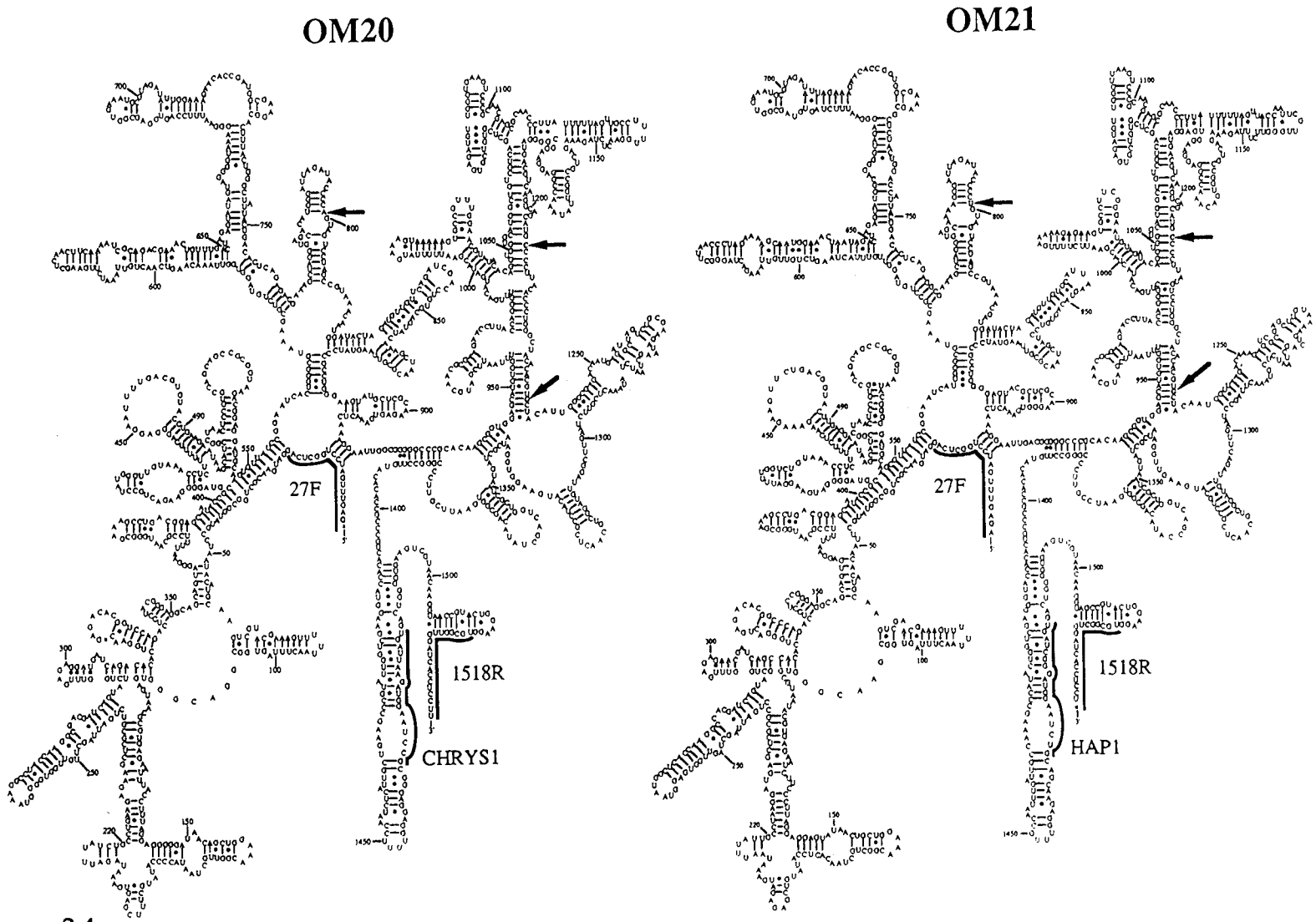


Figure 2.4

HAP1 probes were hybridized to high density colony blots from the clone library to determine the frequency of these clone types (Fig. 2.5). Of 170 clones screened by colony hybridization, 19 hybridized to CHRYS1 (11%), and 24 hybridized to HAP1 (14%). These probes were also hybridized to planktonic RNAs collected from surface seawater at 13 stations on the continental shelf off Cape Hatteras, North Carolina, in March, 1994 (data not shown). Though no obvious pattern could be discerned from the hydrographic data, levels significantly above background were detected for two of 13 samples hybridized to the HAP1 probe.

Phylogenetic probes such as those described here have potential for determining phytoplankton distributions in bulk samples, for measuring the growth status of plankton, and for *in situ* identification of single cells (DeLong et al. 1989, Lim et al. 1993). In phytoplankton ecology, fluorescence *in situ* hybridization (FISH) in conjunction with fluorescently activated cell sorting machines offers possibilitites for evaluating the significance of genetic variation and speciation in phytoplankton populations (Simon et al. 1995). Plastid genes are not subject to sexual recombination, and thus may offer advantages over nuclear genes for some studies of phytoplankton biogeography and speciation, particularly when the divergence of subpopulations is in question.

In the construction of this clone library, the methods used to collect picoplankton and isolate genomic DNA were optimized for prokaryotic organisms. Therefore, phytoplankton species which are sensitive to filtration and lyse on contact with membranes could be selectively overlooked by this approach. Furthermore, the approach used to extract nucleic acids may not recover DNA from some algal species due to interference of algal walls and

Figure 2.5. Hybridization of radiolabelled 16S rRNA oligonucleotide probes CHRYS1 and HAP1 to three colony blots of 170 uncharacterized (columns A through I) and control rDNA clones (columns J through L). The three autoradiograms on the right show the hybridization of the 338R bacterial probe to the same colony blots of uncharacterized and control rDNA clones. Blot 1 = clones 1-105, Blot 2 = clones 106-198, Blot 3 = clones 200-303. The control rDNAs are: JKL1 = OM20, Bacillariophyceae; JKL2 = OM21, Prymnesiophyceae; JKL3 = OM11, SAR86 cluster (gamma-proteobacteria); JKL4 = OCS12, SAR11 cluster (alpha-proteobacteria); JKL5 = SAR420, SAR83 cluster (alpha-proteobacteria); JKL6 = SAR423, marine synechococcus; JKL7 = OM239, prochlorophyte.

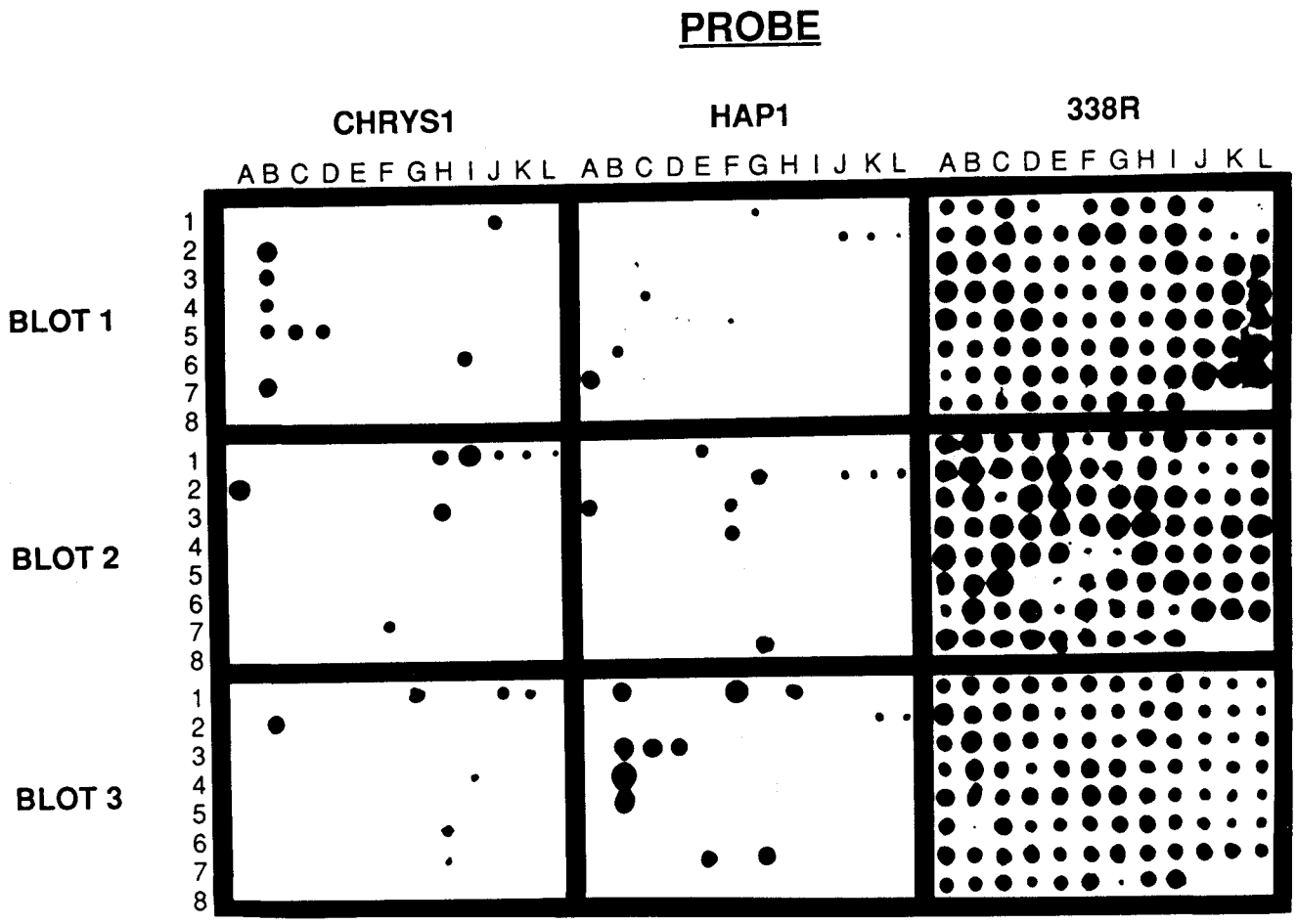

PROBE
CHRYS1
HAP1
338R
A B C D E F G H I J K L
BLOT 1
BLOT 2
BLOT 3
1 2 3 4 5 6 7 8

Figure 2.5

extracellular materials (Coleman and Goff 1991). In addition, the multiplicity of chloroplasts and chloroplast genomes within individual phytoplankton cells makes it difficult to draw inferences from the abundance of plastid clones in the library. *Heterosigma akashiwo* contains approximately 650 chloroplast genomes per cell (Ersland et al. 1981), and many other chromophyte algae contain multiple chloroplast genomes as well. Thus, the results presented here cannot be interpreted as a measurement of the abundance of phytoplankton in this water sample.

The data presented here provide preliminary information that, when gene sequences are known for more chromophytes, might lead to a general genetic method for the identification of eukaryotic picoplankton. Whether this method will offer important advantages over conventional approaches or genetic approaches using other gene markers has yet to be determined. Regardless, these results 1) demonstrate that plastid rRNA genes may be recovered from natural plankton population genomic DNA, 2) demonstrate heterogeneous clusters of clone types similar to those observed previously in prokaryotic organisms, and 3) provide sequence data which can be used for the design of plastid-specific rDNA probes.

2.5 References

Brand, L. E. 1986. Nutrition and culture of autotrophic ultraplankton and picoplankton. *In* Platt, T. & Li W. K. W. [Eds.] *Photosynthetic Picoplankton. Can. Bull. Fish. Aquat. Sci.* 214: 205-33.

Britschgi, T. B. & Giovannoni, S. J. 1991. Phylogenetic analysis of a natural marine bacterioplankton population by rRNA gene cloning and sequencing. *Appl. Environ. Microbiol.* 57:1707-13.

Campbell, L., Shapiro, L. P. & Haugen, E. 1994. Immunochemical characterization of eukaryotic ultraplankton from the Atlantic and Pacific oceans. *J. Plankton Res.* 16:35-51.

Cattolico, R. A. & Loiseaux-de Goër, S. 1989. Analysis of chloroplast evolution and phylogeny: a molecular approach. *In* Green, J. C., Leadbeater, B. S. C. & Diver, W. L [Eds.] *The Chromophyte Algae: Problems and Perspectives.* Oxford, Clarendon Press, pp. 85-100.

Coleman, A. W. & Goff, L. J. 1991. DNA analysis of eukaryotic algal species. *J. Phycol.* 27:463-73.

Delaney, T. P. & Cattolico, R. A. 1989. Chloroplast ribosomal DNA organization in the chromophytic alga *Olisthodiscus luteus. Curr. Genet.* 15:221-9.

DeLong, E. F. 1992. Archaea in coastal marine environments. *Proc. Natl. Acad. Sci. USA* 89:5685-9.

DeLong, E. F., Franks, D. G. & Alldredge, A. L. 1993. Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. *Limnol. Oceanogr.* 38:924-34.

DeLong, E. F., Wickhan, G. S. & Pace, N. R. 1989. Phylogenetic stains: ribosomal RNA-based probes for the identification of single cells. *Science (Wash. D.C.)* 243:1360-3.

Douglas, S. E. & Turner, S. 1991. Molecular evidence for the origin of plastids from a cyanobacterium-like ancestor. *J. Mol. Evol.* 33:267-73.

Ersland, D. R., Aldrich, J. & Cattolico, R. A. 1981. Kinetic complexity, homogeneity, and copy number of chloroplast DNA from the marine alga *Olisthodiscus luteus. Plant Physiol.* 68:1468-73.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783-91.

Fox, G. E., Stackebrandt, E., Hespell, R. B., Gibson, J., Maniloff, J., Dyer, T. A., Wolfe, R. S., Balch, W. E., Tanner, R. S., Magrum, L. J., Zablen, L. B., Blakemore, R., Gupta, R., Bonen, L., Lewis, B. J., Stahl, D. A., Luehrsen, K. R., Chen, K. N. & Woese, C. R. 1980. The phylogeny of prokaryotes. *Science (Wash. D.C.)* 209:457-63.

Fuhrman, J. A., McCallum, K. & Davis, A. A. 1993. Phylogenetic diversity of subsurface marine microbial communities from the atlantic and pacific oceans. *Appl. Environ. Microbiol.* 59:1294-1302.

Giovannoni, S. J. 1991. The polymerase chain reaction. *Sequencing and Hybridization Techniques in Bacterial Systematics.* New York, John Wiley & Sons, pp. 177-201.

Giovannoni, S. J., Britschgi, T. B., Moyer, C. L. & Field, K. G. 1990a. Genetic diversity in Sargasso Sea bacterioplankton. *Nature (Lond.)* 345:60-3.

Giovannoni, S. J., Delong, E. F., Olsen, G. J. & Pace, N. R. 1988a. Phylogenetic group-specific oligodeoxynucleotide probes for the identification of single microbial cells. *J. Bacteriol.* 170:720-6.

Giovannoni, S. J., DeLong, E. F., Schmidt, T. M. & Pace, N. R. 1990b. Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. *Appl. Environ. Microbiol.* 56:2572-5.

Giovannoni, S. J., Turner, S., Olsen, G. J., Barns, S., Lane, D. J. & Pace, N. R. 1988b. Evolutionary relationships among cyanobacteria and green chloroplasts. *J. Bacteriol.* 170:3584-92.

Gutell, R. R. 1993. Collection of small subunit (16S- and 16S-like) ribosomal RNA structures. *Nucleic Acids Res.* 21:3051-4.

Hara, Y. & Chihara, M. 1987. Morphology, ultrastructure and taxonomy of the Raphidophycean alga *Heterosigma akashiwo. Bot. Mag. Tokyo* 100:151-63.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16:111-20.

Kopczynski, E. D., Bateson, M. M. & Ward, D. M. 1994. Recognition of chimeric small-subunit ribosomal DNAs composed of genes from uncultivated microorganisms. *Appl. Environ. Microbiol.* 60:746-8.

Lim, E. L., Amaral, L. A., Caron, D. A. & DeLong, E. F. 1993. Application of rRNA-based probes for observing marine nanoplanktonic protists. *Appl. Environ. Microbiol.* 59:1647-55.

Maidak, B. L., Larsen, N., McCaughey, M. J., Overbeek, R., Olsen, G. J., Fogel, K., Blandy, J. & Woese, C. R. 1994. The Ribosomal Database Project. *Nucleic Acids Res.* 22:3485-7.

Markowicz, Y. & Loiseaux-de Goër, S. 1991. Plastid genomes of the Rhodophyta and Chromophyta constitute a distinct lineage which differs from that of the Chlorophyta and have a composite phylogenetic origin, perhaps like that of the Euglenophyta. *Curr. Genet.* 20:427-30.

Marshall, H. G. & Ranasinghe, J. A. 1989. Phytoplankton distribution along the eastern coast of the U.S.A.-VII. Mean cell concentrations and standing crop. *Continental Shelf Res.* 9:153-64.

Mullins, T. D., Britschgi, T. D., Krest, R. L. & Giovannoni, S. J. 1995. Genetic comparisons reveal the same unknown bacterial lineages in Atlantic and Pacific bacterioplankton communities. *Limnol. Oceanogr.* 40:148-58.

Reysenbach, A. N., Wickham, G. S. & Pace, N. R. 1994. Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus Spring, Yellowstone National Park. *Appl. Environ. Microbiol.* 60:2113-9.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science (Wash. D.C.)* 239:487-91.

Saitou, N. & Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406-25.

Simon, N., LeBot, N., Marie, D., Partensky, F. & Vaulot, D. 1995. Fluorescent *in situ* hybridization with rRNA-targeted oligonucleotide probes to identify small phytoplankton by flow cytometry. *Appl. Environ. Microbiol.* 61:2506-13.

Turner, S., Burger-Wiersma, T., Giovannoni, S. J., Mur, L. R. & Pace, N. R. 1989. The relationship of a prochlorophyte *Prochlorothrix hollandica* to green chloroplasts. *Nature (Lond.)* 337:380-2.

Urbach, E., Robertson, D. L. & Chisholm, S. W. 1992. Multiple evolutionary origins of prochlorophytes within the cyanobacterial radiation. *Nature (Lond.)* 355:267-70.

Woese, C. R. 1987. Bacterial evolution. *Microbiol. Rev.* 51:221-71.

CHAPTER 3

PHYLOGENETIC DIVERSITY OF MARINE COASTAL PICOPLANKTON 16S rRNA GENES CLONED FROM THE CONTINENTAL SHELF OFF CAPE HATTERAS, NORTH CAROLINA

Michael S. Rappé, Paul F. Kemp, and Stephen J. Giovannoni

Submitted to *Limnology and Oceanography,*
American Society of Limnology and Oceanography,
Waco, Texas, in press

3.1 Abstract

The phylogenetic diversity of a continental shelf picoplankton community was examined by analyzing 16S ribosomal RNA (rRNA) genes amplified from environmental DNA with bacterial-specific primers and the polymerase chain reaction (PCR). Picoplankton populations collected from the pycnocline (10 m) over the eastern continental shelf of the United States near Cape Hatteras, NC, served as the source of bulk nucleic acids used in this study. A large proportion of the 169 rDNA clones recovered (33%) were related to plastid 16S rRNA genes, including plastids from both chromophyte and chlorophyte algae. A majority of bacterial gene clones (75% of bacterial clones, 50% of the total) were closely related to rRNA gene lineages that had been discovered previously in clone libraries from open-ocean marine habitats, including the SAR86 cluster (γ Proteobacteria), SAR83, SAR11, and SAR116 clusters (all α Proteobacteria), and the marine Gram-positive cluster (high G + C Gram-positive). Most of the remaining bacterial clones recovered were phylogenetically related to the γ and β subclasses of the Proteobacteria, including an rDNA lineage within the Type I methylotroph clade of the β subclass. The abundance of plastid rDNAs and the lack of cyanobacterial-related clones, as well as the presence of β Proteobacteria, are features of this coastal picoplankton gene clone library which distinguish it from similar studies of oligotrophic open ocean sites. Overall, however, these data indicate that a limited number of as yet uncultured bacterioplankton lineages, related to those previously observed in the open ocean, can account for the majority of cells in this coastal marine bacterioplankton assemblage.

3.2 Introduction

Information on marine microbial diversity has often been derived from enrichment culture studies (e.g. Baumann et al. 1972). In recognition of the biases implicit in microbiological cultivation techniques, there has been a marked shift to reliance on the cloning and sequencing of 16S ribosomal RNA genes directly from naturally-occurring microbial assemblages as a means of assessing microbial diversity (Olsen et al. 1986; Ward et al. 1992; Giovannoni et al. 1995). Although biases associated with the molecular methods applied to questions in microbial ecology are not yet well understood, they appear to be less limiting than those associated with culture-based methods (Ward et al. 1992; Giovannoni et al. 1995).

Pelagic ocean habitats were some of the first environments to be studied with rRNA gene cloning approaches. This choice was not surprising, since ocean systems are both biogeochemically significant and technically tractable for molecular studies (Giovannoni et al. 1990a; Schmidt et al. 1991). Most such studies have characterized bacterioplankton populations from oligotrophic regions of the open ocean, including samples from the surface of the Sargasso Sea in the Atlantic Ocean (Giovannoni et al. 1990a; Fuhrman et al. 1993), the surface of the central Pacific Ocean near Hawaii (Schmidt et al. 1991), and at depths of 100 and 500 meters in the western California Current (Fuhrman et al. 1993). In addition, PCR-amplified 16S rDNAs were analyzed to compare phytodetrital marine snow-associated bacteria with free-living bacterioplankton collected in seawater from the Santa Barbara Channel off the California coast (DeLong et al. 1993). Two general conclusions have been drawn from these studies: 1) the vast majority of 16S rDNAs retrieved from natural, mixed-population bacterioplankton samples do not correspond to the

rRNA gene sequences obtained from cultured marine bacteria (Giovannoni et al. 1995); and 2) though phylogenetically diverse, the majority of rDNA clones obtained from oligotrophic open-ocean bacterioplankton communities fall into one of several distinct phylogenetic groups (Fuhrman et al. 1993; Mullins et al. 1995; Giovannoni et al. 1996a).

Current trends in oceanographic studies emphasize models that treat bacterioplankton as a compartment through which elements and energy flux (reviewed in Ducklow 1994). Typically, such studies measure bacterial processes as community averages, and do not characterize the physiological activity of individual species. These models have great utility, and have served well the need to generally establish the relationship between bacterioplankton and other major elements in marine food webs. With the advent of molecular ecology, it has been recognized that "bacterioplankton" are dominated by a relatively limited subset of phylogenetic groups that are widely distributed and often exhibit clear trends in their vertical distributions in the water column (Mullins et al. 1995; Giovannoni et al. 1996b; Gordon and Giovannoni 1996; Field et al. in press). Such data indicate that bacterioplankton communities are relatively structured, but the influence of community structure and dynamics on the results of system level studies has yet to be established. It seems reasonable to suppose that understanding community dynamics could contribute to the development of a more accurate overview of ecosystem functions.

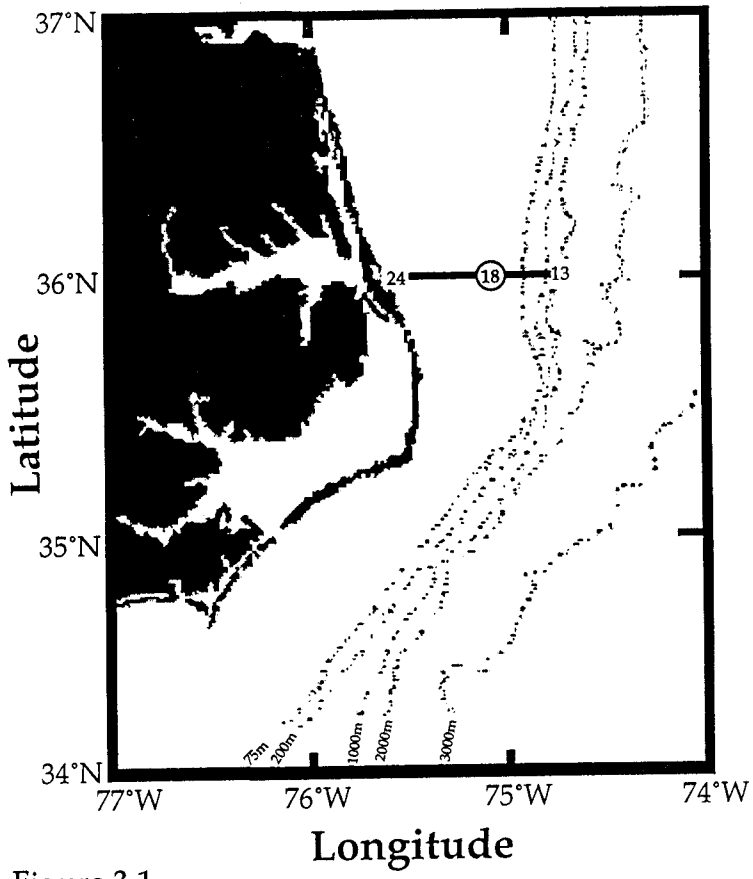
The study we describe here is part of the U.S. Department of Energy's Ocean Margins Program. This program is focused on the eastern continental shelf of the United States near Cape Hatteras, North Carolina; its goal is to study the storage, flux, and transformations of carbon in ocean margin

systems. Although continental shelves account for only a small percentage of the total surface area of the worlds oceans, they may provide a significant sink for atmospheric carbon dioxide and a major pathway for the export of organic matter into deep ocean basins (Walsh et al. 1985; Rowe et al. 1986; Walsh 1991). As a first step in analyzing the role of bacterioplankton community structure on organic carbon transformations in this region, we investigated the phylogenetic diversity of a continental shelf picoplankton community by an analysis of PCR-amplified 16S rRNA genes. Nearly a third of the 169 clones recovered in this library were derived from eukaryotic phytoplankton plastid rRNA genes. In addition to genes from previously identified uncultured bacterioplankton, this study uncovered members of the β subclass of the Proteobacteria that had not been observed previously in 16S rDNA clone libraries constructed from seawater, including a unique lineage within the Type I methylotroph clade.

3.3 Materials and methods

rRNA gene amplification and clone library construction. The construction of this clone library was described previously (Rappé et al. 1995). On 13 May 1993, a picoplankton sample was collected on 0.2 μm filters by filtration of unfiltered seawater taken from a depth of 10 m at station 18 (35°59'N, 75°08'W) of R.V. Gyre cruise 93-G6 (Fig. 3.1). The sonic depth at station 18 was 35 m. Total cellular nucleic acids were isolated from the picoplankton sample by cell lysis with proteinase K and SDS and phenol/chloroform extraction (Giovannoni et al. 1990b). Bacterial and plastid 16S rRNA genes were amplified from the environmental sample of genomic

Figure 3.1. Hydrography of station 18 of R.V. Gyre cruise 93-G6. The station is located approximately 80 km northeast of Cape Hatteras, North Carolina. The panel on the right displays a cross-shelf contour plot of temperature for the transect indicated in the map. Picoplankton samples recovered from a depth of 10 m at station 18 were used as the source of bulk nucleic acids used for study.


37°N
36°N
35°N
34°N
77°W
76°W
75°W
74°W
Latitude
Longitude
24
18
13
75m
200m
1000m
2000m
3000m

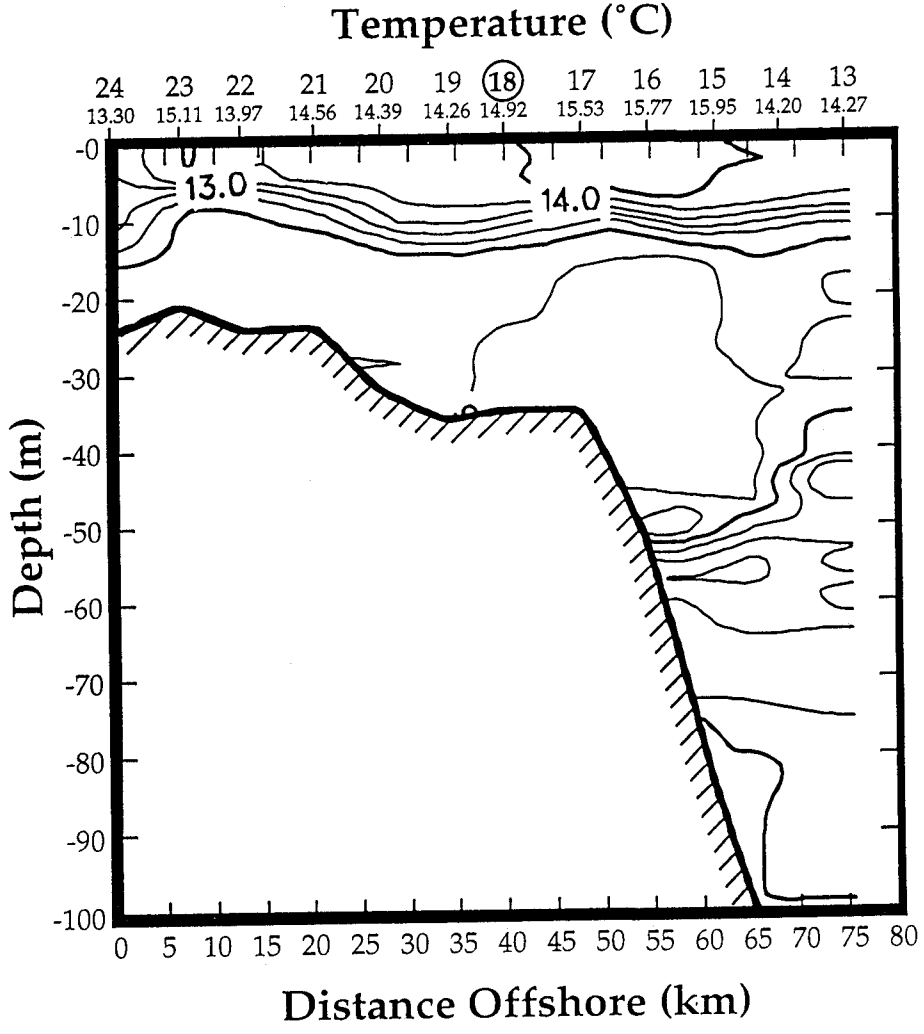

Temperature (°C)
24 23 22 21 20 19 18 17 16 15 14 13
13.30 15.11 13.97 14.56 14.39 14.26 14.92 15.53 15.77 15.95 14.20 14.27
13.0
14.0
-0
-10
-20
-30
-40
-50
-60
-70
-80
-90
-100
Depth (m)
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80
Distance Offshore (km)

Figure 3.1

DNAs using the two general bacterial 16S rRNA primers 27F and 1522R (Giovannoni 1991) and the polymerase chain reaction (Saiki et al. 1988). The blunt-ended rDNA amplification product was inserted into the phagemid vector pBluescript KSII- (Stratagene, La Jolla, California), and the product of this ligation used to transform competent *E. coli* cells. Positive (white colony) transformants were streaked for isolation and stored in stab cultures. Clones in this library were assigned the prefix "OM" (Ocean Margins) and numbered discontinuously from 1-303.

rRNA gene sequencing and phylogenetic analysis. Template plasmid DNAs for sequencing were prepared from overnight cultures of positive transformants by a general alkaline lysis miniprep protocol using commercially available plasmid purification kits, quantified spectrophotometrically, and visualized by agarose gel electrophoresis. A pBluescript plasmid known to contain the correct sized insert of 1.5 kilobases (Kb) was used as a positive size standard while one lacking an insert, isolated from a blue colony, served as a negative control. Plasmid DNAs from clones containing full-length inserts were sequenced (partially or in full) on both strands using an ABI model 373A automated sequencer (Applied Biosystems, Foster City, California), dye-terminator chemistry, and conserved rRNA gene primers (Lane 1991).

DNA sequence data from cloned 16S rDNA genes were manually aligned to bacterial and plastid sequences obtained from the Ribosomal Database Project (RDP, Maidak et al. 1994) and Genbank, using the Genetic Data Environment (GDE) v2.0 sequence analysis software package. Regions of ambiguous alignment were excluded from the analyses. The Phylogeny Inference Package (PHYLIP v3.5; Felsenstein 1989) was used for all

phylogenetic analyses except for maximum likelihood, which used the program fastDNAml distributed by the RDP (Maidak et al. 1994; Olsen et al. 1994). Evolutionary distances were calculated using the program DNADIST and the Kimura 2-parameter model for nucleotide change, with a transition/transversion ratio of 2.0 (Kimura 1980). The program NEIGHBOR was used to reconstruct phylogenetic trees from evolutionary distance data by the neighbor-joining method (Saitou and Nei 1987). Parsimony trees were reconstructed using the program DNAPARS. A total of 100 bootstrapped replicate resampling data sets for DNADIST were generated using SEQBOOT with random sequence addition and global rearrangement (Felsenstein 1985). All phylogenetic trees were edited using the program Treetool, available through the RDP. Secondary structural models were constructed using the program gRNAid v1.4, and were manipulated manually for optimal agreement with published 16S rRNA secondary structures (Gutell 1994). All gene sequences were submitted to the RDP program CHECK_CHIMERA to aid in the detection of chimeric gene artifacts (Maidak et al. 1994). Raw sequence similarities were calculated without distance correction using the program DNADIST, available with PHYLIP v.3.5. For closely related taxa, all positions in the alignment for which both sequences contained a nucleotide were used in the similarity calculation. For more distantly related taxa, hypervariable regions for which alignment was uncertain where excluded from the similarity calculations. Similarities were calculated from partial sequences by considering all available overlapping regions.

Oligonucleotide probes. Taxon-specific oligonucleotide probes were used to partially screen the OM clone library (Table 3.1). Three oligonucleotide probes (CHRYS1, HAP1, SAR86R) were designed from

Table 3.1. Oligonucleotide probes used in this study.

Table 3.1

Name	Probe sequence (5′ - 3′)	Target position*	Specificity	T_H†	Reference
CHRYS1	ACTAATTCTACCTTAGGC	1464-1481	plastid of Bacillariophyceae	40 °C	(Rappé et al. 1995)
HAP1	ACTAGCCCTACCTTAGAC	1464-1481	plastid of Haptophyceae	45 °C	(Rappé et al. 1995)
SAR11A1	AAGCTTTCTCCGTAAAGACTTAT	179-209	portion of SAR11 cluster (α-Proteo.)	37 °C	(Field et al. in press)
SAR83R	CTAATCTAACGCGGGCC	231-247	SAR83 cluster and related α-Proteo.	45 °C	unpubl. data
SAR86R	CTCCCGAAGGAGGTATA	178-194	SAR86 cluster (γ-Proteo.)	45 °C	this study
338R	GCTGCCTCCCGTAGGAGT	338-355	universal for Bacteria	45 °C	(Amann et al. 1990)

**E. coli* numbering as a reference.

†Empirically determined probe stringency hybridization temperature.

preliminary sequence data obtained from the OM clone library. The design and construction of CHRYS1 and HAP1 has been described previously (Rappé et al. 1995). Probe SAR86R was designed to complement the clones OM10, OM11, OM23, OM24, and OM94 at positions corresponding to *E. coli* 16S rRNA positions 178-194 (17 bases, Table 3.2; Brosius et al. 1978). The probes were screened for possible cross-hybridization to unrelated clones by submitting the sequences to the program CHECK_PROBE, available through the RDP, for comparison to sequences in the RDP (Maidak et al. 1996). The probe 338R, universal for the majority of bacterial 16S rDNAs (Amann et al. 1990), was used to check for the presence of rDNA clone inserts in the recombinant plasmids. All probes used in this study were constructed with an Applied Biosystems automated DNA synthesizer and were [^{32}P]-labeled on their 5' termini with T4 polynucleotide kinase as previously described (Giovannoni et al. 1988a).

Colony blot construction and probe hybridization. Colony blots (pinblots) were prepared as previously described (Rappé et al. 1995). Briefly, a Kleinschmidt 96A-PR microtiter plate replicator (M. Kleinschmidt, Palm Springs, California) was used to inoculate membranes overlaid on nutrient agar plates from 96 well microtiter plates containing positive transformants. Inoculated membranes were incubated at 37°C for 6 h on the agar plates. The membranes were then placed in a cell lysis buffer, microwaved, and submerged in a proteinase K solution. Finally, the membranes were baked and UV-crosslinked to fix the DNA to the membrane.

Pinblots were prehybridized in 15 mL Z-hyb buffer (1.0 mM EDTA, 0.50 mM NaH_2PO_4 pH 7.2, 7.0% SDS) for 45-60 min at room temperature. After decanting the pre-hybridization buffer, the blots were hybridized in 6.0 mL Z-

Table 3.2. Alignment of the SAR86R probe target site with the 16S rRNA target site regions from representative Proteobacteria.

Table 3.2

Bacteria/rDNA clone	16S rRNA sequence* (5′-3′, *E. coli* numbering)	# of mismatches
SAR86 CLUSTER	178--TAT ACC TCC TTC GGG AG--194 (SAR86R)†	
OM10	....ttaataccg--TAT ACC TCC TTC GGG AG--aaagaaggc....	0
OM11	....ttaataccg--TAT ACC TCC TTC GGG AG--aaagaaggc....	0
OM23	....ttaataccg--TAT ACC TCC TTC GGG AG--aaagaaggc....	0
OM24	....ttaataccg--TAT ACC TCC TTC GGG AG--aaagaaggc....	0
OM94	....ttaataccg--TAT ACC TCC TTC GGG AG--aaagaaggc....	0
SAR86	....ttaataccg--TAT ACC TCc TTC GGG AG--aaagagggc....	0
MISC. PROTEOBACTERIA		
Bartonella bacilliformis	....ctaatacag--TAT ACG TCC TTC GGG AG--aaagattta....	1
Rhizobium haukuii	....ctaataccg--TAT ACG TCC TTC GGG AG--aaagattta....	1
Phyllobacterium myrsinacearum	....ctaataccg--TAT ACG TCC TTC GGG AG--aaagattta....	1
Thiobacillus sp. str. THI051	....ctaataccg--TAT ACG TCC TTC GGG AG--aaagattta....	1
Rhizobium sp. str. UPM-Ca7	....ctaataccg--TAT ACG TCC TTC GGG AG--aaagattta....	1
Rhizobium loti	....ctaataccg--TAT ACG TCC TTC GGG AG--aaagattta....	1
Rhodomicrobium vannielii	....ctaataccg--TAT ATC TCC TCC GGG AG--aaagattta....	2
Rochalimaea henselae	....ctaataccg--TAT ACG TCC TTA GGG AG--aaagattta....	2
Bartonella vinsonii	....ctaatacag--TAT ACG TCC TTA GGG AG--aaagattta....	2
Bartonella doshiae	....ctaatacag--TAT ACG TCC TTA GGG AG--aaagattta....	2
Marinobacter hydrocarbonoclasticus	....ttaataccg--CAT ACG TCC TAC GGG AG--aaagcaggg....	3
Azospirillum sp. str. DSM1727	....ctaataccg--CAT ACG TCC TAT GGG AG--aaagcaggg....	4
Pseudomonas aeruginosa	....ctaataccg--CAT ACG TCC TGA GGG AG--aaaggggg....	4
MISC. OM CLONES		
OM42	....gtaatacct--TAT ACG CCC TTC GGG GG--aaagattta....	3
OM65	....gtaatacct--TAT ACG CCC TTC GGG GG--aaagattta....	3
OM93	....ctaataccg--CAT ACG TCC TAC GGG AG--aaagcaggg....	3
OM25	....ctaataccg--GAT ACG CCC TTC GGG GG--aaaggtccg....	4
OM38	....ctaataccg--TAT ACG CCC TCC GGG GG--aaaggcttg....	4

*Underlined nucleotides indicate mismatches with the SAR86R target site.

†Sequence corresponds to the SAR86R target site.

hyb buffer containing 200 µL of [^{32}P]-labeled oligonucleotide (25-50 ng of oligonucleotide) for 8-16 h at room temperature. The blots were washed three times for 15 min each in 25 mL wash buffer [0.2X SSPE (150 mM NaCl, 1.0 mM EDTA, 10 mM $NaH_2PO_4 \bullet H_2O$ pH 7.4), 0.1% SDS] at room temperature, and one time for 30 min at the probe hybridization stringency temperature (T_H). The T_H for each probe was determined empirically by washing at successively higher temperatures. Blots were stripped of probe by washing three times in 25-50 mL wash buffer at 70°C for 10 min each wash. All blots were visualized using a Molecular Dynamics PhosphorImager SI (Molecular Dynamics, Sunnyvale, California) as well as autoradiographically with X-ray film.

