CULTURE-INDEPENDENT CHARACTERIZATION OF DOC-TRANSFORMING BACTERIOPLANKTON IN COASTAL SEAWATER

by

XIAOZHEN MOU

(Under the Direction of Robert E. Hodson)

ABSTRACT

Bacterially-mediated transformation of aromatic monomers and organic osmolytes, two important components of the dissolved organic (DOC) pool in coastal seawater, is significant in the biogeochemical cycling of essential elements including carbon, sulfur and nitrogen. A study of bacterioplankton responding to the addition of 20 μM dimethylsulfoniopropionate (DMSP), a sulfur-containing organic osmolyte, suggested that a subset of the bacterial community could degrade DMSP. Cells developing high nucleic acid content in the presence of DMSP included members of α-Proteobacteria (mainly in the *Roseobacter* clade), β-Proteobacteria and γ-Proteobacteria, and a lower number of *Actinobacteria* and *Bacteroidetes*. The relative importance of DMSP-active taxa varied seasonally. Another study tested whether aromatic monomers and organic osmolytes are transformed by specialist bacterial taxa. The taxonomy of coastal bacterioplankton responding to the addition of 100 nM organic osmolyte [DMSP or glycine betaine (GlyB)] or aromatic monomer [para-hydroxybenzoic acid (pHBA) or vanillic acid (VanA)] was determined using incorporation of bromodeoxyuridine (BrdU) to track actively growing cells. 16S rDNA clones of active bacterioplankton indicated that both types of DOC were transformed by bacterial assemblages composed of the same major taxa, including

Actinobacteria, *Bacteroidetes*, α-Proteobacteria (mainly members of the *Roseobacter* clade), β-Proteobacteria, and γ-Proteobacteria (mainly members of *Altermonadaceae*, *Chromatiaceae*, *Oceanospirillaceae* and *Pseudomonadaceae*). The relative abundance of each taxon differed, however. Members of the OM60/241 and OM185, SAR11, SAR86 and SAR116 bacterioplankton groups in the active cell group indicated the ability to transport and metabolize these compounds by these ubiquitous but poorly described environmental clusters. In a final study, the phylogenetic diversity and functional capability of bacterioplankton stimulated by the addition of DMSP or VanA were explored by metagenomic analysis. Metagenomic sequences revealed a similar taxonomic structure for the bacterioplankton enriched with both DOC types, and were generally consistent with the PCR-based 16S rDNA analysis of the same template. Likewise, similar distribution patterns of functional gene categories were found for the two enriched bacterioplankton communities. These results indicate that DMSP and VanA were transformed by metabolic generalists capable of degrading both compounds. These studies provide insights into the response of coastal bacterial communities to changing environmental conditions in an ecological time frame.

INDEX WORDS: DOC degradation, aromatic monomers, organic osmolytes, DMSP, marine bacteria, flow cytometric cell sorting, BrdU incorporation, metagenomics, pyrosequencing

CULTURE-INDEPENDENT CHARACTERIZATION OF DOC-TRANSFORMING BACTERIOPLANKTON IN COASTAL SEAWATER

by

XIAOZHEN MOU

B.S. Qingdao University, China, 1997

M.S. The Institute of Oceanology, Chinese Academy of Sciences, China, 2000

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2006

CULTURE-INDEPENDENT CHARACTERIZATION OF DOC-TRANSFORMING BACTERIOPLANKTON IN COASTAL SEAWATER

by

XIAOZHEN MOU

Major Professor: Robert E. Hodson

Committee: Mary Ann Moran
Brian J. Binder
James T. Hollibaugh
Ming-Yi Sun
Chuanlun L. Zhang

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2006

DEDICATION

This dissertation is dedicated to my mother, A. Mou, who had given me her love and support during her short life. I love you mom!

ACKNOWLEDGEMENTS

I would like to extend my deepest gratitude to my major advisor Robert E. Hodson and direct advisor Mary Ann Moran. They both showed great kindness and patient and provided a tremendous amount of advice, support, guidance and encouragement throughout my dissertation studies. I also appreciate the assistance and support I received from my committee members, Brian Binder, James T. Hollibaugh, Ming-Yi Sun and Chuanlun Zhang. I sincerely thank Wendy Ye, Ed Sheppard, Wade Sheldon, Erin Biers, Shulei Sun, Jim Sullivan and all other members of the Hodson's and Moran's labs, for their friendship and various help during my graduate study.

Special gratitude goes to my husband and daughter, Weidong and Megan Zhao. They have always been my enduring source of support and encouragement throughout this long process.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Marine dissolved organic carbon (DOC) represents the largest active reservoir of the reduced carbon on Earth (6.85×10^{17}g C) and supports a considerable fraction of the oceanic food web (Nagata, 2000). The transformation of DOC is mainly mediated by bacteria. Through the microbial loop, bacteria either mineralize DOC into inorganic forms or transport DOC to higher trophic levels (i.e. bacterivores and larger organisms) (Pomeroy, 1974). Bacterially-mediated DOC transformations ultimately determine the form and amount of nutrients exported to the atmosphere and oceanic reservoirs, thus making them essential to the global cycling of nutrient elements including C, N, P and S (Moran and Hodson, 1994; Hullar et al., 1996; Moran et al., 1999; Raymond and Bauer, 2000). The focus of the research reported herein is the bacterioplankton assemblages responsible for transforming specific DOC compounds in coastal seawater, with particular emphasis on those processing aromatic compounds and organic osmolytes.

The ecological significance and origin of aromatic monomers and organic osmolytes in coastal seawater

Lignin is a complex aromatic polymer produced by all vascular plants as a structural material to strengthen cell walls. Lignin comprises about 25% of the land-based biomass and is the most abundant aromatic compound naturally produced on Earth (Peng et al., 1999). Therefore, its biodegradation is an important component of global carbon cycling. Besides its

ecological significance, the biodegradation of lignin-related aromatic compounds is of great economic interest in the paper industry, particularly in its capacity for ultimately replacing the usage of harmful chemicals and providing a more environmental-friendly method for removing lignin from paper pulp (Messner and Srebotnik, 1994). In salt marshes, especially those along the southeastern U.S. coast, lignin-related compounds primarily originate from *Spartina alterniflora*, a grass that is highly productive (1300 gC $m^{-2}y^{-1}$) (Landin, 1991) and provides more than 80% of the primary production in coastal marsh systems (Pomeroy and Wiegert, 1981). A substantial amount of terrestrially derived lignin is also added to coastal systems through riverine transport (Moran et al., 1991). Due to the stable nature of the aromatic rings, lignin-related aromatic compounds are refractory to most coastal marsh metazoans and animals and have to be converted to more labile forms by fungi and bacteria before being accessed by other organisms in coastal marsh food webs (Benner et al., 1984). In the present study, two compounds that are decay products of vascular plant lignin, vanillic acid (VanA) and *para*-hydroxybenzoic acid (pHBA), were chosen as model aromatic compounds (Harwood and Parales, 1996; Figure 1.1). Each compound contains only one aromatic ring and is less refractory than other polymerized lignin-related compounds.

Organic osmolytes are essential for marine organisms because they stabilize intracellular macromolecules under high osmolality conditions (e.g. salty waters) without interfering with cellular enzyme functions (Kiene et al., 1998). Most marine heterotrophic bacteria cannot synthesize organic osmolytes (Ventosa et al., 1998) but obtain them from the ambient seawater where these essential compounds are released from primary producers (Kiene et al., 2000). In coastal marshes, organic osmolytes are mainly produced by *Spartina*, phytoplankton and benthic algae (Pomeroy and Wiegert, 1981). Because organic osmolytes often contain essential elements

in addition to carbon, such as nitrogen, phosphor and sulfur, they are also important sources of these elements to marine heterotrophic bacteria (Kiene et al., 1999; Roberts, 2005). For example, dimethylsulfoniopropionate (DMSP, a sulfur-containing organic osmolyte) provides up to 15% of the carbon demand and almost all the sulfur demand of marine bacteria (Kiene et al., 2000). One of the degradation products of DMSP is dimethylsulfide (DMS), a sulfur-containing gas that is responsible for approximately 90% of natural atmospheric sulfur emission. Besides its importance in the global sulfur cycle, DMS impacts global climate because its oxidation products can act as cloud condensation nuclei (CCN), which reduce the incoming solar radiation to the surface of the earth by reflection. Glycine betaine (trimethylammonioacetate, GlyB) is a nitrogen-containing analog for DMSP and the most widely adopted organic osmolyte. Studies have suggested that GlyB can also be an important carbon and nitrogen source for marine heterotrophic bacteria (Kiene and Willams, 1998). In this study, DMSP and GlyB served as model marine osmolytes (Fig 1.1).

Bacterial transformation of aromatic DOC

Despite the structural differences among lignin monomers, most are degraded by bacteria through the β-ketoadipate pathway in terrestrial and marine environments, a ortho-cleavage pathway that mediates the aerobic degradation of a wide range of phenolic compounds (Harwood and Parales, 1996; Buchan et al., 2000). In this pathway, polycyclic and homocyclic aromatic compounds are sequentially transformed into catechol or protocatechuate, then β-ketoadipate, and finally intermetidates in the tricarboxylic acid cycle (Figure 1.2).

The β-ketoadipate pathway is genetically conserved and widely used by taxonomically diverse bacteria in both soil and marine environments, including members of *Rhodococcus*, *Streptomyces* (*Actinobacteria*); *Amycolatopsis* (*Actinomycetes*); *Bacillus* (*Firmicutes*); *Sagittula*,

Sphingomonas, *Sulfitobacter* (α-Proteobacteria); *Alcaligenes* (β-Proteobacteria); *Acinetobacter*, *Azotobacter*, *Pseudomonas* (γ-Proteobacteria) and other bacterial genera (Grund et al., 1990; Gauthier et al., 1992; Nishikawa et al., 1998; Hedlund et al., 1999; Buchan et al., 2000; Hedlund and Staley, 2001). In marine systems, the β-ketoadipate pathway has been identified in a number of bacterial taxa including *Firmicutes*, as well as α- and γ-Proteobacteria, but is especially common in the *Roseobacter* clade (Buchan et al., 2000), an ecologically important α-Proteobacteria lineage that comprises approximately 10-20% of bacterial cells in coastal and open ocean surface seawater (Giovannoni and Rappe, 2000).

Alternative pathways of aromatic compound degradation, such as *meta*- and *para*-(gentisate) cleavage, also exist but have received less attention (Harwood and Parales, 1996; Bugg and Winfield, 1998). However, the presence of funcitonal genes for aromatic compound degradation other than those involved in the β-ketoadipate pathway in the genome sequences of *Silicibacter pomeroyi*, *Silicibacter* sp. TM1040, *Pelagibacter ubique* (the model bacterium for SAR11, an α-Proteobacteria lineage that accounts for 25% of oceanic bacterioplankton), *Saccharophagus degradans* (γ-Proteobacteria) and *Magnetococcus* sp. MC-1 (α-Proteobacteria) suggests that these less-well-studied pathways along with other unidentified routes of aromatic compound degradation may be widespread in marine bacteria.

Bacterial transformation of organic osmolytes

Three hypothetical pathways of bacterially mediated DMSP degradation have been proposed (Yoch, 2002): cleavage, demethylation/demethiolation and double demethylation (Figure 1.3). The cleavage pathway is the only route known to produce DMS from DMSP. To date, proteins involved in these pathways are poorly understood and only one gene that converts DMSP demethylation to 3-methiolpropionate has been identified (Howard et al., submitted).

Marine bacteria in the *Roseobacter* clade have been linked to DMSP degradation by a number of culture-dependent and -independent studies under naturally occurring elevated DMSP concentrations (e.g. algal blooms) or following DMSP amendments (Gonzalez et al., 2000; Zubkov et al., 2002; Gonzalez et al., 2003).

The importance of marine bacteria in the *Roseobacter* clade for the DMSP degradation, DMS production and global sulfur cycling has been reviewed in detail elsewhere (Yoch, 2002; Moran et al., 2003). Other marine bacteria belonging to the SAR11 cluster of α-Proteobacteria, and β- and γ-Proteobacteria have also been identified as having the capability for DMSP degradation (Gonzalez et al., 2000; Ansede et al., 2001; Zubkov et al., 2002). However, most previous culture-independent studies have targeted DMSP-degrading bacteria at broad taxonomic levels (phylum or class), which overlooks the intra-group difference and is often too general to provide reliable guidelines for prediction of bacterial function from taxonomy.

The proposed pathway for bacterial degradation of GlyB involves sequential demethylation reactions that generate glycine and formaldehyde (or methylenetetrahydrofolate) as intermediates. Glycine then serves as a substrate for L-serine production (Smith et al., 1988; Figure 1. 4). Few studies have determined the taxonomy of marine bacteria that transform GlyB (Ansede et al., 2001). However, based on *in situ* rate measurements, it has been hypothesized that GlyB-degrading bacteria are composed of taxa similar to those that degrade DMSP (Kiene et al., 1998).

Analytical methods for microbial ecology

The availability of analytical techniques has been one of the limiting factors for improving knowledge of microbial ecology. Classical studies of microbial ecology either have assumed that cultured bacteria represent those found in nature or treated environmental bacteria

as a taxonomically and functionally homogenous group (a black box), thus ignoring the internal dynamics of bacterial community.

Modern phylogenetic studies in microbial ecology rely largely on molecular analysis of 16S rDNA. This gene is unique to prokaryotes (Bacteria and Archaea) and is relatively stable (i.e., rarely horizontally transferred between organisms). Moreover, the 16S rDNA contains regions of sequences that are taxonomically conservative, and therefore allows simple and unambiguous determination of the bacteria phylogenies without cultivation (Woese and Fox, 1977).

Through polymerase chain reaction (PCR) amplification, cloning and sequencing of 16S rDNA, bacterial communities in various marine ecosystems have been found to be far more diverse (1000-fold more) than previously expected from culture-based estimations (Giovannoni et al., 1990; Schmidt et al., 1991; Delong et al., 1993; Fuhrman and Davis, 1997). Cloning and sequencing generates detailed phylogenetic information to the species level; however, it is relatively time consuming and expensive. This is because a large number of clones need to be sequenced before the recovered community can be statistically representative of that found *in situ*. When high phylogenetic resolution is not required, techniques such as terminal restriction fragment length polymorphism (T-RFLP), denaturing gradient gel electrophoresis (DGGE) and fluorescent in situ hybridization (FISH) are often employed to obtain an overview of the bacterial community in a relatively quick and cheap manner.

T-RFLP is a community fingerprinting method that involves amplification of the 16S rDNA with a fluorescently labeled forward or reverse primer following by digestion of the amplicons with a restriction endonuclease. The digested products (i.e., terminal restriction fragments, T-RFs) are separated and quantified on an automated sequence analyzer (Liu et al.,

1997). The taxon richness (revealed by the number of T-RF) and relative abundance (revealed by the fluorescent intensity of each T-RF) are both determined. DGGE is another fingerprinting technique based on electrophoretic separation. When performing DGGE, short lengths of 16S rDNA are PCR-amplified and separated by sequence differences as determined by denaturing behavior on a polyacrylamide gel in a gradient of denaturants (such as urea). After fluorescent staining of DNA, taxon richness and relative abundance are measured by the number and fluorescence intensity of bands on the gel, respectively (Muyzer et al., 1993). In contrast to the two fingerprinting techniques, FISH directly targets the 16S rDNA *in situ* with fluorescently labeled probes, thus allowing visual examination of the bacterial taxon distribution at the individual cell level without the bias potentially introduced by PCR amplification (Amann et al., 1995).

These DNA-based approaches are highly efficient for bacterial phylogenetic analysis but generally lack the ability to make a direct linkage between the presence of bacterial taxa and their *in situ* activity. This is mainly because: 1) except for a few groups that exhibit high coherence between phylogeny and physiology, such as cyanobacteria, metabolic functions cannot be inferred from 16S rDNA; 2) only a limited number of bacterial functional genes are studied well enough to provide reliable DNA primers or probes.

Advanced molecular approaches that permit simultaneous and culture-independent determination of bacterial phylogeny and function have become available recently, including *in situ* PCR (Hodson et al., 1995), microautoradiography-FISH (Cottrell and Kirchman, 2000), stable isotope probing (Boschker et al., 1998; Radajewski et al., 2000), bromodeoxyuridine (BrdU) incorporation-FISH (Pernthaler et al., 2002; Hamasaki et al., 2004), mRNA/rRNA-FISH (Pernthaler and Amann, 2004), flow cytometry (FCM)-FISH (Zubkov et al., 2002), and

metagenomics (Beja et al., 2000; Venter et al., 2004; DeLong et al., 2006). The pros and cons associated with these techniques have been discussed elsewhere (Gray and Head, 2001). To date, these advanced approaches have been employed to demonstrate the taxonomy of bacteria responsible for various DOC-transforming processes in marine systems including acetate and butyrate degradation (Lee et al., 1999), BTEX-compounds degradation (Chen et al., 1999; Dustman, 2002), low and high molecular weight DOM degradation (Cottrell and Kirchman, 2000), methane oxidation (Boschker et al., 1998), nitrification and denitrification (Sullivan, 2002) and DMSP transformation (Zubkov et al., 2002).

In this study, three alternative approaches that enable the characterization of specific DOC-transforming assemblages of bacterioplankton were developed by coupling techniques including PCR, T-RFLP, cloning and sequencing, flow cytometry, BrdU-incorporation followed by immunodection (or immunocaputure), and metagnomics. Compared to the existing methods, our approaches were developed with the idea of providing the highest taxonomic specificity together with a strong direct linkage to ecological functions.