Accession numbers. Nucleotide sequences were filed in Genbank under the following accession numbers: OM1, U70710; OM5, U70715; OM10, U70693; OM23, U70694; OM25, U70678; OM27, U70713; OM38, U70679; OM39, U70716; OM42, U70680; OM43, U70704; OM55, U70681; OM58, U70705; OM59, U70695; OM60, U70696; OM65, U70682; OM75, U70683; OM81, U70717; OM93, U70697; OM110, U70714; OM111, U70718; OM125, U70719; OM133, U70698; OM136, U70684; OM143, U70685; OM153, U70720; OM155, U70686; OM156, U70706; OM164, U70721; OM180, U70707; OM182, U70699; OM185, U70700; OM188, U70687; OM190, U70712; OM231, U70711; OM233, U70701; OM239, U70688; OM241, U70702; OM242, U70689; OM247, U70690; OM252, U70703; OM255, U70722; OM258, U70691; OM270, U70723; OM271, U70708; OM273, U70709; OM283, U70724; OM299, U70692. Identical clone sequences were not filed.

Genbank accession numbers were obtained previously for the following clones (Rappé et al. 1995): OM13, U32667; OM19, U32668; OM20, U32670; OM21, U32671; OM22, U32669.

3.4 Results

Hydrography and biological data. Hydrographic data from R.V. Gyre cruise 93-G6 has been compiled by Verity et al. (1996). At the time of sampling from station 18, the water column was stratified, with warmer, fresher water overriding relic, winter bottom "cold pool" water (Fig. 3.1). This is a common seasonal stratification pattern in this region (e.g. Burrage and Garvine 1988; Chapman and Gawarkiewicz 1993). CTD profiles indicated that the warmer, fresher water was apparently Chesapeake Bay outflow (Verity et al. 1996). The sample used in this study was collected at the pycnocline, and thus was a mixture of Chesapeake Bay outflow and "cold pool" water (Fig. 3.1). Water temperature was 13.7°C and salinity was 30.8-31.0 ‰ at a depth of 10 m below station 18. More detailed discussions of the general hydrography of this region are available (e.g. Sherman et al. 1988, Walsh 1988).

The OM clone library. The 169 clones which constitute the Ocean Margins 16S rDNA clone library described here are those that either hybridized to a taxon-specific oligonucleotide probe or contained a full length (1.5 Kb) rDNA insert. Of 96 clones for which sequence data was obtained, 52 contained unique sequence data. Some clones with lengths of less than ca. 1.5 Kb apparently contained fragments of rRNA genes, which are commonly observed in clone libraries of this type. An analysis of the two rDNA clone groups most closely related to the plastid 16S rRNA genes of the prymnesiophyte *Emiliania huxleyi* and the bacillariophyte *Skeletonema costatum* was described previously (Rappé et al. 1995). Information on an additional 126 clones is provided here.

Phylogenetic diversity of Ocean Margins gene clones. α Proteobacteria. Of 51 OM gene clones related to the α Proteobacteria, 50 were associated with

three previously described 16S rDNA gene clusters -- the SAR83 cluster (Britschgi and Giovannoni 1991), SAR11 cluster (Giovannoni et al. 1990a), and SAR116 cluster (Mullins et al. 1995). Represented by OM42 and OM65 in Fig. 3.2A, OM clones related to the SAR83 cluster comprised 21% of the bacterial clones recovered (Table 3.3). As in previous analyses of SAR83-related clones (Britschgi and Giovannoni 1991; Fuhrman et al. 1993; Mullins et al. 1995), the neighbor-joining and bootstrap analyses in Fig. 3.2A (mask of ca. 450 nucleotides) confirmed this cluster's strong phylogenetic affiliation with *Roseobacter denitrificans* and *Roseobacter litoralis*, members of subgroup α-3 (bootstrap = 99/100 replicates). Members of the genus *Roseobacter* are pink-pigmented heterotrophs that synthesize bacteriochlorophyll *a* aerobically, and have been isolated from marine macrophytic algae (Shiba 1991; Ashen and Goff 1996). The α Proteobacteria subgroup designations follow those proposed by Woese et al. (1984a). Overall sequence similarities for environmental clones in the SAR83 cluster indicated a high level of within-cluster diversity: SAR83 and related clones in the Ocean Margins library had sequence similarities in the range of 91.9-100%, while the clones had sequence similarities to *Roseobacter denitrificans* in the range of 92.4-93.5%. A maximum-likelihood analysis of members of the SAR83 cluster revealed slightly different within-cluster relationships than those proposed in Fig. 3.2A (Fig. 3.3). In particular, the OM42 lineage and SAR83 were recovered as specific relatives, while the OM65 lineage formed an independent line of descent. Together, the two lineages of environmental clones encompassed *Roseobacter denitrificans* and *Roseobacter litoralis*.

The SAR11 cluster comprised 12% of the bacterial rDNA clones recovered (Table 3.3). It appears to be a very cosmopolitan clone type, having

Table 3.3. Frequency and phylogenetic affiliation of 16S rRNA gene clones in the OM library, grouped as in the RDP phylogenetically ordered listing for prokaryotes.

Table 3.3

Phylogenetic affiliation	Sequenced clones*	No. of clones	% of total clones†	% of bacterial clones‡
2.13.1 α-Proteobacteria				
2.13.1.3 Rhodobacter group (SAR83 cluster)	OM42, OM65, OM143, OM247	24	14	21
2.13.1.2 Rickettsia and others (SAR11 cluster)	OM136, OM155, OM188, OM239, OM242, OM258	13	8	12
2.13.1.1 Rhodospirillum rubrum ass. (SAR116 cluster)	OM25, OM38, OM55, OM299	13	8	12
2.13.1.1 Rhodospirillum rubrum ass.	OM75	1	1	1
2.13.3 γ-Proteobacteria				
2.13.3.6.1 Thiothrix nivea subgroup (SAR86 cluster)	OM10, OM23	32	19	29
2.13.3.10.2 Teredinibacter subgroup	OM60, OM241	4	2	3
2.13.3.9.2 Oceanospirillum ass.	OM182, OM185, OM252	3	2	3
2.13.3.10.3 Pseudomonas subgroup	OM93, OM133, OM233	3	2	3
2.13.3.11 Colwellia ass.	OM59	2	1	2
2.13.2 β-Proteobacteria				
2.13.2.7 Methylophilus group	OM43, OM58	6	4	5
2.13.2.2.6 Rubrivivax subgroup	OM180	3	2	3
2.13.2.2.3 Burkholderia subgroup II	OM156	2	1	2
2.6 Flexibacter-Cytophaga-Bacteriodes phylum				
2.6.2.1 Environmental clones	OM273	2	1	2
2.6.1.2.6 Vesiculatum subgroup	OM271	1	1	1
2.15 Gram-positive phylum				
2.15.1.3 Marine clones (Marine Gram-positive cluster)	OM1, OM231	2	1	2
2.9 Planctomyces and relatives				
2.9.1 Planctomyces subdivision (AGG27)	OM190	1	1	1
Possible chimeras or unique clones	OM27, OM110	2	1	2
2.10.2 Chloroplasts and Cyanelles				
Prymnesiophytes	OM21, OM125, OM153	24	14	--
Bacillariophytes	OM13, OM19, OM20, OM22	19	11	--
Prasinophytes	OM5, OM39, OM111, OM255	8	5	--
Cryptophyte	OM283	1	1	--
Unidentified plastids	OM81, OM164, OM270	3	2	--

*Only includes those clones for which a Genbank accession number was obtained.

†Calculated by dividing the number of clones in each group by the total number recovered (n=169).

‡Calculated by dividing the number of clones in each bacterial group by the total number of bacterial clones recovered (n=112).

Figure 3.2. Phylogenetic relationships of α Proteobacteria (A), γ Proteobacteria (B), and β Proteobacteria (C) 16S rDNA clones from the Ocean Margins clone library. The trees were inferred by the neighbor-joining method from a total of ca. 450 nucleotide positions. The number of bootstrap replicates which supported the branching order are shown above the relevant nodes, out of a total of 100 replicate samplings. Values below 50 are not shown. The scale bar corresponds to 10 nucleotide substitutions per 100 sequence positions. The 16S rRNA gene sequences of *Escherichia coli* (α and β Proteobacteria trees) and *Rhodoferax fermentans, Burkholderia cepacia,* and *Rhodocyclus purpureus* (γ Proteobacteria tree) were used as outgroups.

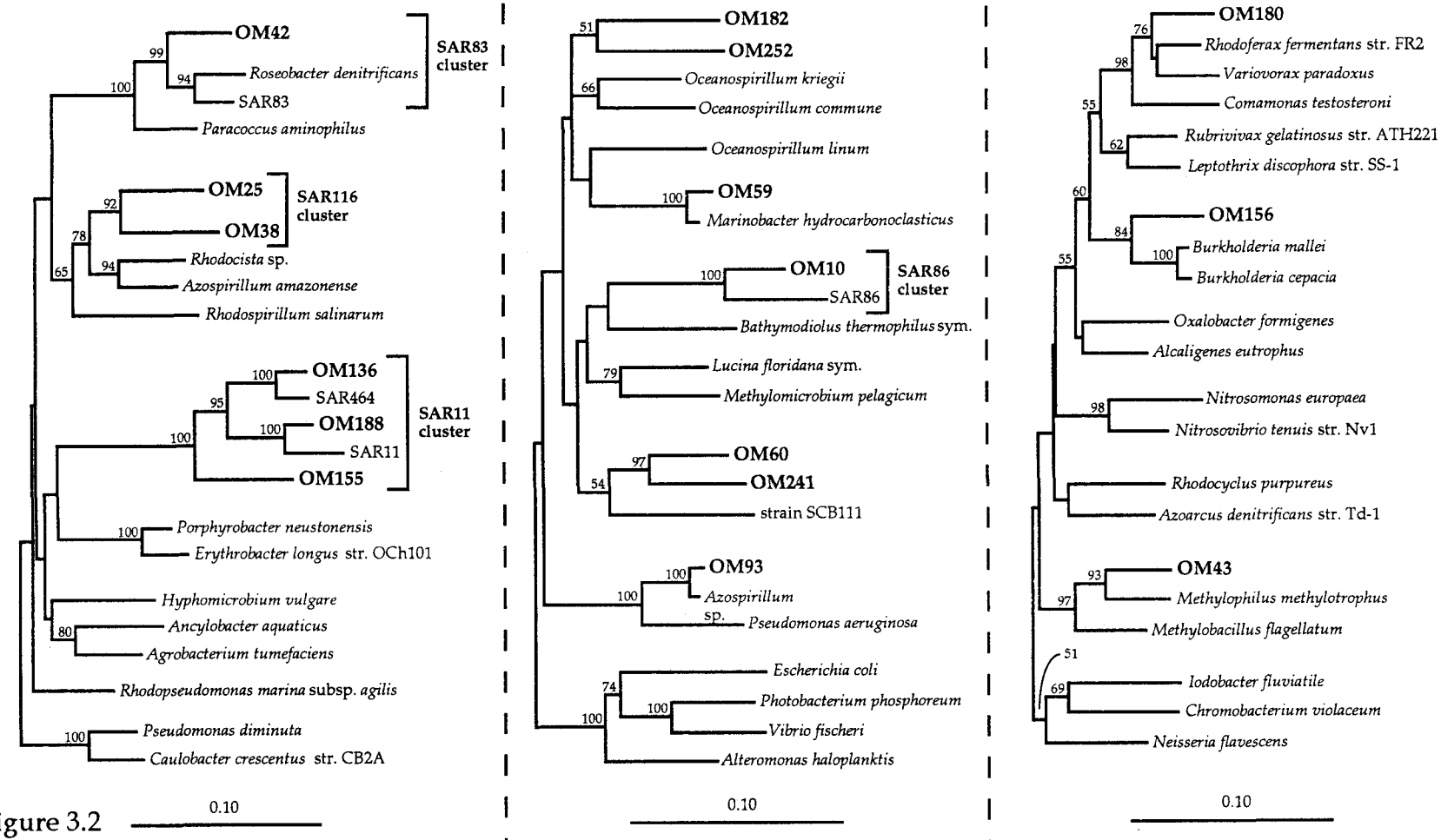

A. α Proteobacteria
OM42
Roseobacter denitrificans
SAR83
SAR83 cluster
Paracoccus aminophilus
OM25
OM38
SAR116 cluster
Rhodocista sp.
Azospirillum amazonense
Rhodospirillum salinarum
OM136
SAR464
OM188
SAR11
OM155
SAR11 cluster
Porphyrobacter neustonensis
Erythrobacter longus str. OCh101
Hyphomicrobium vulgare
Ancylobacter aquaticus
Agrobacterium tumefaciens
Rhodopseudomonas marina subsp. agilis
Pseudomonas diminuta
Caulobacter crescentus str. CB2A
0.10
B. γ Proteobacteria
OM182
OM252
Oceanospirillum kriegii
Oceanospirillum commune
Oceanospirillum linum
OM59
Marinobacter hydrocarbonoclasticus
OM10
SAR86
SAR86 cluster
Bathymodiolus thermophilus sym.
Lucina floridana sym.
Methylomicrobium pelagicum
OM60
OM241
strain SCB111
OM93
Azospirillum sp.
Pseudomonas aeruginosa
Escherichia coli
Photobacterium phosphoreum
Vibrio fischeri
Alteromonas haloplanktis
0.10
C. β Proteobacteria
OM180
Rhodoferax fermentans str. FR2
Variovorax paradoxus
Comamonas testosteroni
Rubrivivax gelatinosus str. ATH221
Leptothrix discophora str. SS-1
OM156
Burkholderia mallei
Burkholderia cepacia
Oxalobacter formigenes
Alcaligenes eutrophus
Nitrosomonas europaea
Nitrosovibrio tenuis str. Nv1
Rhodocyclus purpureus
Azoarcus denitrificans str. Td-1
OM43
Methylophilus methylotrophus
Methylobacillus flagellatum
Iodobacter fluviatile
Chromobacterium violaceum
Neisseria flavescens
0.10

Figure 3.2

Figure 3.3. Phylogenetic relationships within the four most abundant marine bacterioplankton rDNA clusters in the Ocean Margins clone library. The trees were constructed by maximum-likelihood analyses from ca. 425 nucleotide positions. The scale bar corresponds to 10 nucleotide substitutions per 100 sequence positions. The closest phylogenetic relative of each cluster was chosen to root the individual trees, with the exception of the SAR83 cluster. For this group, the closest cultivated relative is included within the cluster. *R., Roseobacter; A., Azospirillum; E., Erythrobacter; P., Porphyrobacter; B., Bathymodiolus; L., Lucina.*

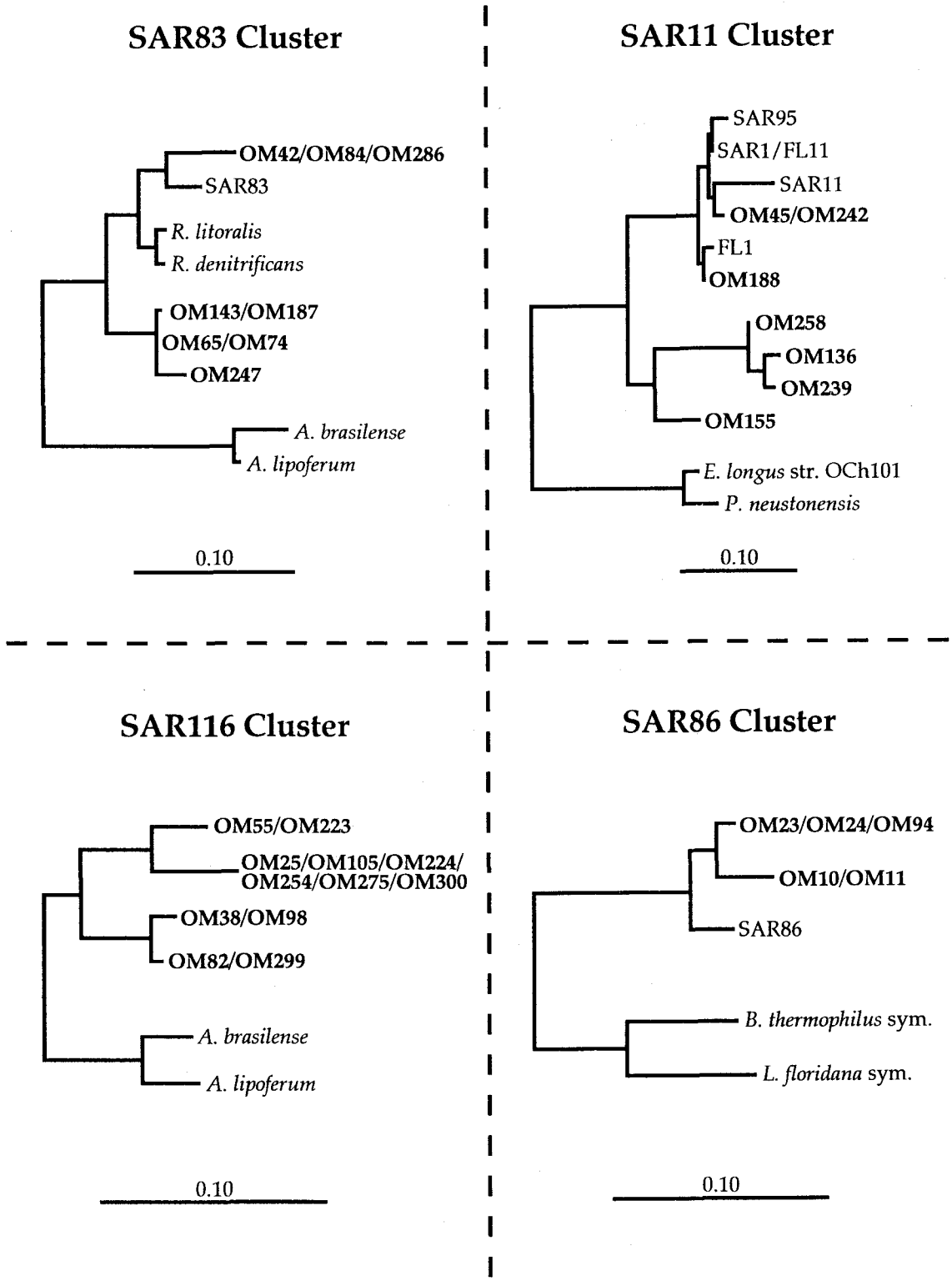

SAR83 Cluster
OM42/OM84/OM286
SAR83
R. litoralis
R. denitrificans
OM143/OM187
OM65/OM74
OM247
A. brasilense
A. lipoferum
0.10
SAR11 Cluster
SAR95
SAR1/FL11
SAR11
OM45/OM242
FL1
OM188
OM258
OM136
OM239
OM155
E. longus str. OCh101
P. neustonensis
0.10
SAR116 Cluster
OM55/OM223
OM25/OM105/OM224/
OM254/OM275/OM300
OM38/OM98
OM82/OM299
A. brasilense
A. lipoferum
0.10
SAR86 Cluster
OM23/OM24/OM94
OM10/OM11
SAR86
B. thermophilus sym.
L. floridana sym.
0.10

Figure 3.3

been recovered from both the Atlantic and Pacific central ocean gyres (Giovannoni et al. 1990a; Schmidt et al. 1991; Fuhrman et al. 1993), as well as a clone library of 32 clones constructed from surface seawater of the Santa Barbara Channel (DeLong et al. 1993). The neighbor-joining analysis in Fig. 3.2A indicated a loose phylogenetic relationship between members of the SAR11 cluster (including OM136, OM155, and OM188 from the Ocean Margins library), *Porphyrobacter neustonensis* and *Erythrobacter longus* str. Och101, members of subgroup α-4 (Fig. 3.2A), though bootstrap analyses and analyses of full length clone sequences did not support this affiliation. In a maximum-likelihood analysis of Ocean Margins SAR11-related clones, the SAR11 cluster exhibited a high level of within-cluster diversity (Fig. 3.3). Ocean Margins rDNA clones related to the SAR11 cluster had sequence similarities ranging from 89.6-100%, with similarities between 79.2-81.0% to the 16S rDNA of *Erythrobacter longus* str. Och101 and 78.6-80.1% to the 16S rDNA of *Porphyrobacter neustonensis.*

Ocean Margins relatives of the SAR116 cluster also comprised 12% of the bacterial rDNA clones recovered in this library. Represented by OM25, OM38, and OM55 in Fig. 3.2A, this cluster of rDNA genes resided within subgroup α-1 of the α Proteobacteria (bootstrap = 65/100 and 78/100 replicates, Fig. 3.2A). Members of the SAR116 cluster were most similar to *Azospirillum lipoferum* (87.5-89.1%), *Azospirillum brasilense* (87.1-88.9%), and *Rhodocista centenaria* (86.3-89.0%), though phylogenetic analyses that included more of the data set could not resolve a nearest cultivated neighbor for this group (mask up to ca. 1100 nucleotide positions; data not shown). Instead, the SAR116 cluster formed a unique line of decent within subgroup α-1 (Fig. 3.2A). As in the SAR83 and SAR11 clusters, SAR116-related clones contained a

high degree of within-cluster sequence variability, with four distinct lineages within the SAR116 cluster apparent (Fig. 3.3). Sequence similarities ranged from 89.6-100% for Ocean Margins rDNA clones related to the SAR116 cluster, while the rDNA sequence of SAR116 was most similar to clone OM55 (98.1%). A unique OM clone related to the α Proteobacteria (OM75) was found to be affiliated with subgroup α-1 and showed highest similarity to *Azospirillum lipoferum* (90.5% similar), but did not show a direct affiliation with the SAR116 cluster (Table 3.3).

The phylogeny of α Proteobacteria depicted in Fig. 3.2A was constructed from ca. 450 nucleotide positions located on the 5′ end of the 16S rRNA gene by the neighbor-joining method. Clones OM25 (SAR116 cluster), OM38 (SAR116 cluster), OM136 (SAR11 cluster), OM155 (SAR11 cluster), and OM75 (unique α Proteobacteria) were sequenced completely and neighbor-joining, parsimony, and maximum-likelihood analyses were performed that included the full data set (mask of ca. 1100 nucleotide positions). The results of these analyses supported the general relationships described above (data not shown). Clone SAR116 was not included in the phylogenetic analyses depicted in Fig. 3.2A and Fig. 3.3 because nucleotide sequence data for this clone starts at *E. coli* position 169, necessitating a much shorter mask in phylogenies that included only sequence data on the 5′ end of the 16S rRNA gene. However, all phylogenetic analyses that included SAR116 in the data set supported the relationships described above (data not shown).

γ Proteobacteria. Forty-four clones recovered in the OM library fell within the γ division of the Proteobacteria, of which 32 (29% of the bacterial clones) were members of the SAR86 cluster (Table 3.3; Mullins et al. 1995). This constituted the largest cluster of bacterial clones recovered in the OM

library and, as in previous analyses, a close cultivated phylogenetic relative of the SAR86 cluster could not be determined (Mullins et al. 1995). The neighbor-joining analysis in Fig. 3.2B indicated a loose relationship between members of the SAR86 cluster and the sulfur-oxidizing bacterial endosymbiont of *Bathymodiolus thermophilus*, a deep sea hydrothermal vent-associated bivalve (Distel et al. 1988). Bootstrap analyses did not support this affiliation, though the same relationship was recovered when a longer mask was employed (data not shown). The SAR86-related clones were most similar to the bacterial symbionts of *Bathymodiolus thermophilus* (86.8-88.7%) and *Lucina floridana* (87.3-88.4%). As in the α Proteobacteria clusters, the SAR86 cluster contained significant within-cluster variability, with similarities between OM10, OM23, and SAR86 in the range of 91.9-97.0% (Fig. 3.3).

Three OM rDNA clones were phylogenetically affiliated with members of the genus *Oceanospirillum* and others in subgroup 3 of the γ Proteobacteria (Table 3.3; γ Proteobacteria subgroups as in Woese et al. 1985). Represented by OM182 in Fig. 3.2B, two Ocean Margins clones were most similar to *Marinomonas vaga* (91.5%), a marine bacteria originally isolated by enrichment culture from surface seawater collected off the coast of Hawaii in the Pacific Ocean (Baumann et al. 1972). OM252, a unique clone type in this library, had highest sequence similarity to the marine bacteria *Oceanospirillum beijerinckii* (89.0%), *Oceanospirillum commune* (88.6%), and clone OM182 (88.4%; Fig. 3.2B).

A set of four clones related to the γ Proteobacteria, represented by OM60 and OM241 in Fig. 3.2B, were most similar to strain SCB111, a marine bacterium isolated from seawater collected off Scripps pier (89.1-89.5% similar; Rehnstam et al. 1993), and the sulfur-oxidizing bacterial endosymbiont of *Lucina floridana* (88.6-90.3%). Though phylogenetically related to each other

(93.1% similar, bootstrap = 97/100 replicates), the phylogenetic relationships of OM60 and OM241 to other members of the γ Proteobacteria were unstable, however, and were not recovered with confidence in the bootstrap analysis depicted in Fig. 3.2B, or in phylogenetic analyses employing mask lengths that encompassed the majority of the rDNA molecule (data not shown).

Four Ocean Margins rDNA clone sequences were highly similar to the 16S rRNA gene sequences from cultured species within the γ Proteobacteria (Table 3.3, Fig. 3.2B). These included OM133 (*Pseudomonas aeruginosa* str. NIH18, 99.3% similar), OM93 (*Azospirillum* sp. str. DSM1727, 99.0% similar), OM59 (*Marinobacter hydrocarbonoclasticus*, 97.0% similar), and OM233 (*Pseudomonas flavescens* str. B62, 95.5% similar). It should be noted that *Azospirillum* sp. str. DSM1727 is not phylogenetically related to other members of the genus *Azospirillum* (α Proteobacteria) and is obviously misclassified (Fig. 3.2A and B; Xia et al. 1994).

β Proteobacteria. Eleven 16S rDNA clones from the OM clone library (10% of the bacterial clones recovered) were related to the β Proteobacteria (Table 3.3). One set of clones, represented by OM43 in Fig. 3.2C, comprised 5% of the bacterial clones recovered and grouped within a clade of Type I methylotrophs related to subgroup β-3 of the β Proteobacteria (Bootstrap = 97/100 replicates; Tsuji et al. 1990; β Proteobacteria subgroup designations as in Woese et al. 1984b). In particular, the rDNA gene sequence of OM43 was most similar to *Methylophilus methylotrophus* str. AS1, an obligate Type I methylotrophic bacterium isolated from activated sludge (92.6% similar; Jenkins et al. 1987). This relationship was also supported by the bootstrap analysis in Fig. 3.2C, which grouped OM43 and *Methylophilus methylotrophus* str. AS1 in 93 out of 100 replicate samplings.

The next most abundant set of clones related to the β Proteobacteria are represented by OM180 in Fig. 3.2C. Comprising 3% of the bacterial 16S rDNAs recovered (3 clones), clone OM180 phylogenetically affiliated with *Rhodoferax fermentans* (91.8% similar), a phototrophic purple non-sulfur bacteria isolated from ditch water and activated sludge (Hiraishi et al. 1991, Hiraishi 1994), and *Variovorax paradoxus* (90.9% similar), a heterotrophic bacterium isolated from soil (Willems et al. 1991). *Rhodoferax fermentans* and *Variovorax paradoxus* are members of the family *Comamonadaceae* within subgroup β-1 of the β Proteobacteria and are 94.0% similar in 16S rRNA gene sequence, but do not appear to share many phenotypic traits (Willems et al. 1991). In this analysis, the *Comamonadaceae* (*Rhodoferax fermentans, Variovorax paradoxus,* and *Comamonas testosteroni*) were recovered as a monophyletic group within the β Proteobacteria in 98 out of 100 bootstrap replicates, with the inclusion of clone OM180 (Fig. 3.2C).

Two OM 16S rDNA clones (2% of the bacterial clones) are represented by OM156 in Table 3.3 and Fig. 3.2C. In the neighbor-joining analysis depicted in Fig. 3.2C, OM156 affiliated with members of the genus *Burkholderia* of the β-2 subgroup (bootstrap = 84/100 replicates), including *Bukholderia mallei* (91.2% similar) and *Burkholderia cepacia* (91.9% similar). *Bukholderia mallei* and *Burkholderia cepacia* (96.1% similar) demonstrated a robust affiliation in Fig. 3.2C, forming a monophyletic cluster in 100 out of 100 bootstrap replicate samplings.

Flexibacter-cytophaga-bacteroides. Three OM clones were related to the flexibacter-cytophaga-bacteroides (FCB) line of descent (Table 3.3, Fig. 3.4; Paster et al. 1985). Two OM clones, represented by OM273 in Fig. 3.4, were highly similar to the sequence of BDA1-14, an rDNA clone fragment recovered

Figure 3.4. Unrooted phylogenetic tree showing relationships of OM rDNA clones related to the flexibacter-cytophaga-bacteroides phylum, and the Planctomycetales. The tree was constructed by the neighbor-joining method from 425 nucleotide positions. The scale bar corresponds to 10 nucleotide substitutions per 100 sequence positions. A close phylogenetic relationship between the FCB phylum and Planctomycetales is not implied by this tree; they are included in the same figure out of convenience.

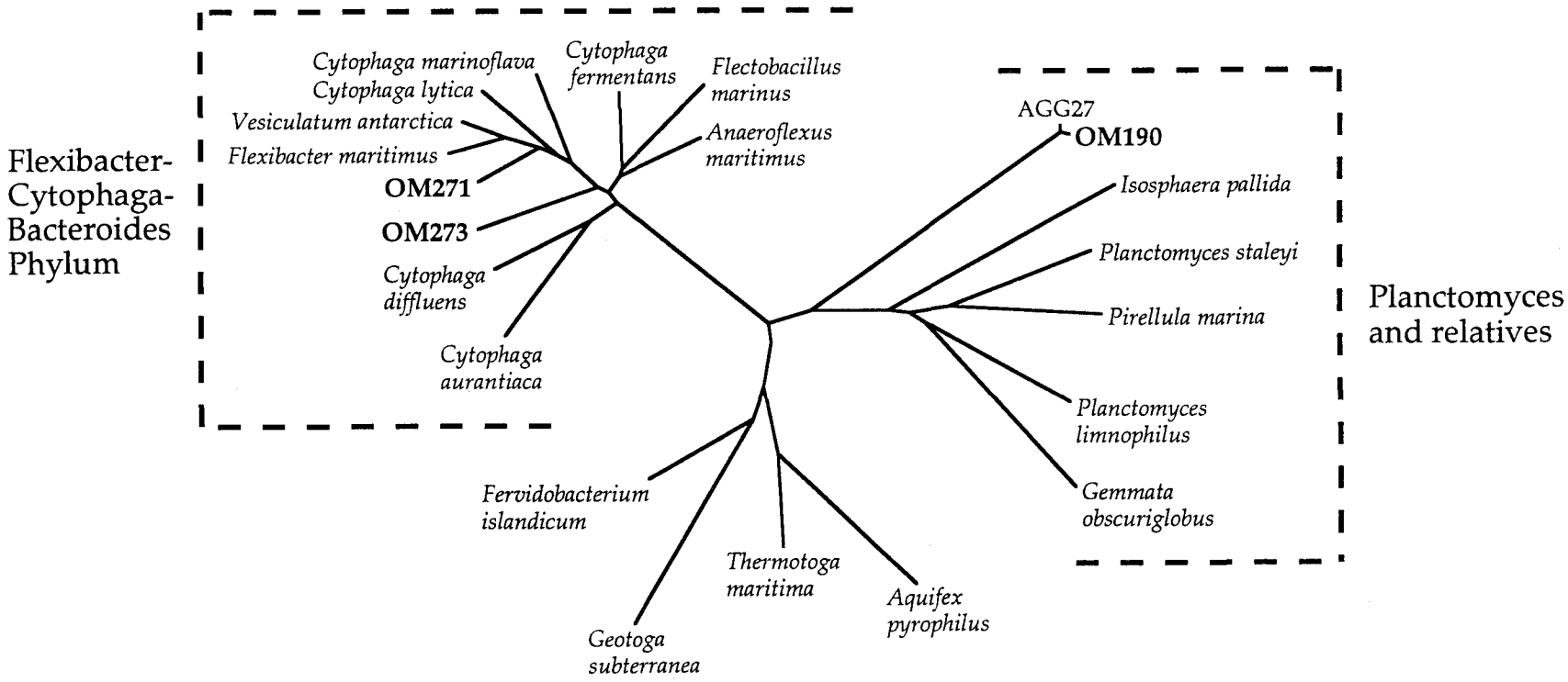

Flexibacter-
Cytophaga-
Bacteroides
Phylum
Cytophaga marinoflava
Cytophaga lytica
Vesiculatum antarctica
Flexibacter maritimus
OM271
OM273
Cytophaga
diffluens
Cytophaga
aurantiaca
Cytophaga
fermentans
Flectobacillus
marinus
Anaeroflexus
maritimus
AGG27
OM190
Isosphaera pallida
Planctomyces staleyi
Pirellula marina
Planctomyces
limnophilus
Gemmata
obscuriglobus
Planctomyces
and relatives
Fervidobacterium
islandicum
Thermotoga
maritima
Aquifex
pyrophilus
Geotoga
subterranea
0.10

Figure 3.4

in a PCR-generated rDNA library from Sargasso Sea surface water (98.1% similar for 242 base positions of overlap; Fuhrman et al. 1993). Phylogenetic analyses indicated these clones formed a unique lineage within the FCB phylum (Fig. 3.4), distantly related to *Microscilla sericea* str. SIO-7 (83.4.% similar), *Cytophaga lytica* (83.2% similar), and *Cytophaga marinoflava* (83.0% similar). Phylogenetic analyses of clone OM271, a unique clone type in the OM library, placed it within the cytophaga subgroup of the FCB phylum (Fig. 3.4), of which a number of marine bacterial isolates are members (Gherna and Woese 1992). OM271 was most similar to *Flexibacter maritimus* (88.9%), a marine fish pathogen (Wakabayashi et al. 1984), and *Vesiculata antarctica* (88.8%), a pigmented, gas vacuolate bacterium isolated from Antarctic marine water (Irgens et al. 1989).

Gram-positive. Two Ocean Margins clones (OM1 and OM231 in Table 3.3) were highly similar to a cluster of 9 fragmentary rDNA clone sequences of ca. 105 nucleotides in length recovered by Fuhrman et al. (1993) in rDNA clone libraries constructed from Pacific (100 m) and Atlantic (10 m) seawater samples (95.2-99.0 % similar). Employing a mask of ca. 1100 nucleotides, neighbor-joining and bootstrap analyses of clones OM1 and SAR432, an rDNA clone 95.8% similar to OM1 recovered from 80 m in the Sargasso Sea (D. Gordon, unpubl. data), placed this cluster of clones within the high G + C Gram-positive phylum (Fig. 3.5). This "marine Gram-positive" cluster formed a unique, deeply branching lineage within the high G + C division of the Gram-positive bacteria and showed low sequence similarity (<81.0%) to any complete or nearly complete rRNA gene sequence recovered from a cultured bacterium available through Genbank or the RDP. A secondary structural model was constructed from the rDNA gene sequence of clone OM1 to check

Figure 3.5. Phylogenetic affiliations of 16S rDNA clones OM1 and SAR432 with other members of the high G + C Gram-positive phylum. The tree was constructed by the neighbor-joining method from ca. 1100 nucleotide positions. The scale bar corresponds to 10 nucleotide substitutions per 100 sequence positions. The 16S rRNA gene sequences of *Aquifex pyrophilus, Geotoga subterranea, Fervidobacterium islandicum,* and *Thermotoga maritima* were used as an outgroup to position the root.

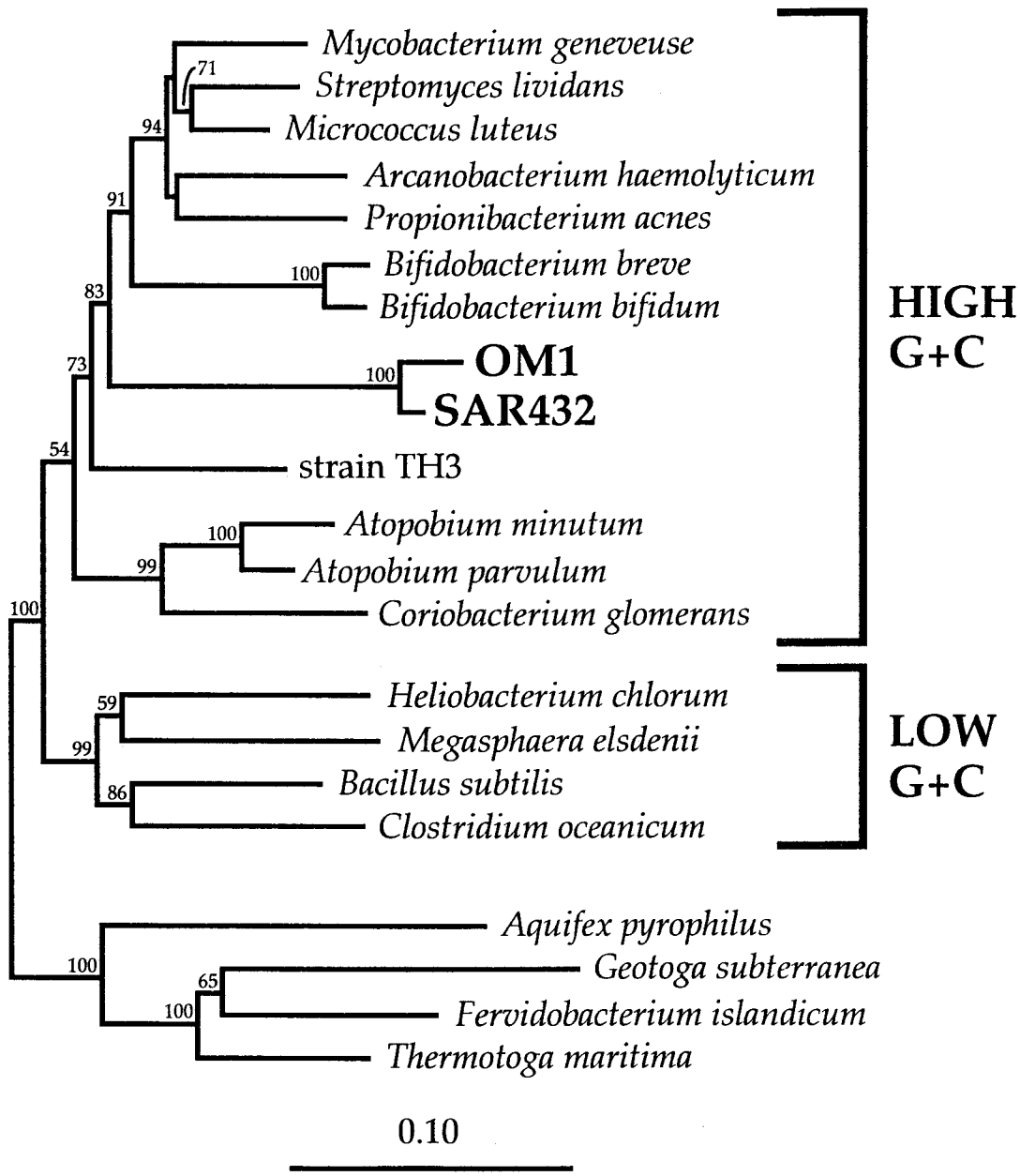

Mycobacterium geneveuse
Streptomyces lividans
Micrococcus luteus
Arcanobacterium haemolyticum
Propionibacterium acnes
Bifidobacterium breve
Bifidobacterium bifidum
OM1
SAR432
strain TH3
Atopobium minutum
Atopobium parvulum
Coriobacterium glomerans
Heliobacterium chlorum
Megasphaera elsdenii
Bacillus subtilis
Clostridium oceanicum
Aquifex pyrophilus
Geotoga subterranea
Fervidobacterium islandicum
Thermotoga maritima
71
94
91
100
83
73
100
54
100
99
59
99
86
100
65
100
HIGH G+C
LOW G+C
0.10

Figure 3.5

for base-pairing idosyncrasities due to chimeric gene formation (Kopczynski et al. 1994), as well as to identify signatures unique to the high G + C Gram-positive phylum (Fig. 3.6). Clone OM1 contained 5 of 6 sequence signatures for the high G + C Gram-positive phylum as defined by Woese (1987), including an "A" residue at *E. coli* position 906, a "C" residue at position 955, a "C" residue at position 1207, a "G" residue at position 1229, and a "G" residue at position 1410 (Fig. 3.6; numbering based on the *E. coli* 16S rRNA gene; Brosius et al. 1978). At position 1198, clone OM1 contained a "G" residue instead of the high G + C signature "A" residue, though it is noted that a "G" residue at this position is the bacterial consensus, and is found in a minor percentage (<15%) of high G + C Gram-positive bacteria. A portion of the helix in variable region 9 (V9, positions 1410-1435, 1466-1490; Dams et al. 1988) was also reported by Woese as a high G + C Gram-positive rRNA signature (Woese 1987). In the secondary structural model for clone OM1,nucleotide sequence and base pairing within this helix was consistent with the high G + C Gram-positive signature (Fig. 3.6).