Flow cytometric analysis (FCM) analyzes liquid samples at a high speed (1000 cell s^{-1}) and provides insights into the single-cell performance in community-level processes by simultaneously measuring multiple parameters including cell numbers, forward scatters (related to the size or volume of cells), side scatters (related to the internal structure, surface texture, shap and size of cells) and fluorescence intensity (Shapiro, 2000; Vives-Rego et al., 2000; Brehm-Stecher and Johnson, 2004). In conjunction with a sorting unit, defined FCM populations can be physically separated and then subjected to further analyses. In our study, FCM sorting was used to distinguish and separate bacteria that were actively responding to DOC addition from the bulk community for subsequent phylogenetic analysis.

BrdU is a halogenated molecular analog for thymidine. It is incorporated into newly synthesized DNA in active cells and readily immunodetected or immunocaptured with high specificity and sensitivity (Borneman, 1999; Urbach et al., 1999). In our study, BrdU incorporation was used as a proxy for bacterial activity to help distinguish bacteria stimulated by DOM addition from those that were unresponsive.

Metagenomics, also called environmental genomics or community genomics, is the direct analysis of microbial genomic material in an mixed environmental sample, as opposed to a pure culture, and provides a comprehensive culture-independent view of the phylogenetic structure and metabolic capability of *in situ* microbial communities (Handelsman, 2004). The growing interest in metagenomics in microbial ecology is reflected in the increasing number of research papers (Tyson et al., 2004; Venter et al., 2004; Tringe et al., 2005; DeLong et al., 2006; Edwards et al., 2006) and literature reviews (Handelsman, 2004; Riesenfeld et al., 2004; Edwards and Rohwer, 2005; Steele and Streit, 2005; Ward, 2006; Wilmes and Bond, 2006). Even though metagenomics is a powerful tool, its wide application is hindered by the high cost and relatively long processing time associated with the sequencing step, which mostly employs the traditional Sanger approach. In our study, we explored metagenomic analyses using a high throughput pyrosequencing approach, which generates short sequences (100 bp average size) but is comparatively fast and affordable.

Sampling site description

The Sapelo Island Microbial Observatory (SIMO) site is within the boundaries of the Georgia Coastal Ecosystems Long Term Ecological Research (LTER) site 6 (Figure 1.5) on Sapelo Island, GA. The Sapelo Island ecosystem is a relatively pristine coastal marsh ecosystem that is mostly vegetated with *Spartina alterniflora*, a salt marsh cord grass that dominates the

southeastern U. S. coast. The marshes and tidal creeks on Sapelo Island receive no riverine input and are semidiurnally flooded by coastal waters. The primary autochthonous sources of DOC to the marsh tidal creeks are *Spartina* and phytoplankton as well as benthic microalgae. In our study, all water samples were taken from the surface of the Dean Creek (one of the major tidal creeks on Sapelo Island) during mid-tides in an area surrounded by tall-form *Spartina*.

Hypothesis

It has been revealed that coastal bacterial communities are highly dynamic in terms of both phylogenetic composition and functional capability. However, the distribution of specific metabolic activities among phylogenetic groups is not clear. First, we hypothesized that the functional assemblage that transforms specific DOC compounds is a subset of the total bacterial community whose composition varies seasonally. The DOC supply in coastal marshes includes a variety of organic compounds with different chemical structures that ultimately determine their bioavailability to the microbial community. Based on the hypothesis stated above bacterial assemblages responsible for transforming distinct DOC classes (e.g. aromatic compounds and aliphatic organic osmolytes) are expected to be different in composition. It has been widely accepted that organisms in nature are highly adapted and competitive (Cohan, 2001). Therefore, many marsh bacteria may have the capability (either intrinsic or gained by horizontal gene transfer) to transform commonly encountered DOC. Based on this consideration, we further hypothesized that functional assemblages specialized on individual DOC components may overlap, although their overall composition will be distinct from each other.

Chapter overview

To identify the functional members of aromatic monomer- and organic osmolyte-transforming assemblages in coastal marsh waters and to test the hypotheses stated above, three culture-independent studies were conducted and reported in Chapters 2, 3 and 4, and summarized in Chapter 5.

In Chapter 2, a microcosm study was undertaken to characterize the phylogenetic diversity of marsh bacterioplankton involved in DMSP transformation over a thirteen-month period. The cells that responded to DMSP addition based on increased cell size and nucleic acid content were assumed to be DMSP-transforming cells and identified and separated from the bulk community by FCM cell sorting. The phylogenetic structure of the DMSP-transforming assemblages was determined by T-RFLP, cloning and sequencing the 16S rDNA of sorted cells. Seasonal variation of the DMSP-transforming bacterial assemblage was also monitored to capture the temporal variation.

In Chapter 3, the study was expanded to model DOC compounds of both organic osmolytes and aromatic compounds using a slightly different approach. BrdU incorporation was used as a proxy of bacterial metabolic activity. Cells with stimulated metabolic activity in the presence of single added DOC compounds (DMSP and GlyB, organic osmolytes; pHBA and VanA, lignin-related aromatic monomers) were distinguished and separated from the bulk community by FCM sorting and subsequently characterized by 16S rDNA analyses including T-RFLP, cloning and sequencing. The extent of overlap between the functional assemblages was measured at various levels of phylogenetic resolution.

In Chapter 4, a more comprehensive view of the phylogenetic structure and metabolic capability of functional assemblages transforming aromatic compounds or organic osmolytes

was pursued with metagenomics coupled to BrdU-incorporation. BrdU-incorporation was used to label the cells that were responsive to DMSP or VanA addition. By immunochemical purification, the BrdU-labeled genomic DNA was separated from the bulk community genomic DNA and served as the substrate for metagenomic analysis via the pyrosequencing method. Bacterial phylogeny was also determined by PCR-based 16S rDNA analyses including T-RFLP and clone library sequencing. In addition, because the metagenomic sequences provided by pyrosequencing are short (averaging 100 bp in length), the reliability of the phylogenetic and functional prediction from short sequences was evaluated by *in silico* simulations of the accuracy of BLAST annotations for short sequence fragments generated from known 16S rDNA and functional genes.

REFERENCES

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

Ansede, J.H., Friedman, R., and Yoch, D.C. (2001) Phylogenetic analysis of culturable dimethyl sulfide-producing bacteria from a *Spartina*-dominated salt marsh and estuarine water. *Appl Environ Microbiol* **67**: 1210-1217.

Beja, O., Aravind, L., Koonin, E.V., Suzuki, M.T., Hadd, A., Nguyen, L.P. et al. (2000) Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea. *Science* **289**: 1902-1906.

Benner, R., Newell, S.Y., Maccubbin, A.E., and Hodson, R.E. (1984) Relative contributions of bacteria and fungi to rates of degradation of lignocellulosic detritus in salt-marsh sediments. *Appl Environ Microbiol* **48**: 36-40.

Borneman, J. (1999) Culture-independent identification of microorganisms that respond to specified stimuli. *Appl Environ Microbiol* **65**: 3398-3400.

Boschker, H.T.S., Nold, S.C., Wellsbury, P., Bos, D., de Graaf, W., Pel, R. et al. (1998) Direct linking of microbial populations to specific biogeochemical processes by C-13-labelling of biomarkers. *Nature* **392**: 801-805.

Brehm-Stecher, B.F., and Johnson, E.A. (2004) Single-cell microbiology: tools, technologies, and applications. *Microbiol Mol Biol Rev* **68**: 538-559.

Buchan, A., Collier, L.S., Neidle, E.L., and Moran, M.A. (2000) Key aromatic-ring-cleaving enzyme, protocatechuate 3,4- dioxygenase, in the ecologically important marine *Roseobacter* lineage. *Appl Environ Microbiol* **66**: 4662-4672.

Bugg, T.D.H., and Winfield, C.J. (1998) Enzymatic cleavage of aromatic rings: mechanistic aspects of the catechol dioxygenases and later enzymes of bacterial oxidative cleavage pathways. *Nat Prod Rep***15**: 513-530.

Chen, F., Dustman, W.A., and Hodson, R.E. (1999) Microscopic detection of the toluene dioxygenase gene and its expression inside bacterial cells in seawater using prokaryotic *in situ* PCR. *Hydrobiologia* **401**: 131-138.

Cohan, F.M. (2001) Bacterial species and speciation. *Syst Biol* **50**: 513-524.

Cottrell, M.T., and Kirchman, D.L. (2000) Natural assemblages of marine Proteobacteria and members of the *Cytophaga-Flavobacter* cluster consuming low- and high- molecular-weight dissolved organic matter. *Appl Environ Microbiol* **66**: 1692-1697.

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

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

Dustman, W.A. (2002) Development and application of *in situ* PCR techniques to investigate microbial community functional and phylogenetic diversity at the individual cell level. Ph. D Dissertation. University of Georgia, Athens GA.