Planctomycetales. One clone from the OM library, OM190, was found to be closely related to the environmental clone AGG27, a 16S rDNA clone PCR-amplified from phytodetrital marine aggregates collected from the Pacific Ocean off the coast of California (99.6% similar; DeLong et al. 1993). Neighbor-joining analyses indicated these two clones formed an extremely deep evolutionary branch related to the Planctomycetale line of decent (Fig. 3.4; Schlesner and Stackebrandt 1986, Ward et al. 1995). Though the phylogenetic tree in Fig. 3.4 was constructed from only ca. 425 nucleotides, the unique position of OM190 and AGG27 did not change when a longer mask of ca. 1100 nucleotides was employed in neighbor-joining, parsimony, and

Figure 3.6. Proposed secondary structural model of clone OM1 16S rRNA. Arrows indicate signature nucleotides of the Gram-positive high G + C 16S rRNA gene lineage, and the boxed area indicates a conserved region within variable region 9 (V9) of the bacterial 16S rRNA unique to members of this phylum. Numbers refer to nucleotide positions in the *E. coli* 16S rRNA gene (Brosius et al. 1978). Target sites for the primers 27F and 1522R are indicated.

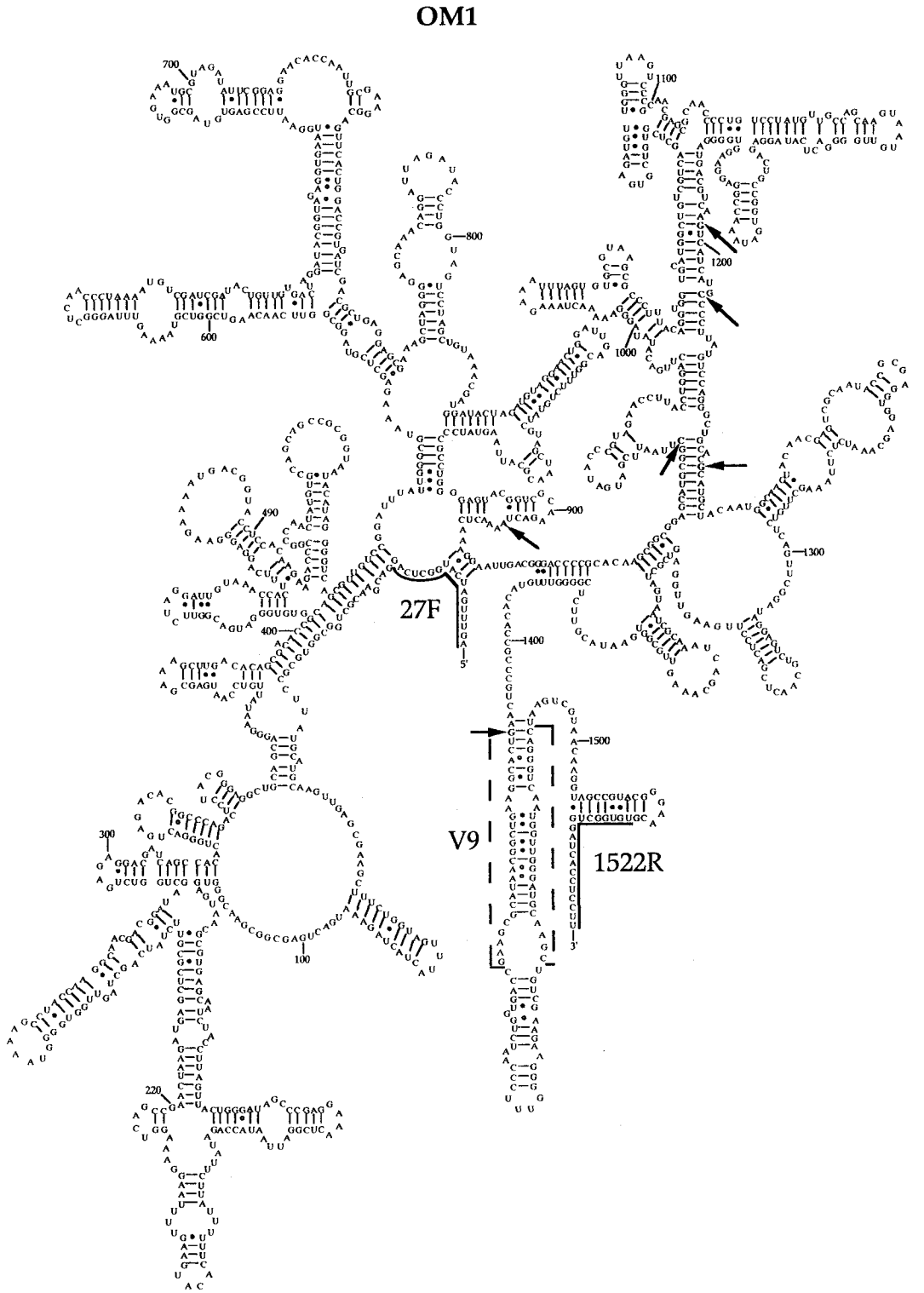

OM1
27F
V9
1522R
100
200
300
400
490
600
700
800
900
1000
1100
1200
1300
1400
1500

Figure 3.6

maximum-likelihood analyses (data not shown). Clone OM190 had no sequence similarities greater than 78-79% when compared to 16S rRNA gene sequences from cultured bacteria available in Genbank and the RDP, the lowest of any clone in the Ocean Margins library. A signature sequence analysis comparing clones OM190 and AGG27, members of the order Planctomycetales, and corresponding bacterial consensus nucleotides revealed OM190 and AGG27 contained 6 of 8 Planctomycetale nucleotide signatures (Table 3.4; Woese 1987). OM190 and AGG27 did not contain the single base insertion after *E. coli* position 983 unique to members of the Planctomycetales, and had the bacterial consensus nucleotide at *E. coli* position 955, one of the two positions were it lacked a Planctomycetale signature base (Table 3.4). Because of its unique phylogenetic position and lack of similarity to the 16S rRNA gene sequences of cultured bacteria, a secondary structural model for the OM190 rDNA gene product was constructed (Fig. 3.7). No unusual base pairing was observed within conserved helices, although the cloned molecule was lacking ca. 90 bases on its 5′ end, including the 27F PCR primer site. For clone OM190, the secondary structural element at variable region 2 (V2, positions 140-223; Dams et al. 1988) of the bacterial 16S rRNA was most similar to that of *Isosphaera pallida* and *Planctomyces limnophilus,* though the Planctomycetales order as a whole are variable in this region.

Plastids. Fifty-five rDNA clones in the Ocean Margins library, or 33% of the total recovered, branched within the oxygenic phototroph line of descent. In particular, these clones affiliated with eukaryotic plastid 16S rRNA genes, which form a monophyletic clade within the cyanobacteria (Douglas and Turner 1991, Giovannoni et al. 1988b). As previously reported, the two most abundant plastid rDNA groups in the OM library were closely related to

Table 3.4. 16S rRNA signature sequence analysis of OM190 and AGG27.

Table 3.4

Position	Consensus*	Planctomyces and relatives*	OM190	AGG27
353	A	U	U	U
570	G	U	A	A
933	G	A	A	A
955	U	C	U	U
983:1	–†	U	–	–
1109	C	A	A	A
1384	C	U	U	U
1410	A	G	G	G
1520	G	C	C	n.d.‡

*From Woese 1987.

†–, no nucleotide at this position.

‡n.d., no data available.

Figure 3.7. Proposed secondary structural model of clone OM190 16S rRNA. Arrows indicate signature nucleotides of the Planctomyces 16S rRNA gene lineage. Numbers and primers are as in Fig. 3.6.

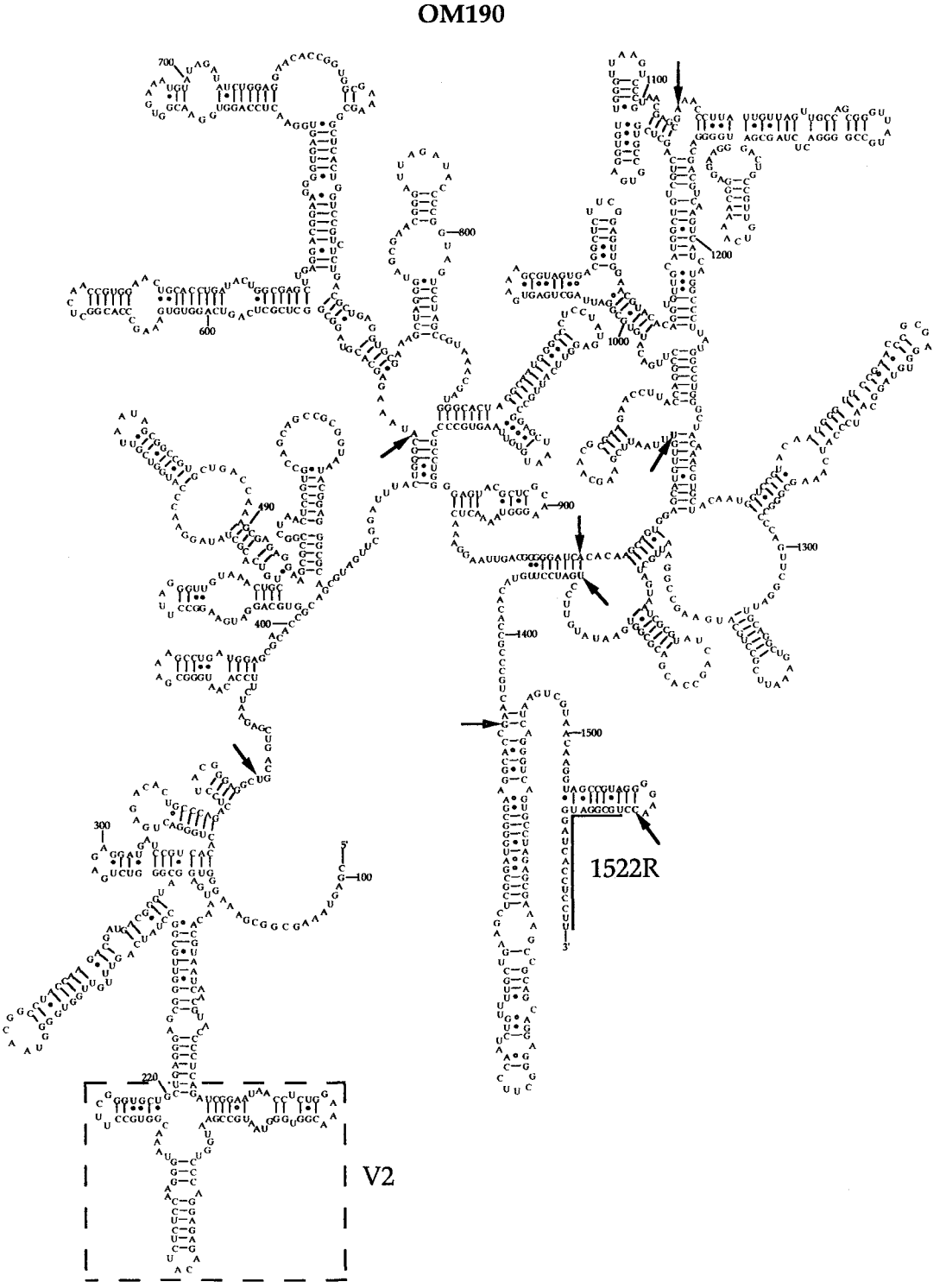


Figure 3.7

the plastid 16S rRNA gene sequences from the Prymnesiophyceae (24 clones) and Bacillariophyceae (19 clones; Rappé et al. 1995). In addition, OM clones tentatively related to the Prasinophyceae (8 clones; V. Huss, pers. comm.), and the Cryptophyceae (1 clone) were discovered (Table 3.3). A phylogenetically close cultivated neighbor was not determined for three unique OM rDNA clones that fall within the plastid 16S rDNA line of descent.

Oligonucleotide probe hybridizations. Of 213 transformants included on the colony blots, 44 were later determined to be not of full length and thus were not included in the library analysis. The taxon-specific oligonucleotide probes listed in Table 3.1 were hybridized to the pinblots to aid in identifying clones belonging to potentially abundant rDNA groups in the OM library (data not shown). In addition to the clones identified by sequencing, 25 clones were identified by the SAR86R probe, 21 by the HAP1 probe, 13 by the SAR83R probe, 13 by the CHRYS1 probe, and 4 by the SAR11A1 probe. Clones identified by probe hybridization contributed to the total number for each clone type listed in Table 3.3. All other clone groups were identified solely by partial or complete rDNA sequencing.

Coverage of the OM clone library. The term coverage (C) describes the fraction of a clone library of infinite size that would be sampled by a clone library of smaller size (Good 1953; Giovannoni et al. 1995). Thus, coverage estimates emphasize the fraction of genes (cells), rather than taxa, "covered" by a library. Coverage values are derived from the equation:

$$C = 1 - (n_1/N)$$

where n_1 is the number of unique clones and N is the total number of clones examined. Assuming clones which are >97% similar are identical, the coverage for the 169 clones analyzed here was 95%. This number compares favorably with numbers published for smaller clone libraries from seawater (Mullins et al. 1995). By its definition, the coverage estimate provided above pertains only to a hypothetical clone library of infinite size, derived from the same PCR product used in the construction of the clone library described here. Thus, any groups that were excluded by our choice of PCR primers (e.g. Archaea, Eukarya) are not included in the coverage analysis.

Chimeric and unique rDNA genes. All OM clones were closely examined for the possibility that they were chimeric gene artifacts. First, all rDNA gene sequences obtained in this study were submitted to the RDP program CHECK_CHIMERA. This revealed evidence of chimeric gene formation in the partial gene sequence obtained for rDNA clone OM110, a unique clone type in the OM library. Approximately 390 nucleotides of sequence information was determined from the 5' end of this clone, corresponding to *E. coli* positions 8-400. CHECK_CHIMERA indicated a fragment of 140 nucleotides on the 5' end of the sequence was closely related to the γ subclass of the Proteobacteria, including the marine isolate SCB111 and clones OM60 and OM241. The remaining fragment of 250 nucleotides was phylogenetically related to the α Proteobacteria and, in particular, to species in the genus *Azospirillum*.

OM clones were analyzed further in two ways: first, secondary structures were generated for all sequenced OM rDNA clones to confirm their predicted 16S rRNA gene products could be folded into consensus secondary structures consistent with the general topological features of other established

bacterial 16S rRNAs (data not shown except were noted; Gutell 1994). In addition, phylogenies and sequence similarities were derived from different regions of the cloned 16S rDNAs to aid in the detection of chimeras (data not shown; Robinson-Cox et al. 1995; Field et al. 1997). A phylogenetic analysis of the nearly complete rDNA sequence for clone OM27 indicated it was related to the Proteobacteria, but did not affiliate with any of its major subclasses (data not shown). A CHECK_CHIMERA analysis did not provide any evidence for chimera formation, and failed to reveal a potential break point in the molecule. A secondary structural analysis of the entire rDNA sequence for OM27 revealed an odd "bulge" caused by an insertion of six nucleotides after position 246 of the *E. coli* 16S rRNA gene. To examine the proposition that this position might be a potential break point, separate phylogenies were generated for the 240 nucleotides on the 5′ end of the molecule, and ca. 1000 nucleotides on the 3′ end. Due to the short length of the fragment on the 5′ end of the molecule, we could not resolve its phylogenetic position with confidence, though it did not strongly affiliate with any of the major Proteobacterial subclasses. This fragment contained the secondary structural element at V2 of the bacterial 16S rRNA expected for the α Proteobacteria (data not shown, Woese 1987). The 3′ end of the molecule was of similar phylogenetic position to that of the entire rDNA sequence for clone OM27; it did not reside within any of the major subclasses of the Proteobacteria. The predicted secondary structural model for the 3′ fragment of clone OM27 contained the expected α Proteobacteria secondary structural element at V3 of the bacterial 16S rRNA (Woese 1987). Phylogenetic analyses employing masks that incorporated variable lengths of the 5′ and 3′ end of the rDNA sequence failed to reveal any other potential break points in the molecule. In each case, OM27 appeared as

an independent line of descent within the class Proteobacteria. Based on the above analyses, clone OM27 was either a phylogenetically unique member of the Proteobacteria, or a chimeric gene artifact of a type that our current techniques could not resolve.

3.5 Discussion

Microbial ecology is moving rapidly into a new era in which focus is shifting to the importance of uncultured species and the role of community composition in ecosystem function. The study described here, which follows this trend, was initiated under the auspices of the Department of Energy Ocean Margins Program with the long term goal of understanding the carbon cycle in the local region of the eastern continental shelf near Cape Hatteras, NC. Our immediate objectives were 1) to determine whether the bacterioplankton in this region of the continental shelf were similar to those which had been found previously in open-ocean samples, and 2) to develop probes for the detection of key species in ecological studies planned for the future. Although we found much overlap between coastal and open-ocean bacterioplankton, we also identified important differences. Among the most important observations was the discovery of β proteobacterial genes closely allied to known taxa of obligate methylotrophic bacteria.

The evidence of methylotrophs was a surprise, since none had been found in previous clone libraries from marine systems, and there was nothing about the study site to suggest that it, in particular, should be the locus of unusual levels of C1 metabolism. The presence of these genes originating from a clade of Type I methylotrophs does not necessarily imply that C1 metabolism

was a prominent feature of the carbon cycle at the study site. There is no firm evidence that the organisms from which the genes originated are indeed methylotrophs or that, if they were methylotrophs, they were actively engaged in C1 metabolism. Nonetheless, the novel genes branched from within the methylotroph clade, rather than forming an outgroup. The phylogenetic tree in Fig. 3.1 shows the position of OM43 relative to *Methylophilus methylotrophus* and *Methylobacillus flagellatum*, but the clade also includes *Methylobacillus glycogenes*, *Methylomonas methanolica*, and *Methylomonas methylovora*. These five species are all obligate methylotrophs which use the RuMP pathway for C1 assimilation and are unable to oxidize methane.

Although one aspiration of studies such as this is the discovery of new phylogenetic groups, from an ecological standpoint, the finding that relatively few taxa dominate the 16S rDNA gene distribution in this clone library is encouraging, since it suggests that a tractably small set of oligonucleotide probes might be sufficient for characterizing community ecology in future studies. A majority of the bacterial rDNAs recovered in this clone library were related to previously detected rDNA lineages cloned from open-ocean marine habitats. In fact, nearly three fourths of the bacterial fraction of the OM library were affiliated with four gene clusters - the SAR86 cluster (γ Proteobacteria), SAR83 cluster, SAR116 cluster, and SAR11 cluster (all α Proteobacteria). In addition, a small number of clones in this library were related to three rDNA lineages previously recovered from seawater, including a lineage within the flexibacter-cytophaga-bacteriodes line of descent, the high G + C Gram-positives, and a phylogenetically distinct lineage remotely associated with the Planctomycetales. The fact that these lineages have been recovered in independent investigations of marine prokaryotic diversity from both

oligotrophic open-ocean habitats as well as this eutrophic continental shelf seawater sample indicates that they are widespread in pelagic marine systems and potentially important members of bacterioplankton communities.

rDNA gene clusters are related sets of genes which are more similar to each other than to any rRNA gene sequences from a cultivated organism (Giovannoni et al. 1996a), and are commonly recovered in rDNA clone libraries constructed from environmental samples (Giovannoni et al. 1990a; DeLong 1992; DeLong et al. 1993; Fuhrman et al. 1993). From a phylogenetic perspective, gene clusters resemble gene clades from cultivated organisms, but because gene cluster sequences come from uncultivated organisms, it is not known whether they represent separate cellular lineages. An analysis of the 16S rRNA gene sequences from the seven *rrn* operons of *E. coli* strain PK3 (Cilia et al. 1996) revealed inter-operon similarities between 98.9%-100.0%, indicating that the divergence between *rrn* operons of the same organism may explain the shallowest branches in an environmental gene cluster, but is not sufficient to explain the high degree of sequence divergence characteristic of the gene clusters described here. Recently, we have reported evidence that the genetic diversity within the SAR11 gene cluster includes gene lineages which exhibit different depth-specific distributions (Field et al. 1997), but little information is currently available on the ecological significance of the many gene clusters recovered in this analysis.

In this study, we determined complete sequences from a large number (16) of gene clones and multiple partial gene sequences for representatives of gene clusters that had been identified in previous studies. An alternative strategy would have been to be identify larger numbers of clones with fragmentary gene sequences, thereby increasing the chances of identifying

evolutionarily novel gene clones. Our choice of rDNA gene sequencing strategy was dictated by our long-term goal of understanding the ecology of marine bacterioplankton by applying oligonucleotide probes and other genetic methods to large numbers of environmental samples. For this purpose, complete gene sequences offer the important advantage of providing the experimentalist with a larger number of potential variable regions to be considered in the design of oligonucleotide probes. Since the process by which substitutions are accumulated in gene lineages is random, the occurrence of suitable probe sites for ecological studies cannot be guaranteed through the inspection of only a small domain of the 16S ribosomal RNA molecule. In principle, the greater number of potential probe hybridization sites made available by complete sequences can lead to the selection of probes which have a higher specificity, and also to the design of multiple hybridization probes, which can increase the resolution of *in situ* hybridization studies based on the use of fluorescent rDNA probes. A second reason for determining complete gene sequences is that they permit comparisons between data obtained by different investigators. For example, previous investigations of 16S rDNA clones from the Atlantic and Pacific Oceans (Schmidt et al. 1991; Furhman et al. 1993) examined parts of the 16S ribosomal RNA gene which did not overlap, preventing direct comparisons between these data sets. A third reason for determining complete gene sequences is that they offer a higher degree of phylogenetic resolution. This attribute of longer nucleotide sequences is particularly important when the subject in question is a lineage which branches deeply from within the bacterial phylogenetic tree and displays uncertain affiliations with known phylogenetic groups, as in the case of SAR202 (Giovannoni et al. 1996b), SAR406 (Gordon and Giovannoni 1996), the

marine Gram-positive group (Fuhrman et al. 1993), and the AGG27/OM190 lineage (DeLong et al. 1993).

There are also compelling reasons to determine multiple gene sequences from members of a single gene cluster. In such cases, a single gene sequence is usually sufficient to determine the phylogenetic position of the gene cluster relative to known microbial groups. However, for the purpose of designing hybridization probes, multiple gene sequences offer powerful advantages because they permit the selection of probe hybridization sites that are either common to many members of a gene cluster, or to specific sub-branches within a phylogenetic cluster. In the case of the SAR11 cluster, these advantages appear to be ecologically significant (Field et al. 1997).

The methods employed here are not without limitations and potential biases of their own. Although such concerns are widely expressed, no methodological biases have yet been described that would invalidate the broad conclusions reached by studies such as these. Indeed, it is particular interesting to note the general congruence between 16S rDNA sequencing studies which differed significantly in the methodologies employed (Giovannoni et al. 1995; Mullins et al. 1995). We have recently reported biases which can be introduced by PCR in cases where product concentrations rise to high values during late rounds of replication, but these biases serve to increase, rather than decrease, the diversity of genes in PCR products derived from natural systems (Suzuki and Giovannoni 1996). It is important to note that even if this clone library accurately reflected the relative abundance of rRNA gene types in the original sample, the abundance of rRNA genes and organism abundance in the environment may differ due to variations in rRNA gene copy number and genome size between taxa (Farrelly et al. 1995).

The composition of this clone library was decidedly different than those from previous open-ocean studies, although there was substantial overlap. Given the locale of our sampling site, the frequency of plastid genes and lack of prokaryotic oxygenic phototrophs in this clone library did not come as a great suprise. In many coastal pelagic ecosystems, the biomass of eukaryotic phytoplankton often outweighs that of prokaryotic oxygenic phototrophs by several orders of magnitude (e.g. Murphy and Haugen 1985; Glover et al. 1986; Marshall and Ranasinghe 1989). Data from the cruise our sample originated from showed that, in addition to outweighing prokaryotic oxygenic phototrophs by 2-3 orders of magnitude, eukaryotic phytoplankton were often numerically more abundant than their cyanobacterial counterparts (Verity et al. 1996). Other factors may have contributed to the abundance of plastid rRNA genes in this clone library, including the absence of a pre-filtration step in our sample collection, the multiplicity of chloroplasts in many phytoplankton taxa, and the multiplicity of plastid genomes within these plastids. As in previous studies of this type, there appeared to be little overlap between rDNA clone sequences and those of cultured marine bacteria, with only four clones in the Ocean Margins library described here showed greater than 95.0% similarity to the rRNA gene sequences of cultured bacteria.

Investigations using oligonucleotide probes for specific taxa have begun to reveal pronounced temporal and spacial patterns in the distributions of some bacterial groups in pelagic ocean environments (Giovannoni et al. 1996b; Gordon and Giovannoni 1996). The significance of these observations and their generality with respect to the complex hydrology of the Cape Hatteras area cannot be established without investigating a large number of water

samples. By identifying potentially abundant bacterioplankton in this region, we have taken the first step towards this goal.

3.6 References

Amann, R. I., B. J. Binder, R. J. Olsen, S. W. Chisholm, R. Devereux, and D. A. Stahl. 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl. Environ. Microbiol. **56**:1919-25.

Ashen, J. B., and L. J. Goff. 1996. Molecular identification of a bacterium associated with gall formation in the marine red alga Prionitis lanceolata. J. Phycol. **32**:286-97.

Baumann, L., P. Baumann, M. Mandel, and R. D. Allen. 1972. Taxonomy of aerobic marine eubacteria. J. Bacteriol. **110**:402-29.

Britschgi, T. B., and S. J. Giovannoni. 1991. Phylogenetic analysis of a natural marine bacterioplankton population by rRNA gene cloning and sequencing. Appl. Environ. Microbiol. **57**:1707-13.

Brosius, J., M. L. Palmer, J. P. Kennedy, and H. F. Noller. 1978. Complete nucleotide sequence of a 16S ribosomal RNA gene from Escherichia coli. Proc. Natl. Acad. Sci. USA **75**:4801-5.

Burrage, D. M., and R. W. Garvine. 1988. Summertime hydrography at the shelfbreak front in the Middle Atlantic Bight. J. Phys. Oceanogr. **18**:1309-19.

Chapman, D. C., and G. Gawarkiewicz. 1993. On the establishment of the seasonal pycnocline in the Middle Atlantic Bight. J. Phys. Oceanogr. **23**:2487-92.

Cilia, V., B. Lafay, and R. Christen. 1996. Sequence heterogeneities among 16S ribosomal RNA sequences, and their effect on phylogenetic analyses at the species level. Mol. Biol. Evol. **13**:451-61.

Dams, E., L. Hendricks, Y. Van de Peer, J. Neefs, G. Smits, I. Vandenbempt, and R. De Wachter. 1988. Compilation of small ribosomal subunit RNA sequences. Nucleic Acids Res. **16**:87-183.

DeLong, E. F. 1992. Archaea in coastal marine environments. Proc. Natl. Acad. Sci. USA **89**:5685-9.

DeLong, E. F., D. G. Franks, and A. L. Alldredge. 1993. Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. Limnol. Oceanogr. **38**:924-34.

Distel, D. L., D. J. Lane, G. J. Olsen, S. J. Giovannoni, B. Pace, N. R. Pace, D. A. Stahl, and H. Felbeck. 1988. Sulfur-oxidizing bacterial endosymbionts: analysis of phylogeny and specificity by 16S rRNA sequences. J. Bacteriol. **170**:2506-10.

Douglas, S. E., and S. Turner. 1991. Molecular evidence for the origin of plastids from a cyanobacterium-like ancestor. J. Mol. Evol. **33**:267-73.

Ducklow, H. W. 1994. Modeling the microbial food web. Microb. Ecol. **28**:303-19.

Farrelly, V., F. A. Rainey, and E. Stackebrandt. 1995. Effect of genome size and rrn gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. Appl. Environ. Microbiol. **61**:2798-2801.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783-91.

Felsenstein, J. 1989. PHYLIP - phylogeny inference package (v3.5). Cladistics **5**:164-6.

Field, K. G., D. Gordon, M. S. Rappé, E. Urbach, K. Vergin, and S. J. Giovannoni. 1997. Diversity and depth-specific distribution of SAR11 cluster rRNA genes from marine planktonic bacteria. Appl. Environ. Microbiol. **63**:63-70.

Fuhrman, J. A., K. McCallum, and A. A. Davis. 1993. Phylogenetic diversity of subsurface marine microbial communities from the atlantic and pacific oceans. Appl. Environ. Microbiol. **59**:1294-1302.

Gherna, R., and C. R. Woese. 1992. A partial phylogenetic analysis of the "Flavobacter-Bacteroides" phylum: basis for taxonomic restructuring. System. Appl. Microbiol. **15**:513-21.

Giovannoni, S. J. 1991. The polymerase chain reaction, p. 175-203. In E. Stackebrandt and M. Goodfellow [eds.], Nucleic acid techniques in bacterial systematics. Wiley.

Giovannoni, S. J., T. B. Britschgi, C. L. Moyer, and K. G. Field. 1990a. Genetic diversity in Sargasso Sea bacterioplankton. Nature (Lond.) **345**:60-3.

Giovannoni, S. J., E. F. DeLong, G. J. Olsen, and N. R. Pace. 1988a. Phylogenetic group-specific oligodeoxynucleotide probes for the identification of single microbial cells. J. Bacteriol. **170**:720-6.

Giovannoni, S. J., E. F. DeLong, T. M. Schmidt, and N. R. Pace. 1990b. Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. Appl. Environ. Microbiol. **56**:2572-5.

Giovannoni, S. J., T. D. Mullins, and K. G. Field. 1995. Microbial diversity in oceanic systems: rRNA approaches to the study of unculturable microbes, p. 217-48. In I. Joint [ed.], Molecular ecology of aquatic microbes. Springer-Verlag.

Giovannoni, S. J., M. S. Rappé, D. Gordon, E. Urbach, M. Suzuki, and K. G. Field. 1996a. Ribosomal RNA and the evolution of bacterial diversity. In [eds.], Evolution of microbial life. Cambridge. In press.

Giovannoni, S. J., M. S. Rappé, K. L. Vergin, and N. L. Adair. 1996b. 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the Green Non-Sulfur bacteria. Proc. Natl. Acad. Sci. USA. **93**:7979-84.

Giovannoni, S. J., S. Turner, G. J. Olsen, S. Barns, D. J. Lane, and N. R. Pace. 1988b. Evolutionary relationships among cyanobacteria and green chloroplasts. J. Bacteriol. **170**:3584-92.

Glover, H. E., L. Campbell, and B. B. Prézelin. 1986. Contribution of Synechoccus spp. to size-fractionated primary productivity in three water masses in the Northwest Atlantic Ocean. Mar. Biol. **91**:193-203.

Good, I. J. 1953. The population frequencies of species and the estimation of the population parameters. Biometrika **40**:237-64.

Gordon, D. A., and S. J. Giovannoni. 1996. Stratified microbial populations related to Chlorobium and Fibrobacter detected in the Atlantic and Pacific oceans. Appl. Environ. Microbiol. **62**:1171-7.

Gutell, R. R. 1994. Collection of small subunit (16S- and 16S-like) ribosomal RNA structures: 1994. Nucleic Acids Res. **22**:3502-7.

Hiraishi, A. 1994. Phylogenetic affiliations of Rhodoferax fermentans and related species of phototrophic bacteria as determined by automated 16S rDNA sequencing. Curr. Microbiol. 28:25-9.

Hiraishi, A., Y. Hoshino, and T. Satoh. 1991. Rhodoferax fermentans gen. nov., sp. nov., a phototrophic purple nonsulfur bacterium previously referred to as the "Rhodocyclus gelatinosus-like" group. Arch. Microbiol. 155:330-6.

Irgens, R., I. Suzuki, and J. T. Staley. 1989. Gas vacuolate bacteria obtained from marine waters of Antarctica. Curr. Microbiol. 18:261-5.

Jenkins, O., D. Byrom, and D. Jones. 1987. Methylophilus: a new genus of methanol-utilizing bacteria. Int. J. Syst. Bacteriol. 37:446-8.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16:111-20.

Kopczynski, E. D., M. M. Bateson, and D. M. Ward. 1994. Recognition of chimeric small-subunit ribosomal DNAs composed of genes from uncultivated microorganisms. Appl. Environ. Microbiol. 60:746-8.

Lane, D. J. 1991. 16S/23S rRNA sequencing, p. 115-48. In E. Stackebrandt and M. Goodfellow [eds.], Nucleic acid techniques in bacterial systematics. Wiley.

Maidak, B. L., N. Larsen, M. J. McCaughey, R. Overbeek, G. J. Olsen, K. Fogel, J. Blandy, and C. R. Woese. 1994. The ribosomal database project (RDP). Nucleic Acids Res. 22:3485-7.

Marshall, H. G., and J. A. Ranasinghe. 1989. Phytoplankton distribution along the eastern coast of the U.S.A.-VII. Mean cell concentrations and standing crop. Continental Shelf Res. 9:153-64.

Mullins, T. D., T. B. Britschgi, R. L. Krest, and S. J. Giovannoni. 1995. Genetic comparisons reveal the same unknown bacterial lineages in Atlantic and Pacific bacterioplankton communities. Limnol. Oceanogr. 40:148-58.

Murphy, L. S., and E. M. Haugen. 1985. The distribution and abundance of phototrophic ultraplankton in the North Atlantic. Limnol. Oceanogr. 30:47-58.

Olsen, G. J., H. Matsuda, R. Hagstrom, and R. Overbeek. 1994. FastDNAml: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. Comput. Appl. Biosci. 10:41-8.

Olsen, G. J., D. J. Lane, S. J. Giovannoni, and N. R. Pace. 1986. Microbial ecology and evolution: a ribosomal RNA approach. Ann. Rev. Microbiol. **40**:337-65.

Paster, B. J., W. Ludwig, W. G. Weisburg, E. Stackebrandt, R. B. Hespell, C. M. Hahn, H. Reichenbach, K. O. Stetter, and C. R. Woese. 1985. A phylogenetic grouping of the bacteroides, cytophagas, and certain flavobacteria. System. Appl. Microbiol. **6**:34-42.

Rappé, M. S., P. F. Kemp, and S. J. Giovannoni. 1995. Chromophyte plastid 16S ribosomal RNA genes found in a clone library from Atlantic ocean seawater. J. Phycol. **31**:979-88.

Rehnstam, A.-S., S. Bäckman, D. C. Smith, F. Azam, and Å. Hagström. 1993. Blooms of sequence-specific culturable bacteria in the sea. FEMS Microbiol. Ecol. **102**:161-6.

Robinson-Cox, J. F., M. M. Bateson, and D. M. Ward. 1995. Evaluation of nearest-neighbor methods for detection of chimeric small-subunit rRNA sequences. Appl. Environ. Microbiol. **61**:1240-5.

Rowe, G., S. Smith, P. Falkowski, T. Whitledge, R. Theroux, W. Phoel, and H. Ducklow. 1986. Do continental shelves export organic matter? Nature (Lond.) **324**:559-61.

Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science (Wash. D.C.) **239**:487-91.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406-25.

Schlesner, H., and E. Stackebrandt. 1986. Assignment of the genera Planctomyces and Pirella to a new family Planctomycetaceae fam. nov. and description of the order Planctomycetales ord. nov. Syst. Appl. Microbiol. **8**:174-6.

Schmidt, T. M., E. F. DeLong, and N. R. Pace. 1991. Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. J. Bacteriol. **173**:4371-8.

Sherman, K., M. Grosslein, D. Mountain, D. Busch, J. O'Reilly, and R. Theroux. 1988. The continental shelf ecosystem off the northeast coast of the United

States, p. 279-337. In H. Postma and J. J. Zijlstra [eds.], Continental shelves. Elsevier.

Shiba, T. 1991. Roseobacter litoralis gen. nov., sp. nov., and Roseobacter denitrificans sp. nov., aerobic pink-pigmented bacteria which contain bacteriochlorophyll a. System. Appl. Microbiol. **14**:140-5.

Suzuki, M., and S. J. Giovannoni. 1996. Mechanisms of bias in the amplification of 16S rRNA genes by PCR. Appl. Environ. Microbiol. **62**:625-30.

Tsuji, K., H. C. Tsien, R. S. Hanson, S. R. DePalma, R. Scholtz, and S. LaRoche. 1990. 16S ribosomal RNA sequence analysis for determination of phylogenetic relationship among methylotrophs. J. Gen. Microbiol. **136**:1-10.

Verity, P. G., G. A. Paffenhöfer, D. Wallace, E. Sherr, and B. Sherr. 1996. Composition and biomass of plankton in spring on the Cape Hatteras shelf, with implications for carbon flux. Continental Shelf Res. **16**:1087-116.

Wakabayashi, H., M. Hikida, and K. Masumura. 1986. Flexibacter maritimus sp. nov., a pathogen of marine fishes. Int. J. Syst. Bacteriol. **36**:396-8.

Walsh, J. J. 1988. On the nature of continental shelves. Academic Press.

Walsh, J. J. 1991. Importance of continental margins in the marine biogeochemical cycling of carbon and nitrogen. Nature (Lond.) **350**:53-5.

Walsh, J. J., E. T. Premuzic, J. S. Gaffney, G. T. Rowe, G. Harbottle, R. W. Stoenner, W. L. Balsam, P. R. Betzer, and S. A. Macko. 1985. Organic storage of CO_2 on the continental slope off the mid-Atlantic bight, the southeastern Bering Sea, and the Peru coast. Deep-Sea Res. **32**:853-83.

Ward, D. M., M. M. Bateson, R. Weller, and A. L. Ruff-Roberts. 1992. Ribosomal RNA analysis of microorganisms as they occur in nature. Adv. Microb. Ecol. **12**:219-86.

Ward, N., F. A. Rainey, E. Stackebrandt, and H. Schlesner. 1995. Unraveling the extent of diversity within the order Planctomycetales. Appl. Environ. Microbiol. **61**:2270-5.

Willems, A., J. De Ley, M. Gillis, and K. Kersters. 1991. Comamonadaceae, a new family encompassing the acidovorans rRNA complex, including Variovorax paradoxus gen. nov., comb. nov., for Alcaligenes paradoxus (Davis 1969). Int. J. Syst. Bacteriol. **41**:445-50.

Woese, C. R. 1987. Bacterial evolution. Microbiol. Rev. **51**:221-71.

Woese, C. R., E. Stackebrandt, W. G. Weisburg, B. J. Paster, M. T. Madigan, V. J. Fowler, C. M. Hahn, P. Blanz, R. Gupta, K. H. Nealson, and G. E. Fox. 1984a. The phylogeny of purple bacteria: the alpha subdivision. System. Appl. Microbiol. **5**:315-326.

Woese, C. R., W. G. Weisburg, C. M. Hahn, B. J. Paster, L. B. Zablen, B. J. Lewis, T. J. Macke, W. Ludwig, and E. Stackebrandt. 1985. The phylogeny of purple bacteria: the gamma subdivision. System. Appl. Microbiol. **6**:25-33.

Woese, C. R., W. G. Weisburg, B. J. Paster, C. M. Hahn, R. S. Tanner, N. R. Krieg, H.-P. Koops, H. Harms, and E. Stackebrandt. 1984b. The phylogeny of purple bacteria: the beta subdivision. System. Appl. Microbiol. **5**:327-36.

Xia, Y., T. M. Embley, and A. G. O'Donnell. 1994. Phylogenetic analysis of Azospirillum by direct sequencing of PCR amplified 16S rDNA. System. Appl. Microbiol. **17**:197-201.

CHAPTER 4

PHYLOGENETIC DIVERSITY OF ULTRAPLANKTON RECOVERED IN ENVIRONMENTAL NUCLEIC ACID SAMPLES FROM THE PACIFIC AND ATLANTIC COASTS OF THE UNITED STATES

Michael S. Rappé, Marcelino T. Suzuki, Kevin Vergin, and Stephen J. Giovannoni

submitted to *Applied and Environmental Microbiology*,
American Society for Microbiology,
Washington, D.C.

4.1 Abstract

The scope of marine phytoplankton diversity is in many respects uncertain because, like bacteria, these organisms sometimes lack defining morphological characteristics and can be a challenge to grow in culture. Here we report the recovery of phylogenetically diverse plastid SSU rRNA gene clones were recovered from natural plankton populations collected in the Pacific Ocean off the mouth of Yaquina Bay, Oregon (OCS clones), and from the eastern continental shelf of the United States off Cape Hatteras, North Carolina (OM clones). The SSU rDNA clone libraries were prepared by amplifying rRNA genes from nucleic acids isolated from the plankton samples, and cloning them into plasmid vectors. The PCR primers used for the amplification reactions were designed to be specific for bacterial SSU rRNA genes; however, plastid genes share a common phylogenetic origin with bacteria and were common in both SSU rRNA gene clone libraries. A combination of restriction fragment length polymorphism analyses, nucleic acid sequencing, and taxon-specific oligonucleotide probe hybridizations revealed that 54 of the 116 SSU rRNA gene clones in the Oregon Coast Study (OCS) were of plastid origin. Collectively, the clones from the OCS and OM libraries formed at least eight unique lineages within the plastid radiation, including gene lineages related to the Bacillariophyceae, Cryptophyceae, Prymnesiophyceae, Chrysophyceae, and Prasinophyceae, as well as a number of unique clones for which close phylogenetic neighbors could not be identified with confidence. Only a group of two OCS rDNA clones showed close identity to the plastid SSU rRNA gene sequence of a cultured organism (*Emiliania huxleyi* (Lohmann) Hay et Mohler, 99.8% similar). The remaining clones could not be identified to the genus or species level. Although cryptic

species are not as prevalent among phytoplankton as among their bacterial counterparts, this genetic survey nonetheless uncovered significant new information about phytoplankton diversity, presumably because it circumvented the pitfalls of systematics based on morphology alone.

4.2 Introduction

Eukaryotic phytoplankton in the ultraplankton size range (cells < 5-10 μm in diameter) are recognized as important members of the photosynthetic assemblage in the world's oceans (Hoepffner and Haas 1990, Johnson and Sieburth 1982, Joint 1986, Murphy and Haugen 1985). These small cells make up a majority of the eukaryotic plankton cells present in the phytoplankton mass in many, if not most pelagic marine habitats including the oligotrophic open-ocean and ocean margin environments (Courties et al. 1994).

The identification of ultraphytoplankton, either in natural communities or clonal cultures, is often hindered by their small size and lack of taxonomically informative morphological features. In addition, many of the smaller plankton cells are difficult to or have thus far completely evaded culture. The taxonomic identification of small eukaryotic phytoplankton often must involve the laboreous techniques of electron microscopy, HPLC, or immunological cross-reactivity. Identification of individual cells at the class level is often difficult at best, and is rarely accomplished with the ease and speed neccessary to analyze large numbers of field samples.

More recently, the identification of small eukaryotic phytoplankton has benefited from the application of molecular techniques. Many of the more recent descriptions of phytoplankton taxa include sequence analyses of

conserved genes such as the nuclear-encoded 18S rRNA gene (Andersen et al. 1993). As well as aiding in the placement of the newly described microorganism in a phylogenetic and taxonomic framework, the use of conserved gene sequences provides information that can be employed in the construction of taxon-specific oligonucleotide probes (Knauber et al. 1996, Lange et al. 1996, Miller and Scholin 1996, Simon et al. 1995).

Nucleic acid probing has become a common approach in assessing the diversity of microorganisms in natural communities (Amann et al. 1995). In pelagic marine environments, the focus has been on prokaryotic plankton, including organisms from the domains *Bacteria* and *Archaea*. The findings have been dramatic - many novel bacterioplankton and archaea have been described by only their rDNA sequences, most having thus far evaded culuture in the laboratory. Oligonucleotide probes based on SSU rRNA gene sequence signatures are now being designed for widespread application to ecological problems, including determining the distribution of prokaryotic plankton in their natural habitats (Alfreider et al. 1996, DeLong et al. 1994, Giovannoni et al. 1996, Gordon and Giovannoni 1996). rRNA probes constructed from taxonomic groups within the ultraplankton size class can be used in future studies to identify cells rapidly in field samples (Simon et al. 1995).

This study was undertaken to examine the diversity of bacterioplankton in coastal Oregon seawater. The Oregon coast represents an unusual area of high productivity with wind-driven coastal upwelling processes seasonally dominating the region's oceanography (Huyer 1977, Huyer 1983). Similar hydrographical events occur in both the northern and southern hemispheres along eastern ocean boundaries (Barber and Smith 1981). The high nutrient availability in these regions results in high productivity, such that they may be

responsible for half the world's catch of fish even though they represent only 0.1% of the world's surface area (Ryther 1969). High phytoplankton abundance and elevated levels of primary production are characteristic of the Oregon coast during upwelling events (Hood et al. 1992, Small and Menzies 1981). The fortuitous recovery of plastid-derived SSU rRNA gene clones from two clone libraries prepared from natural plankton populations off the Pacific and Atlantic coasts of the United States allowed us to assess the phylogenetic diversity of phytoplankton present in these locales independent of the ability to culture cells present in the original samples. Herein, we show that most of the plastid-dervided rDNA clones recovered belong to several classes of eukaryotic algae which are known to contain marine phytoplankton members (Chrétiennot-Dinet et al. 1993). However, several rDNA clones did not affiliate with any currently available plastid SSU rRNA gene sequences.

4.3 Materials and methods

Construction of the Oregon Coast Study (OCS) rDNA clone library has been partially described in (Suzuki et al. 1997), while construction of the Ocean Margins (OM) rDNA clone library has been described by Rappé and co-workers (Rappé et al. 1995, Rappé et al. 1997). Herein, we provide details of the construction and analysis of the OCS clone library not provided elsewhere.

Sample collection and nucleic acid isolation. On 28 April 1993, a 16 L plankton sample was collected from a depth of 10 m at a station located 8 km west of Yaquina Bay, Oregon (44°39.1'N, 124°10.6'W). Water depth was ca. 50 m at the sampling site. The water was pre-screened through 10 μm nylon mesh and immediately transported in autoclaved polyethylene carboys to the

laboratory, where the plankton sample was collected by filtration onto 0.2 μm polysulfone filters (Supor-200, Gelman Inc., Ann Arbor, MI). Total cellular nucleic acids were isolated by cell lysis with proteinase K and SDS, and phenol/chloroform extraction as previously described (Giovannoni et al. 1990, Suzuki et al. 1997).

Clone library construction. SSU rRNA genes were amplified from the environmental sample of genomic DNAs using two general bacterial SSU rRNA primers, *Pfu* DNA polymerase (Stratagene, La Jolla, CA), and the polymerase chain reaction (PCR) (Saiki et al. 1988) as previously described (Suzuki et al. 1997). The amplification products from six reactions were pooled and inserted into the *Sma*I restriction site of the phagemid vector pBluescript KSII- (Stratagene) by blunt-end ligation as previously described (Giovannoni et al. 1990, Suzuki et al. 1997), and used to transform competent *E. coli* XL1-blue cells (Stratagene). Positive (white colony morphotype) transformants were streaked for isolation, and stored in both stab cultures and DMSO. Clones were numbered discontinuously from 1-182 and assigned the prefix OCS (Oregon Coast Study).

Restriction fragment length polymorphism (RFLP) analysis. All gene clones containing full-length inserts were characterized by *Hae*III RFLPs as previously described (Suzuki et al. 1997). Briefly, insert-bearing plasmids, isolated by alkaline lysis, were used as templates in the amplification of environmental clone rDNAs by the PCR employing the same primers as used previously to amplify the rRNA genes. Since all of the PCR reactions yielded similar amounts of product, 7 μl of non-purified PCR products were digested with 3 units of the restriction endonuclease *Hae*III (Promega) for 2 h. The restriction fragments were resolved by gel electrophoresis in 3% NuSieve low

melting point agarose (FMC, Rockland, ME) in 1X TAE and stained with ethidium bromide (0.5 $\mu g \cdot mL^{-1}$).

rRNA gene sequencing and phylogenetic analyses. Template plasmid DNAs for sequencing were prepared by alkaline lysis as previously described (Suzuki et al. 1997). Clone sequences submitted to Genbank were sequenced on both stands with either an ABI model 377 or 373a automated sequencer (Applied Biosystems, Foster City, CA), dye-terminator chemistry, and conserved bacterial SSU rRNA primers (Lane 1991). DNA sequence data from cloned SSU rDNA genes were manually aligned to bacterial and plastid sequences obtained from Genbank and the Ribosomal Database Project (RDP) (Maidak et al. 1994) using the Genetic Data Environment (GDE) v2.2 sequence analysis software package.

Evolutionary trees were constructed by distance, maximum parsimony, and maximum likelihood methods. Each phylogenetic analysis employed conservative phylogenetic masks which only included regions of unambiguous alignment. In addition, all phylogenetic analyses were repeated multiple times with different taxa employed as an outgroup. The Phylogeny Inference Package (PHYLIP v3.5c) (Felsenstein 1989) was used for phylogenies constructed from evolutionary distance matrices. Evolutionary distances were calculated from pair-wise sequence similarities with the Kimura 2-parameter model for nucleotide change and a transition/transversion ratio of 2.0 (Kimura 1980) using the program DNADIST. The program NEIGHBOR was used to reconstruct phylogenetic trees from the evolutionary distance matrices by the neighbor-joining method (Saitou and Nei 1987). Relative confidence in monophyletic groups withing each phylogenetic analysis was estimated by

bootstrap analyses, which included 100 replicate resampled data sets with random sequence addition and global rearrangement (Felsenstein 1985).

Test versions of PAUP* written by David L. Swofford was used for all maximum parsimony, maximum likelihood, and LogDet phylogenetic analyses. Maximum parsimony and maximum parsimony bootstrap analyses were performed by a heuristic search method, with tree bisection-reconnection and MULPARS options. Maximum likelihood calculations employed nucleotide frequencies estimated from the data, with a transition/trasversion ratio of 2.0. Bootstrap analyes were performed as above.

Secondary structure analysis and identification of chimeras. Several methods were employed to verify the integrity of the sequence data and aid in the detection of chimeric gene artifacts. First, the program gRNAid v1.4, provided by Shannon Whitmore (Mentor Graphics, Wilsonville, OR) was used to construct secondary structural models for all sequenced SSU rDNAs. These models were manipulated manually for optimal agreement with published SSU rRNA secondary structures (Gutell 1994), and the nature and location of mutations was analyzed. Second, all gene sequences were submitted to the RDP program CHECK_CHIMERA (Maidak et al. 1994), which can be fairly proficient at detecting chimeric gene artifacts created from parent sequences that are no greater than ca. 84% similar (Robinson-Cox et al. 1995). Phylogenetic trees were also constructed and compared from different regions within single rRNA clones, under the assumption that if two regions of an rDNA clone originated from different parental rRNA genes, they will separate into different lineages that reflect the phylogeny of their different parent rRNA sequences (Field et al. 1997, Robinson-Cox et al. 1995).

Oligonucleotide probe hybridizations. To help in identifying unique SSU rDNA clones, taxon-specific oligonucleotide probes were used to screen a dot blot of 32 clones representing 23 RFLP patterns. For the selected clones, PCR amplicons from the RFLP analysis were purified with a Qiaquick-spin PCR purification kit (Qiagen, Chatsworth, CA). For each amplicon, 30 ng of product were resuspended in 180 µl of TE buffer. Products were denatured by the addition of 20 µl of 2.0 N NaOH and incubated 10 min at room temperature prior to blotting on Zetaprobe nylon membrane (BioRad). Each well was washed with 200 µl 2x SSPE (1x SSPE: 0.18 M NaCl, 10 mM $NaPO_4$, pH7.2, 1mM EDTA) to neutralize the NaOH. The membrane was dried under vacuum at 80°C for 15 min, exposed to 120 Joule/m^2 of 260 nm UV radiation, and stored desiccated before probing.

The following oligonucleotide probes were hybridized to the OCS unknown dotblot: 338R (T_H=45°) (Amann et al. 1990), SAR83R (T_H=45°) (Rappé et al. 1997), SAR11A1 (T_H= 37°C) (Field et al. 1997), SAR86 (T_H=45°C) (Rappé et al. 1997), CHRYS1 (T_H=45°C) (Rappé et al. 1995), and HAP1 (T_H=45°C) (Rappé et al. 1995). The T_H (probe hybridization stringency temperature) for each probe was determined empirically by washing at successively higher temperatures. All probes used in this study were constructed with an Applied Biosystems automated DNA synthesizer and were [^{32}P]-labeled on their 5' termini with T4 polynucleotide kinase as previously described (Giovannoni et al. 1988).

Prior to each probe hybridization, the OCS unknown blot was prehybridized in 15 mL Z-hyb buffer (1.0 mM EDTA, 0.50 mM NaH_2PO_4 pH 7.2, 7.0% SDS) for 45-60 min at room temperature. After decanting the pre-hybridization buffer, the blot was hybridized in 6.0 mL Z-hyb buffer

containing 200 µL of [32P]-labeled oligonucleotide (25-50 ng of oligonucleotide) for 8-16 h at room temperature. The blot was washed three times for 15 min each in 25 mL wash buffer [0.2X SSPE (150 mM NaCl, 1.0 mM EDTA, 10 mM NaH_2PO_4 H_2O pH 7.4), 0.1% SDS] at room temperature, and one time for 30 min at the probe T_H. The blot was stripped of probe by washing three times in 25-50 mL wash buffer at 70°C for 10 min each wash. Individual probe hybridizations were visualized using a Molecular Dynamics PhosphorImager SI (Molecular Dynamics, Sunnyvale, CA) as well as autoradiographically with X-ray film.

LH-PCR genescan. Genomic DNA was extracted as described above from two 4 liter seawater subsamples taken from the same sample used in the construction of the OCS clone library. The first subsample received no treatment, whereas the second subsample was screened through a 0.8 µm polycarbonate membrane (Poretics, Livermore, CA). 10 ng of purified genomic DNA from each subsample was used as template for PCR. The forward primer, 27F (5'-AGAGTTTGATCMTGGCTCAG-3'), was 5'-end-labeled with the phophoramidite dye 6-FAM, supplied by Applied Biosystems Inc. (ABI). The general bacterial primer 338R (5'-GCTGCCTCCCGTAGGAGT-3') was employed as the reverse primer. Labeled PCR products were purified using Qiaquick-spin columns (Qiagen). Ten fmoles of DNA were discriminated by Long Ranger (FMC, Rockland, ME) PAGE in an ABI 377 automated DNA sequencer (ABI) in Genescan® mode. Genescan estimates the lengths of bands in the gel, measures the integrated fluorescence emission by the bands, and outputs an electropherogram in which the bands are represented by peaks and the integrated fluorescence of each band is the area under the peaks. Since the integrated fluorescence increased linearly with concentration up to 50 fmol of

PCR products (data not shown), it was assumed that the relative proportion of the integrated fluorescence of each peak corresponded to the proportion of each amplicon in the PCR products. We then calculated the ratio between the integrated fluorescence of each of the peaks and the total integrated fluorescence of all peaks.

Accession numbers. Nucleotide sequences were filed in Genbank under the following accession numbers: OCS20, AF001654; OCS31, AF001655; OCS50, AF001656; OCS54, AF001657; OCS56, AF001658; OCS162, AF001659; OCS182, AF001660; OM134, AF001661.

Genbank accession numbers were obtained previously for the following environmental clone rDNAs used in this study: OM20 (U32670) and OM21 (U32671) (Rappé et al. 1995); OM5 (U70715), OM39 (U70716), OM81 (U70717), OM111 (U70718), OM125 (U70719), OM153 (U70720), OM164 (U70721), OM255 (U70722), OM270 (U70723), OM283 (U70724) (Rappé et al. 1997).

4.4 Results

The OCS clone library. Of 182 putative positive transformants (white colonies) in the OCS library, 116 clones were shown to have inserts of the correct size (1.5 Kb) by gel electrophoresis of both plasmids and amplicons derived from plasmids (data not shown). *Hae*III RFLP analyses revealed 37 different banding patterns distributed among the OCS clones, with 19 OCS clones containing unique *Hae*III RFLP patterns. Except where noted below, a minimum of two clones from all RFLP types containing two or more clones, as well as all clones with unique RFLP banding patterns, were sequenced. Cursory phylogenetic analyses identified 54 OCS SSU rDNA clones (8 different

RFLP banding patterns) as members of the plastid line of descent. An additional clone with a unique RFLP banding pattern was identified as plastid-related based on its hybridization with a taxon-specific nucleotide probe.

Prymnesiophyceae. SSU rRNA gene sequences from prymnesiophyte algae plastids formed a monophyletic cluster, strongly supported by both neighbor joining (100 of 100 replicates) and maximum parsimony (100 of 100 replicates) bootstrap analyses, within the line of descent that includes rhodophyte plastids and plastids derived from secondary symbioses (Fig. 4.1). Sequenced clones from two RFLP types within the OCS library (RFLP types I and IX) were found to be related to prymnesiophyte algae plastid SSU rRNA genes (Figs. 4.1, 4.2; Table 4.1). RFLP type I, represented by OCS31 in Figs. 4.1 and 4.2, consisted of 30 clones, or 56% of the plastid SSU rDNA clones recovered in the OCS clone library. OCS31 was most closely related to OM21, an SSU rDNA clone recovered off the coast of Cape Hatteras, North Carolina (97.6% similar for 1470 nucleotide positions). These two environmental clones form a monophyletic cluster in the phylogenies depicted in Figs. 4.1 and 4.2, though bootstrap support for this cluster was low for full length sequences under maximum parsimony criteria (69 of 100 replicates).

RFLP type IX consisted of 3 clones in the OCS library (6% of the plastid clones; Table 4.1), and is represented by OCS50 in Fig. 4.2. OCS50 was closely related to the plastid SSU rRNA gene sequence available for the prymnesiophyte *Emiliania huxleyi* (Lohmann) Hay et Mohler (99.8% similar for 955 nucleotide positions). Bootstrap support for the monophyletic clustering of OCS50 and *Emiliania huxleyi* was high in both the neighbor joining (100 of 100

Table 4.1. Frequency and phylogenetic affiliation of plastid-derived rDNA clones in the OCS library.

Table 4.1

RFLP type	Phylogenetic affiliation	No. recovered	Representative clones*	Sequence length†	Genbank access. #
I	Prymnesiophyceae	30	OCS31	1472	AF001655
IX		3	OCS50	956	AF001656
III	Bacillariophyceae	9	OCS54	1478	AF001657
XI		2	OCS56	554	AF001658
XXX		1	OCS129	n.d.‡	
V	Cryptophyceae	6	OCS20	1479	AF001654
XVII	Prasinophyceae	2	OCS162	1490	AF001659
XXXVI		1	OCS182	1482	AF001660

*identical sequences were not filed in Genbank.

†includes sequence for the primers used in the amplification of environmental nucleic acids

‡n.d. = no sequence was determined.

Figure 4.1. Phylogenetic relationships between OCS and OM plastid rDNA clones and the plastid SSU rRNA gene sequences from cultured organisms. This phylogenetic tree was constructed by the neighbor joining method using a LogDet matrix as input (Lockhart et al. 1994). A conserved mask of 987 nucleotides spanning the entire length of the plastid SSU rRNA gene was included in the analyses. Bootstrap values (100 replications) generated with the neighbor joining method and the Kimura 2-parameter model for nucleotide change are shown above relevent nodes, while values from a maximum parsimony bootstrap analysis are shown below relevent nodes. All sequences except *Corethron criophilum* (Medlin, unpubl. data), *Cyanidium caldarium* (Giovannoni, unpubl. data), *Closterium ehrenbergii, Chlamydomonas parkeae, Yamagishiella unicocca,* and *Pyramimonas parkeae* (Kim et al. 1994) are available from the Genbank or RDP database.

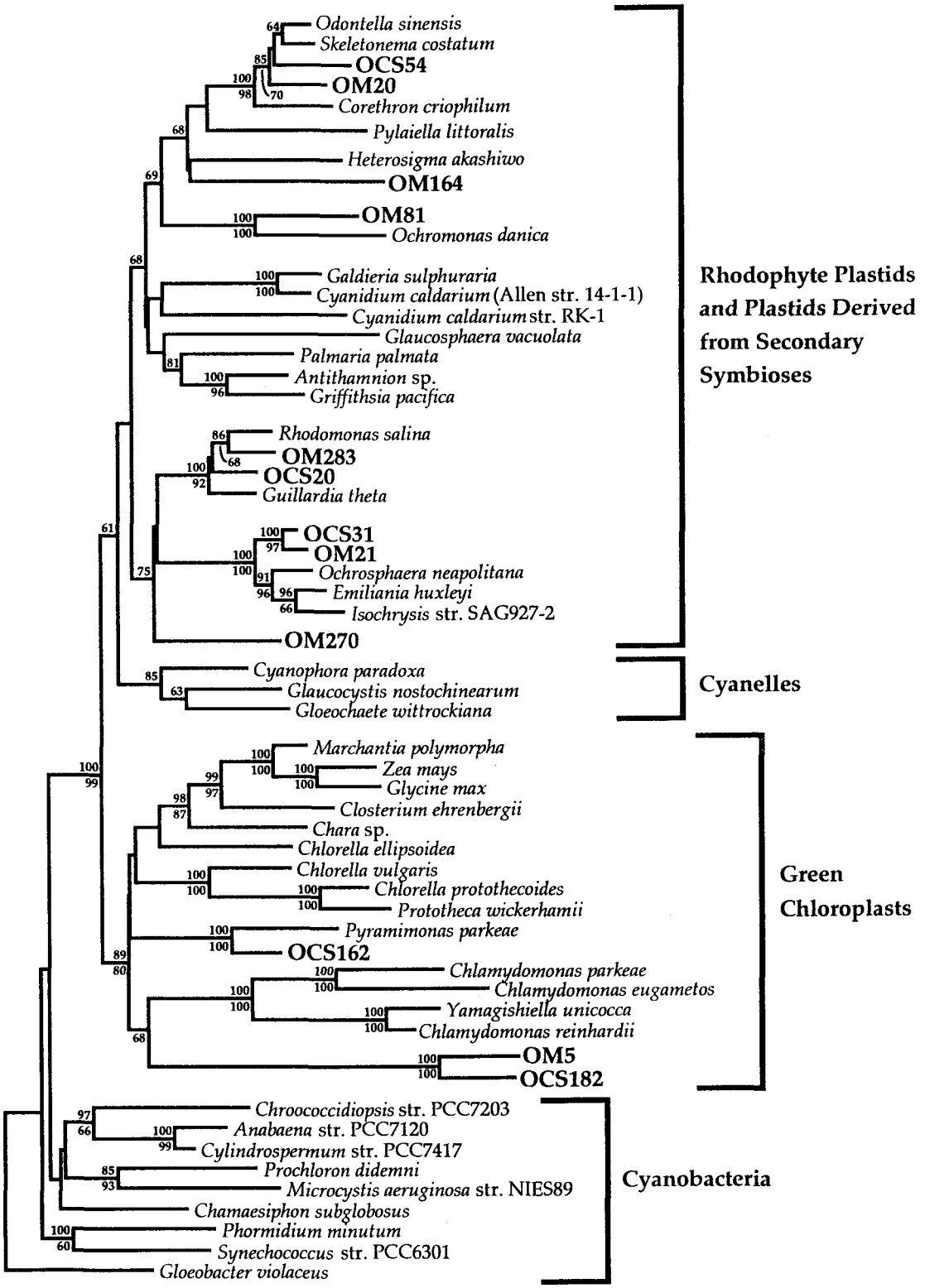

Odontella sinensis
Skeletonema costatum
OCS54
OM20
Corethron criophilum
Pylaiella littoralis
Heterosigma akashiwo
OM164
OM81
Ochromonas danica
Galdieria sulphuraria
Cyanidium caldarium (Allen str. 14-1-1)
Cyanidium caldarium str. RK-1
Glaucosphaera vacuolata
Palmaria palmata
Antithamnion sp.
Griffithsia pacifica
Rhodomonas salina
OM283
OCS20
Guillardia theta
OCS31
OM21
Ochrosphaera neapolitana
Emiliania huxleyi
Isochrysis str. SAG927-2
OM270
Cyanophora paradoxa
Glaucocystis nostochinearum
Gloeochaete wittrockiana
Marchantia polymorpha
Zea mays
Glycine max
Closterium ehrenbergii
Chara sp.
Chlorella ellipsoidea
Chlorella vulgaris
Chlorella protothecoides
Prototheca wickerhamii
Pyramimonas parkeae
OCS162
Chlamydomonas parkeae
Chlamydomonas eugametos
Yamagishiella unicocca
Chlamydomonas reinhardii
OM5
OCS182
Chroococcidiopsis str. PCC7203
Anabaena str. PCC7120
Cylindrospermum str. PCC7417
Prochloron didemni
Microcystis aeruginosa str. NIES89
Chamaesiphon subglobosus
Phormidium minutum
Synechococcus str. PCC6301
Gloeobacter violaceus
Rhodophyte Plastids and Plastids Derived from Secondary Symbioses
Cyanelles
Green Chloroplasts
Cyanobacteria

Figure 4.1

Figure 4.2. Phylogenetic position of OCS and OM plastid rDNA clone sequences related to the Prymnesiophyceae. This phylogenetic tree was constructed by the maximum likelihood method (fDNAml) (Olsen et al. 1994). Bootstrap values (100 replicates) for both distance analyses (above the node) and maximum parsimony (below the node) are shown. A mask of 750 nucleotides corresponding to the 5' end of the plastid SSU rRNA gene was employed in the analysis. The plastid SSU rRNA gene sequences from the cryptophyte algae *Guillardia theta* and *Rhodomonas salina* were used to root the phylogenetic tree.

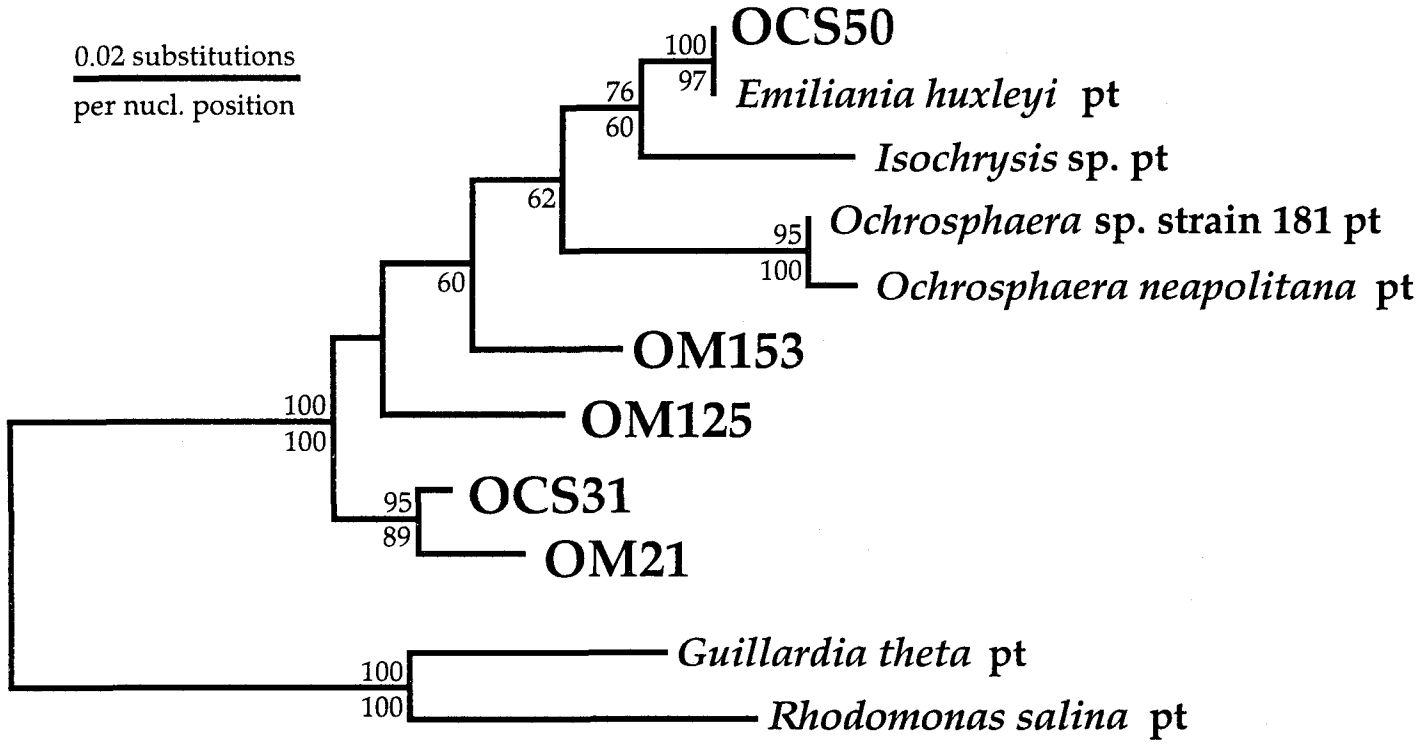

0.02 substitutions
per nucl. position
100
97
OCS50
76
60
Emiliania huxleyi pt
Isochrysis sp. pt
62
95
100
Ochrosphaera sp. strain 181 pt
Ochrosphaera neapolitana pt
60
OM153
100
100
OM125
95
89
OCS31
OM21
100
100
Guillardia theta pt
Rhodomonas salina pt

Figure 4.2

replicates) and maximum parsimony (97 of 100 replicates) bootstrap analyses (Fig. 4.2).

Bacillariophyceae. SSU rDNA gene sequences from bacillariophyte algae plastids formed a monophyletic cluster within the heterokont algae line of descent that was strongly supported by both the neighbor joining (100 of 100 replicates) and maximum parsimony (98 of 100 replicates) bootstrap analyses depicted in Fig. 4.1. Three RFLP types in the OCS library (RFLP types III, XI, and XXX) were related to bacillariophyte plastid SSU rRNA genes (Table 4.1). RFLP type III consisted of 9 clones (17% of the OCS plastid clones), and is represented by OCS54 in Fig. 4.1. Though not closely affiliated with any plastid SSU rRNA gene sequences available from cultured diatoms, OCS54 clearly emerged within the bacillariophyte SSU rDNA clade (Fig. 4.1). OCS56, representing RFLP type XI (2 clones, 4% of the OCS plastid clones), also did not specifically affiliate with any currently available bacillariophyte plastid SSU rRNA gene sequences, but clearly branched within the bacillariophyte SSU rDNA clade as well (data not shown). OCS129, a unique clone in the OCS library (RFLP type XXX), was identified as related to bacillariophyte plastid SSU rRNA genes by its hybridization with the bacillariophyte-specific probe CHRYS1 (data not shown), and was not sequenced.

Cryptophyceae. One RFLP type within the OCS library (RFLP type V; 6 clones or 11% of the plastid clones recovered) was related to cryptophyte algae SSU rRNA genes (Table 4.1). Represented by OCS20, RFLP type V branched within the cryptophyte plastid SSU rRNA gene clade, specifically affiliating with the cryptophyte *Rhodomonas salina* (95.9% similar for 1479 positions; Fig. 4.1). Together with OM283, a unique clone from the OM SSU rDNA clone library, OCS20 formed a monophyletic clade with the *Rhodomonas salina* and

Guillardia theta plastid SSU rRNA genes within the rhodophyte plastid and plastids derived from secondary symbioses line of descent in the plastid SSU rRNA gene tree depicted in Fig. 4.1. This clade received high bootstrap support in both the neighbor joining (100 of 100 replications) and maximum parsimony (92 of 100 replications) bootstrap analyses (Fig. 4.1).

Prasinophyceae. Three clones representing two RFLP types in the OCS clone library formed two separate lineages with the Chlorophyta (green chloroplast) line of descent (RFLP types XVII and XXXVI, Table 4.1). Two clones with RFLP type XVII were represented by OCS162 in Figs. 4.1 and 4.3 (Prasinophyte I lineage). OCS162 showed a close phylogenetic affiliation with the plastid SSU rRNA gene sequence from the prasinophyte algae *Pyramimonas parkeae* (Kim et al. 1994), well supported by both neighbor joining (100 of 100 replicates) and maximum parsimony (100 of 100 replicates) bootstrap analyses of full length SSU rRNA gene sequences (92.8% similar for 1414 positions; Fig. 4.1). Phylogenetic positioning of the prasinophyte clade represented by OCS162 and *Pyramimonas parkeae* within the green chloroplast line of descent could not be determined with confidence from the available green plastid SSU rRNA gene sequence data set (Fig. 4.3). Instead, this lineage appeared to diverge very early in the evolution of the green chloroplast line of descent under a wide range of phylogenetic reconstructions utilizing the full length SSU rRNA gene sequence, including maximum likelihood (Fig. 4.3), LogDet (Fig. 4.1), distance (Kimura 2-parameter model with neighbor joining; Figs. 4.1, 4.3), and maximum parsimony analyses (Figs. 4.1, 4.3).

A secondary structural model for the rRNA gene sequence of OCS162 was constructed to check for base-pairing idosyncrasities due to sequencing errors and chimeric gene formation (Kopczynski et al. 1994), as well as to

Figure 4.3. Phylogenetic position of representative OCS and OM rDNA clones related to the Chlorophyta line of descent. This phylogenetic tree was constructed by the maximum likelihood method (fDNAml) (Olsen et al. 1994). Bootstrap values (100 replicates) for both neighbor joining (above the node) and maximum parsimony (below the node) are shown. A conserved mask of 1133 nucleotides spanning the entire length of the plastid SSU rRNA gene was included in the analysis. The gene sequences of *Cyanophora paradoxa, Ochrosphaera neapolitana, Emiliania huxleyi,* and *Isochrysis* str. SAG927-2 were used to root the phylogenetic tree. All sequences except *Closterium ehrenbergii, Chlamydomonas parkeae, Yamagishiella unicocca, Gonium pectorale,* and *Pyramimonas parkeae* (Kim et al. 1994) are available from the Genbank or RDP database.

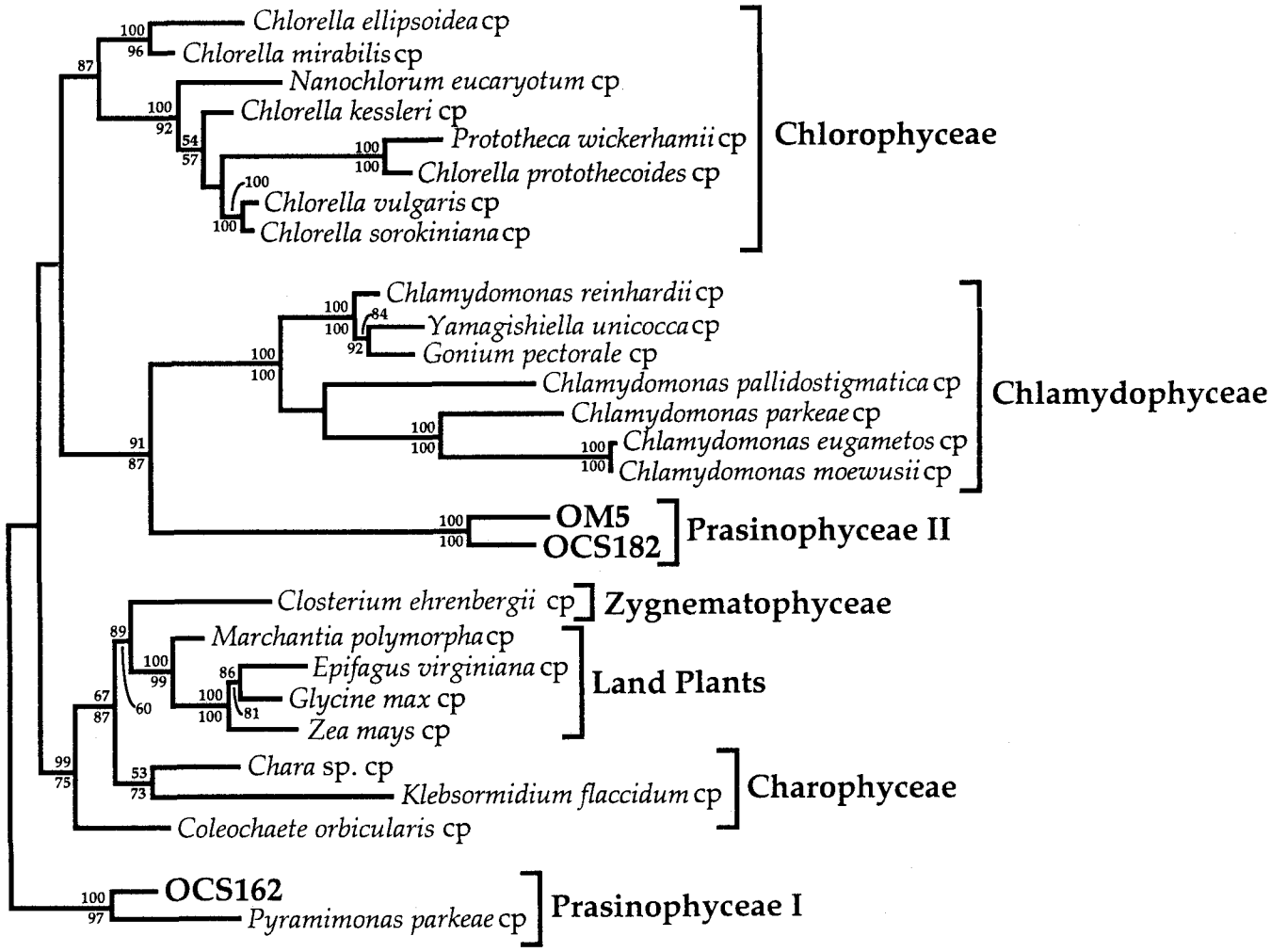

Chlorella ellipsoidea cp
Chlorella mirabilis cp
Nanochlorum eucaryotum cp
Chlorella kessleri cp
Prototheca wickerhamii cp
Chlorella protothecoides cp
Chlorella vulgaris cp
Chlorella sorokiniana cp
Chlorophyceae
Chlamydomonas reinhardii cp
Yamagishiella unicocca cp
Gonium pectorale cp
Chlamydomonas pallidostigmatica cp
Chlamydomonas parkeae cp
Chlamydomonas eugametos cp
Chlamydomonas moewusii cp
Chlamydophyceae
OM5
OCS182
Prasinophyceae II
Closterium ehrenbergii cp
Zygnematophyceae
Marchantia polymorpha cp
Epifagus virginiana cp
Glycine max cp
Zea mays cp
Land Plants
Chara sp. cp
Klebsormidium flaccidum cp
Charophyceae
Coleochaete orbicularis cp
OCS162
Pyramimonas parkeae cp
Prasinophyceae I
0.10 substitutions
per nucl. position

Figure 4.3

identify signatures unique to the prasinophyte lineage defined by OCS162 and *Pyramimonas parkeae* (Fig. 4.4). The majority of nucleotide substitutions between OCS162 and *Pyraminonas parkeae* were compensatory base changes that preserved the integrity of the SSU rRNA secondary structure across regions of double-stranded base pairing, or occurred in highly variable loop regions (Fig. 4.4). Though clearly related phylogenetically, OC162 and *Pyramimonas parkeae* showed significant structural variation in two hypervariable regions of the bacterial SSU rRNA secondary structure [variable regions as in (Dams et al. 1988)]. Variable region two of the *Pyramimonas parkeae* plastid SSU rRNA secondary structural model contained an insertion of 27 nucleotides relative to that of OCS162. In addition, the secondary structural model for OCS162 contained a 9 nucleotide deletion in variable region seven relative to *Pyramimonas parkeae*. The structural variation observed in variable region two for these two related SSU rRNA genes is consistent with structural variation observed in other other groups of closely related bacterial rRNA genes (Giovannoni et al. 1996, Wright et al. 1997). However, structural variation in variable region seven of the bacterial SSU rRNA gene has been observed in only a limited number of the major bacterial phyla (Woese 1987). Deletions in this region of the bacterial SSU rRNA secondary structure appear to be common in the plastid SSU rRNA gene lineages we have examined.

A unique clone in the OCS library, OCS182 (RFLP type XXXVI), was phylogenetically affiliated with OM5, a SSU rDNA gene clone recovered in the OM SSU rDNA clone library (Figs. 4.1, 4.3). The LogDet analysis shown in Figure 4.1 revealed the clade defined by OCS182 and OM5 was a deeply branching member of the green plastid line of descent. A maximum likelihood phylogenetic analysis which included full length sequences and more

Figure 4.4. Proposed secondary structural model for clone OCS162 SSU rRNA. Boxed areas indicate variable regions in the SSU rRNA gene product distinguishing OCS162 from *Pyramimonas parkeae.* Variable regions of the bacterial SSU rRNA are numbered as in Dams et al. (1988). Numbers refer to nucleotide positions in the *E. coli* SSU rRNA gene (Brosius et al. 1978). Target sites for the primers 27F and 1522R are indicated.