Edwards, R.A., and Rohwer, F. (2005) Viral metagenomics. *Nature Rev Microbiol* **3**: 504-510.

Edwards, R.A., Rodriguez-Brito, B., Wegly, L., Haynes, M., Breitbart, M., Peterson, D.M. et al. (2006) Using pyrosequencing to shed light on deep mine microbial ecology under extreme hydrogenologic conditions. *BMC Genomics* **7**: 57.

Fuhrman, J.A., and Davis, A.A. (1997) Widespread Archaea and novel Bacteria from the deep sea as shown by 16S rRNA gene sequences. *Mar Ecol Prog Ser* **150**: 275-285.

Gauthier, M.J., Lafay, B., Christen, R., Fernandez, L., Acquaviva, M., Bonin, P., and Bertrand, J.C. (1992) *Marinobacter-Hydrocarbonoclasticus* gen. nov., sp. nov., a new, extremely halotolerant, hydrocarbon-degrading marine bacterium. *Int J Syst Bacteriol* **42**: 568-576.

Giovannoni, S.J., and Rappe, M. (2000) Evolution, diversity, and molecular ecology of marine prokaryotes. In *Microbial Ecology of the Oceans*. D. L. Kirchman (ed). Wiley-Liss, Inc., pp. 47-84.

Giovannoni, S.J., Britschgi, T.B., Moyer, C.L., and Field, K.G. (1990) Genetic diversity in sargasso sea bacterioplankton. *Nature* **345**: 60-63.

Gonzalez, J.M., Simo, R., Massana, R., Covert, J.S., Casamayor, E.O., Pedros-Alio, C., and Moran, M.A. (2000) Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. *Appl Environ Microbiol* **66**: 4237-4246.

Gonzalez, J.M., Covert, J.S., Whitman, W.B., Henriksen, J.R., Mayer, F., Scharf, B. et al. (2003) *Silicibacter pomeroyi* sp. nov. and *Roseovarius nubinhibens* sp. nov., dimethylsulfoniopropionate-demethylating bacteria from marine environments. *Int Syst Evol Micr* **53**: 1261-1269.

Gray, N.D., and Head, I.M. (2001) Linking genetic identity and function in communities of uncultured bacteria. *Environ Microbiol* **3**: 481-492.
Grund, E., Knorr, C., and Eichenlaub, R. (1990) Catabolism of benzoate and monohydroxylated benzoates by *Amycolatopsis* and *Streptomyces* spp. *Appl Environ Microbiol* **56**: 1459-1464.

Hamasaki, K., Long, R.A., and Azam, F. (2004) Individual cell growth rates of marine bacteria, measured by bromodeoxyuridine incorporation. *Aquat Microb Ecol* **35**: 217-227.

Handelsman, J. (2004) Metagenomics: Application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* **68**: 669-685.

Harwood, C.S., and Parales, R.E. (1996) The β-ketoadipate pathway and the biology of self-identity. *Annu Rev Microbiol* **50**: 553-590.

Hedlund, B.P., and Staley, J.T. (2001) *Vibrio cyclotrophicus* sp. nov., a polycyclic aromatic hydrocarbon (PAH)-degrading marine bacterium. *Int J Syst Evol Micr* **51**: 61-66.

Hedlund, B.P., Geiselbrecht, A.D., Bair, T.J., and Staley, J.T. (1999) Polycyclic aromatic hydrocarbon degradation by a new marine bacterium, *Neptunomonas naphthovorans* gen. nov., sp. nov. *Appl Environ Microbiol* **65**: 251-259.

Hodson, R.E., Dustman, W.A., Garg, R.P., and Moran, M.A. (1995) *In situ* PCR for visualization of microscale distribution of specific genes and gene products in prokaryotic communities. *Appl Environ Microbiol* **61**: 4074-4082.

Howard, E., Henriken, J. R., Buchan, A., Reisch, C. R., Bürgmann, H., Welsh, R. et al. Bacterial taxa limiting sulfur flux from the ocean. *Submitted*.

Hullar, M., Fry, B., Peterson, B., and Wright, R. (1996) Microbial utilization of estuarine dissolved organic carbon: a stable isotope tracer approach tested by mass balance. *Appl Environ Microbiol* **62**: 2489-2493.

Kiene, R.P., and Willams, L.P.H. (1998) Glycine betaine uptake, retention, and degradation by microorganisms in seawater. *Limnol Oceanogr* **43**: 1592-1603.

Kiene, R.P., Willams, L.P.H., and Walker, J.E. (1998) Seawater microorganisms have a high affinity glycine betaine uptake system which also recognizes dimethylsulfoniopropionate. *Aquat Microb Ecol* **15**: 39-51.

Kiene, R.P., Linn, L.J., and Bruton, J.A. (2000) New and important roles for DMSP in marine microbial communities. *J Sea Res* **43**: 209-224.

Kiene, R.P., Linn, L.J., Gonzalez, J., Moran, M.A., and Bruton, J.A. (1999) Dimethylsulfoniopropionate and methanethiol are important precursors of methionine and protein-sulfur in marine bacterioplankton. *Appl Environ Microbiol* **65**: 4549-4558.

Landin, M.C. (1991) Growth habits and other considerations of smooth cordgrass, *Spartina alterniflora* Loisel. In *Spartina Workshop Record*. Mumford, T.F., Jr., P. Peyton, J.R. Sayce, and S. Harbell (eds).Washington Sea Grant Program, University of Washington, Seattle, pp. 15-20.

Lee, N., Nielsen, P.H., Andreasen, K.H., Juretschko, S., Nielsen, J.L., Schleifer, K.H., and Wagner, M. (1999) Combination of fluorescent in situ hybridization and microautoradiography -

a new tool for structure-function analyses in microbial ecology. *Appl Environ Microbiol* **65**: 1289-1297.

Liu, W.T., Marsh, T.L., Cheng, H., and Forney, L.J. (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Appl Environ Microbiol* **63**: 4516-4522.

Messner, K., and Srebotnik, E. (1994) Biopulping: An overview of developments in an environmentally safe paper-making technology. *FEMS Microb Rev* **13**: 351-364.

Moran, M.A., and Hodson, R.E. (1994) Dissolved humic substances of vascular plant-origin in a coastal marine-environment. *Limnol Oceanogr* **39**: 762-771.

Moran, M.A., Sheldon, W.M., and Sheldon, J.E. (1999) Biodegradation of riverine dissolved organic carbon in five estuaries of the southeastern United States. *Estuaries* **22**: 55-64.

Moran, M.A., Gonzalez, J.M., and Kiene, R.P. (2003) Linking a bacterial taxon to sulfur cycling in the sea: studies of the marine *Roseobacter* group. *Geomicrobiol J* **20**: 375-388.

Moran, M.A., Pomeroy, L.R., Sheppard, E.S., Atkinson, L.P., and Hodson, R.E. (1991) Distribution of terrestrially derived dissolved organic-matter on the southeastern United States continental shelf. *Limnol Oceanogr* **36**: 1134-1149.

Muyzer, G., de Waal, E., and Uitterlinden, A. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59**: 695-700.

Nagata, T. (2000) Production mechanisms of dissolved organic matter. In *Microbial ecology of the oceans*. Kirchman, D. (ed): Wiley-Liss Inc., pp. 121-152.

Peng, X., Masai, E., Katayama, Y., and Fukuda, M. (1999) Characterization of the meta-cleavage compound hydrolase gene involved in degradation of the lignin-related biphenyl structure by *Sphingomonas paucimobilis* SYK-6. *Appl Environ Microbiol* **65**: 2789-2793.

Pernthaler, A., and Amann, R. (2004) Simultaneous fluorescence *in situ* hybridization of mRNA and rRNA in environmental bacteria. *Appl Environ Microbiol* **70**: 5426-5433.

Pernthaler, A., Pernthaler, J., Schattenhofer, M., and Amann, R. (2002) Identification of DNA-synthesizing bacterial cells in coastal North Sea plankton. *Appl Environ Microbiol* **68**: 5728-5736.

Pomeroy, L. (1974) The oceans food web, a changing paradigm. *BioScience* **24**: 499-504.

Pomeroy, L.R., and Wiegert, R.G. (1981) The ecology of a salt marsh. Springer-Verlag, New York, New York, USA. .

Radajewski, S., Ineson, P., Parekh, N.R., and Murrell, J.C. (2000) Stable-isotope probing as a tool in microbial ecology. *Nature* **403**: 646-649.

Raymond, P.A., and Bauer, J.E. (2000) Bacterial consumption of DOC during transport through a temperate estuary. *Aquat Microb Ecol* **22**: 1-12.

Riesenfeld, C.S., Schloss, P.D., and Handelsman, J. (2004) Metagenomics: Genomic analysis of microbial communities. *Annu Rev Genet* **38**: 525-552.