OCS162

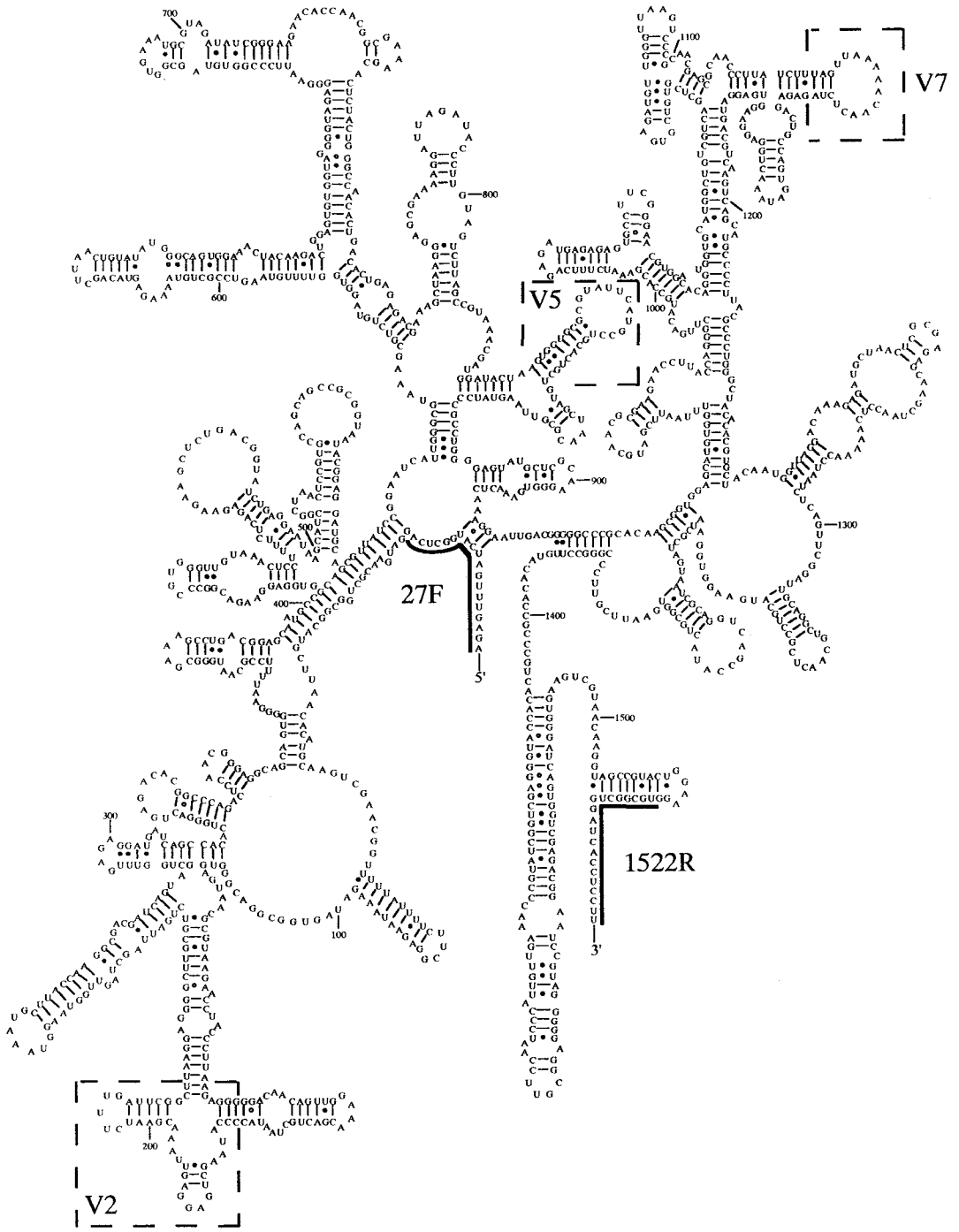


Figure 4.4

members of the green plastid line of descent showed the OCS182/OM5 lineage to affiliate with the Chlamydophyceae class of single-celled algae (Prasinophyte II lineage, Fig. 4.3). This affiliation was only weakly supported by both neighbor joining (<60 of 100 replicates) and maximum parsimony (68 of 100 replicates) bootstrap analyses (Fig. 4.3), though it was also recovered in a LogDet analysis of the green plastid data set (data not shown). As in the Prasinophyte I lineage described above, the Prasinophyte II lineage appeared to diverge early in the evolutionary history of the green chloroplasts. Phylogenetic analyses of six members of the OCS182/OM5 clade revealed the clones fell into three closely related though well-defined lineages (Fig. 4.5). One lineage consisted of the Oregon Coast clone (OCS182) alone, while two of the lineages consisted of OM clones from the continental shelf off Cape Hatteras, North Carolina (Fig. 4.5). The OCS182/OM5 clade was determined to be closely related to prasinophyte algae plastids by phylogenetic analyses with an unpublished prasinophyte plastid SSU rRNA gene sequence (Huss 1997).

The OM clone library. In addition to the OM plastid rDNA clones described above and in Rappé et al. (1995), three clones in the OM SSU rDNA clone library formed unique lineages within the plastid radiation defined by the rhodophyte plastids and plastids derived from secondary symbioses. OM81 and OM164 both affiliated with rhodophyte plastids and plastids derived from secondary symbioses in the LogDet analysis shown in Fig. 4.1. OM164 did not specifically affiliate with any classes of algae containing members with available plastid SSU rRNA gene sequences (Fig. 4.1). However, OM81 associated with the chrysophyte alga *Ochromonas danica* (91.4% similar for 1272 nucleotide positions), a relationship supported by both

Figure 4.5. Phylogenetic relationships between OCS and OM clones related to the Prasinophyceae II lineage, constructed by the maximum likelihood method (fDNAml) (Olsen et al. 1994). This analysis included a mask of 382 nucleotides corresponding to the 5' end of the plastid SSU rRNA gene. Bootstrap replicates (100 resamplings) were performed by distance methods (above the node) or maximum parsimony (below the node). Bootstrap values below 50% are not shown. Chloroplast SSU rRNA sequences from the charophytes *Chara* sp. and *Klebsormidium flaccidum* were used to root the tree.

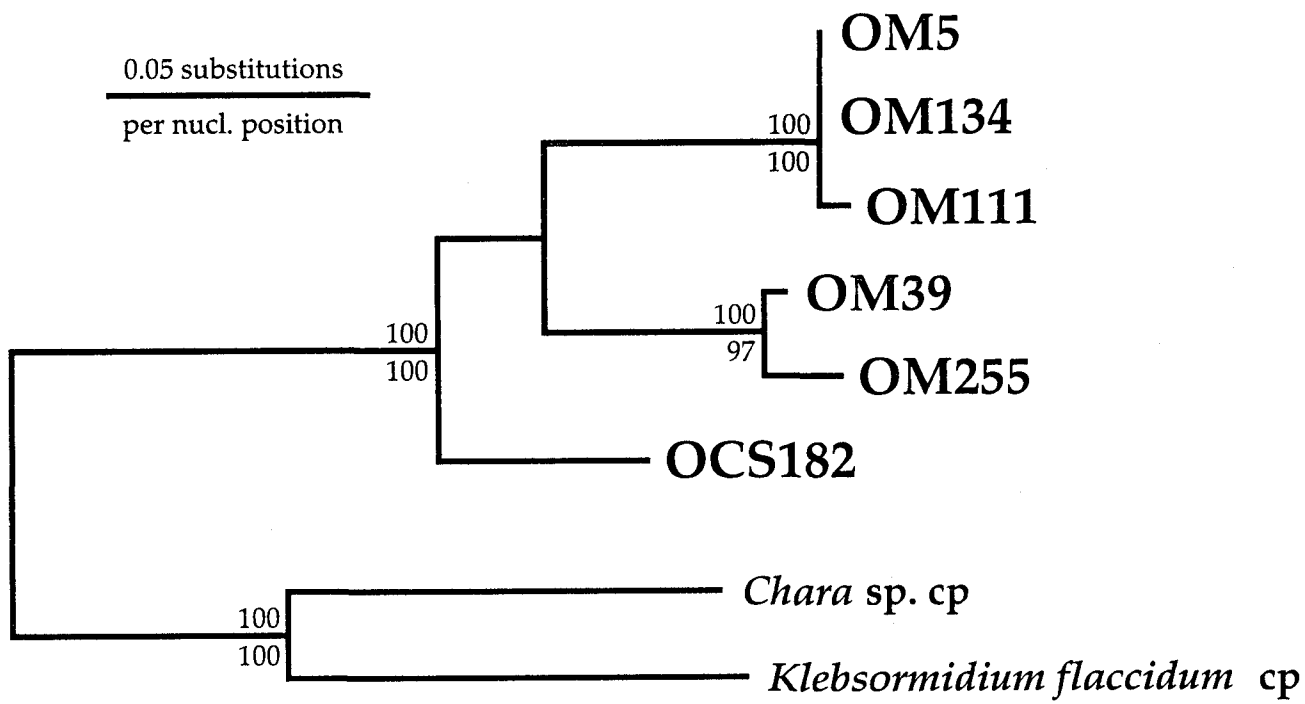

0.05 substitutions
per nucl. position
OM5
OM134
OM111
OM39
OM255
OCS182
Chara sp. cp
Klebsormidium flaccidum cp
100
100
100
100
97
100
100
100

Figure 4.5

neighbor joining (100 of 100 replicates) and maximum parsimony (100 of 100 replicates) bootstrap proportions employing the entire length of the SSU rRNA gene (Fig. 4.1). In the LogDet analysis depicted in Fig. 4.1, OM270 occupied one lineage in a polytomy which also included the prymnesiophye and crytophyte plastid SSU rRNA lines of descent.

Chimeric gene product. One OCS rDNA clone, preliminarily identified as plastid-related, was identified as a chimeric gene product when sequence obtained from both the 5' and 3' regions of the clone was examined. Phylogenetic analyses of ca. 400 nucleotides on the 3' end of clone OCS22 revealed this portion of the clone to be closely related to bacillariophyte plastid SSU rRNA genes (data not shown). Sequence obtained from the 5' end of OCS22 (ca. 400 nucleotides) was indentical to an OCS rRNA gene clone related to the alpha subdivision of the Proteobacteria (OCS126, data not shown).

LH-PCR. The results of LH-PCR for the untreated and filtered subsamples are shown in Fig. 4.6. The most striking result was the disappearance of the peaks corresponding to sizes of plastid SSU rRNA genes in water filtered through 0.8 μm membranes (Fig. 4.6). Peaks of 318, 319 and 326 base pairs in length were observed in the unfiltered water, but were not detected in the filtered water. The peak of 317 bp corresponded to sizes of both alpha proteobacteria and plastids (Suzuki et al. in prep). We conclude that filtration through a 0.8 μm membrane excluded eukaryotic plastid DNA from the sample.

Figure 4.6. Electropherogram displays of "bacterial" diversity in unfiltered (A) and filtered (B) seawater from a depth of 10 m, 8 km off the coast of Yaquina Head, Oregon. SSU rRNA genes were amplified with the 27F/338R primer pair, and the amplicons were separated by natural length polymorphisms. Peaks E, F, G, and I were removed by filtration through a 0.8 um Nucleopore filter. Peaks E, F, and G have been identified as plastid genes (prymnesiophytes, bacillariophytes, and prasinophytes, respectively). Peak I is unidentified.

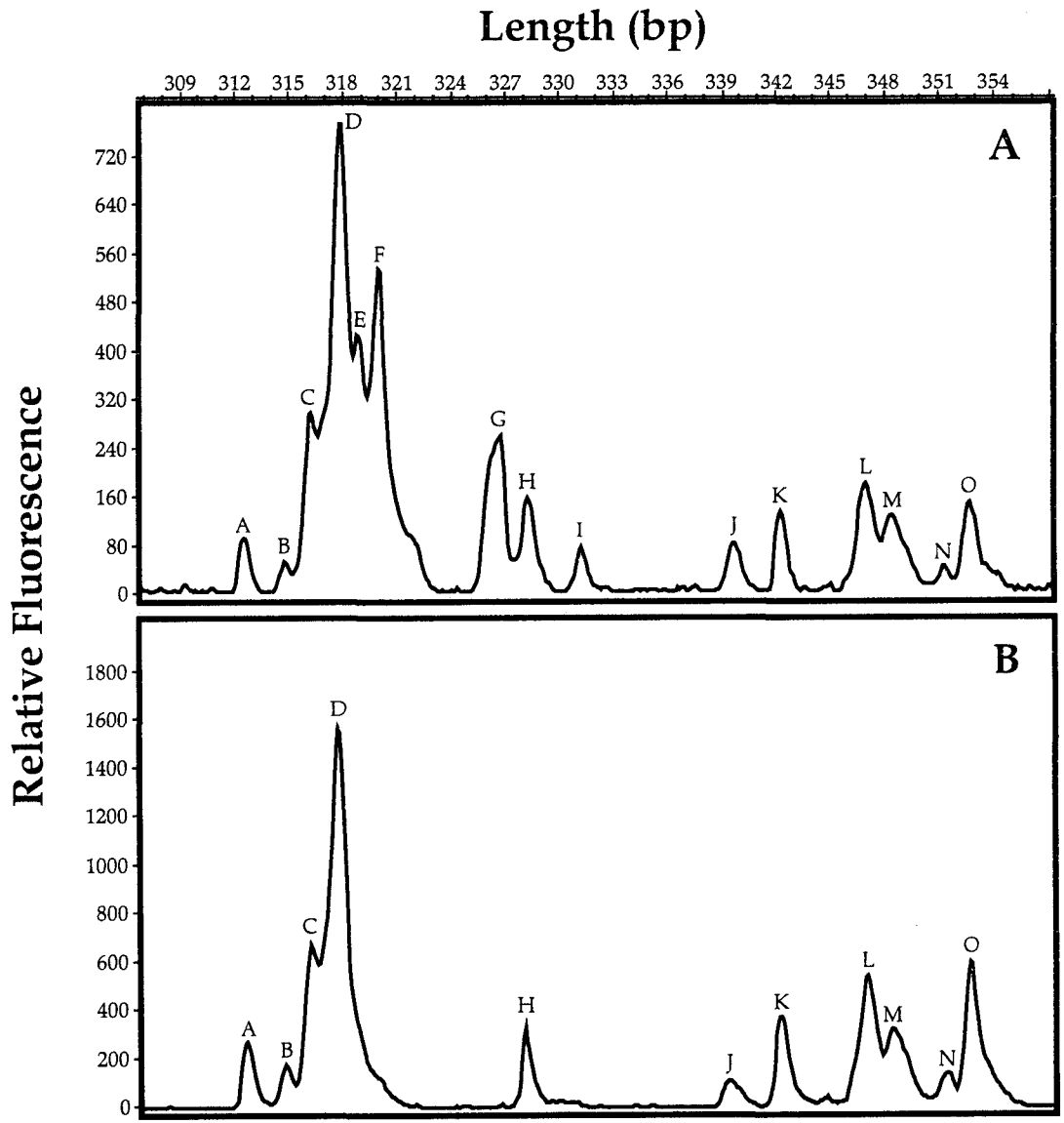

Length (bp)
309 312 315 318 321 324 327 330 333 336 339 342 345 348 351 354
Relative Fluorescence
A
720 640 560 480 400 320 240 160 80 0
A B C D E F G H I J K L M N O
B
1800 1600 1400 1200 1000 800 600 400 200 0
A B C D H J K L M N O

Figure 4.6

4.5 Discussion

The plastid genes that we described here and those that we described previously were discovered in investigations that had been designed to survey bacterioplankton diversity. Other studies using similar experimental designs have previously reported the presence of bacillariophyte plastid SSU rRNA genes recovered from phytodetrietal aggregates (DeLong et al. 1993). Nonetheless, we were surprised to find so many plastid SSU rRNA genes since numerous phylogenetic surveys of bacterioplankton diversity have reported few plastid genes. In the two clone libraries we describe here we have identified 109 plastid genes belonging to five major groups of algae, including bacillariophytes (diatoms), prymnesiophytes (coccolithophores), cryptophytes, chrysophytes, and prasinophytes. Although many examples of identical genes were found, in other cases the data support the presence of gene clusters which most probably represent assemblages of related phytoplankton species.

The contrast between the relative absence of plastid genes from clone libraries obtained from mid-ocean samples and the large number of such genes found in these environmental rDNA clone libraries from coastal sites may be attributed to several causes. Continental shelf regions are known to be areas of high nutrient availability and high primary productivity. Although marine *Synechococcus* and prochlorophytes are common on continental shelves, the ratio of these unicellular prokaryotic phytoplankton species to their eukaryotic counterparts generally decreases as nutrient levels and primary production increase. Thus, the observation of large numbers of phytoplankton plastids in these clone libraries may result directly from the higher proportion of phytoplankton nucleic acids in samples from continental shelf regions. A second factor which may contribute to the high proportion of phytoplankton

genes is the high copy of plastid organelle genomes. Some phytoplankton plastids contain up to 650 copies of small circular genomes per plastid, each usually containing 2 ribosomal operons per genome (Ersland et al. 1981). In contrast, prokaryotes frequently possess from one to only ten copies of their ribosomal operons. Although it appears likely that the large numbers of plastid genes observed in these clone libraries are in part due to the higher abundance of phytoplankton on the continental shelves, the ratios between the observed numbers of plastid genes and bacterial genes may be highly skewed towards the former as a consequence of these differences in copy number.

As with phylogenetic surveys of bacterial diversity, this analysis of phytoplankton diversity uncovered evidence of novel groups of phytoplankton that had not been recognized previously on the basis of culture and morphological studies. We observed two phylogenetically distinct lineages of Prasinophyceae. Prasinophytes are small flagellated cells that possess chlorophyll *a* and chlorophyll *b*. These organisms are thought to be primitive species because of their simple cell morphology, a feature that also makes them relatively difficult to identify in seawater samples, with the consequence that the marine members of this group are not routinely identified below the family level. These results are of interest to phytoplankton ecologists because they show the presence of at least two prasinophyte groups in seawater and provide information that could be used to design phylogenetic oligonucleotide probes for studies aimed at exploring the ecological roles of these putative species.

This report is the first to show the position of two prasinophyte plastid SSU rRNA gene lineages in phylogenetic trees. In our analysis, which is based upon complete prasinophyte SSU rRNA genes, these plastids appear to be

most closely allied with the Chlamydophyceae. However, divergence between Prasinopyceae Group I and Prasinophyceae Group II is so deep that they appear polyphyletic in Figure 4.3. An estimate of statistical confidence by bootstrap replication failed to resolve relationships between the two prasinophyte groups; however, clearly these are highly divergent lineages which warrant independent study.

Clone OM270 was unique among the genes investigated; it was phylogenetically allied to the rhodophyte plastids and plastids derived by secondary symbioses. The phylogenetic position of this deeply branching clone has interesting evolutionary implications. Bootstrap replicates provide only modest support for its position in the plastid phylogenetic tree. In Figure 4.1 and other analyses, OM270 appears as the outgroup to the prymnesiophytes and cryptophytes. The prymnesiophytes are heterkont flagellates exhibiting the photosynthetic pigments chlorophyll *a* and *c*. The Cryptophyceae are an enigmatic group that appear to be the result of a sequential series of symbioses involving the capture of a rhodophyte phytoplankton cell by a eukaryotic cell, resulting in a cell configuration includes a eukaryotic nucleus and a nucleomorph. The nucleomorph resides within an internal membrane that also encloses a plastid containing chlorophyll *a* and *c* and phycobilliproteins. Given this complicated phylogenetic picture, it would of considerable interest to examine the phenotype of the lineage represented by OM270. Analysis of the OM270 sequence in a secondary structure failed to uncover any evidence that this unusual gene was a result of cloning or sequencing artifacts (data not shown).

Emiliania huxleyi is a cultured marine prymnesiophyte, which is thought to represent the large blooms of coccolithophorids which are abundant

primary producers in many regions of the earth's oceans. We observed a single gene, OCS50, that allied very closely with the published SSU rRNA gene sequence from the *Emiliania huxleyi* plastid. However, published studies suggest that the various morphotypes of *Emiliania huxleyi* have identical plastid rRNA gene sequences. In addition to OCS50, we observed multiple plastid sequences related to the prymnesiophytes but clearly distinct from *Emiliania* and *Ochrosphaera* species.

Figure 4.6 provides an example of the utility of plastid rDNA sequence data for ecological studies. Figure 4.6 shows the distribution of three major phytoplankton groups in a coastal water sample. The different plastid types are resolved by length heterogeneities in the 5' region of their plastid SSU rDNAs (LH-PCR). Such measurements are highly reproducible and make it possible to rapidly survey the phylogenetic diversity within samples for comparative purposes (Suzuki et al. in prep).

The picture of phytoplankton diversity emerging from this study is that eukaryotic phytoplankton are better understood than their counterparts, the prokaryotic plankton, but nonetheless contain many members that have not been identified previously by scientific investigations. There are several possible explanations for these observations. The most obvious of which is that some of the novel phytoplankton lineages we present here may correspond to cultured microorganisms that have not been studied by chloroplast rDNA gene sequencing. Alternatively, some phytoplankton may be very difficult to identify either by morphology or through cultivation.

In the eyes of many practitioners, the survey of novel microbes by ribosomal RNA gene cloning and sequencing is not an ultimate goal but instead a stage of progress that may ultimately lead to the understanding of

species that have not yielded to other means of study. Ribosomal RNA genes provide the information needed to design molecular probes that can be used to study the distributions of organisms in nature at a gross level by the hybridization of probes to nucleic acids or at the cellular level by the hybridization of fluorescently labelled probes to single cells at the microscopic level.

4.6 References

Alfreider, A., J. Pernthaler, R. Amann, B. Sattler, F.-O. Glöckner, A. Wille and R. Psenner. 1996. Community analysis of the bacterial assemblages in the winter cover and pelagic layers of a high mountain lake by in situ hybridization. Appl. Environ. Microbiol. 62: 2138-2144.

Amann, R. I., B. J. Binder, R. J. Olsen, S. W. Chisholm, R. Devereux, and D. A. Stahl. 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl. Environ. Microbiol. 56: 1919-1925.

Amann, R. I., W. Ludwig and K.-H. Schleifer. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microb. Rev. 59: 143-169.

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

Barber, R. T. and R. L. Smith. 1981. Coastal upwelling ecosystems. In Analysis of Marine Ecosystems (ed. A. R. Longhurst), pp. 31-68. Academic Press, New York.

Brosius, J., M. L. Palmer, P. J. Kennedy and H. F. Noller. 1978. Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*. Proc. Natl. Acad. Sci. USA. 75: 4801-4805.

Chrétiennot-Dinet, M.-J., A. Sournia, M. Ricard and C. Billard. 1993. A classification of the marine phytoplankton of the world from class to genus. Phycologia. 32: 159-179.

Courties, C., A. Vaquer, M. Troussellier, J. Lautier, M. J. Chrétiennot-Dinet, J. Neveux, C. Machado and H. Claustre. 1994. Smallest eukaryotic organism. Nature. 370: 255.

Dams, E., L. Hendriks, Y. Van de Peer, J.-M. Neefs, G. Smits, I. Vandenbempt and R. De Wachter. 1988. Compilation of small ribosomal subunit RNA sequences. Nucleic Acids Res. 16: r87-r173.

DeLong, E. F., K. Y. Wu, B. B. Prézelin and R. V. M. Jovine. 1994. High abundance of Archaea in Antarctic marine picoplankton. Nature. 371: 695-697.

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

Felsenstein, J. 1989. PHYLIP-phylogeny inference package (v3.5). Cladistics. 5: 164-166.

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

Giovannoni, S. J., E. F. DeLong, G. J. Olsen and N. R. Pace. 1988. Phylogenetic group-specific oligodeoxynucleotide probes for identification of single microbial cells. J. Bacteriol. 170: 720-726.

Giovannoni, S. J., E. F. DeLong, T. M. Schmidt and N. R. Pace. 1990. Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. Appl. Environ. Microbiol. 56: 2572-2575.

Giovannoni, S. J., M. S. Rappé, K. Vergin and N. Adair. 1996. 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the Green Non-Sulfur bacteria. Proc. Natl. Acad. Sci. USA. 93: 7979-7984.

Gordon, D. A. and S. J. Giovannoni. 1996. Detection of stratified microbial populations related to *Chlorobium* and *Fibrobacter* species in the Atlantic and Pacific oceans. Appl. Environ. Microbiol. 62: 1171-1177.

Gutell, R. R. 1994. Collection of small subunit (16S and 16S-like) ribosomal RNA structures: 1994. Nucleic Acids Res. 22: 3502-3507.

Hoepffner, N. and L. W. Haas. 1990. Electron microscopy of nanoplankton from the North Pacific central gyre. J. Phycol. 26: 421-439.

Hood, R. R., S. Neuer and T. J. Cowles. 1992. Autotrophic production, biomass and species composition at two stations across an upwelling front. Mar. Ecol. Prog. Ser. 83: 221-232.

Huyer, A. 1977. Seasonal variation in temperature, salinity, and density over the continental shelf off Oregon. Limnol. Oceanogr. 22: 442-453.

Huyer, A. 1983. Coastal upwelling in the California Current system. Prog. Oceanogr. 12: 259-284.

Johnson, P. W. and J. M. Sieburth. 1982. In-situ morphology and occurrence of eucaryotic phototrophs of bacterial size in the picoplankton of estuarine and oceanic waters. J. Phycol. 18: 318-327.

Joint, I. R. 1986. Physiological ecology of picoplankton in various oceanographic provinces. In Photosynthetic picoplankton (ed. T. Platt and W. K. W. Li), pp. 287-309. Can. Bull. Fish. Aquat. Sci., Ottawa.

Kim, Y.-S., H. Oyaizu, S. Matsumoto, M. M. Watanabe and H. Nozaki. 1994. Chloroplast small-subunit ribosomal RNA gene sequence from *Chlamydomonas parkeae* (Chlorophyta): molecular phylogeny of a green alga with a peculiar pigment composition. Eur. J. Phycol. 29: 213-217.

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

Knauber, D. C., E. S. Berry and M. W. Fawley. 1996. Ribosomal RNA-based oligonucleotide probes to identify marine green ultraphytoplankton. J. Euk. Microbiol. 43: 89-94.

Kopczynski, E. D., M. M. Bateson and D. M. Ward. 1994. Recognition of chimeric small-subunit ribosomal DNAs composed of genes from uncultivated microorganisms. Appl. Environ. Microbiol. 60: 746-748.

Lane, D. J. 1991. 16S/23S rRNA sequencing. In Nucleic acid techniques in bacterial systematics (ed. E. Stackebrandt and M. Goodfellow), pp. 115-175. John Wiley and Sons, New York.

Lange, M., L. Guillou, D. Vaulot, N. Simon, R. I. Amann, W. Ludwig and L. K. Medlin. 1996. Identification of the class Prymnesiophyceae and the genus *Phaeocystis* with ribosomal RNA-targeted nucleic acid probes detected by flow cytometry. J. Phycol. 32: 858-868.

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

Maidak, B. L., N. Larsen, M. J. McCaughey, R. Overbeek, G. J. Olsen, K. Fogel, J. Blandy and C. R. Woese. 1994. The Ribosomal Database Project. Nucleic Acids Res. 22: 3485-3487.

Miller, P. E. and C. A. Scholin. 1996. Identification of cultured *Pseudo-Nitzschia* (Bacillariophyceae) using species-specific LSU rRNA-targeted fluorescent probes. J. Phycol. 32: 646-655.

Murphy, L. S. and E. M. Haugen. 1985. The distribution and abundance of phototrophic ultraplankton in the North Atlantic. Limnol. Oceanogr. 30: 47-58.

Olsen, G. J., H. Matsuda, R. Hagstrom and R. Overbeek. 1994. fastDNAml: a tool for construcion of phylogenetic trees of DNA sequences using maximum likelihood. Comput. Appl. Biosci. 10: 41-48.

Rappé, M. S., P. F. Kemp and S. J. Giovannoni. 1995. Chromophyte plastid 16S ribosomal RNA genes found in a clone library from Atlantic Ocean seawater. J. Phycol. 31: 979-988.

Rappé, M. S., P. F. Kemp and S. J. Giovannoni. 1997. Phylogenetic diversity of marine coastal picoplankton 16S rRNA genes cloned from the continental shelf off Cape Hatteras, North Carolina. Limnol. Oceanogr. in press.

Robinson-Cox, J. F., M. M. Bateson and D. M. Ward. 1995. Evaluation of nearest-neighbor methods for detection of chimeric small-subunit rRNA sequences. Appl. Environ. Microbiol. 61: 1240-1245.

Ryther, J. H. 1969. Photosynthesis and fish production in the sea. Science. 166: 72-76.

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

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

Simon, N., N. LeBot, D. Marie, F. Partensky and D. Vaulot. 1995. Fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes to identify small phytoplankton by flow cytometry. Appl. Environ. Microbiol. 61: 2506-2513.

Small, L. F. and D. W. Menzies. 1981. Patterns of primary productivity and biomass in a coastal upwelling region. Deep Sea Res. 28: 123-149.

Suzuki, M., M. S. Rappé and S. J. Giovannoni. in prep. Estimating coastal picoplankton community structure by LH-PCR. Appl. Environ. Microbiol.

Suzuki, M. T., M. S. Rappé, Z. W. Haimberger, H. Winfield, N. Adair, J. Ströbel and S. J. Giovannoni. 1997. Bacterial diversity among SSU rRNA gene clones and cellular isolates from the same seawater sample. Appl. Environ. Microbiol. 63: 983-989.

Woese, C. R. 1987. Bacterial evolution. Microbiol. Rev. 51: 221-271.

Wright, T. D., K. Vergin, P. W. Boyd and S. J. Giovannoni. 1997. A novel _-subdivision proteobacterial lineage stratified in the lower ocean surface layer. Appl. Environ. Microbiol. 63: 1441-1448.

CHAPTER 5

DIVERSITY OF BACTERIOPLANKTON OVER THE CONTINENTAL SHELF OFF OREGON, USA, DETERMINED BY SSU rRNA GENE CLONING, SEQUENCING, AND LH-PCR

Michael S. Rappé, Marcelino T. Suzuki, Kevin Vergin, and Stephen J. Giovannoni

to be submitted to *Applied and Environmental Microbiology*,
American Society for Microbiology,
Washington, D.C.

5.1 Abstract

To identify members of the bacterioplankton population inhabiting the western continental shelf of the United States, a SSU rRNA gene clone library was constructed from plankton populations collected from a depth of 10 m in the Pacific Ocean, 8 km west of the mouth of Yaquina Bay, Oregon. SSU rRNA gene clones in the Oregon Coast Study (OCS) clone library were polymerase chain reaction-amplified with general bacterial primers from nucleic acids isolated from the plankton samples, and cloned into the vector pBluescript KSII-. Plastid SSU rRNA genes comprised 47% of the 116 SSU rRNA gene clones recovered, while 60% of the bacterial clones were affiliated with SSU rRNA gene clones previously recovered from open-ocean and coastal seawater environments, including the SAR86, SAR11, and SAR116 clusters of the Proteobacteria, the *Roseobacter* clade of the alpha Proteobacteria, the SAR406 cluster, and the BDA1-5 (marine Gram-positive) cluster. In addition, SSU gene clones related to beta Proteobacteria constituted 16% of the bacterial SSU rRNA gene clones recovered. The remaining bacterial clones were identified as members of the alpha and gamma subgroups of the Proteobacteria. A Length Heterogeneity PCR (LH-PCR) electropherogram of the PCR product cloned in this study closely resembled the distribution of clone types recovered in the OCS library, indicating no biases were apparent in the blunt-end ligation procedure. The distribution of phylotypes in the OCS clone library provides further evidence that a limited number of as yet uncultured bacterioplankton may dominate coastal as well as open ocean pelagic seawater environments. In addition, the recovery of beta Proteobacteria SSU rRNA gene clones supports our earlier findings that beta Proteobacteria may constitute a

significant fraction of the bacterioplankton population in coastal seawater environments.

5.2 Introduction

Over the past two decades, the roles heterotrophic bacteria play in marine food webs and oceanic biogeochemical cycles has received considerable attention (e.g. see Azam and Cho 1987, Fuhrman et al. 1989, Sherr and Sherr 1988). It is now widely recognized that heterotrophic bacterioplankton are important components of marine systems - both as a sink and/or link for carbon in marine planktonic food webs, and for their role in nutrient regeneration in the water column (Cole et al. 1988, Ducklow et al. 1986). Several recent studies indicate that a wide range of phenotypic and genotypic diversity exists within and between heterotrophic bacterioplankton populations, and the bacterial fraction of planktonic ecosystems appears to be no less structurally dynamic than is commonly observed in the eukaryotic photoautotrophic and heterotrophic fractions of the marine planktonic community (Fuhrman et al. 1994, Lee and Fuhrman 1991, Martinez et al. 1996, Rehnstam et al. 1993, Smith et al. 1995). Methodological constraints, however, have led to a lack of consideration for individual species dynamics within the heterotrophic bacterioplankton communities in systems level studies. The effect that spatial and temporal variation in bacterioplankton population structure has on the cycling of carbon and associated biogenic elements in the world's oceans is largely unknown (Azam et al. 1995, Kemp 1995).

In order to evaluate the significance of quantitative and qualitative changes in the structure of bacterioplankton populations, a knowledge of the

types of organisms that can be potentially most abundant in a given seawater sample is necessary. As with most ecosystems, the diversity of organisms within marine bacterioplankton populations has traditionally been assayed by culture-dependent methods (e.g. see references (Baumann et al. 1972, Kita-Tsukamoto et al. 1993, ZoBell 1946). It has long been known, however, that numbers of bacteria in seawater estimated by CFU counts are orders of magnitude lower than those estimated by epifluorescence direct counts (Ferguson et al. 1984, Hobbie et al. 1977, Jannasch and Jones 1959, Kogure et al. 1979). A significant proportion of cells visualized microscopically appear to be viable, but do not form visible colonies on plates (Staley and Konopka 1985, though see Zweifel and Hagström 1995 and Choi et al. 1996). It is not currently known how much of the information derived from pure cultures of marine bacteria can be related to the activity of bacterioplankton *in situ*.

In an effort to circumvent selective biases associated with cultivation methodology, many microbial ecologists have turned to molecular methods to survey the diversity of bacteria in their natural environment (Amann et al. 1995, Olsen et al. 1986, Ward et al. 1992). The most widely used approach involves the cloning and sequencing of SSU ribosomal RNA genes directly from naturally-occurring microbial assemblages, and has provided marine microbial ecologists a powerful means of assaying the diversity of bacterioplankton independent of the culturability of cells (reviewed in Giovannoni et al. 1995). Though biases associated with molecular analyses of bacterial diversity in the environment have not been thoroughly evaluated, they appear to be less limiting than those associated with culture-based methods (Giovannoni et al. 1995, Ward et al. 1992), and may in fact be controllable in some situations (Suzuki and Giovannoni 1996).

To date, natural samples of bacterioplankton have been characterized by SSU rRNA gene cloning and sequencing from a fairly narrow range of oligotrophic regions of the world's oceans (Britschgi and Giovannoni 1991, Field et al. 1997, Fuhrman et al. 1993, Giovannoni et al. 1990a, Giovannoni et al. 1996b, Gordon and Giovannoni 1996, Mullins et al. 1995, Schmidt et al. 1991, Wright et al. 1997). Two studies, however, have analyzed coastal bacterioplankton samples, including a clone library of 18 bacterioplankton SSU rRNA genes cloned from the Santa Barbara Channel off the California coast (DeLong et al. 1993) and 112 bacterioplankton SSU rRNA gene clones from the eastern continental shelf of the Unites States off Cape Hatteras, North Carolina (Rappé et al. 1997). A majority of the SSU rRNA gene clones retrieved from these samples are not phylogenetically affiliated with rRNA gene sequences obtained from cultured marine bacteria, and instead fall into one of several distinct phylogenetic groups (Giovannoni et al. 1996, Giovannoni et al. 1995, Mullins et al. 1995).

We addressed two questions by constructing a SSU rRNA gene clone library from surface seawater collected over the western continental shelf of the Pacific Ocean off Oregon, USA. First, we sought to determine if the SSU rRNA gene clones recovered in a clone library constructed from a natural bacterioplankton sample corresponded to the SSU rRNA genes obtained from bacterial isolates cultured from the same seawater sample. The answer to this question, provided in (Suzuki et al. 1997), was that very little overlap was observed in the SSU rRNA genes recovered from the cultures and gene clones. In this study, we determined the extent to which the bacterial SSU rRNA gene clones recovered in this clone library overlapped rRNA gene clones obtained previously from marine bacterioplankton samples. The majority of bacterial

phylotypes in the OCS clone library closely resembled clone groups recovered from both the Pacific and Atlantic Oceans, including the eastern continental shelf of the United States. In addition, the recovery of related beta Proteobacteria SSU rRNA gene clones from both the eastern and western continental shelves of the United States indicated beta Proteobacteria may constitute a significant fraction of the bacterioplankton population in coastal pelagic seawater environments.

5.3 Materials and methods

SSU rRNA gene amplification and clone library construction.

Construction of the Oregon Coast Study (OCS) SSU rRNA gene clone library has been described previously (Rappé et al. submitted, Suzuki et al. 1997). Briefly, a 16 L plankton sample was collected on 28 April 1993 from a depth of 10 m at a station located 8 km west of Yaquina Bay, Oregon (44°39.1'N, 124°10.6'W). The seawater sample was pre-screened through 10 µm nylon mesh and collected on 0.2 µm polysulfone filters (Supor-200, Gelman Inc., Ann Arbor, MI). Total cellular nucleic acids were isolated by cell lysis with proteinase K and SDS, and phenol/chloroform extraction as peviously described (Giovannoni et al. 1990b). Bacterial and plastid SSU rRNA genes were amplified from the genomic DNA sample using two general bacterial SSU rRNA primers (27F and 1522R), *Pfu* DNA polymerase (Stratagene, La Jolla, CA), and the polymerase chain reaction (PCR) (Saiki et al. 1988) as previously described (Suzuki et al. 1997). The blunt-ended amplification product was inserted into the phagemid vector pBluescript KSII- (Stratagene) and the product of this ligation was used to transform competent *E. coli* cells.

Positive transformants were streaked for isolation and stored in both stab cultures and 7% DMSO. Clones were numbered discontinuously from 1-182 and assigned the prefix "OCS".

RFLP analysis. All gene clones containing full-length inserts were characterized by *Hae*III RFLPs as previously described (Rappé et al. submitted, Suzuki et al. 1997). Briefly, environmental clone rDNAs were amplified from insert-bearing plasmids by the PCR employing the two general bacterial primers used above. PCR products were digested with the restriction endonuclease *Hae*III (Promega, Madison, WI) and resolved by gel electrophoresis in low melting point agarose (FMC, Rockland, ME).

Gene sequencing and phylogenetic analyses. Plasmid DNAs were sequenced as described in Suzuki et al. (1997), using either an ABI model 377 or 373a automated sequencer (Applied Biosystems Inc., Foster City, CA), dye-terminator chemistry, and conserved universal and bacterial SSU rRNA primers (Lane 1991). Sequence data from cloned SSU rDNA genes were manually aligned with bacterial and plastid sequences obtained from Genbank, the Ribosomal Database Project (RDP) (Maidak et al. 1994), and the ARB database (Ludwig and Strunk 1997) using the Genetic Data Environment (GDE) v2.2 sequence analysis software package (Steve Smith, Millipore, Bedford, MA).

Evolutionary trees were constructed as described in Rappé et al. (submitted). Briefly, distance, maximum parsimony, and maximum likelihood methods were employed to determine the phylogenetic affiliations of OCS rRNA gene clone sequences. Each phylogenetic analysis employed a sequence mask that varied in its level of conservation depending on the individual dataset being analyzed. These masks are listed in each corresponding figure

legend. In addition, all phylogenetic analyses were repeated several times with different taxa employed as outgroups (data not shown). Evolutionary distances were calculated from pair-wise sequence similarities with the Kimura 2-parameter model for nucleotide change and a transition/transversion ratio of 2.0 (Kimura 1980) using the program DNADIST available with the Phylogeny Inference Package (PHYLIP v3.5c) (Felsenstein 1989). The program NEIGHBOR was used to reconstruct phylogenetic trees from the evolutionary distance matrices by the neighbor-joining method (Saitou and Nei 1987). Bootstrap analyses which included 100 replicate resampled data sets with random sequence addition and global rearrangement were used to estimate the relative confidence in monophyletic groups withing each phylogenetic analysis (Felsenstein 1985).

Maximum parsimony, maximum likelihood, and LogDet phylogenetic analyses were performed using the program PAUP* (Swofford 1997). Maximum parsimony and maximum parsimony bootstrap analyses were performed by a heuristic search method, with tree bisection-reconnection and MULPARS options. Maximum likelihood calculations employed nucleotide frequencies estimated from the data, with a transition/transversion ratio of 2.0. Bootstrap analyses were performed as above. Secondary structure generation and detection of chimeric SSU rRNA gene clones were performed as in Rappé et al. (submitted).

Oligonucleotide probe hybridizations. As described in Rappé et al. (submitted), a dot blot of 32 clones representing 23 RFLP patterns was constructed and screened with radiolabeled taxon-specific oligonucleotide probes to aid in identifying OCS SSU rRNA gene clones. The following oligonucleotide probes were hybridized to the OCS unknown dot blot:

SAR83R (T_H=45°C) (Rappé et al. 1997), SAR11A1 (T_H= 37°C) (Field et al. 1997), SAR86 (T_H=45°C) (Rappé et al. 1997), CHRYS1 (T_H=45°C) (Rappé et al. 1995), and HAP1 (T_H=45°C) (Rappé et al. 1995). The T_H listed was the hybridization stringency temperature used for each oligonucleotide probe. Probe hybridizations were visualized using a Molecular Dynamics PhosphorImager SI (Molecular Dynamics, Sunnyvale, CA) as well as autoradiographically with X-ray film.

LH-PCR. The PCR product cloned to produce the OCS clone library was used as a template for LH-PCR (Suzuki et al. in prep). In a final volume of 100 µl, the reaction contained 0.2 mM of pre-mixed dNTPs (Stratagene), 1.5 mM MgCl, 5% acetamide, 1X *Taq* DNA polymerase reaction buffer (Promega), and 2.5 units of *Taq* DNA polymerase (Promega). All reactions used the Ampliwax hotstart protocol (Perkin Elmer Cetus, Norwalk, CT) in a PLT100 thermal cycler (MJ Research Inc., Watertown, MA) and were amplified for 16 cycles of 1 min denaturation (96°C), 1 min annealing (55°C), and 3 min extension (72°C). The number of cycles was optimized to avoid PCR bias described by Suzuki and coworkers (Suzuki et al. in prep, Suzuki and Giovannoni 1996). The forward primer in the LH-PCR reaction (27F; 5'-AGAGTTTGATCMTGGCTCAG-3') was 5'-end-labeled with the phosphoramidite dye 6-FAM, and was supplied by Applied Biosystems Inc. (ABI). 338R (5'-GCTGCCTCCCGTAGGAGT-3') was used as the reverse primer in the LH-PCR reactions (Amann et al. 1990). Labelled PCR products were purified using Qiaquick-spin columns (Qiagen, Chatsworth, CA), and 10 fmoles of DNA were discriminated by Long Ranger (FMC) PAGE in an ABI model 377 automated DNA sequencer (ABI) in Genescan mode. Genescan estimates the sizes of bands in the gel, and measures the integrated

fluorescence emission of the bands. The output of Genescan is an electropherogram in which the bands are represented by peaks and the integrated fluorescence of each band is the area under the peaks. For PCR product concentrations less than 50 fmoles, it was assumed that the relative proportion of the integrated fluorescence of each peak corresponded to the proportion of each amplicon in the PCR products (Suzuki et al. in prep). The relative abundance of each amplicon was estimated as the ratio between the integrated fluorescence of each peak and the total integrated fluorescence of all peaks.

5.4 Results

A total of 116 full-length SSU rRNA gene clones were recovered in the OCS clone library, and were distributed among 37 different *Hae*III RFLP banding patterns (Table 5.1). Nineteen of the 37 RFLP types consisted of a single OCS rRNA gene clone. Analyses of nine RFLP types (55 clones) in the OCS library related to plastid SSU rRNA genes, including one chimeric gene clone, were described previously (Rappé et al. submitted).

Alpha Proteobacteria. Thirteen different RFLP types representing 22 OCS SSU rRNA gene clones (38% of the bacterial clones recovered) were related to the alpha subclass of the Proteobacteria (Table 5.1). Of the 22 alpha Proteobacteria gene clones, 18 were affiliated with SSU rRNA gene clones previously recovered from seawater samples, including members of the SAR11 cluster (Giovannoni et al. 1990a), the SAR116 cluster (Mullins et al. 1995), and SSU rRNA gene clones related to SAR83 and other members of the *Roseobacter* clade (Britschgi and Giovannoni 1991).

Table 5.1. Frequency and phylogenetic affiliation of OCS SSU rRNA gene clones.

Table 5.1

Phylogenetic affiliation	RFLP type	Clones recovered	Representative clones*	GenBank access. No.
Alpha Proteobacteria				
SAR116 cluster	VII	4	OCS28	AF001636
	VI	2	OCS24	AF001637
	XXIX	1	OCS126	AF001638
SAR11 cluster	XV	2	OCS138	AF001639
	XVI	2	OCS143	U75266
			OCS154	AF001640
	XXIII	1	OCS53	AF001641
	XXXIV	1	OCS180	AF001642
Roseobacter clade	XII	2	OCS84	U78943
	XIV	2	OCS122	U78945
	XXI	1	OCS19	U78942
Unique alpha	XXXVII	2	OCS116	U78944
	XX	1	OCS14	AF001643
	XXVIII	1	OCS124	AF001644
Beta Proteobacteria				
Zoogloea ramigera ATCC25935	VIII	3	OCS7	AF001645
	XIII	2	OCS111	AF001646
Type I methylotroph clade	XXXIII	1	OCS178	AF001647
Variovorax paradoxus	XXVII	1	OCS98	AF001648
Unique beta	XXIV	1	OCS66	AF001649
	XXV	1	OCS67	‡
Gamma Proteobacteria				
SAR86 cluster	II	13	OCS44	AF001650
	XVIII	1	OCS5	AF001651
sulfur-oxidizing symbionts	IV	9	OCS2	‡
Pseudomonas putida DSM291	XXXV	1	OCS181	U78946
Miscellaneous				
BDA1-5 cluster	X	2	OCS155	AF001652
SAR406 cluster	XXXII	1	OCS146	AF001653
Unidentified	XXVI	1	OCS93	‡
	XXXI	1	OCS131	‡
Chimera	XIX	1	OCS12	U75252
	XXII	1	OCS22	‡
Plastids	†	54	†	†

*identical sequences were not filed in Genbank.
†data included in (Rappé et al. in prep.).
‡no sequence was submitted to Genbank.

Sequenced clones from RFLP types XII, XIV, and XXI within the OCS clone library were found to be related to SAR83 and related SSU rRNA gene clones, a group of SSU rRNA gene clones orginally described from surface seawater samples in the Sargasso Sea (Figs. 5.1, 5.2) (Britschgi and Giovannoni 1991). Represented by OCS84 in Fig. 5.1 and OCS19, OCS84, and OCS122 in Fig. 5.2, SAR83-related SSU rRNA genes consisted of 5 clones or 9% of the bacterial SSU rDNA clones recovered in the OCS clone library (Table 5.1). The phylogenetic affiliations of SAR83-related clones from the OCS clone library were first described in Suzuki et al. (1997). OCS84 was most closely related to OM65, a SSU rDNA clone recovered off the coast of Cape Hatteras, North Carolina in the OM SSU rRNA gene clone library (96.0% similar for 421 nucleotide positions) (Rappé et al. 1997). OCS19 was also most closely related to an rDNA clone recovered in the OM clone library, OM42 (100% similar for 376 nucleotide positions. OCS122 was most closely related to SAR102, a unique SSU rRNA gene clone recovered from surface water of the Sargasso Sea (100% similar for 397 nucleotide postions) (Mullins et al. 1995). As reported in earlier studies, this complex group of environmental clones is affiliated with members of the genus *Roseobacter* of the alpha-3 subclass of the Proteobacteria (Fig. 5.2) (Britschgi and Giovannoni 1991, Fuhrman et al. 1993, Mullins et al. 1995, Rappé et al. 1997). Subclasses within the alpha Proteobacteria are as in Woese et al. (Woese et al. 1984). Validly described species in the genus *Roseobacter* synthesize bacteriochlorophyll *a* aerobically and have been isolated from marine environments, including the phycosphere of dinoflagellates (Lafay et al. 1995) and marine macrophytic algae (Ashen and Goff 1996, Shiba 1991). Also related to this group are several other marine seawater isolates represented in Fig. 5.2, including *Marinosulfonomonas methylotropha,* a restricted

Figure 5.1. Inferred phylogenetic relationships among SSU rRNA gene clones recovered off the coast of Oregon, USA, with other members of the domain *Bacteria*. This dendogram was reconstructed by a neighbor joining analysis of evolutionary distances, calculated from pair-wise sequence similarities with the Kimura 2-parameter model for nucleotide change. A mask of 300 nucleotide positions located on the 5' end of the SSU rRNA gene was used in this analysis. Bootstrap proportions over 60% that supported the branching order, calculated by the neighbor joining method from 100 total replicates, are shown above the corresponding nodes. "SAR"-SSU rRNA gene clones from the Sargasso Sea; "OM"-SSU rRNA gene clones from the eastern continental shelf off Cape Hatteras, North Carolina, USA; "BAL"- isolated bacterial strains from the Baltic Sea (Pinhassi et al. 1997).

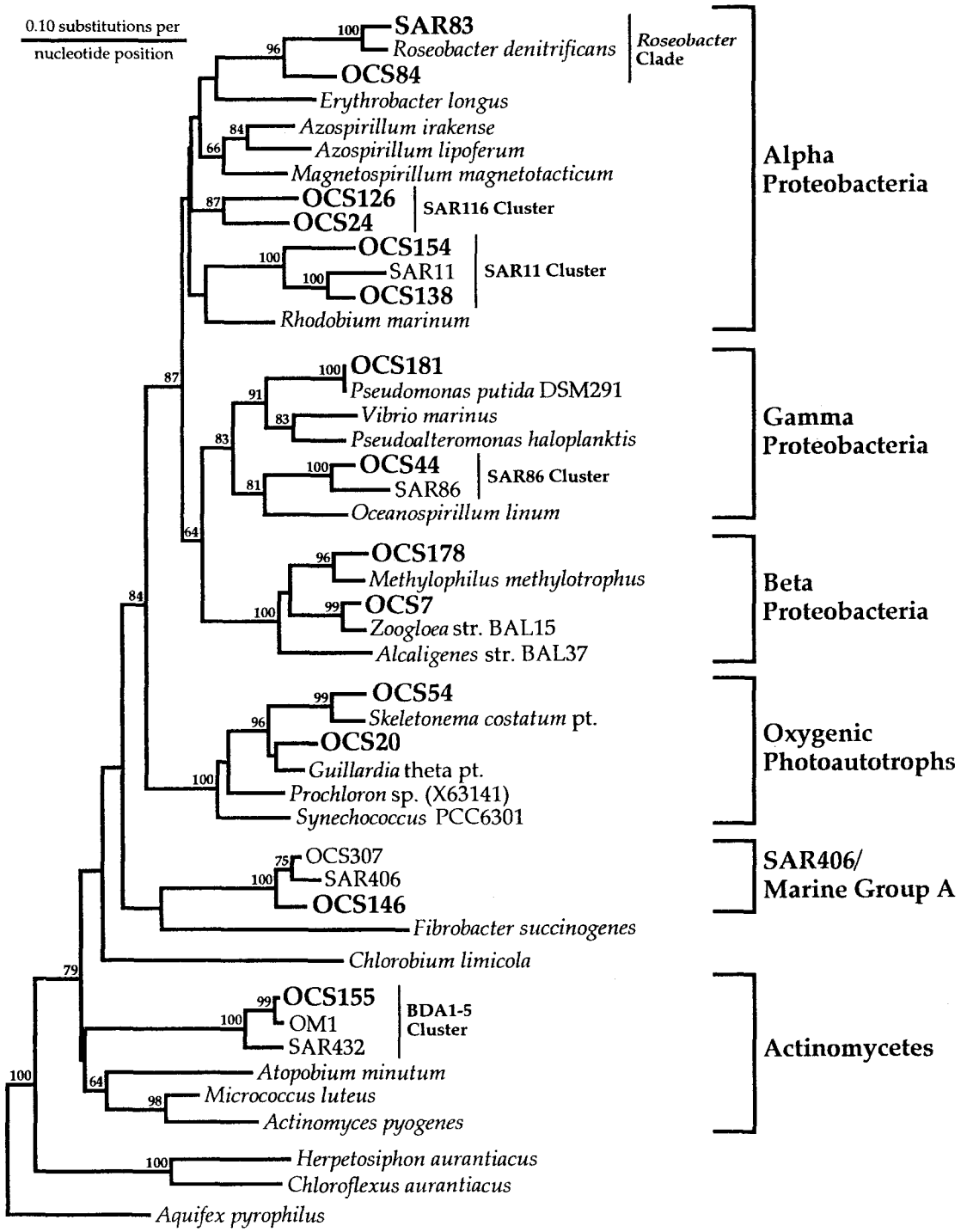

0.10 substitutions per nucleotide position
SAR83
Roseobacter denitrificans
Roseobacter Clade
OCS84
Erythrobacter longus
Azospirillum irakense
Azospirillum lipoferum
Magnetospirillum magnetotacticum
OCS126
OCS24
SAR116 Cluster
OCS154
SAR11
OCS138
SAR11 Cluster
Rhodobium marinum
Alpha Proteobacteria
OCS181
Pseudomonas putida DSM291
Vibrio marinus
Pseudoalteromonas haloplanktis
OCS44
SAR86
SAR86 Cluster
Oceanospirillum linum
Gamma Proteobacteria
OCS178
Methylophilus methylotrophus
OCS7
Zoogloea str. BAL15
Alcaligenes str. BAL37
Beta Proteobacteria
OCS54
Skeletonema costatum pt.
OCS20
Guillardia theta pt.
Prochloron sp. (X63141)
Synechococcus PCC6301
Oxygenic Photoautotrophs
OCS307
SAR406
OCS146
SAR406/ Marine Group A
Fibrobacter succinogenes
Chlorobium limicola
OCS155
OM1
SAR432
BDA1-5 Cluster
Atopobium minutum
Micrococcus luteus
Actinomyces pyogenes
Actinomycetes
Herpetosiphon aurantiacus
Chloroflexus aurantiacus
Aquifex pyrophilus

Figure 5.1

Figure 5.2. Phylogenetic diversity of OCS SSU rRNA gene clones related to the alpha subclass of the Proteobacteria. This dendogram was reconstructed by the neighbor joining method from evolutionary distances derived as in Fig. 5.1, from a mask of 340 nucleotides located on the 5' end of the SSU rRNA gene. Bootstrap proportions over 60% that supported the branching order, calculated by the neighbor joining method from 100 total replicates, are shown above the corresponding nodes. The SSU rRNA gene sequence of *Escherichia coli* was used as an outgroup in this analysis.

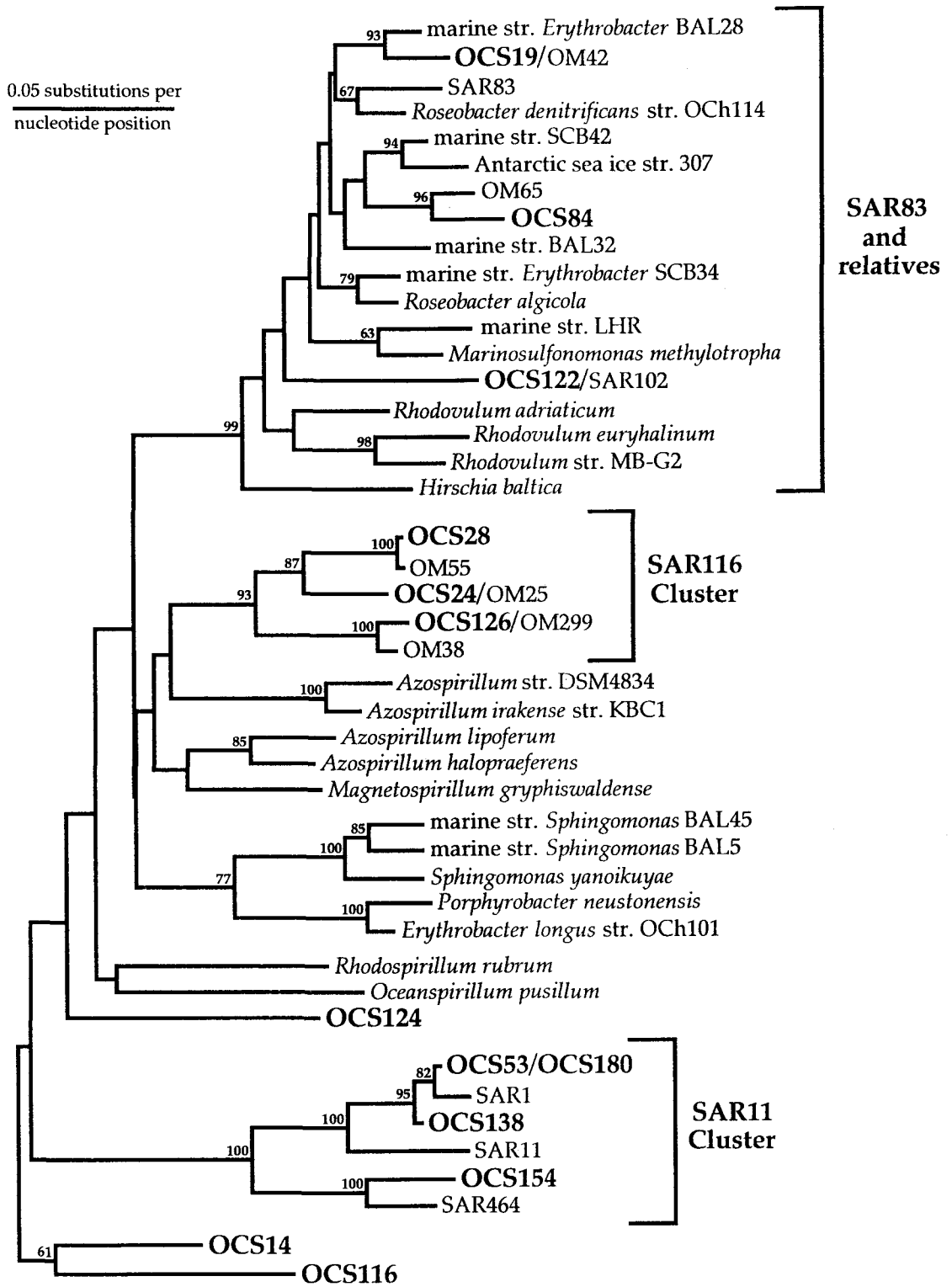

0.05 substitutions per
nucleotide position
marine str. Erythrobacter BAL28
OCS19/OM42
SAR83
Roseobacter denitrificans str. OCh114
marine str. SCB42
Antarctic sea ice str. 307
OM65
OCS84
marine str. BAL32
marine str. Erythrobacter SCB34
Roseobacter algicola
marine str. LHR
Marinosulfonomonas methylotropha
OCS122/SAR102
Rhodovulum adriaticum
Rhodovulum euryhalinum
Rhodovulum str. MB-G2
Hirschia baltica
SAR83 and relatives
OCS28
OM55
OCS24/OM25
OCS126/OM299
OM38
SAR116 Cluster
Azospirillum str. DSM4834
Azospirillum irakense str. KBC1
Azospirillum lipoferum
Azospirillum halopraeferens
Magnetospirillum gryphiswaldense
marine str. Sphingomonas BAL45
marine str. Sphingomonas BAL5
Sphingomonas yanoikuyae
Porphyrobacter neustonensis
Erythrobacter longus str. OCh101
Rhodospirillum rubrum
Oceanspirillum pusillum
OCS124
OCS53/OCS180
SAR1
OCS138
SAR11
OCS154
SAR464
SAR11 Cluster
OCS14
OCS116

Figure 5.2

facultative methylotroph which is able to grow on methanesulfonic acid as its sole source of carbon and energy (Holmes et al. 1997), marine bacteria of the genus *Rhodovulum* (Hiraishi and Ueda 1994, Hiraishi and Ueda 1995), the marine strain LHR, a bacterium capable of cleaving dimethylsulfoniopropionate to form dimethylsulfide (Ledyard et al. 1993), several isolates from Oregon coast seawater (Suzuki et al. 1997), an isolate from Sargasso Sea seawater (Suzuki et al. 1997), isolates from Santa Barbara Channel and Baltic Sea seawater (Pinhassi et al. 1997), and isolates from Antarctic Sea ice (Bowman et al. 1997).

Three RFLP types in the OCS clone library (RFLP types VI, VII, and XXIX) consisting of 7 total clones (12% of the bacterial clones; Table 5.1) were related to the SAR116 cluster of the alpha Proteobacteria first described by Mullins and coworkers (Mullins et al. 1995). Represented by OCS24 and OCS126 in Fig. 5.1 and OCS24, OCS28, and OCS126 in Fig. 5.2, the SAR116 cluster SSU rRNA gene clones recovered in the OCS clone library were closely related to SSU rRNA gene sequences recovered off the eastern continental shelf of the United States in the OM clone library (Rappé et al. 1997). These similarities included OCS24 and OM25 (99.7% similar for 1474 nucleotide positions), OCS28 and OM55 (99.5% similar for 582 nucleotide positions), and OCS126 and OM299 (99.4% similar for 362 nucleotide positions). The gene sequence of clone SAR116 was most closely related to OCS28 (97.3% similar for 473 nucleotide positions) and OM55 (97.8% similar for 451 nucleotide positions), but was not included in Fig. 5.2 due to its lack of sequence on the 5' end of the gene clone. OCS clones affiliated with the SAR116 cluster were <90% similar to their closest cultivated phylogenetic neighbors, members of the genus *Azospirillum* of the alpha-1 subclass of the Proteobacteria (Fig. 5.2).

Members of this genus are Gram-negative, nitrogen-fixing rods commonly associated with non-leguminous plants.

Four RFLP types in the OCS clone library (RFLP types XV, XVI, XXIII, and XXXIV) were related to the SAR11 cluster of the alpha Proteobacteria, first described from environmental clones recovered from the surface of the Sargasso Sea (Table 5.1) (Giovannoni et al. 1990a). Represented by OCS138 and OCS154 in Fig. 5.1 and OCS53, OCS138, OCS154, and OCS180 in Fig. 5.2, SAR11 cluster SSU rRNA gene clones recovered in the OCS clone library consisted of 6 total clones (10% of the bacterial clones; Table 5.1). A more detailed analysis of gene clones related to the SAR11 cluster, including one clone recovered in the OCS clone library, has been described previously (Field et al. 1997). SAR11-related OCS rRNA gene clones affiliated with two distinct subclusters of the SAR11 gene cluster (Fig. 5.2). OCS53, OCS138, and OCS180 were highly similar to each other (98.9-99.8%), and were also highly similar to Sargasso Sea clones SAR1 and SAR95 (98.0-99.1%) and FL11, an environmental SSU rRNA gene clone recovered from seawater collected from the Santa Barbara Channel (98.1-99.3% similar) (DeLong et al. 1993). OCS154 was only 91.0-91.1% similar to OCS53, OCS138, and OCS180, and was phylogenetically affiliated with SAR464 (94.9% similar for 432 positions) and OM136 (96.8% similar for 432 positions).

The three remaining OCS RFLP types related to the alpha subclass of the Proteobacteria (RFLP types XX, XXVIII, and XXXVII) are represented by OCS14, OCS124, and OCS116 in Fig. 5.2, respectively. The phylogenetic affilations of OCS14, OCS116, and OCS124 varied with respect to the taxa included in the analyis, the choice of outgroup, and the method of phylogenetic reconstruction employed, and thus were not determined with

confidence from the relatively small amount of sequence data obtained from these clones (data not shown).

Beta Proteobacteria. Six different RFLP types in the OCS clone library (9 clones total) were related to SSU rRNA genes from members of the beta subclass of the Proteobacteria (Table 5.1). Represented by OCS7 in Fig. 5.1 and OCS7 and OCS111 in Fig. 5.3, five clones from two RFLP types (VIII and XIII) in the OCS clone library were related to *Zoogloea ramigera* ATCC25935, an aerobic chemoorganotrophic bacteria isolated from activated sludge and wastewater (Fig. 5.3). OCS7 and OCS111 were 98.9% similar for 498 nucleotide positions, but were identical in sequence over the region included in the analysis shown in Fig. 5.3. A marine bacterium related to *Zoogloea ramigera* ATCC25935 has recently been isolated from the Baltic Sea (Pinhassi et al. 1997), though SSU rRNA gene sequence comparisons indicated this marine isolate is more closely related to environmental clones recovered from a dentoalveolar abscess than to OCS gene clones in this clade (Fig. 5.3). OCS7 and OCS111 were 92.8-95.8% similar to *Zoogloea ramigera* ATCC25935 and the Baltic Sea isolate *Zoogloea* str. BAL15. In addition to this marine isolate, bacteria related to *Zoogloea ramigera* ATCC25935 have been isolated from an oligotrophic arctic lake in Alaska, USA (Fig. 5.3) (Bahr et al. 1996).

RFLP types XXIV and XXV in the OCS clone library consisted of two clones which had identical sequences in the 5' end of the SSU rRNA gene, and are represented by OCS66 in Fig. 5.3. OCS66 was most similar to a unique clone in the Ocean Margins clone library, OM156 (91.9% similar for 458 nucleotide positions) (Rappé et al. 1997). Together, these two clones loosely affiliated with the genus *Burkholderia*, a group of organisms included within the original rRNA homology group II pseudomonads (87.7-89.9% similar for

Figure 5.3. Phylogenetic diversity of OCS SSU rRNA gene clones related to the beta subclass of the Proteobacteria. This dendogram was reconstructed as in Fig. 5.1 from a mask of 306 nucleotides located on the 5' end of the SSU rRNA gene. Bootstrap proportions over 60% that supported the branching order, calculated by the neighbor joining method from 100 total replicates, are shown above the corresponding nodes. The SSU rRNA gene sequence of *Escherichia coli* was used as an outgroup in this analysis.

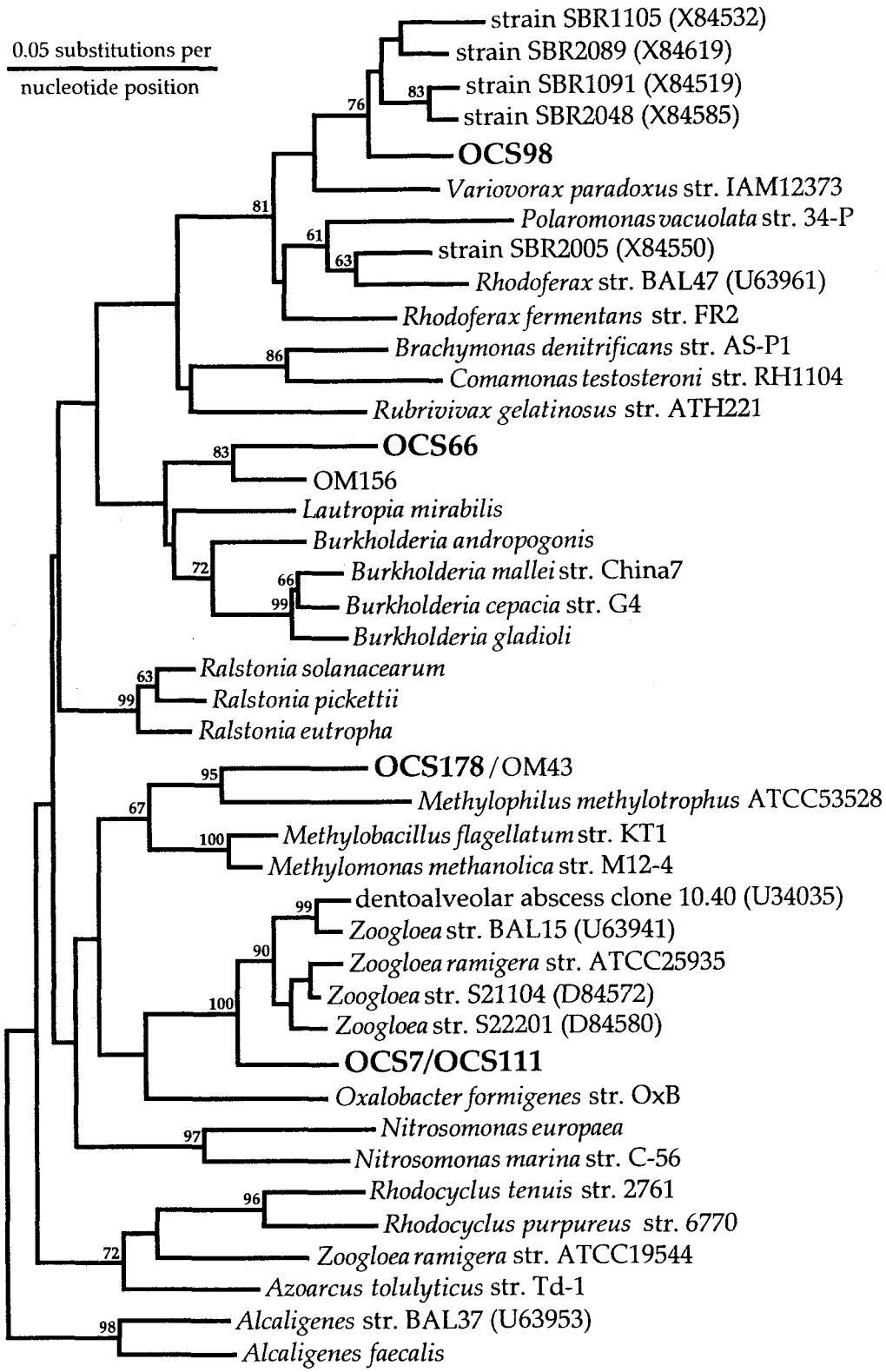

0.05 substitutions per
nucleotide position
strain SBR1105 (X84532)
strain SBR2089 (X84619)
strain SBR1091 (X84519)
strain SBR2048 (X84585)
OCS98
Variovorax paradoxus str. IAM12373
Polaromonas vacuolata str. 34-P
strain SBR2005 (X84550)
Rhodoferax str. BAL47 (U63961)
Rhodoferax fermentans str. FR2
Brachymonas denitrificans str. AS-P1
Comamonas testosteroni str. RH1104
Rubrivivax gelatinosus str. ATH221
OCS66
OM156
Lautropia mirabilis
Burkholderia andropogonis
Burkholderia mallei str. China7
Burkholderia cepacia str. G4
Burkholderia gladioli
Ralstonia solanacearum
Ralstonia pickettii
Ralstonia eutropha
OCS178 / OM43
Methylophilus methylotrophus ATCC53528
Methylobacillus flagellatum str. KT1
Methylomonas methanolica str. M12-4
dentoalveolar abscess clone 10.40 (U34035)
Zoogloea str. BAL15 (U63941)
Zoogloea ramigera str. ATCC25935
Zoogloea str. S21104 (D84572)
Zoogloea str. S22201 (D84580)
OCS7/OCS111
Oxalobacter formigenes str. OxB
Nitrosomonas europaea
Nitrosomonas marina str. C-56
Rhodocyclus tenuis str. 2761
Rhodocyclus purpureus str. 6770
Zoogloea ramigera str. ATCC19544
Azoarcus tolulyticus str. Td-1
Alcaligenes str. BAL37 (U63953)
Alcaligenes faecalis
76
83
81
61
63
86
83
72
66
99
63
99
95
67
100
99
90
100
97
96
72
98

Figure 5.3

454 nucleotide positions) (Yabuuchi et al. 1992). Identification of a close cultivated phylogenetic neighbor to the OCS66/OM156 gene lineage was not resolved in the analysis depicted in Fig. 5.3, however.

OCS98, a unique clone in the OCS clone library (RFLP type XXVII), was related to *Variovorax paradoxus*, a heterotrophic beta Proteobacteria isolated from soil (Fig. 5.3). Several relatives of this genus have recently been cultivated from phosphate- and non-phosphate-removing sequencing batch reactors (Bond et al. 1995), with which OCS98 formed a monophyletic clade moderately supported by the bootstrap analyses depicted in Fig. 5.3 (94.1-95.2% similar for 369 nucleotide positions). A second unique clone in the OCS clone library, OCS178 (RFLP type XXXIII), was closely related to a SSU rRNA gene clone recovered in the OM clone library from the eastern continental shelf of the United States, OM43 (99.4% similar for 504 nucleotide postitions) (Rappé et al. 1997). These two clones grouped within a clade of Type I methylotrophic bacteria of the beta Proteobacteria, affiliating in particular with *Methylophilus methylotrophus* str. AS1, an obligate Type I methylotrophic bacterium isolated from activated sludge (Fig. 5.3) (90.7% similar for 496 nucleotide positions). Recently, environmental rRNA gene clones related to the Type I methylotrophs of the beta Proteobacteria have also been recovered from an oligotrophic lake in arctic Alaska, USA (Bahr et al. 1996).

Gamma Proteobacteria. Four different RFLP types representing 24 OCS SSU rRNA gene clones (41% of the bacterial clones recovered) were related to the gamma subclass of the Proteobacteria (Table 5.1). RFLP type II, consisting of 13 clones represented by OCS44 in Fig. 5.1 and Fig. 5.4, and RFLP type XVIII, consisted of the unique clone OCS5 in Fig. 5.4, were affiliated with the SAR86 cluster, an SSU rRNA gene clone cluster first recovered from Sargasso

Figure 5.4. Dendogram depicting relationships between gene clones related to the SAR86 cluster of the gamma Proteobacteria. This neighbor joining analysis was performed with the program ARB and was constructed from non-overlapping gene sequences based on their relationships to taxa with full length gene sequences included in the analysis. The SSU rRNA gene sequence of *Escherichia coli* was used to root the tree.

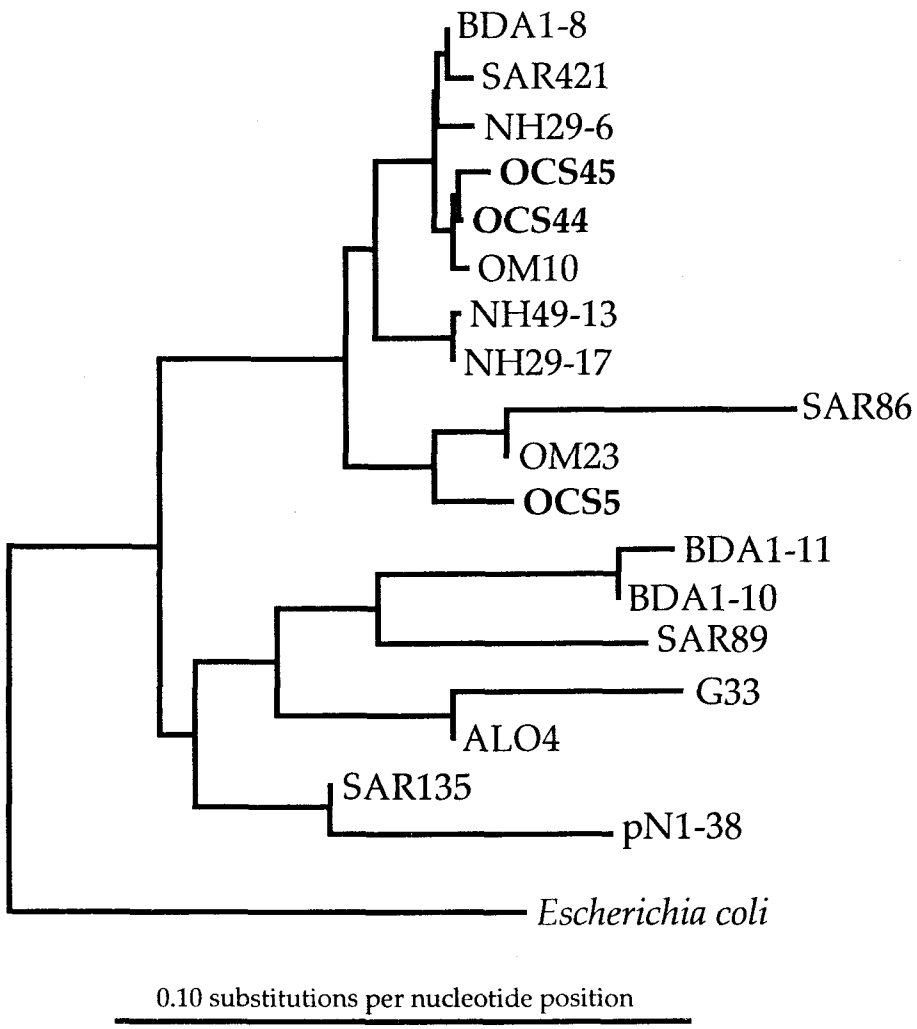

BDA1-8
SAR421
NH29-6
OCS45
OCS44
OM10
NH49-13
NH29-17
SAR86
OM23
OCS5
BDA1-11
BDA1-10
SAR89
G33
ALO4
SAR135
pN1-38
Escherichia coli
0.10 substitutions per nucleotide position

Figure 5.4

Sea bacterioplankton (Mullins et al. 1995) and the Pacific and Atlantic Oceans (Fuhrman et al. 1993). These two RFLP types made up 24% of the bacterial gene clones recovered in this library. OCS44 and OCS5 affiliated with two distinct lineages within the SAR86 cluster, and were 94.9% similar for 1524 nucleotide positions (Fig. 5.4). Each clone, however, was highly similar to environmental clones recovered in the OM clone library (OCS44 and OM10, 99.7% similar for 1527 nucleotide positions; OCS5 and OM23, 99.7% similar for 799 nucleotide positions) (Rappé et al. 1997). As in previous analyses of this gene clone cluster, no SSU rRNA gene sequence is yet available from a cultured bacterium that is >90% similar to full-length gene sequences from members of the SAR86 cluster. A secondary structural analysis of sequence variation within the SAR86 cluster, superimposed on a proposed secondary structure model for clone OCS44, is shown in Fig. 5.5. Nearly all of the substitutions differentiating lineages within the SAR86 cluster are located in hypervariable regions dispersed throughout the SSU rRNA molecule (Fig. 5.5). In addition, the majority of substitutions occurring in stem regions of the SAR86 secondary structure involved reciprocal base changes across helices (Fig. 5.5).

The partial SSU rRNA gene sequence determined for a unique clone in the OCS clone library, OCS181 (RFLP type XXVII), was identical to that of *Pseudomonas putida* DSM291, a member of the genus *Pseudomonas* of the gamma Proteobacteria (100% similar for 477 nucleotide positions; Table 5.1). This SSU rDNA clone was also closely related to the most abundant group of bacterial isolates recovered from the same seawater sample as used in construction of the OCS clone library (isolate R2A30, 95.5% similar for 354 nucleotide positions) (Suzuki et al. 1997). A recent study by Moore and coworkers

Figure 5.5. Proposed secondary structural model of clone OCS44 SSU rRNA. Numbers refer to nucleotide positions in the *Escherichia coli* SSU rRNA gene. Positions at which variation is observed between OCS44 and either OCS5 or SAR421 are indicated. Target sites for the probe SAR86R (*E. coli* positions 178-194) and the primers 27F (*E. coli* positions 8-27), 1522R (*E. coli* positions 1522-1541), and 338R (*E. coli* positions 338-355) are included.