Roberts, M.F. (2005) Organic compatible solutes of halotolerant and halophilic microorganisms. *Saline Sys* **1:5**: doi: 10.1186/1746-1448-1181-1185.

Schmidt, T.M., Delong, E.F., and Pace, N.R. (1991) Analysis of a marine picoplankton community by 16s ribosomal-RNA gene cloning and wequencing. *J of Bacteriol* **173**: 4371-4378.

Shapiro, H.M. (2000) Microbial analysis at the single-cell level: tasks and techniques. *J Microbiol Meth* **42**: 3-16.

Smith, L., Pocard, J., Bernard, T., and Le Rudulier, D. (1988) Osmotic control of glycine betaine biosynthesis and degradation in *Rhizobium meliloti*. *J Bacteriol* **170**: 3142-3149.

Steele, H.L., and Streit, W.R. (2005) Metagenomics: Advances in ecology and biotechnology. *FEMS Microbiol Lett* **247**: 105-111.

Sullivan, J.B. (2002) Molecular characterization of aquatic nitrogen-cycling bacterial communities. Ph. D dissertation, University of Georgia.

Tringe, S.G., von Mering, C., Kobayashi, A., Salamov, A.A., Chen, K., Chang, H.W. et al. (2005) Comparative metagenomics of microbial communities. *Science* **308**: 554-557.

Tyson, G.W., Chapman, J., Hugenholtz, P., Allen, E.E., Ram, R.J., Richardson, P.M. et al. (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428**: 37-43.

Urbach, E., Vergin, K.L., and Giovannoni, S.J. (1999) Immunochemical detection and isolation of DNA from metabolically active bacteria. *Appl Environ Microbiol* **65**: 1207-1213.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A. et al. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66-74.

Ventosa, A., Nieto, J.J., and Oren, A. (1998) Biology of moderately halophilic aerobic bacteria. *Microb Mol Biol Rev Mineral Geochem* **62**: 504- 544.

Vives-Rego, J., Lebaron, P., and Nebe-von Caron, G. (2000) Current and future applications of flow cytometry in aquatic microbiology. *FEMS Microbiol Rev* **24**: 429-448.

Ward, N. (2006) New directions and interactions in metagenomics research. *FEMS Microbiol Ecol* **55**: 331-338.

Wilmes, P., and Bond, P.L. (2006) Metaproteomics: studying functional gene expression in microbial ecosystems. *Trends Microbiol* **14**: 92-97.

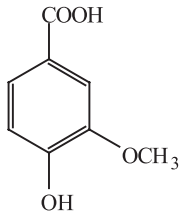
Woese, C.R., and Fox, G.E. (1977) The phylogenetic structure of the procaryotic domain: the primary kingdoms. *Proc Natl Acad Sci* **74**: 5088-5090.

Yoch, D.C. (2002) Dimethylsulfoniopropionate: Its sources, role in the marine food web, and biological degradation to dimethylsulfide. *Appl Environ Microbiol* **68**: 5804-5815.

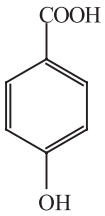
Zubkov, M.V., Fuchs, B.M., Archer, S.D., Kiene, R.P., Amann, R., and Burkill, P.H. (2002) Rapid turnover of dissolved DMS and DMSP by defined bacterioplankton communities in the stratified euphotic zone of the North Sea. *Deep-Sea Res PT II* **49**: 3017-3038.

Figure 1.1. Chemical structures of the lignin-related aromatic monomers (vanillic acid and para-hyproxybenzoic acid) and organic osmolytes (dimethylsulfoniopropionate and glycine betaine) generally found in the DOC pool of marsh-dominated coastlines.

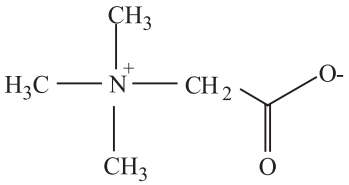
Figure 1.1



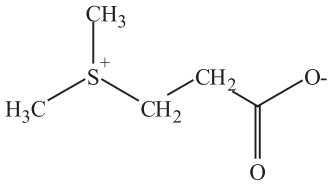
Vanillic Acid
(VanA)



para-Hydroxybenzoic Acid
(pHBA)



Glycine Betaine
(GlyB)



Dimethylsulfoniopropionate
(DMSP)

Figure 1.2. The protocatechuate branch of the β-ketoadipate pathway for the aerobic microbial degradation of lignin-related aromatic monomers. The fraction illustrating the pathway of VanA degradation to aliphatic compounds is highlighted with grey. Gene designations are shown in italics. Adapted from Buchan et al., 2000.

Fig. 1.2

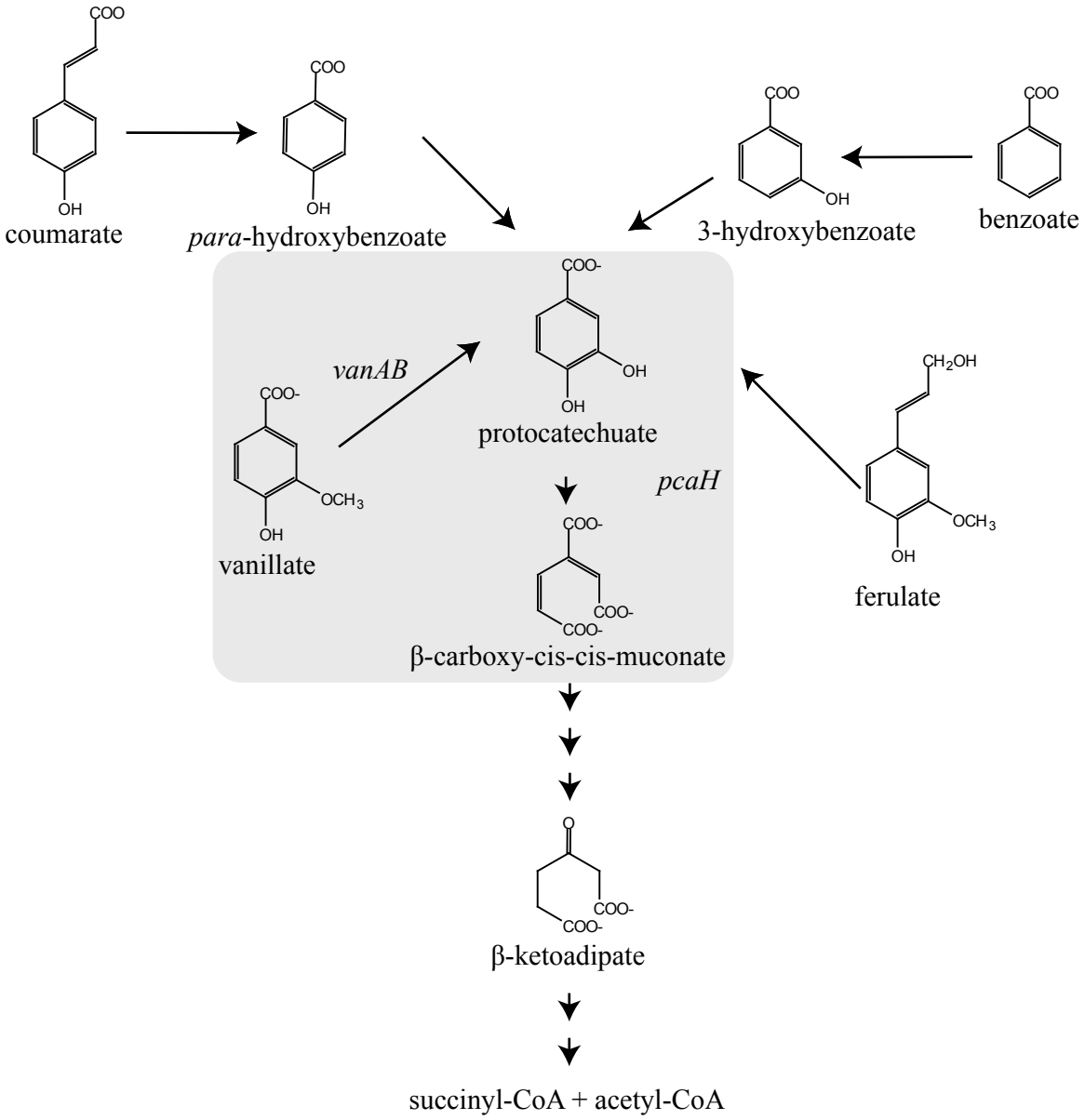

coumarate
para-hydroxybenzoate
3-hydroxybenzoate
benzoate
vanAB
protocatechuate
pcaH
vanillate
ferulate
β-carboxy-cis-cis-muconate
β-ketoadipate
succinyl-CoA + acetyl-CoA

Figure 1.3. The three hypothesized biochemical pathways for the catabolism of DMSP. DMSP can be degraded by the cleavage pathway, which hydrolyzes the C-3 carbon of DMSP and produces DMS and acrylate. Alternatively, DMSP is demethylated at the DMS moiety and produces 3-methiopropionate (MMPA). MMPA may be further demethylated to 3-mercaptoprpionate (MPA) via the double demethylation pathway or demethiolated to acryate and methanethiol (MeSH) via the demethylation/demethiolation pathway. The gene for the first step demethylation is shown in italics. MPA may be further degraded to acrylate by eliminating hydrogen sulfide (H_2S). Adapted from Yoch, 2002.