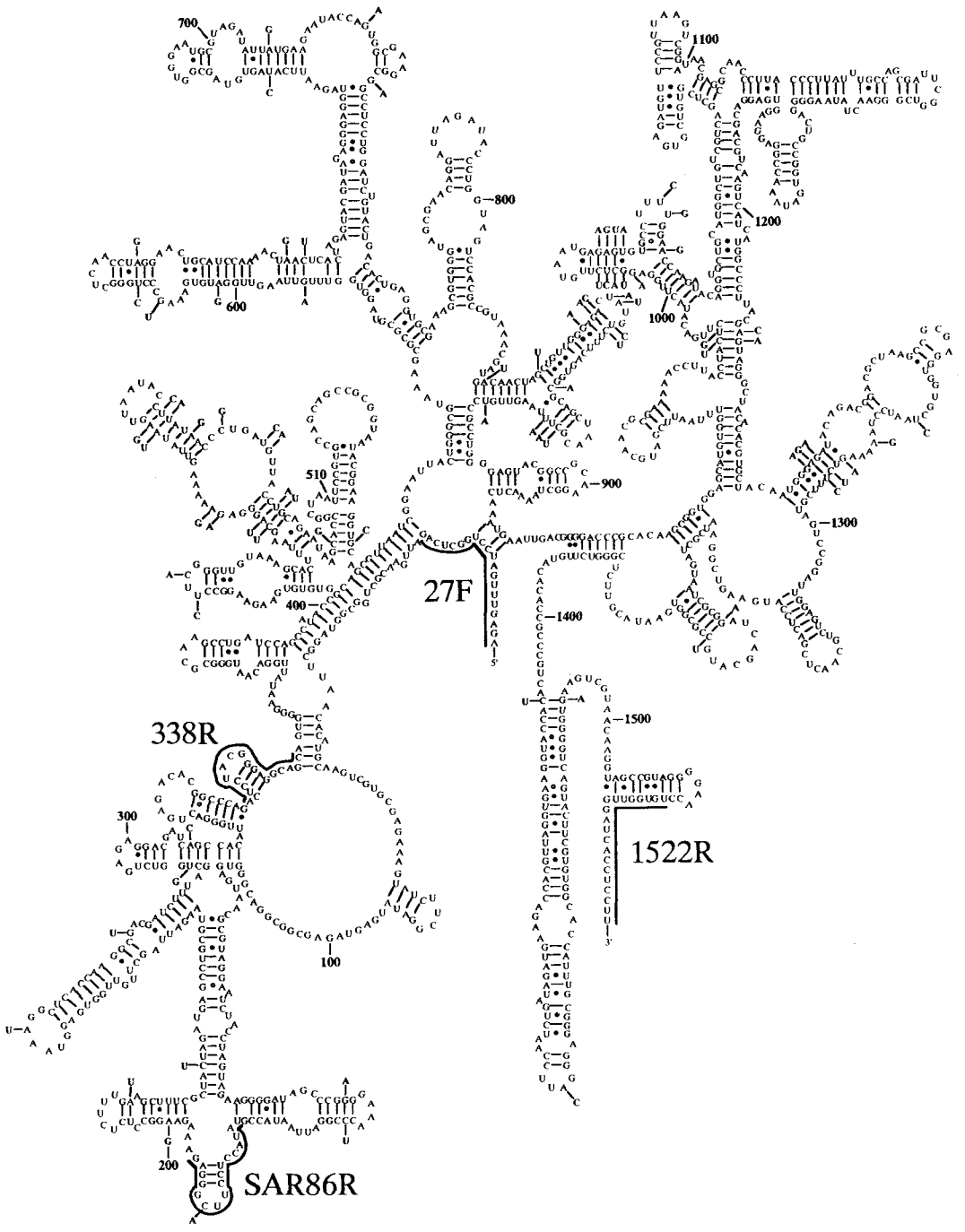


Figure 5.5

(Moore et al. 1996) describing intrageneric relationships of species within the genus *Pseudomonas* based on SSU rRNA sequence analyses afforded us the opportunity to estimate the phylgenetic affiliations of OCS181 and R2A30 within this genus. Both phylogenetic and signature sequence analyses of 497 nucleotides located on the 5' end of the SSU rRNA gene clone OCS181 strongly supported its affiliation with the "*Pseudomonas putida* lineage" within the genus *Pseudomonas* described by Moore and coworkers (data not shown). Similar analyses of a 330 nucleotide fragment of isolate R2A30 indicated it belonged to the "*Pseudomonas fluorescens* lineage" within the genus *Pseudomonas* (data not shown).

One RFLP type in the OCS clone library (RFLP type IV), consisting of 9 clones, was related to sulfur-oxidizing symbionts of the gamma Proteobacteria (Table 5.1). Though free-living bacteria related to sulfur-oxidizing symbionts have been recovered from marine sediments (Gray and Herwig 1996), partial gene sequences obtained from OCS clones related to sulfur-oxidizing symbionts were identical to the SSU rRNA gene sequence reported for the gill symbiont of the bivalve *Solemya reidi* (Cary 1994), which we used as a positive size standard in PCR reactions from the environmental nucleic acid samples. Though a negative control was included in the reactions, it did not contain a visible product band in PCR reactions used to obtain the SSU rRNA gene amplicon cloned for this study. It is possible, however, that the positive size standard contaminated the target PCR amplicon. An oligonucleotide probe specific for *Solemya reidi* symbiont SSU rRNA (Cary 1994) did not yield a positive signal when hybridized to nucleic acids isolated from bacterioplankton collected in the Pacific Ocean off the coast of Oregon, USA (data not shown).

Gram-positive bacteria. Two clones representing RFLP type X in the OCS library affiliated with a unique lineage within the Actinomycete (high G+C Gram-positive) line of descent (Table 5.1). Represented by OCS155 in Fig. 5.1, these two gene clones showed a close phylogenetic affiliation with SSU rRNA gene sequences obtained from members of the BDA1-5 (marine Gram-positive) cluster first described by Fuhrman and coworkers from fragmentary SSU rRNA gene sequences recovered in both Pacific and Atlantic Ocean samples (Fuhrman et al. 1993) and by Rappé and coworkers from full-length gene sequences recovered off the eastern continental shelf of the United States (Rappé et al. 1997). No SSU rRNA gene sequences from cultured Actinomycetes are currently available in the public sequence databases that were >82% similar to OCS155. A comparison of BDA1-5 cluster SSU rRNA gene clones with environmental SSU rRNA gene clones related to the Actinomycetes recovered from marine sediments (Gray and Herwig 1996), a peat bog (Rheims et al. 1996), and Australian forested soil (Liesack and Stackebrandt 1992, Stackebrandt et al. 1993) did not reveal a close phylogenetic neighbor to the BDA1-5 cluster (data not shown).

SAR406 Cluster. A unique clone in the OCS clone library, OCS146 (RFLP type XXXII), was related to the marine group A (Fuhrman et al. 1993) and SAR406 (Gordon and Giovannoni 1996) SSU rRNA gene lineages (Fig. 5.1). NH16-12, the longest SSU rRNA gene clone fragment related to marine group A analyzed by Fuhrman and coworkers, was 94.9% similar to OCS146 over 428 nucleotide positions. NH16-12 did not appear in Fig. 5.1 due to its lack of sequence on the 5' end of the SSU rRNA molecule. OCS146 was 94.7% similar over the entire length of the SSU rRNA molecule (1456 nucleotides) to both SAR406, recovered from a depth of 80 m in the Sargasso Sea, and OCS307,

recovered from a 120 m depth off the coast of Oregon, USA (Gordon and Giovannoni 1996). This gene cluster did not affiliate with any of the major bacterial phyla in the phylogenetic analysis of partial gene sequences depicted in Fig. 5.1 or in analyses employing sequence data from the entire length of the SSU rRNA molecule, and only loosely affiliated with members of the genus *Fibrobacter* and the green sulfur bacteria (Fig. 5.1 and data not shown).

Chimeric gene products. One OCS rDNA clone, OCS12, was preliminarily identified as related to the SAR11 cluster, but was identified as a chimeric gene product when sequence was examined from both the 5' and 3' regions of the clone. This gene clone was first described by Field and coworkers (Field et al. 1997). Briefly, a fragment approximately 1000 nucleotides in length on the 5' end of OCS12 was phylogenetically affiliated with SAR1 and SAR95, members of the SAR11 cluster of the alpha Proteobacteria. A fragment approximately 500 nucleotides in length on the 3' end of OCS12 was phylogenetically affiliated with OM43 (98.2% similar for 500 nucleotide positions), a Type I methylotroph of the beta Proteobacteria described above. In a similar manner, a second gene clone in the OCS clone library, OCS22, was identified as a chimera when sequence from both the 5' and 3' ends of the molecule was obtained. OCS22 has been previously described by Rappé and coworkers (Rappé et al. submitted). Briefly, a 400 nucleotide fragment located on the 5' end of the OCS22 gene clone was identical to OCS126, a member of the SAR116 cluster of the alpha Proteobacteria, while phylogenetic analyses of a 400 nucleotide fragment located on the 3' end of OCS22 indicated it was closely related to bacillariophyte plastid SSU rRNA genes.

LH-PCR. Fig. 5.6 shows the results of a LH-PCR electropherogram from the PCR product cloned to produce the OCS clone library, separated by natural length polymorphisms between the 27F and 338R SSU rRNA gene primer pair. The integrated area under each peak of the electropherogram was used to determine the relative abundance of each fragment length (Fig. 5.6A) (Suzuki et al. in prep). The relative abundance of different fragment lengths in the OCS clone library (Fig. 5.6C) and the bacterial isolate collection recovered from the same seawater sample (Fig. 5.6B) (Suzuki et al. 1997) are also shown. The relative distribution of fragment lengths in the OCS clone library closely resembled the distribution obtained from the PCR product cloned in this study, indicating no biases were apparent in the blunt-end ligation procedure. The relative distribution of fragment lengths obtained from the culture collection did not resemble either the distribution obtained from the clone library or the PCR product we cloned.

5.5 Discussion

The dominance of this clone library by rRNA gene clones affiliated with the SAR86 cluster, SAR116 cluster, SAR11 cluster, BDA1-5 cluster, marine group A/SAR406 cluster, and clones related to the *Roseobacter* clade indicates that they are potentially important members of bacterioplankton communities in eutrophic coastal pelagic marine systems as well as the oligotrophic open-ocean. This came as a surprise, as the site we sampled in this study is characterized by a markedly different trophic structure than that found in the oligotrophic open-ocean and is an area of unusually high productivity influenced by the Columbia River plume, coastal river effluent, and seasonal

Figure 5.6. Relative abundance of different fragment lengths in the PCR amplicon cloned in this study (A), bacterial isolates recovered from the same seawater sample (B), and clones recovered in the OCS clone library (C). Panel A was measured from an LH-PCR electropherogram, while panels B and C were measured from sequence data. The genes were separated by natural length polymorphisms between the 27F and 338R SSU rRNA gene primer pair.

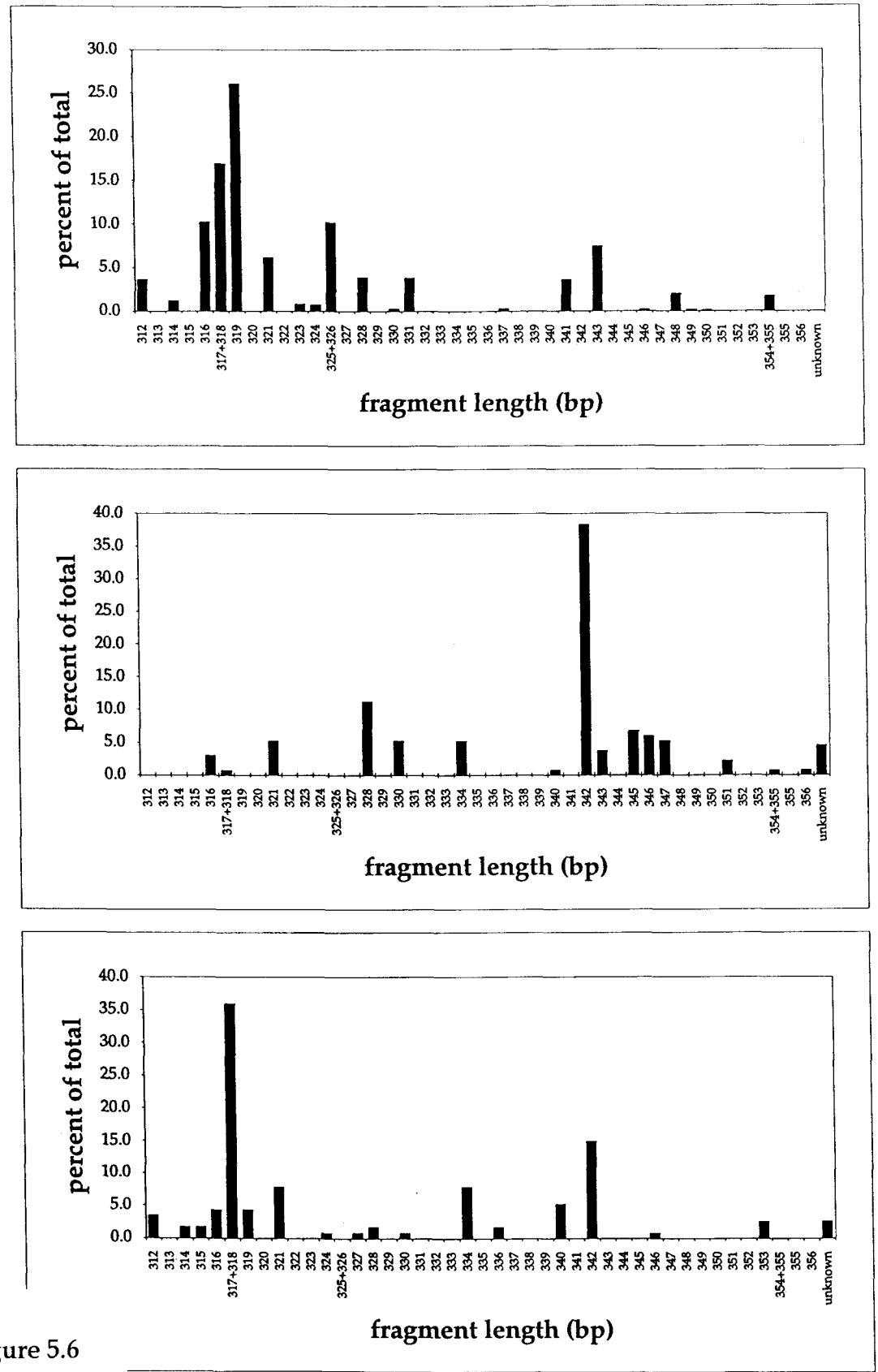

percent of total
30.0
25.0
20.0
15.0
10.0
5.0
0.0
312 313 314 315 316 317+318 319 320 321 322 323 324 325+326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354+355 355 356 unknown
fragment length (bp)
percent of total
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
312 313 314 315 316 317+318 319 320 321 322 323 324 325+326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354+355 355 356 unknown
fragment length (bp)
percent of total
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
312 313 314 315 316 317+318 319 320 321 322 323 324 325+326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354+355 355 356 unknown
fragment length (bp)

Figure 5.6

wind-driven coastal upwelling. It is interesting to note that, in addition to its cosmopolitan distribution in the world's oceans, gene clones related to the SAR11 cluster have also been recovered from an oligotrophic lake in arctic Alaska, USA (Bahr et al. 1996) and from permanently anaerobic marine sediments (Rochelle et al. 1994).

The recovery of SSU rRNA gene clones affiliated with the beta subdivision of the Proteobacteria is unique to the two clone libraries constructed from the eastern and western continental shelves of the United States, as no beta Proteobacteria have been recovered from oligotrophic open-ocean clone libraries. It is plausible that the bacterioplankton represented by these gene lineages are adapted to continental shelf or coastal seawater environments. It is interesting to note that gene clones related to obligate Type I methylotrophic bacteria were recovered from both continental shelves, though no microorganisms related to this clade have been cultured from seawater. In addition, the recovery of rRNA gene clones related to the genus *Zoogloea* is potentially significant, as members of this genus are capable of forming flocculent masses of cells embedded in gelatinous matrices. The formation of these cellular aggregates may have significant implications for organic matter cycling in this environment.

In addition to beta Proteobacteria, plastid SSU rRNA gene clones were recovered in previous studies of bacterioplankton diversity from the eastern continental shelf of the United States as well as the Santa Barbara Channel, so their presence in this clone library was anticipated. Several factors, outlined by Rappé and coworkers (1997), probably contributed to the abundance of plastid SSU rRNA gene clones and the lack of prokaryotic oxygenic phototroph SSU rRNA gene clones in this library. These include the relative

abundances of eukaryotic phytoplankton compared to prokaryotic oxygenic phototrophs in this sample, the multiplicity of chloroplasts in many phytoplankton taxa, and the multiplicity of plastid genomes within these plastids.

The SSU rRNA gene clusters we describe here come from uncultivated organisms, and it is not known to what extent they represent separate cellular lineages or sequence heterogeneity among ribosomal operons of a single strain. The extent of sequence divergence between separate lineages within gene clone clusters recovered from marine environments is frequently up to 10%, and appears to be much larger than the level of sequence divergence obtained from the limited number of examples where separate ribosomal operons have been sequenced from the same microorganism (e.g. see Cilia et al. 1996). Thus, inter-operon heterogeneity probably explains only a minor fraction of the divergence observed in the environmental SSU rRNA gene clusters. In addition, we employed *Pfu* DNA polymerase in the amplification of the SSU rRNA genes analyzed in this study. This thermostable enzyme contains a 3' to 5' proofreading activity, reducing the number of mis-incorporated nucleotides in our PCR reactions. The ecological significance of these gene clone clusters has yet to be elucidated, though we have reported oligonucleotide probe hybridization data which indicates the genetic diversity within the SAR11 gene cluster includes gene lineages which exhibit different depth-specific distributions (Field et al. 1997).

Additional oligonucleotide probe hybridization data from our laboratory indicate that SSU rRNA genes from other uncultured bacterioplankton display pronounced temporal and spatial distribution patterns (Giovannoni et al. 1996b, Gordon and Giovannoni 1996, Wright et al.

1997). In this study, we determined full-length SSU rRNA sequences for three unique members of the SAR116 cluster, two unique members of the SAR86 cluster, and a unique member of the marine group A/SAR406 cluster. By determining complete gene sequences for individual lineages within environmental gene clone clusters, oligonucleotide probes can be designed to be specific for many members of a gene cluster, or to individual lineages within a cluster. This level of resolution was neccessary to detect the niche specialization we observed in subclusters of the SAR11 cluster (Field et al. 1997). Niche specializations within the other marine pelagic environmental gene clone clusters described here are anticipated.

Of the approximately 74% of the marine bacterial species described in the systematic literature that are represented by SSU rRNA gene sequences in sequence databases, the median similarity between these and rRNA gene clones retrieved from natural samples of bacterioplankton is 87% (Giovannoni et al. 1996a). Distant phylogenetic relationships such as these are not a reliable basis from which to predict the phenotype of an unknown organism. Thus, the ecological roles and physiological activities of the vast majority of phylotypes recovered in this clone library are currently unknown. The tools of molecular genetics yields a glimpse of the diversity, abundance, and relative distributions of bacteria in their natural habitat, but the cultivation of members of dominant environmental SSU rRNA gene clone clusters will ultimately be neccessary to gain insight into the physiological properties of these marine bacteria.

5.6 References

Amann, R. I., W. Ludwig and K.-H. Schleifer. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microb. Rev. 59: 143-169.

Ashen, J. B. and L. J. Goff. 1996. Molecular identification of a bacterium associated with gall formation in the marine red alga *Prionitis lanceolata*. J. Phycol. 32: 286-297.

Azam, F. and B. Cho. 1987. Bacterial utilization of organic matter in the sea. In Ecology of microbial communities (ed. M. Fletcher, T. R. G. Gray, and J. G. Jones), pp. 261-281. Symp. Soc. Gen. Microbiol.

Azam, F., D. C. Smith, R. A. Long and G. F. Steward. 1995. Bacteria in oceanic carbon cycling as a molecular problem. In Molecular ecology of aquatic microbes (ed. I. Joint), pp. 39-54. Springer-Verlag, Berlin.

Bahr, M., J. E. Hobbie and M. L. Sogin. 1996. Bacterial diversity in an arctic lake: a freshwater SAR11 cluster. Aquat. Microb. Ecol. 11: 271-277.

Baumann, L., P. Baumann, M. Mandel and R. D. Allen. 1972. Taxonomy of aerobic marine bacteria. J. Bacteriol. 110: 402-429.

Bond, P. L., P. Hugenholtz, J. Keller and L. L. Blackall. 1995. Bacterial community structures of phosphate-removing and non-phosphate-removing activated sludges from sequencing batch reactors. Appl. Environ. Microbiol. 61: 1910-1916.

Bowman, J. P., S. A. McCammon, M. V. Brown and T. A. McMeekin. 1997. unpublished sequences in Genbank.

Britschgi, T. B. and S. J. Giovannoni. 1991. Phylogenetic analysis of a natural marine bacterioplankton population by rRNA gene cloning and sequencing. Appl. Environ. Microbiol. 57: 1707-1713.

Cary, S. C. 1994. Vertical transmission of a chemoautotrophic symbiont in the protobranch bivalve, *Solemya reidi*. Mol. Mar. Biol. Biotechnol. 3: 121-130.

Choi, J. W., E. B. Sherr and B. F. Sherr. 1996. Relation between presence-absence of a visible nucleoid and metabolic activity in bacterioplankton cells. Limnol. Oceanogr. 41: 1161-1168.

Cilia, V., B. Lafay and R. Christen. 1996. Sequence heterogeneities among 16S ribosomal RNA sequences, and their effect on phylogenetic analyses at the species level. Mol. Biol. Evol. 13: 451-461.

Cole, J. J., S. Findlay and M. L. Pace. 1988. Bacterial production in fresh and saltwater ecosystems: a cross-system overview. Mar. Ecol. Prog. Ser. 43: 1-10.

DeLong, E. F., D. G. Franks and A. L. Alldredge. 1993. Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. Limnol. Oceanogr. 38: 924-934.

Ducklow, H. W., D. Purdie, P. J. L. Williams and J. M. Davies. 1986. Bacterioplankton: a sink for carbon in a coastal marine plankton community. Science. 232: 865-867.

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

Felsenstein, J. 1989. PHYLIP-phylogeny inference package (v3.5). Cladistics. 5: 164-166.

Ferguson, R. L., E. N. Buckley and A. V. Palumbo. 1984. Response of marine bacterioplankton to differential filtration and confinement. Appl. Environ. Microbiol. 47: 49-55.

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

Fuhrman, J. A., S. H. Lee, Y. Masuchi, A. A. Davis and R. M. Wilcox. 1994. Characterization of marine prokaryotic communities via DNA and RNA. Microb. Ecol. 28: 133-145.

Fuhrman, J. A., K. McCallum and A. A. Davis. 1993. Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific oceans. Appl. Environ. Microbiol. 59: 1294-1302.

Fuhrman, J. A., T. D. Sleeter, C. A. Carlson and L. Proctor. 1989. Dominance of bacterial biomass in the Sargasso Sea and its ecological implications. Mar. Ecol. Prog. Ser. 57: 207-217.

Giovannoni, S. J., T. B. Britschgi, C. L. Moyer and K. G. Field. 1990a. Genetic diversity in Sargasso Sea bacterioplankton. Nature. 345: 60-63.

Giovannoni, S. J., E. F. DeLong, T. M. Schmidt and N. R. Pace. 1990b. Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. Appl. Environ. Microbiol. 56: 2572-2575.

Giovannoni, S. J., M. S. Rappé, D. Gordon, E. Urbach, M. Suzuki and K. G. Field. 1996a. Ribosomal RNA and the evolution of bacterial diversity. In Evolution of microbial life (ed. D. M. Roberts, P. Sharp, G. Alderson, and M. Collins), pp. 63-85. Cambridge University Press.

Giovannoni, S. J., M. S. Rappé, K. Vergin and N. Adair. 1996b. 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the Green Non-Sulfur bacteria. Proc. Natl. Acad. Sci. USA. 93: 7979-7984.

Giovannoni, S. J., T.D. Mullins and K.G. Field. 1995. Microbial diversity in oceanic systems: rRNA approaches to the study of unculturable microbes. In Molecular ecology of aquatic microbes (ed. I. Joint), pp. 217-248. Springer-Verlag, New York.

Gordon, D. A. and S. J. Giovannoni. 1996. Detection of stratified microbial populations related to *Chlorobium* and *Fibrobacter* species in the Atlantic and Pacific oceans. Appl. Environ. Microbiol. 62: 1171-1177.

Gray, J. P. and R. P. Herwig. 1996. Phylogenetic analysis of the bacterial communities in marine sediments. Appl. Environ. Microbiol. 62: 4049-4059.

Hiraishi, A. and Y. Ueda. 1994. Intrageneric structure of the genus *Rhodobacter*: transfer of *Rhodobacter sulfidophilus* and related marine species to the genus *Rhodovolum* gen. nov. Int. J. Syst. Bacteriol. 44: 15-23.

Hiraishi, A. and Y. Ueda. 1995. Isolation and characterization of *Rhodovulum strictum* sp. nov. and some other purple nonsulfur bacteria from colored blooms in tidal and seawater pools. Int. J. Syst. Bacteriol. 45: 319-326.

Hobbie, J. E., R. J. Daley and S. Jasper. 1977. Use of nuclepore filters for counting bacteria by fluorescence microscopy. Appl. Environ. Microbiol. 33: 1225-1228.

Holmes, A. J., D. P. Kelly, S. C. Baker, A. S. Thompson, P. De Marco, E. M. Kenna and J. C. Murrell. 1997. *Methylosulfonomonas methylovora* gen. nov., sp. nov., and *Marinosulfonomonas methylotropha* gen. nov., sp. nov.: novel methylotrophs able to grow on methanesulfonic acid. Arch. Microbiol. 167: 46-53.

Jannasch, H. W. and G. E. Jones. 1959. Bacterial populations in seawater as determined by different methods of enumeration. Limnol. Oceanogr. 4: 128-139.

Kemp, P. F. 1995. Can we estimate bacterial growth rates from ribosomal RNA content? In Molecular ecology of aquatic microbes (ed. I. Joint), pp. 279-302. Springer-Verlag, Berlin.

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

Kita-Tsukamoto, K., H. Oyaizu, K. Nanba and U. Simidu. 1993. Phylogenetic relationships of marine bacteria, mainly members of the family *Vibrionaceae*, determined on the basis of 16S rRNA sequences. Int. J. Syst. Bacteriol. 43: 8-19.

Kogure, K., U. Simidu and N. Taga. 1979. A tentative direct microscopic method for counting living marine bacteria. Can. J. Microbiol. 25: 415-420.

Lafay, B., R. Ruimy, C. Rausch de Traubenberg, V. Breittmayer, M. J. Gauthier and R. Christen. 1995. *Roseobacter algicola* sp. nov., a new marine bacterium isolated from the phycosphere of the toxin-producing dinoflagellate *Prorocentrum lima*. Int. J. Syst. Bacteriol. 45: 290-296.

Lane, D. J. 1991. 16S/23S rRNA sequencing. In Nucleic acid techniques in bacterial systematics (ed. E. Stackebrandt and M. Goodfellow), pp. 115-175. John Wiley and Sons, New York.

Ledyard, K. M., E. F. DeLong and J. W. H. Dacey. 1993. Characterization of a DMSP-degrading bacterial isolate from the Sargasso Sea. Arch. Microbiol. 160: 312-318.

Lee, S. and J. A. Fuhrman. 1991. Spatial and temporal variation of natural bacterioplankton assemblages studied by total genomic DNA cross-hybridization. Limnol. Oceanogr. 36: 1277-1287.

Liesack, W. and E. Stackebrandt. 1992. Occurence of novel groups of the domain *Bacteria* as revealed by analysis of genetic material isolated from an Australian terrestrial environment. J. Bacteriol. 174: 5072-5078.

Ludwig, W. and O. Strunk. 1997. Available on the WWW from http://www.mikro.biologie.tu-muenchen.de.

Maidak, B. L., N. Larsen, M. J. McCaughey, R. Overbeek, G. J. Olsen, K. Fogel, J. Blandy and C. R. Woese. 1994. The Ribosomal Database Project. Nucleic Acids Res. 22: 3485-3487.

Martinez, J., D. C. Smith, G. F. Steward and F. Azam. 1996. Variability in ectohydrolytic enzyme activites of pelagic marine bacteria and its significance for substrate processing in the sea. Aquat. Microb. Ecol. 10: 223-230.

Moore, E. R. B., M. Mau, A. Arnscheidt, E. Bottger, R. A. Hutson, M. D. Collins, Y. Van De Peer, R. De Wachter and K. N. Timmis. 1996. The determination and comparison of the 16S rRNA gene sequences of species of the genus *Pseudomonas* (sensu stricto) and estimation of the natural intrageneric relationships. Syst. Appl. Microbiol. 19: 478-492.

Mullins, T. D., T. B. Britschgi, R. L. Krest and S. J. Giovannoni. 1995. Genetic comparisons reveal the same unknown bacterial lineages in Atlantic and Pacific bacterioplankton communities. Limnol. Oceanogr. 40: 148-158.

Olsen, G. J., D. L. Lane, S. J. Giovannoni and N. R. Pace. 1986. Microbial ecology and evolution: a ribosomal RNA approach. Ann. Rev. Microbiol. 40: 337-365.

Pinhassi, J., U. L. Zweifel and Å. Hagström. 1997. unpublished sequences in Genbank.

Rappé, M. S., P. F. Kemp and S. J. Giovannoni. 1995. Chromophyte plastid 16S ribosomal RNA genes found in a clone library from Atlantic Ocean seawater. J. Phycol. 31: 979-988.

Rappé, M. S., P. F. Kemp and S. J. Giovannoni. 1997. Phylogenetic diversity of marine coastal picoplankton 16S rRNA genes cloned from the continental shelf off Cape Hatteras, North Carolina. Limnol. Oceanogr. in press.

Rappé, M. S., M. T. Suzuki, K. Vergin and S. J. Giovannoni. in prep. Phylogenetic diversity of ultraplankton plastid SSU rRNA genes recovered in environmental nucleic acid samples from the Pacific and Atlantic coasts of the United States. Appl. Environ. Microbiol.

Rehnstam, A. S., S. Bäckman, D. C. Smith, F. Azam and Å. Hagström. 1993. Blooms of sequence-specific culturable bacteria in the sea. FEMS Microbiol. Ecol. 102: 161-166.

Rheims, H., C. Spröer, F. A. Rainey and E. Stackebrandt. 1996. Molecular biological evidence for the occurence of uncultured members of the

actinomycete line of descent in different environments and geographical locations. Microbiol. 142: 2863-2870.

Rochelle, P. A., B. A. Cragg, J. C. Fry, R. J. Parkes and A. J. Weightman. 1994. Effect of sample handling on estimation of bacterial diversity in marine sediments by 16S rRNA gene sequence analysis. FEMS Microbiol. Ecol. 15: 215-226.

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

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

Schmidt, T. M., E. F. DeLong and N. R. Pace. 1991. Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. J. Bacteriol. 173: 4371-4378.

Sherr, E. B. and B. F. Sherr. 1988. Role of microbes in pelagic food webs: A revised concept. Limnol. Oceanogr. 33: 1225-1227.

Shiba, T. 1991. *Roseobacter litoralis* gen. nov., sp. nov., and *Roseobacter denitrificans* sp. nov., aerobic pink-pigmented bacteria which contain bacteriochlorophyll *a*. System. Appl. Microbiol. 14: 140-145.

Smith, D. C., G. F. Steward, R. A. Long and F. Azam. 1995. Bacterial mediation of carbon fluxes during a diatom bloom in a mesocosm. Deep Sea Res. II. 42: 75-97.

Stackebrandt, E., W. Liesack and B. M. Göebel. 1993. Bacterial diversity in a soil sample from a subtropical Australian environment as determined by 16S rRNA analysis. FASEB J. 7: 232-236.

Staley, J. T. and A. Konopka. 1985. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. Ann. Rev. Microbiol. 39: 321-346.

Suzuki, M., M. S. Rappé and S. J. Giovannoni. in prep. Estimating coastal picoplankton community structure by LH-PCR. Appl. Environ. Microbiol.

Suzuki, M. T. and S. J. Giovannoni. 1996. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Appl. Environ. Microbiol. 62: 625-630.

Suzuki, M. T., M. S. Rappé, Z. W. Haimberger, H. Winfield, N. Adair, J. Ströbel and S. J. Giovannoni. 1997. Bacterial diversity among SSU rRNA gene clones and cellular isolates from the same seawater sample. Appl. Environ. Microbiol. 63: 983-989.

Ward, D. M., M. M. Bateson, R. Weller and A. L. Ruff-Roberts. 1992. Ribosomal RNA analysis of microorganisms as they occur in nature. Adv. Microb. Ecol. 12: 219-286.

Woese, C. R., E. Stackebrandt, W. G. Weisburg, B. J. Paster, M. T. Madigan, V. J. Fowler, C. M. Hahn, P. Blanz, R. Gupta, K. H. Nealson and G. E. Fox. 1984. The phylogeny of purple bacteria: the alpha subdivision. Syst. Appl. Microbiol. 5: 315-326.

Wright, T. D., K. Vergin, P. W. Boyd and S. J. Giovannoni. 1997. A novel δ-subdivision proteobacterial lineage stratified in the lower ocean surface layer. Appl. Environ. Microbiol. 63: 1441-1448.

Yabuuchi, E., Y. Kosako, H. Oyaizu, I. Yano, H. Hotta, Y. Hashimoto, T. Ezaki and M. Arakawa. 1992. Proposal of *Burkholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov. Microbiol. Immunol. 36: 1251-1275.

ZoBell, C. E. 1946. Marine microbiology, a monograph on hydrobacteriology. Chronica Botanica, Waltham, Mass.

Zweifel, U. L. and Å. Hagström. 1995. Total counts of marine bacteria include a large fraction of non-nucleoid-containing bacteria (ghosts). Appl. Environ. Microbiol. 61: 2180-2185.

CHAPTER 6.
SUMMARY

In this study, 285 SSU rRNA gene clones were analyzed from samples collected over the eastern and western continental shelves of the United States. In contrast to SSU rRNA gene clone libraries constructed from the oligotrophic open-ocean, a large number of clones recovered in this study (109) were derived from the plastid genomes of several recognizable groups of eukaryotic phytoplankton, including members of the Bacillariophyceae, Prymnesiophyceae, Cryptophyceae, Chrysophyceae, and Prasinophyceae. Also recovered were several unique plastid SSU rRNA gene clones for which close cultivated phylogenetic neighbors could not be identified.

A vast majority (72%) of the bacterial rDNAs recovered in the OCS and OM clone libraries were related to previously detected rDNA lineages cloned from oligotrophic open-ocean marine habitats, including the SAR86 cluster (γ Proteobacteria), SAR83 cluster, SAR116 cluster, and SAR11 clusters (all α Proteobacteria), the BDA1-5 cluster (Actinomycetes), the marine group A/SAR406 cluster, a lineage within the cytophaga-flavobacteria-bacteriodes phylum, and a phylogenetically distinct lineage remotely associated with the Planctomycetes. The fact that these lineages have been recovered in independent investigations of marine prokaryotic diversity make them widespread in seawater and potentially important members of bacterioplankton communities. The picture of environmental SSU rRNA gene clone diversity obtained from bacterioplankton populations studied to date, including work described here, is shown graphically in Figure 6.1. Focusing on these abundant taxa is an obvious choice for addressing community

Figure 6.1. Frequency of the most common marine bacterioplankton 16S rRNA gene clusters, calculated from a combined data set of 442 environmental 16S rRNA gene clones from pelagic marine samples. This data set is composed of all published pelagic marine 16S rRNA gene clone libraries, including the OCS and OM clone libraries. SAR6 and SAR7 are environmental clones from the marine *Synechococcus/Prochlorococcus* cluster.

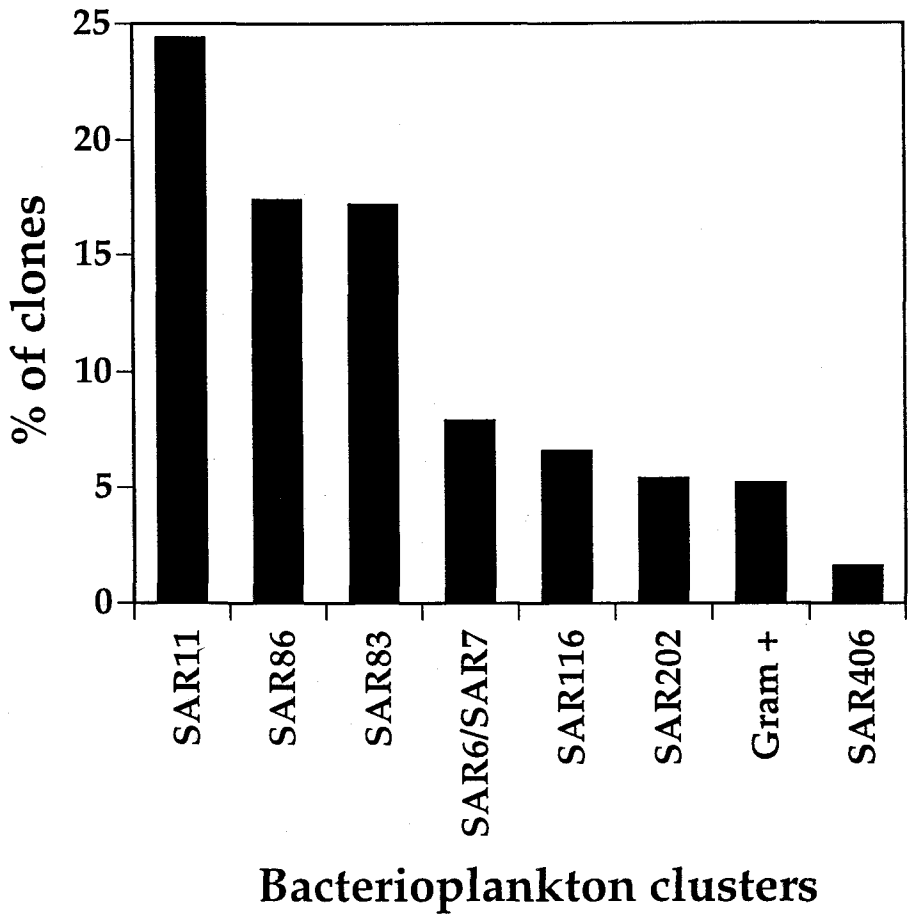

25
20
15
10
5
0
% of clones
SAR11
SAR86
SAR83
SAR6/SAR7
SAR116
SAR202
Gram +
SAR406
Bacterioplankton clusters

Figure 6.1

dynamics in what now appears to be a wide range of pelagic marine habitats including productive, nutrient-rich coastal environments.

This study also uncovered members of the β subclass of the Proteobacteria that had not been observed previously in 16S rDNA clone libraries constructed from seawater, including a unique lineage closely allied to known taxa of obligate Type I methylotrophic bacteria recovered in both clone libraries. This clade of β Proteobacteria use the RuMP pathway for C1 assimilation and are unable to oxidize methane. No microorganisms related to this clade have been cultured from seawater. The evidence of methylotrophs was a surprise, since there was nothing about our study sites to suggest that they, in particular, should be the locus of unusual levels of C1 metabolism. The presence of these genes originating within a clade of Type I methylotrophs does not necessarily imply that C1 metabolism was a prominent feature of the carbon cycle in these coastal regions. Regardless, the recovery of SSU rRNA gene clones related to this clade is an exciting discovery, and is sure to provoke further investigation. The recovery of rRNA gene clones related to the genus *Zoogloea* is also potentially significant, as members of this genus are capable of forming flocculent masses of cells embedded in gelatinous matrices. The formation of these cellular aggregates may have significant implications for organic matter cycling in this environment. We believe it is plausible that bacterioplankton affiliated with the β Proteobacteria are adapted to continental shelf or coastal seawater environments.

In addition to providing significant insight into continental shelf bacterioplankton community structure, the vast amount of nucleotide sequence data from rDNA genes cloned from the eastern and western continental shelves of the United States is currently being employed in the construction of

taxon-specific oligonucleotide probes, thus enabling studies which focus on the ecology of these uncultured bacteria in natural settings or which seek to isolate them into culture. Overlap between environmental rDNA clones from vastly different regions of the world's oceans suggests that carbon flux in bacterioplankton communities might be addressed by focusing on the activities of a few key microbial groups.

BIBLIOGRAPHY

Alfreider, A., J. Pernthaler, R. Amann, B. Sattler, F.-O. Glöckner, A. Wille and R. Psenner. 1996. Community analysis of the bacterial assemblages in the winter cover and pelagic layers of a high mountain lake by in situ hybridization. Appl. Environ. Microbiol. 62: 2138-2144.

Amann, R. I., B. J. Binder, R. J. Olson, S. W. Chisholm, R. Devereux and D. A. Stahl. 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl. Environ. Microbiol. 56:1919-1925.

Amann, R. I., W. Ludwig and K.-H. Schleifer. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microb. Rev. 59:143-169.

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

Ashen, J. B. and L. J. Goff. 1996. Molecular identification of a bacterium associated with gall formation in the marine red alga *Prionitis lanceolata*. J. Phycol. 32:286-297.

Azam, F. and B. Cho. 1987. Bacterial utilization of organic matter in the sea. In Ecology of microbial communities (ed. M. Fletcher, T. R. G. Gray, and J. G. Jones), pp. 261-281. Symp. Soc. Gen. Microbiol.

Azam, F., T. Fenchel, J. G. Field, J. S. Gray, L. A. Meyer-Reil and F. Thingstad. 1983. The ecological role of water-column microbes in the sea. Mar. Ecol. Prog. Ser. 10:257-263.

Azam, F., D. C. Smith, R. A. Long and G. F. Steward. 1995. Bacteria in oceanic carbon cycling as a molecular problem. In Molecular ecology of aquatic microbes (ed. I. Joint), pp. 39-54. Springer-Verlag, Berlin.

Bahr, M., J. E. Hobbie and M. L. Sogin. 1996. Bacterial diversity in an arctic lake: a freshwater SAR11 cluster. Aquat. Microb. Ecol. 11:271-277.

Barber, R. T. and R. L. Smith. 1981. Coastal upwelling ecosystems. In Analysis of marine ecosystems (ed. A. R. Longhurst), pp. 31-68. Academic Press, New York.

Barns, S. M., R. E. Fundyga, M. W. Jeffries and N. R. Pace. 1994. Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. Proc. Natl. Acad. Sci. USA. 91:1609-1613.

Baumann, L., P. Baumann, M. Mandel and R. D. Allen. 1972. Taxonomy of aerobic marine bacteria. J. Bacteriol. 110:402-429.

Billen, G., C. Joiris, L. Meyer-Reil and H. Lindeboom. 1990. Role of bacteria in the North Sea ecosystem. Neth. J. Sea Res. 26:265-293.

Bond, P. L., P. Hugenholtz, J. Keller and L. L. Blackall. 1995. Bacterial community structures of phosphate-removing and non-phosphate-removing activated sludges from sequencing batch reactors. Appl. Environ. Microbiol. 61:1910-1916.

Bowman, J. P., S. A. McCammon, M. V. Brown and T. A. McMeekin. 1997. Unpublished sequences in Genbank.

Brand, L. E. 1986. Nutrition and culture of autotrophic ultraplankton and picoplankton. In Photosynthetic picoplankton (ed. T. Platt and W. K. W. Li), pp. 205-233. Can. Bull. Fish. Aquat. Sci., Ottawa.

Britschgi, T. B. and S. J. Giovannoni. 1991. Phylogenetic analysis of a natural marine bacterioplankton population by rRNA gene cloning and sequencing. Appl. Environ. Microbiol. 57:1707-1713.

Brosius, J., M. L. Palmer, P. J. Kennedy and H. F. Noller. 1978. Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*. Proc. Natl. Acad. Sci. USA. 75:4801-4805.

Burrage, D. M. and R. W. Garvine. 1988. Summertime hydrography at the shelfbreak front in the Middle Atlantic Bight. J. Phys. Oceanogr. 18:1309-1319.

Campbell, L., L. P. Shapiro and E. Haugen. 1994. Immunochemical characterization of eukaryotic ultraplankton from the Atlantic and Pacific oceans. J. Plankton Res. 16:35-51.

Carlson, C. A., H. W. Ducklow and A. F. Michaels. 1994. Annual flux of dissolved organic carbon from the euphotic zone in the northwestern Sargasso Sea. Nature. 371:405-408.

Cary, S. C. 1994. Vertical transmission of a chemoautotrophic symbiont in the protobranch bivalve, *Solemya reidi*. Mol. Mar. Biol. Biotechnol. 3:121-130.

Cattolico, R. A. and S. Loiseaux-de Göer. 1989. Analysis of chloroplast evolution and phylogeny: a molecular approach. In The chromophyte algae: problems and prospectives (ed. J. C. Green, B. S. C. Leadbeater, and W. L. Diver), pp. 85-100. Clarendon Press, Oxford.

Chapman, D. C. and G. Gawarkiewicz. 1993. On the establishment of the seasonal pycnocline in the Middle Atlantic Bight. J. Phys. Oceanogr. 23:2487-2492.

Choi, J. W., E. B. Sherr and B. F. Sherr. 1996. Relation between presence-absence of a visible nucleoid and metabolic activity in bacterioplankton cells. Limnol. Oceanogr. 41:1161-1168.

Chrétiennot-Dinet, M.-J., A. Sournia, M. Ricard and C. Billard. 1993. A classification of the marine phytoplankton of the world from class to genus. Phycologia. 32:159-179.

Cilia, V., B. Lafay and R. Christen. 1996. Sequence heterogeneities among 16S ribosomal RNA sequences, and their effect on phylogenetic analyses at the species level. Mol. Biol. Evol. 13:451-461.

Cole, J. J., S. Findlay and M. L. Pace. 1988. Bacterial production in fresh and saltwater ecosystems: a cross-system overview. Mar. Ecol. Prog. Ser. 43:1-10.

Coleman, A. W. and L. J. Goff. 1991. DNA analysis of eukaryotic algal species. J. Phycol. 27:463-473.

Courties, C., A. Vaquer, M. Troussellier, J. Lautier, M. J. Chrétiennot-Dinet, J. Neveux, C. Machado and H. Claustre. 1994. Smallest eukaryotic organism. Nature. 370:255.

Dams, E., L. Hendriks, Y. Van de Peer, J.-M. Neefs, G. Smits, I. Vandenbempt and R. De Wachter. 1988. Compilation of small ribosomal subunit RNA sequences. Nucleic Acids Res. 16:r87-r173.

Delaney, T. P. and R. A. Cattolico. 1989. Chloroplast ribosomal DNA organization in the chromophytic alga *Olisthodiscus luteus*. Curr. Genet. 15:221-229.

DeLong, E. F. 1992. Archaea in coastal marine environments. Proc. Natl. Acad. Sci. USA. 89:5685-5689.

DeLong, E. F., D. G. Franks and A. L. Alldredge. 1993. Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. Limnol. Oceanogr. 38:924-934.

DeLong, E. F., G. S. Wickham and N. R. Pace. 1989. Phylogenetic stains: Ribosomal RNA-based probes for the identification of single cells. Science. 243:1360-1363.

DeLong, E. F., K. Y. Wu, B. B. Prézelin and R. V. M. Jovine. 1994. High abundance of Archaea in Antarctic marine picoplankton. Nature. 371:695-697.

Distel, D. L., D. L. Lane, G. J. Olsen, S. J. Giovannoni, B. Pace, N. R. Pace, D. A. Stahl and H. Felbeck. 1988. Sulfur-oxidizing bacterial endosymbionts: Analysis of phylogeny and specificity by 16S rRNA sequences. J. Bacteriol. 170:2506-2510.

Douglas, S. E. and S. Turner. 1991. Molecular evidence for the origin of plastids from a cyanobacterium-like ancestor. J. Mol. Evol. 33:267-273.

Ducklow, H. W. 1994. Modeling the microbial food web. Microb. Ecol. 28:303-319.

Ducklow, H. W. and C. A. Carlson. 1992. Ocean bacterial production. Adv. Mar. Biol. 12:113-181.

Ducklow, H. W., D. Purdie, P. J. L. Williams and J. M. Davies. 1986. Bacterioplankton: a sink for carbon in a coastal marine plankton community. Science. 232:865-867.

Ersland, D. R., J. Aldrich and R. A. Cattolico. 1981. Kinetic complexity, homogeneity, and copy number of chloroplast DNA from the marine alga *Olisthodiscus luteus*. Plant Physiol. 68:1468-1473.

Farrelly, V., F. A. Rainey and E. Stackebrandt. 1995. Effect of genome size and rrn gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. Appl. Environ. Microbiol. 61:2798-2801.

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

Felsenstein, J. 1989. PHYLIP-phylogeny inference package (v3.5). Cladistics. 5:164-166.

Ferguson, R. L., E. N. Buckley and A. V. Palumbo. 1984. Response of marine bacterioplankton to differential filtration and confinement. Appl. Environ. Microbiol. 47:49-55.

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

Fox, G. E., E. Stackebrandt, R. B. Hespell, J. Gibson, J. Maniloff, T. A. Dyer, R. S. Wolfe, W. E. Balch, R. S. Tanner, L. J. Magrum, L. B. Zablen, R. Blakemore, R. Gupta, L. Bonen, B. J. Lewis, D. A. Stahl, K. R. Luehrsen, K. N. Chen and C. R. Woese. 1980. The phylogeny of prokaryotes. Science. 209:457-463.

Fuhrman, J. A., S. H. Lee, Y. Masuchi, A. A. Davis and R. M. Wilcox. 1994. Characterization of marine prokaryotic communities via DNA and RNA. Microb. Ecol. 28:133-145.

Fuhrman, J. A., K. McCallum and A. A. Davis. 1993. Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific oceans. Appl. Environ. Microbiol. 59:1294-1302.

Fuhrman, J. A., T. D. Sleeter, C. A. Carlson and L. Proctor. 1989. Dominance of bacterial biomass in the Sargasso Sea and its ecological implications. Mar. Ecol. Prog. Ser. 57:207-217.

Gherna, R. and C. R. Woese. 1992. A partial phylogenetic analysis of the "Flavobacter-Bacteroides" phylum: basis for taxonomic restructuring. System. Appl. Microbiol. 15:513-521.

Giovannoni, S. J. 1991. The polymerase chain reaction. In Nucleic acid techniques in bacterial systematics (ed. E. Stackebrandt and M. Goodfellow), pp. 177-201. John Wiley & Sons, New York.

Giovannoni, S. J., T. B. Britschgi, C. L. Moyer and K. G. Field. 1990a. Genetic diversity in Sargasso Sea bacterioplankton. Nature. 345:60-63.

Giovannoni, S. J., E. F. DeLong, G. J. Olsen and N. R. Pace. 1988. Phylogenetic group-specific oligodeoxynucleotide probes for identification of single microbial cells. J. Bacteriol. 170:720-726.

Giovannoni, S. J., E. F. DeLong, T. M. Schmidt and N. R. Pace. 1990b. Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. Appl. Environ. Microbiol. 56:2572-2575.

Giovannoni, S. J., M. S. Rappé, D. Gordon, E. Urbach, M. Suzuki and K. G. Field. 1996a. Ribosomal RNA and the evolution of bacterial diversity. In Evolution of microbial life (ed. D. M. Roberts, P. Sharp, G. Alderson, and M. Collins), pp. 63-85. Cambridge University Press.

Giovannoni, S. J., M. S. Rappé, K. Vergin and N. Adair. 1996b. 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the Green Non-Sulfur bacteria. Proc. Natl. Acad. Sci. USA. 93:7979-7984.

Giovannoni, S. J., T.D. Mullins and K.G. Field. 1995. Microbial diversity in oceanic systems: rRNA approaches to the study of unculturable microbes. In Molecular ecology of aquatic microbes (ed. I. Joint), pp. 217-248. Springer-Verlag, New York.

Giovannoni, S. J., S. Turner, G. J. Olsen, S. Barns, D. J. Lane and N. R. Pace. 1988. Evolutionary relationships among cyanobacteria and green chloroplasts. J. Bacteriol. 170:3584-3592.

Glover, H. E., L. Campbell and B. B. Prézelin. 1986. Contribution of *Synechococcus* spp. to size-fractionated primary productivity in three water masses in the Northwest Atlantic Ocean. Mar. Biol. 91:193-203.

Good, I. J. 1953. The population frequencies of species and the estimation of the population parameters. Biometrika. 40:237-264.

Gordon, D. A. and S. J. Giovannoni. 1996. Detection of stratified microbial populations related to *Chlorobium* and *Fibrobacter* species in the Atlantic and Pacific oceans. Appl. Environ. Microbiol. 62:1171-1177.

Gray, J. P. and R. P. Herwig. 1996. Phylogenetic analysis of the bacterial communities in marine sediments. Appl. Environ. Microbiol. 62:4049-4059.

Gutell, R. R. 1993. Collection of small subunit (16S and 16S-like) ribosomal RNA structures. Nucleic Acids Res. 21:3051-3054.

Gutell, R. R. 1994. Collection of small subunit (16S and 16S-like) ribosomal RNA structures: 1994. Nucleic Acids Res. 22:3502-3507.

Hales, B. A., C. Edwards, D. A. Ritchie, G. Hall, R. W. Pickup and J. R. Saunders. 1996. Isolation and identification of methanogen-specific DNA

from blanket bog peat by PCR amplification and sequence analysis. Appl. Environ. Microbiol. 62:668-675.

Hara, Y. and M. Chihara. 1987. Morphology, ultrastructure and taxonomy of the Raphidophycean alga *Heterosigma akashiwo*. Bot. Mag. Tokyo. 100:151-163.

Hiraishi, A. 1994. Phylogenetic affiliations of *Rhodoferax fermentans* and related species of phototrophic bacteria as determined by automated 16S rDNA sequencing. Curr. Microbiol. 28:25-29.

Hiraishi, A., Y. Hoshino and T. Satoh. 1991. *Rhodoferax fermentans* gen. nov., sp. nov., a phototrophic purple nonsulfur bacterium previously referred to as the "*Rhodocyclus gelatinosus*-like" group. Arch. Microbiol. 155:330-336.

Hiraishi, A. and Y. Ueda. 1994. Intrageneric structure of the genus *Rhodobacter*: transfer of *Rhodobacter sulfidophilus* and related marine species to the genus *Rhodovolum* gen. nov. Int. J. Syst. Bacteriol. 44:15-23.

Hiraishi, A. and Y. Ueda. 1995. Isolation and characterization of *Rhodovulum strictum* sp. nov. and some other purple nonsulfur bacteria from colored blooms in tidal and seawater pools. Int. J. Syst. Bacteriol. 45:319-326.

Hobbie, J. E., R. J. Daley and S. Jasper. 1977. Use of nuclepore filters for counting bacteria by fluorescence microscopy. Appl. Environ. Microbiol. 33:1225-1228.

Hoepffner, N. and L. W. Haas. 1990. Electron microscopy of nanoplankton from the North Pacific central gyre. J. Phycol. 26:421-439.

Holmes, A. J., D. P. Kelly, S. C. Baker, A. S. Thompson, P. De Marco, E. M. Kenna and J. C. Murrell. 1997. *Methylosulfonomonas methylovora* gen. nov., sp. nov., and *Marinosulfonomonas methylotropha* gen. nov., sp. nov.: novel methylotrophs able to grow on methanesulfonic acid. Arch. Microbiol. 167:46-53.

Hood, R. R., S. Neuer and T. J. Cowles. 1992. Autotrophic production, biomass and species composition at two stations across an upwelling front. Mar. Ecol. Prog. Ser. 83:221-232.

Huyer, A. 1977. Seasonal variation in temperature, salinity, and density over the continental shelf off Oregon. Limnol. Oceanogr. 22:442-453.

Huyer, A. 1983. Coastal upwelling in the California Current system. Prog. Oceanogr. 12:259-284.

Irgens, R. L., I. Suzuki and J. T. Staley. 1989. Gas vacuolate bacteria obtained from marine waters of Antarctica. Curr. Microbiol. 18:261-265.

Jannasch, H. W. and G. E. Jones. 1959. Bacterial populations in seawater as determined by different methods of enumeration. Limnol. Oceanogr. 4:128-139.

Jenkins, O., D. Byrom and D. Jones. 1987. *Methylophilus*: a new genus of methanol-utilizing bacteria. Int. J. Syst. Bacteriol. 37:446-448.

Johnson, P. W. and J. M. Sieburth. 1982. In-situ morphology and occurrence of eucaryotic phototrophs of bacterial size in the picoplankton of estuarine and oceanic waters. J. Phycol. 18:318-327.

Joint, I. R. 1986. Physiological ecology of picoplankton in various oceanographic provinces. In Photosynthetic picoplankton (ed. T. Platt and W. K. W. Li), pp. 287-309. Can. Bull. Fish. Aquat. Sci., Ottawa.

Kemp, P. F. 1995. Can we estimate bacterial growth rates from ribosomal RNA content? In Molecular ecology of aquatic microbes (ed. I. Joint), pp. 279-302. Springer-Verlag, Berlin.

Kim, Y.-S., H. Oyaizu, S. Matsumoto, M. M. Watanabe and H. Nozaki. 1994. Chloroplast small-subunit ribosomal RNA gene sequence from *Chlamydomonas parkeae* (Chlorophyta): molecular phylogeny of a green alga with a peculiar pigment composition. Eur. J. Phycol. 29:213-217.

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

Kita-Tsukamoto, K., H. Oyaizu, K. Nanba and U. Simidu. 1993. Phylogenetic relationships of marine bacteria, mainly members of the family *Vibrionaceae*, determined on the basis of 16S rRNA sequences. Int. J. Syst. Bacteriol. 43:8-19.

Knauber, D. C., E. S. Berry and M. W. Fawley. 1996. Ribosomal RNA-based oligonucleotide probes to identify marine green ultraphytoplankton. J. Euk. Microbiol. 43:89-94.

Kogure, K., U. Simidu and N. Taga. 1979. A tentative direct microscopic method for counting living marine bacteria. Can. J. Microbiol. 25:415-420.

Kopczynski, E. D., M. M. Bateson and D. M. Ward. 1994. Recognition of chimeric small-subunit ribosomal DNAs composed of genes from uncultivated microorganisms. Appl. Environ. Microbiol. 60:746-748.

Lafay, B., R. Ruimy, C. Rausch de Traubenberg, V. Breittmayer, M. J. Gauthier and R. Christen. 1995. *Roseobacter algicola* sp. nov., a new marine bacterium isolated from the phycosphere of the toxin-producing dinoflagellate *Prorocentrum lima.* Int. J. Syst. Bacteriol. 45:290-296.

Lane, D. J. 1991. 16S/23S rRNA sequencing. In Nucleic acid techniques in bacterial systematics (ed. E. Stackebrandt and M. Goodfellow), pp. 115-175. John Wiley and Sons, New York.

Lange, M., L. Guillou, D. Vaulot, N. Simon, R. I. Amann, W. Ludwig and L. K. Medlin. 1996. Identification of the class Prymnesiophyceae and the genus *Phaeocystis* with ribosomal RNA-targeted nucleic acid probes detected by flow cytometry. J. Phycol. 32:858-868.

Ledyard, K. M., E. F. DeLong and J. W. H. Dacey. 1993. Characterization of a DMSP-degrading bacterial isolate from the Sargasso Sea. Arch. Microbiol. 160:312-318.

Lee, S. and J. A. Fuhrman. 1991. Spatial and temporal variation of natural bacterioplankton assemblages studied by total genomic DNA cross-hybridization. Limnol. Oceanogr. 36:1277-1287.

Liesack, W. and E. Stackebrandt. 1992. Occurence of novel groups of the domain *Bacteria* as revealed by analysis of genetic material isolated from an Australian terrestrial environment. J. Bacteriol. 174:5072-5078.

Lim, E. L., L. A. Amaral, D. A. Caron and E. F. DeLong. 1993. Application of rRNA-based probes for observing marine nanoplanktonic protists. Appl. Environ. Microbiol. 59:1647-1655.

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

Ludwig, W. and O. Strunk. 1997. Available on the WWW from http://www.mikro.biologie.tu-muenchen.de.

Maidak, B. L., N. Larsen, M. J. McCaughey, R. Overbeek, G. J. Olsen, K. Fogel, J. Blandy and C. R. Woese. 1994. The Ribosomal Database Project. Nucleic Acids Res. 22:3485-3487.

Markowicz, Y. and S. Loiseaux-de Goër. 1991. Plastid genomes of the Rhodophyta and Chromophyta constitute a distinct lineage which differs from that of the Chlorophyta and have a composite phylogenetic origin, perhaps like that of the Euglenophyta. Curr. Genet. 20:427-430.

Marshall, H. G. and J. A. Ranasinghe. 1989. Phytoplankton distribution along the eastern coast of the U.S.A.-VII. Mean cell concentrations and standing crop. Continental Shelf Res. 9:153-164.

Martinez, J., D. C. Smith, G. F. Steward and F. Azam. 1996. Variability in ectohydrolytic enzyme activites of pelagic marine bacteria and its significance for substrate processing in the sea. Aquat. Microb. Ecol. 10:223-230.

Miller, P. E. and C. A. Scholin. 1996. Identification of cultured *Pseudo-Nitzschia* (Bacillariophyceae) using species-specific LSU rRNA-targeted fluorescent probes. J. Phycol. 32:646-655.

Moore, E. R. B., M. Mau, A. Arnscheidt, E. Bottger, R. A. Hutson, M. D. Collins, Y. Van De Peer, R. De Wachter and K. N. Timmis. 1996. The determination and comparison of the 16S rRNA gene sequences of species of the genus *Pseudomonas* (sensu stricto) and estimation of the natural intrageneric relationships. Syst. Appl. Microbiol. 19:478-492.

Moran, N. A., M. A. Munson, P. Baumann and H. Ishikawa. 1993. A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. Proc. Royal Societ. London. B253:167-171.

Mullins, T. D., T. B. Britschgi, R. L. Krest and S. J. Giovannoni. 1995. Genetic comparisons reveal the same unknown bacterial lineages in Atlantic and Pacific bacterioplankton communities. Limnol. Oceanogr. 40:148-158.

Murphy, L. S. and E. M. Haugen. 1985. The distribution and abundance of phototrophic ultraplankton in the North Atlantic. Limnol. Oceanogr. 30:47-58.

Ohkuma, M. and T. Kudo. 1996. Phylogenetic diversity of the intestinal bacterial community in the termite *Reticulitermes speratus*. Appl. Environ. Microbiol. 62:461-468.

Olsen, G. J., D. L. Lane, S. J. Giovannoni and N. R. Pace. 1986. Microbial ecology and evolution: a ribosomal RNA approach. Ann. Rev. Microbiol. 40:337-365.

Olsen, G. J., H. Matsuda, R. Hagstrom and R. Overbeek. 1994. fastDNAml: a tool for construcion of phylogenetic trees of DNA sequences using maximum likelihood. Comput. Appl. Biosci. 10:41-48.

Paster, B. J., F. E. Dewhirst, S. M. Cooke, V. Fussing, L. K. Poulsen and J. Breznak. 1996. Phylogeny of not-yet cultivated spirochetes from termite guts. Appl. Environ. Microbiol. 62:347-352.

Paster, B. J., W. Ludwig, W. G. Weisburg, E. Stackebrandt, R. B. Hespell, C. M. Hahn, H. Reichenbach, K. O. Stetter and C. R. Woese. 1985. A phylogenetic grouping of the Bacteroides, Cytophagas, and certain Flavobacteria. Syst. Appl. Microbiol. 6:34-42.

Pinhassi, J., U. L. Zweifel and Å. Hagström. 1997. Unpublished sequences in Genbank.

Pomeroy, L. R. 1974. The ocean's food web, a changing paradigm. BioScience. 24(9):499-504 340.

Rappé, M. S., P. F. Kemp and S. J. Giovannoni. 1995. Chromophyte plastid 16S ribosomal RNA genes found in a clone library from Atlantic Ocean seawater. J. Phycol. 31:979-988.

Rappé, M. S., P. F. Kemp and S. J. Giovannoni. 1997. Phylogenetic diversity of marine coastal picoplankton 16S rRNA genes cloned from the continental shelf off Cape Hatteras, North Carolina. Limnol. Oceanogr. in press.

Rappé, M. S., M. T. Suzuki, K. Vergin and S. J. Giovannoni. in prep. Phylogenetic diversity of ultraplankton plastid SSU rRNA genes recovered in environmental nucleic acid samples from the Pacific and Atlantic coasts of the United States.

Rehnstam, A. S., S. Bäckman, D. C. Smith, F. Azam and Å. Hagström. 1993. Blooms of sequence-specific culturable bacteria in the sea. FEMS Microbiol. Ecol. 102:161-166.

Reysenbach, A. N., G. S. Wickham and N. R. Pace. 1994. Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus Spring, Yellowstone National Park. Appl. Envron. Microbiol. 60:2113-2119.

Rheims, H., C. Spröer, F. A. Rainey and E. Stackebrandt. 1996. Molecular biological evidence for the occurence of uncultured members of the actinomycete line of descent in different environments and geographical locations. Microbiol. 142:2863-2870.

Robinson-Cox, J. F., M. M. Bateson and D. M. Ward. 1995. Evaluation of nearest-neighbor methods for detection of chimeric small-subunit rRNA sequences. Appl. Environ. Microbiol. 61:1240-1245.

Rochelle, P. A., B. A. Cragg, J. C. Fry, R. J. Parkes and A. J. Weightman. 1994. Effect of sample handling on estimation of bacterial diversity in marine sediments by 16S rRNA gene sequence analysis. FEMS Microbiol. Ecol. 15:215-226.

Rowe, G., S. Smith, P. Falkowski, T. Whitledge, R. Theroux, W. Phoel and H. Ducklow. 1986. Do continental shelves export organic matter? Nature. 324:559-561.

Ryther, J. H. 1969. Photosynthesis and fish production in the sea. Science. 166:72-76.

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

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

Schlesner, H. and E. Stackebrandt. 1986. Assignment of the genera *Planctomyces* and *Pirella* to a new family *Planctomycetaceae* fam. nov. and description of the order *Planctomycetales* ord. nov. Syst. Appl. Microbiol. 8:174-176.

Schmidt, T. M., E. F. DeLong and N. R. Pace. 1991. Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. J. Bacteriol. 173:4371-4378.

Sherman, K., M. Grosslein, D. Mountain, D. Busch, J. O'Reilly and R. Theroux. 1988. The continental shelf ecosystem off the northeast coast of the United States. In Continental shelves (ed. H. Postma and J. J. Zijlstra), pp. 279-337. Elsevier.

Sherr, E. B. and B. F. Sherr. 1988. Role of microbes in pelagic food webs: A revised concept. Limnol. Oceanogr. 33:1225-1227.

Shiba, T. 1991. *Roseobacter litoralis* gen. nov., sp. nov., and *Roseobacter denitrificans* sp. nov., aerobic pink-pigmented bacteria which contain bacteriochlorophyll *a*. System. Appl. Microbiol. 14:140-145.

Simon, N., N. LeBot, D. Marie, F. Partensky and D. Vaulot. 1995. Fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes to identify small phytoplankton by flow cytometry. Appl. Environ. Microbiol. 61:2506-2513.

Small, L. F. and D. W. Menzies. 1981. Patterns of primary productivity and biomass in a coastal upwelling region. Deep Sea Res. 28:123-149.

Smith, D. C., G. F. Steward, R. A. Long and F. Azam. 1995. Bacterial mediation of carbon fluxes during a diatom bloom in a mesocosm. Deep Sea Res. II. 42:75-97.

Stackebrandt, E., W. Liesack and B. M. Göebel. 1993. Bacterial diversity in a soil sample from a subtropical Australian environment as determined by 16S rRNA analysis. FASEB J. 7:232-236.

Staley, J. T. and A. Konopka. 1985. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. Ann. Rev. Microbiol. 39:321-346.

Suzuki, M., M. S. Rappé and S. J. Giovannoni. in prep. Estimating coastal picoplankton community structure by LH-PCR.

Suzuki, M. T. and S. J. Giovannoni. 1996. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Appl. Environ. Microbiol. 62:625-630.

Suzuki, M. T., M. S. Rappé, Z. W. Haimberger, H. Winfield, N. Adair, J. Ströbel and S. J. Giovannoni. 1997. Bacterial diversity among SSU rRNA gene clones and cellular isolates from the same seawater sample. Appl. Environ. Microbiol. 63:983-989.

Tsuji, K., H. C. Tsien, R. S. Hanson, S. R. DePalma, R. Scholtz and S. LaRoche. 1990. 16S ribosomal RNA sequence analysis for determination of phylogenetic relationship among methylotrophs. J. Gen. Microbiol. 136:1-10.

Turner, S., T. Burger-Wiersma, S. J. Giovannoni, L. R. Mur and N. R. Pace. 1989. The relationship of a prochlorophyte, *Prochlorothrix hollandica*, to green chloroplasts. Nature. 337:380-382.

Urbach, E., D. L. Robertson and S. W. Chisholm. 1992. Multiple evolutionary origins of prochlorophytes within the cyanobacterial radiation. Nature. 355:267-270.

Verity, P. G., G.-A. Paffenhöfer, D. Wallace, E. Sherr and B. Sherr. 1996. Composition and biomass of plankton in spring on the Cape Hatteras shelf, with implications for carbon flux. Continental Shelf Res. 16:1087-1116.

Wakabayashi, H., M. Hikida and K. Masumura. 1986. *Flexibacter maritimus* sp. nov., a pathogen of marine fishes. Int. J. Syst. Bacteriol. 36:396-398.

Walsh, J. J. 1988. On the nature of continental shelves. Academic Press.

Walsh, J. J. 1991. Importance of continental margins in the marine biogeochemical cycling of carbon and nitrogen. Nature. 350:53-55.

Walsh, J. J., E. T. Premuzic, J. S. Gaffney, G. T. Rowe, G. Harbottle, R. W. Stoenner, W. L. Balsam, P. R. Betzer and S. A. Macko. 1985. Organic storage of CO2 on the continental shelf slope off the mid-Atlantic bight, the southeastern Bering Sea, and the Peru coast. Deep Sea Res. 32:853-883.

Ward, D. M., M. M. Bateson, R. Weller and A. L. Ruff-Roberts. 1992. Ribosomal RNA analysis of microorganisms as they occur in nature. Adv. Microb. Ecol. 12:219-286.

Ward, D. M., R. Weller and M. M. Bateson. 1990. 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. Nature. 345:63-65.

Ward, N., F. A. Rainey, E. Stackebrandt and H. Schlesner. 1995. Unraveling the extent of diversity within the order *Planctomycetales*. Appl. Environ. Microbiol. 61:2270-2275.

Willems, A., J. De Ley, M. Gillis and K. Kersters. 1991. *Comamonadaceae*, a new family encompassing the acidovorans rRNA complex, including *Variovorax paradoxus*, gen. nov., comb. nov., for *Alcaligenes paradoxus* (Davis 1969). Int. J. Syst. Bacteriol. 41:445-450.

Williamson, P. and P. M. Holligan. 1990. Ocean productivity and climate change. TREE. 5:299-303.

Woese, C. R. 1987. Bacterial evolution. Microbiol. Rev. 51:221-271.

Woese, C. R., O. Kandler and M. L. Wheelis. 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. USA. 87:4576-4579.

Woese, C. R., E. Stackebrandt, W. G. Weisburg, B. J. Paster, M. T. Madigan, V. J. Fowler, C. M. Hahn, P. Blanz, R. Gupta, K. H. Nealson and G. E. Fox. 1984. The phylogeny of purple bacteria: the alpha subdivision. Syst. Appl. Microbiol. 5:315-326.

Woese, C. R., W. G. Weisburg, C. M. Hahn, B. J. Paster, L. B. Zablen, B. J. Lewis, T. J. Macke, W. Ludwig and E. Stackebrandt. 1985. The phylogeny of purple bacteria: the gamma subdivision. Syst. Appl. Microbiol. 6:25-33.

Woese, C. R., W. G. Weisburg, B. J. Paster, C. M. Hahn, R. S. Tanner, N. R. Krieg, H.-P. Koops, H. Harms and E. Stackebrandt. 1984. The phylogeny of purple bacteria: the beta subdivision. Syst. Appl. Microbiol. 5:327-336.

Wright, T. D., K. Vergin, P. W. Boyd and S. J. Giovannoni. 1997. A novel _-subdivision proteobacterial lineage stratified in the lower ocean surface layer. Appl. Environ. Microbiol. 63:1441-1448.

Xia, Y., T. M. Embley and A. G. O'Donnell. 1994. Phylogenetic analysis of *Azospirillum* by direct sequencing of PCR amplified 16S rDNA. System. Appl. Microbiol. 17:197-201.

Yabuuchi, E., Y. Kosako, H. Oyaizu, I. Yano, H. Hotta, Y. Hashimoto, T. Ezaki and M. Arakawa. 1992. Proposal of *Burkholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov. Microbiol. Immunol. 36:1251-1275.

ZoBell, C. E. 1946. Marine microbiology, a monograph on hydrobacteriology. Chronica Botanica, Waltham, Mass.

Zweifel, U. L. and Å. Hagström. 1995. Total counts of marine bacteria include a large fraction of non-nucleoid-containing bacteria (ghosts). Appl. Environ. Microbiol. 61:2180-2185.

APPENDIX

NOVEL SSU RRNA GENE SEQUENCES DETERMINED IN THIS STUDY

1. Ocean Margins SSU rRNA gene clone library

OM1
Genbank accession #U70710
1519 nucleotides
Marine Gram-positive cluster; high G+C Gram-positive phylum

OM5
Genbank accession #U70715
1485 nucleotides
Prasinophyceae II lineage; prasinophyte algae plastids

OM10
Genbank accession #U70693
1527 nucleotides
SAR86 cluster; gamma Proteobacteria

OM13
Genbank accession #U32667
697 nucleotides
Bacillariophyte algae plastids

OM19
Genbank accession #U32668
699 nucleotides
Bacillariophyte algae plastids

OM20
Genbank accession #U32670
1474 nucleotides
Bacillariophyte algae plastids

OM21
Genbank accession #U32671
1472 nucleotides
Prymnesiophyte algae plastids

OM22
Genbank accession #U32669
700 nucleotides
Bacillariophyte algae plastids

OM23
Genbank accession #U70694
755 nucleotides
SAR86 cluster; gamma Proteobacteria

OM25
Genbank accession #U70678
1477 nucleotides
SAR116 cluster; alpha Proteobacteria

OM27
Genbank accession #U70713
1338 nucleotides
Unidentified Proteobacteria

OM38
Genbank accession #U70679
1479 nucleotides
SAR116 cluster; alpha Proteobacteria

OM39
Genbank accession #U70716
698 nucleotides
Prasinophyceae II lineage; prasinophyte algae plastids

OM42
Genbank accession #U70680
587 nucleotides
Roseobacter clade; alpha Proteobacteria

OM43
Genbank accession #U70704
1533 nucleotides
Type I methylotroph clade; beta Proteobacteria

OM55
Genbank accession #U70681
583 nucleotides
SAR116 cluster; alpha Proteobacteria

OM58
Genbank accession #U70705
357 nucleotides
Type I methylotroph clade; beta Proteobacteria

OM59
Genbank accession #U70695
644 nucleotides
Marinobacter sp.; gamma Proteobacteria

OM60
Genbank accession #U70696
1525 nucleotides
Gamma Proteobacteria

OM65
Genbank accession #U70682
587 nucleotides
Roseobacter clade; alpha Proteobacteria

OM75
Genbank accession #U70683
1491 nucleotides
Alpha Proteobacteria

OM81
Genbank accession #U70717
1464 nucleotides
Unidentified algal plastids

OM93
Genbank accession #U70697
900 nucleotides
Pseudomonas sp.; gamma Proteobacteria

OM110
Genbank accession #U70714
390 nucleotides
Probable chimera

OM111
Genbank accession #U70718
402 nucleotides
Prasinophyceae II lineage; prasinophyte algae plastids

OM125
Genbank accession #U70719
1085 nucleotides
Prymnesiophyte algae plastids

OM133
Genbank accession #U70698
420 nucleotides
Pseudomonas sp.; gamma Proteobacteria

OM134
Genbank accession #AF001661
750 nucleotides
Prasinophyceae II lineage; prasinophyte algae plastids

OM136
Genbank accession #U70684
587 nucleotides
SAR11 cluster; alpha Proteobacteria

OM143
Genbank accession #U70685
411 nucleotides
Roseobacter clade; alpha Proteobacteria

OM153
Genbank accession #U70720
1126 nucleotides
Prymnesiophyte algae plastids

OM155
Genbank accession #U70686
1304 nucleotides
SAR11 cluster; alpha Proteobacteria

OM156
Genbank accession #U70706
642 nucleotides
Beta Proteobacteria

OM164
Genbank accession #U70721
1478 nucleotides
Unidentified algal plastids

OM180
Genbank accession #U70707
689 nucleotides
Beta Proteobacteria

OM182
Genbank accession #U70699
644 nucleotides
Gamma Proteobacteria

OM185
Genbank accession #U70700
316 nucleotides
Gamma Proteobacteria

OM188
Genbank accession #U70687
588 nucleotides
SAR11 cluster; alpha Proteobacteria

OM190
Genbank accession #U70712
1470 nucleotides
Planctomycetes

OM231
Genbank accession #U70711
432 nucleotides
Marine Gram-positive cluster; high G+C Gram-positive phylum

OM233
Genbank accession #U70701
263 nucleotides
Pseudomonas sp.; gamma Proteobacteria

OM239
Genbank accession #U70688
277 nucleotides
SAR11 cluster; alpha Proteobacteria

OM241
Genbank accession #U70702
1491 nucleotides
Gamma Proteobacteria

OM242
Genbank accession #U70689
365 nucleotides
SAR11 cluster; alpha Proteobacteria

OM247
Genbank accession #U70690
359 nucleotides
Roseobacter clade; alpha Proteobacteria

OM252
Genbank accession #U70703
693 nucleotides
Gamma Proteobacteria

OM255
Genbank accession #U70722
412 nucleotides
Prasinophyceae II lineage; prasinophyte algae plastids

OM258
Genbank accession #U70691
439 nucleotides
SAR11 cluster; alpha Proteobacteria

OM270
Genbank accession #U70723
1475 nucleotides
Unidentified algal plastids

OM271
Genbank accession #U70708
570 nucleotides
Flexibacter-cytophaga-bacteroides phylum

OM273
Genbank accession #U70709
730 nucleotides
Flexibacter-cytophaga-bacteroides phylum

OM283
Genbank accession #U70724
1482 nucleotides
Cryptophyte algae plastids

OM299
Genbank accession #U70692
363 nucleotides
SAR116 cluster; alpha Proteobacteria

2. Oregon Coast Study SSU rRNA gene clone library

OCS5
Genbank accession #AF001651
1526 nucleotides
SAR86 cluster; gamma Proteobacteria

OCS7
Genbank accession #AF001645
1525 nucleotides
Zoogloea sp.; beta Proteobacteria

OCS12
Genbank accession #U75252
1474 nucleotides
Probable chimera

OCS14
Genbank accession #AF001643
658 nucleotides
Alpha Proteobacteria

OCS19
Genbank accession #U78942
377 nucleotides
Roseobacter clade; alpha Proteobacteria

OCS20
Genbank accession #AF001654
1479 nucleotides
Cryptophyte algae plastids

OCS24
Genbank accession #AF001637
1476 nucleotides
SAR116 cluster; alpha Proteobacteria

OCS28
Genbank accession #AF001636
1478 nucleotides
SAR116 cluster; alpha Proteobacteria

OCS31
Genbank accession #AF001655
1472 nucleotides
Prymnesiophyte algae plastids

OCS44
Genbank accession #AF001650
1527 nucleotides
SAR86 cluster; gamma Proteobacteria

OCS50
Genbank accession #AF001656
956 nucleotides
Prymnesiophyte algae plastids

OCS53
Genbank accession #AF001641
437 nucleotides
SAR11 cluster; alpha Proteobacteria

OCS54
Genbank accession #AF001657
1478 nucleotides
Bacillariophyte algae plastids

OCS56
Genbank accession #AF001658
554 nucleotides
Bacillariophyte algae plastids

OCS66
Genbank accession #AF001649
500 nucleotides
Beta Proteobacteria

OCS84
Genbank accession #U78943
420 nucleotides
Roseobacter clade; alpha Proteobacteria

OCS98
Genbank accession #AF001648
499 nucleotides
Variovorax sp.; beta Proteobacteria

OCS111
Genbank accession #AF001646
497 nucleotides
Zoogloea sp.; beta Proteobacteria

OCS116
Genbank accession #U78944
466 nucleotides
Alpha Proteobacteria

OCS122
Genbank accession #U78945
448 nucleotides
Roseobacter clade; alpha Proteobacteria

OCS124
Genbank accession #AF001644
474 nucleotides
Alpha Proteobacteria

OCS126
Genbank accession #AF001638
1478 nucleotides
SAR116 cluster; alpha Proteobacteria

OCS138
Genbank accession #AF001639
436 nucleotides
SAR11 cluster; alpha Proteobacteria

OCS143
Genbank accession #U75266
294 nucleotides
SAR11 cluster; alpha Proteobacteria

OCS146
Genbank accession #AF001653
1553 nucleotides
Marine group A/SAR406 cluster

OCS154
Genbank accession #AF001640
437 nucleotides
SAR11 cluster; alpha Proteobacteria

OCS155
Genbank accession #AF001652
513 nucleotides
Marine Gram-positive cluster; high G+C Gram-positive phylum

OCS162
Genbank accession #AF001659
1490 nucleotides
Prasinophyceae I lineage; prasinophyte algae plastids

OCS178
Genbank accession #AF001647
505 nucleotides
Type I methylotroph clade; beta Proteobacteria

OCS180
Genbank accession #AF001642
437 nucleotides
SAR11 cluster; alpha Proteobacteria

OCS181
Genbank accession #U78946
497 nucleotides
Pseudomonas sp.; gamma Proteobacteria

OCS182
Genbank accession #AF001660
1482 nucleotides
Prasinophyceae II lineage; prasinophyte algae plastids