Fig. 1.3

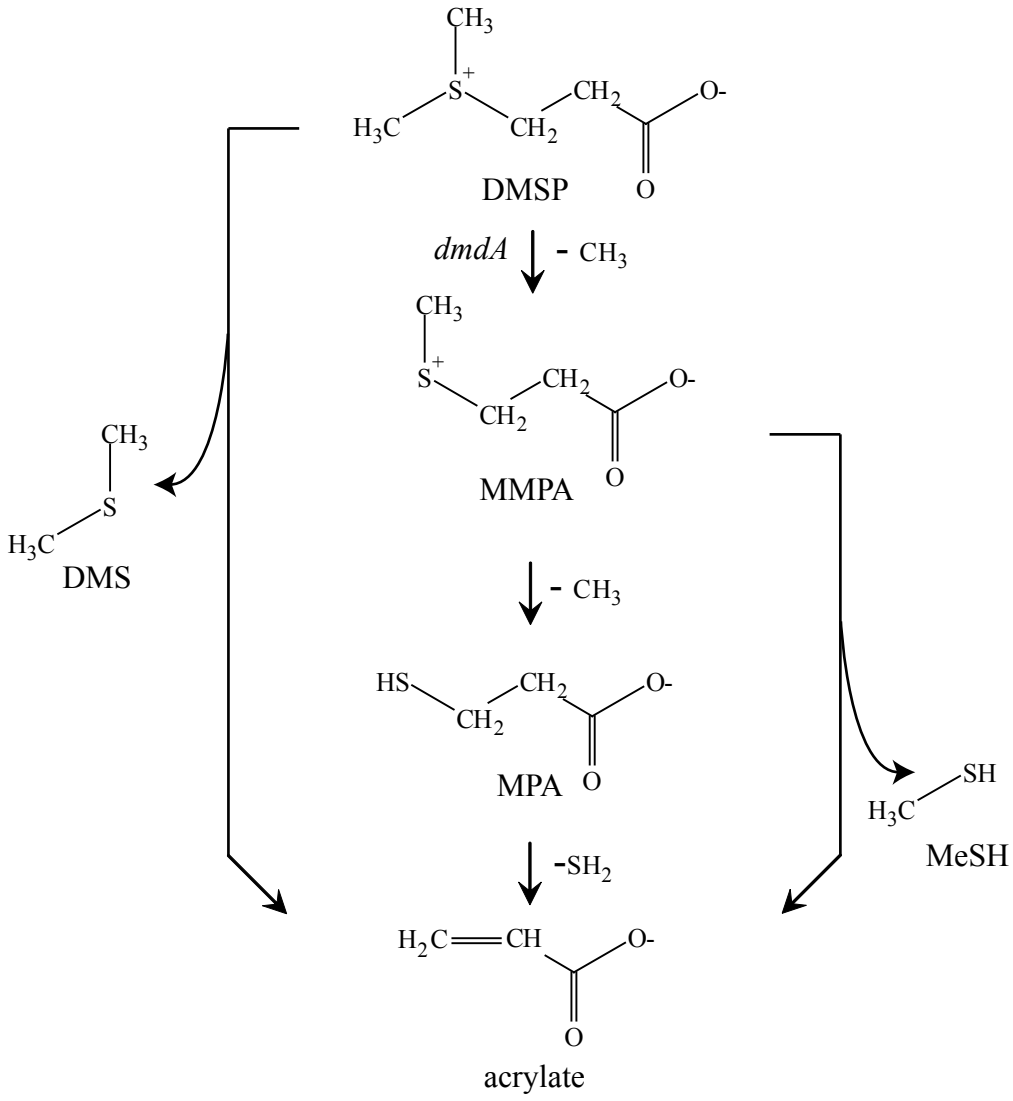
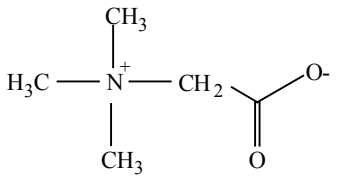
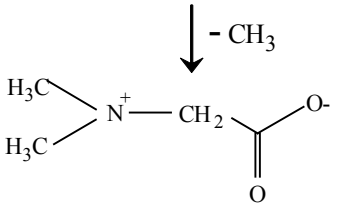

DMSP
dmdA
- CH3
MMPA
- CH3
MPA
-SH2
acrylate
DMS
MeSH
Cleavage
Double Demethylation
Demethylation/Demethiolation

Figure 1.4. The proposed pathway for bacterial degradation of GlyB, which involves a series of demethylations. Adapted from Smith et al., 1988.

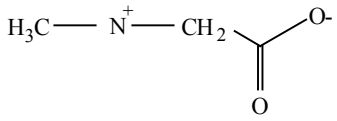
Fig. 1.4



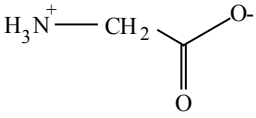
glycine betaine



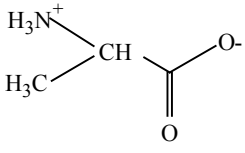
dimethylglycine



monomethylglycine



glycine

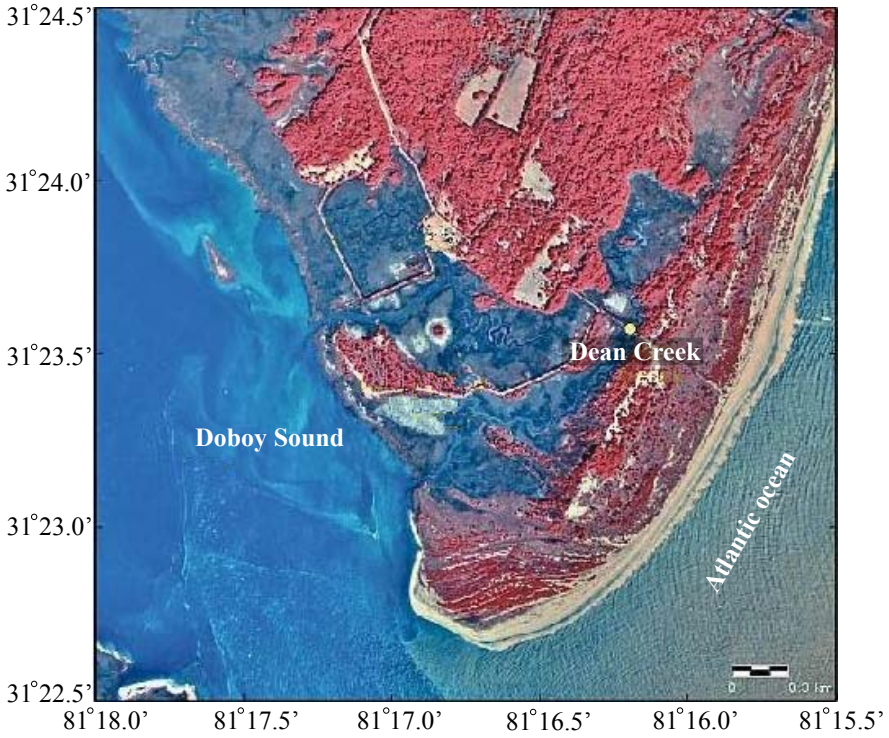


serine

Figure 1.5. The location of the primary sampling site for these studies on Sapelo Island, GA. Obtained from the SIMO website, http://www.simo.marsci.uga.edu/public_db/studysites.htm.

Fig. 1.5


Georgia
Sapelo Island


31°24.5'
31°24.0'
31°23.5'
31°23.0'
31°22.5'
81°18.0'
81°17.5'
81°17.0'
81°16.5'
81°16.0'
81°15.5'
Dean Creek
Doboy Sound
Atlantic ocean

CHAPTER 2

FLOW-CYTOMETRIC CELL SORTING AND SUBSEQUENT MOLECULAR ANALYSES FOR CULTURE-INDEPENDENT IDENTIFICATION OF BACTERIOPLANKTON INVOLVED IN DIMETHYLSULFONIOPROPIONATE TRANSFORMATIONS[2]

[1]Mou, X., M. A. Moran, R. Stepanauskas, J. M. Gonzalez, and R. E. Hodson. 2005. *Applied and Environmental Microbiology*. 71: 1405-1416.

Reprinted here with permission of publisher.

ABSTRACT

Marine bacterioplankton transform dimethylsulfoniopropionate (DMSP) into the biogeochemically important and climatically active gas dimethylsulfide. In order to identify specific bacterial taxa mediating DMSP processing in a natural marine ecosystem, we amended water samples from a southeastern U.S. salt marsh with 20 μM DMSP and tracked community shifts with fluorescence-activated cell sorting (FACS) coupled to 16S rDNA analyses. In two out of four seasons studied, DMSP amendments induced the formation of distinct bacterioplankton populations with elevated nucleic acid (NA) content within 24 h, indicative of cells actively utilizing DMSP. The 16S rDNAs of the cells with and without elevated NA content were analyzed following cell sorting and PCR amplification with sequencing and terminal restriction fragment length polymorphism approaches. Compared to cells in the control FCM populations, bacteria with elevated NA content in the presence of DMSP were relatively enriched in taxa related to *Loktanella*, *Oceanicola*, and *Sulfitobacter* (*Roseobacter* lineage, α-Proteobacteria); *Caulobacter* (α-Proteobacteria); and *Brachymonas* and *Xenophilus* (β-Proteobacteria) in the May-02 sample and to *Ketogulonicigenium* (*Roseobacter* lineage, α-Proteobacteria) and novel γ-Proteobacteria in the Sept-02 sample. Our study suggests that diverse bacterioplankton participate in the metabolism of DMSP in coastal marine systems and that their relative importance varies temporally.

INTRODUCTION

Dimethylsulfoniopropionate [DMSP; $(CH_3)_2$-S-CH_2-CH_2COOH] is an osmoprotectant synthesized by marine algae and vascular plants (25, 26, 33). As a consequence of senescence, grazing, and viral activity, intracellular DMSP can be released from algal and plant cells into the water column, where it undergoes microbial transformations. Bacterially mediated DMSP transformations can generate over 90% of oceanic dimethylsulfide (DMS, H_3C-S-CH_3), and DMS accounts for at least half of the global biogenic sulfur flux to the atmosphere (30, 31). Along with its role as a precursor for DMS, DMSP is a ubiquitous source of organic matter in marine surface waters and can provide up to 15% of the carbon requirement of heterotrophic bacteria and fulfill most of the bacterial cellular sulfur demand (14, 29, 39).

Previous studies have made significant progress in identifying which bacterial taxa mediate DMSP degradation in natural marine communities. Culturing approaches have identified DMSP-utilizing bacteria in the *Roseobacter* lineage (α-Proteobacteria) and in the genera *Alcaligenes* (β-Proteobacteria) and *Pseudomonas* (γ-Proteobacteria) (1, 5, 6, 12, 21, 22, 38). Culture-independent studies of surface waters from estuaries and open ocean sites with high DMSP turnover rates have revealed DMSP-utilizing bacteria in the *Bacteroidetes*, *Roseobacter*, β-Proteobacteria, and γ-Proteobacteria lineages (1, 13, 16, 24, 35, 39, 40). Most of these field studies, however, have targeted a broad taxonomic grouping (at the phylum or class level), and few have resolution at the level of genus or below.

Flow cytometry (FCM) discriminates between populations within complex bacterial communities based on fluorescence and size differences among the cells. In conjunction with a sorting unit (fluorescence-activated cell sorting, FACS), defined FCM populations can be physically separated and then subjected to further taxonomic analysis (2, 3, 23, 28, 32, 36).

However, the combined methods are challenging to use, primarily because of the low bacterioplankton densities typical of seawater and the large number of sorted cells needed for subsequent DNA-based analysis.

Coastal salt marshes have one of the highest natural DMS emission rates and represent an active DMSP-degrading ecosystem (4). In this study, a culture-independent approach was employed to identify bacteria involved in DMSP degradation at the species level. We amended water from a salt marsh tidal creek with DMSP and identified cells responding to the addition by FACS analysis of cells with elevated nucleic acid (NA) content. Methods were optimized to obtain DNA template from as few as 25,000 FCM sorted preserved bacterial cells. Terminal restriction fragment length polymorphism (T-RFLP) analysis and sequencing of 16S rDNA in the sorted populations were used to identify bacterial taxa involved in DMSP degradation.

MATERIALS AND METHODS

Sample collection and processing. Surface water samples were collected from Dean Creek (a salt marsh tidal creek) on Sapelo Island, Georgia, in acid-washed Nalgene carboys at 4-month intervals from May 2002 to May 2003 (samples May-02, Sept-02, Jan-03, and May-03). Once collected, water samples were stored in the dark at 4°C and processed within 3 days. To isolate a bacterioplankton size fraction and exclude bacteriovores, water samples were sequentially filtered through 47-mm-diameter, 3.0- and 1.0-μm-pore-size polycarbonate filters (Poretics Products, Livermore, Calif.) under pressures lower than 250 mm Hg. The filtrate was then amended with a mixture of inorganic nitrogen and phosphorus (5 μM NH_4Cl, 5 μM $NaNO_3$, and 1 μM NaH_2PO_4, final concentrations) and incubated in the dark at 25°C with 100-rpm shaking for 48 h to establish carbon-limited conditions. Microcosms were established by adding 150 ml

of the preincubated water to each of six 250-ml Erlenmeyer flasks. DMSP (20 μM, final concentration) was added to three of the microcosms, and the remaining three microcosms were used as no-addition controls. Flasks were incubated in the dark at 25°C with 100-rpm shaking for 48 h. Five-milliliter subsamples from each microcosm were collected at 0, 1, 3, 6, 12, 24, and 48 h after DMSP amendments; preserved with 1% paraformaldehyde (final concentration) at 4°C for 1 h; and then stored in the dark at 20°C until flow-cytometric analysis. All glassware was ashed at 550°C prior to use.

FCM. Flow-cytometric analysis was performed with a MoFlo flow cytometer (DakoCytomation, Glostrup, Denmark). Prior to running on the instrument, preserved bacterioplankton were stained with Sybr Green II (1:5,000 dilution of the commercial stock; Molecular Probes Inc.) in the dark at room temperature for 20 min (18) and then mixed with an internal bead standard (1-μm-diameter yellow-green fluorescent beads; Fluoresbrite YG Microspheres; Polysciences, Warrington, Pa.). A sterilized phosphate-buffered saline (PBS) solution (Puraflo; Dakocytomation, Fort Collins, Colo.) served as the sheath fluid. Data acquisition was triggered by green fluorescence (FL1). All signals were collected with logarithmic amplification.

All FCM populations were gated initially on the FCM cytogram of one DMSP replicate after 24 h of incubation, and the gating frames were kept fixed through all the measurements within a single sample date. Gate notation was based on the value of FL1, which is related to the cell NA content. The FCM populations that exhibited high FL1 values were designated the "high-NA-content groups" (HNAs), and the others were designated the "low-NA-content groups" (LNAs). The FCM populations observed for each sample are shown in Figure 2.1. Because the FCM population gates chosen were specific to each date, the gate notation refers to different ranges of fluorescence and side scatter for each sample date. Flow sorting of selected

FCM populations was performed on a MoFlo sorter in the "purify 1-2 drop" mode. Two FCM populations were sorted at one time. The sorted cells were collected in sterile 1.5-ml polypropylene test tubes containing 1 ml of PBS solution. Sorting was terminated when the sorted cells were over 500,000 counts for each FCM population or at 30 min.

DNA extraction and PCR amplifications. Sorted cells from each FCM population were filtered onto a 25-mm-diameter, 0.2-μm-pore-size polycarbonate filter (Poretics Products) and washed three times with 2 ml of PBS followed by 2 ml of deionized water. The filter was air dried, and its edge was trimmed off. One quarter of the trimmed filter was stored at 20°C, while the other three quarters were cut into pieces and transferred into a 0.5-ml tube. Ten microliters of Lyse-N-Go PCR reagent (Pierce Inc., Rockford, Ill.) was added to the tube to lyse the bacterial cells on the filter according to the manufacturer's protocol. A second lysis step was conducted by adding 82 μl of autoclaved water to the same tube and subjecting the tube to three freeze-thaw cycles. The resulting lysate served as the DNA template for PCR amplification.

16S rDNA amplification was carried out with Ready-To-Go PCR beads (Amersham Pharmacia, Piscataway, N.J.) with 0.4 μM concentrations of both forward (8F, 5'-AGAGTTTGATCCTGGCTCAG-3') and reverse (1492R, 5'-T ACGGYTACCTTGTTACGACTT-3') primers. A touchdown PCR program was performed with annealing temperature sequentially decreasing from 62 to 52°C by 1°C per cycle, followed by 15 cycles at 52°C. In each cycle, denaturing (at 95°C), annealing (at 62 to 52°C), and extension (at 72°C) steps were of 1-min duration. An initial 3-min denaturation and final 10-min extension step were also included in the PCR program. PCR amplification was confirmed by electrophoresis on ethidium bromide-stained 1% agarose gels. The amplicons were excised from

the gel and cleaned with the QIAquick gel extraction kit (Qiagen, Valencia, Calif.) before subsequent T-RFLP and sequencing analysis.

T-RFLP analysis. Direct PCR amplification of 16S rDNA from sorted cells with a 6-carboxyfluorescein (FAM)-labeled forward primer was inconsistent, probably due to the difficulty in binding the FAM-labeled primer to the low-copy-number templates. Instead, 16S rDNAs of sorted cells were amplified first with nonlabeled primers 8F and 1492R as described above and then reamplified with 0.4 μM FAM-labeled 8F and nonlabeled 785R (5'-CTACCA GGGTATCTAATCC-3') primers with two Ready-To-Go PCR beads in 50-μl volume reaction mixtures with a single annealing temperature of 59°C. PCR amplification was confirmed by electrophoresis on ethidium bromide-stained 1% agarose gels. The reamplified amplicons were excised from the gel, cleaned with the QIAquick gel extraction kit, and digested with CfoI restriction enzyme (Roche, Indianapolis, Ind.) at 37°C for 3 h, after which an ethanol precipitation was performed. The restricted amplicons were resuspended in 12 μl of deionized formamide plus 0.7 μl of DNA fragment length standard (GeneScan 2500 TAMRA; Applied Biosystems, Warrington, United Kingdom). The terminal restriction fragment (T-RF) lengths were determined on an ABI Prism 310 genetic analyzer (Applied Biosystems, Foster City, Calif.) in GeneScan mode. T-RFLP data were obtained successfully for most sorted bacterial FCM populations. The exceptions were as follows: (i) one set of the May-02 triplicates was used up in the process of optimizing the conditions for DNA extraction and PCR amplification, (ii) some FCM populations were too small to provide enough template for PCR amplification of 16S rDNA (Figure 2.1), and (iii) one FCM population (Sept-02, LNA1) failed to amplify even though cell numbers were relatively high.

Identification of bacteria associated with responsive T-RFs. 16S rDNAs directly amplified from sorted cells were used to construct clone libraries of 16S rDNA for selected samples (May-02 and Sept-02). Libraries were constructed using a TA cloning kit (Invitrogen Corporation, Carlsbad, Calif.) with the pCR 2.1 vector according to the manufacturer's instructions.

Sequencing was carried out on an ABI PRISM 310 genetic analyzer in sequencing mode, with the BigDye terminator cycle-sequencing kit (PE Biosystems, Foster City, Calif.). 8F served as the sequencing primer. The average length of usable sequence was 400 bp. Bacterial identities were determined using the GenBank BLASTN search (http://www.ncbi.nlm.nih.gov/BLAST/) and the ribosomal database project (RDP)-II sequence match program (http://rdp.cme.msu.edu/html/analyses.html).

The bacteria associated with T-RFs showing significant differences between control and DMSP-amended microcosms (responsive T-RFs) were putatively assigned using the following process. Initially, 20 clones were randomly selected from each library for sequencing, and the length of the T-RF for each clone was calculated based on sequence data with TRFLPtools, a Visual Basic program for Microsoft Excel (13, 34). If any responsive T-RFs remained unidentified, a clone pool approach was used to rapidly screen the 16S rDNA library. Briefly, five random clones were pooled prior to plasmid extraction with the QIAprep Spin Miniprep kit (Qiagen Inc.). T-RFLP analysis was carried out by the method described previously with 500 ng of extracted plasmid DNAs as templates for the FAM-labeled PCR amplification. If any matching T-RFs were found, each of the five clones was analyzed individually to identify the specific clone(s) of interest in the pooled sample. This process continued until the identities of all T-RFs were assigned or more than 200 clones were screened. Then, clones putatively responsible

for responsive T-RFs were coinjected with community amplicons in a single T-RFLP analysis. Finally, the inserts of the clones of interest were sequenced to determine taxonomic identity.

T-RFLP analysis of small populations. We determined whether apparent bacterial community structures were significantly affected by the number of cells used in the T-RFLP analysis, since sorted FCM populations differed in cell number by as much as 25-fold (Table 2.1). The total number of cells in the May-02 water sample was determined by epifluorescence microscopy counting after 4', 6'-diamidino-2-phenylindole (DAPI) staining. Aliquots containing various numbers of cells (2×10^9, 2×10^7, 2×10^6, 2×10^5, and 2×10^3) were filtered onto 25-mm-diameter, 0.2-μm-pore-size polycarbonate filters. PCR and T-RFLP analysis of the 16S rDNA were then performed according to the procedure described above.

Statistical analysis. A t-test for two samples of unequal variance was performed to compare total bacterial abundance as well as abundance in each of the FCM populations between control and DMSP microcosms. The t-test was also performed to compare the relative area of T-RFs between the sorted FCM populations in control and DMSP-amended microcosms to determine responsive T-RFs. The confidence interval was set at 95%, and significant differences were reported when $p < 0.05$.

A hierarchical cluster analysis was performed using Primer v5 (Primer-E Ltd, Plymouth, United Kingdom) to make quantitative comparisons of T-RFLP profiles among bacterial FCM populations. Before analysis, T-RFLP output data were standardized as described previously (34), except that relative peak area instead of peak height was used as a proxy for the relative abundance of bacterial taxa associated with each T-RF. T-RFs with lengths greater than 600 bp or relative peak areas less than 2% of total area were excluded from the analysis.

Nucleotide sequence accession numbers. The GenBank accession numbers for 16S rDNA sequences determined in this study are AF547399, AF547405 to AF547409, AF547413 to AF547415, AF547417, AF547418, AF547421, AF547422, AF547424, AF547425, AF547427, AY476738 to AY476741, AY476749 to AY476751, AY476763, AY476764, AY476767, AY476769 to AY476771, AY476774 to AY476776, AY476778 to AY476781, AY476784, AY476785, AY476794, AY476795, AY476797, AY476799 to AY476802, AY476804, and AY476805.

RESULTS

FCM analysis of sorted populations. Significant growth of bacterioplankton was observed over the 48-h incubation in microcosms amended with 20 μM DMSP but not in the control microcosms for all four experiments (May-02, Sept-02, Jan-03, and May-03), indicating stimulated bacterial activity following the DMSP amendments. At the 24-h time point, DMSP amended microcosms had cell numbers 2.9-, 32-, 1.4-, and 2.1-fold higher than did control microcosms in the May-02, Sept-02, Jan-03, and May-03 experiments, respectively (Figure 2.2). DMSP amendments resulted in the development of a cluster of cells with high NA content within 24 h in the May-02 and Sept-02 microcosms but not in the Jan-03 or May-03 microcosms (Figure 2.1; Table 2.1). In the May-02 microcosms, HNA and LNA3 groups accounted for 28.2% and 34.6%, respectively, of the total assemblage following DMSP addition; in the control microcosms these populations contained only 2.1% and 16.8% of the total cells, respectively. Likewise for Sept-02, the HNA group accounted for 9.7% of total cells in DMSP-amended microcosms but 2.1% in the controls (Table 1; there was no LNA3 gate defined in the Sept-02 sample). The greatest difference between control and DMSP-amended microcosms in cell NA

content (FL1) and to some extent cell size (side scatter) was found after 24 h of incubation for these two experiments. At other time points (0, 3, 6, 12, and 48 h), scatter plots of DMSP-amended microcosms generally resembled those of control microcosms, even though total cell number was highest at 48 h in some cases (Figure 2.2). No NA content or size changes were observed in Jan-03 and May-03 experiments over the course of the incubation, despite the increase in cell numbers. Therefore, flow sorting and subsequent molecular analysis were performed only on subsamples taken 24 h after DMSP amendments.

T-RFLP and sequence analysis of sorted populations. T-RFLP analyses of the replicate flasks for a given sample date were highly similar as revealed by cluster analysis (Figure 2.3). Thus, there was low variability among the replicate incubations within microcosms and consistency in results of molecular analysis. The cluster analysis of T-RFLP data also grouped all samples based on collection date, regardless of the sorting gates used. For the May-02 and Sept-02 samples (those in which a bacterial cell NA-size response to DMSP amendments was evident), FCM populations of control and DMSP amended samples clustered separately with high dissimilarity values. However, for Jan-03 and May-03 samples (those in which little or no bacterial cell NA-size response to DMSP additions was evident), populations from control and DMSP amended microcosms were intermingled.

The same major T-RFs were typically found in all FCM populations within a sample, but significant differences in relative abundance in response to DMSP amendments were evident for some T-RFs (identified as responsive T-RFs) (Figure 2.4). In the May-02 samples (Figure 2.4a), there were seven responsive T-RFs, six of which were putatively identified from the 16S rDNA clone library. They fell into five major phylogenetic groups that are commonly found in coastal marine environments: α-Proteobacteria (mainly the *Roseobacter* lineage), β-Proteobacteria, γ-

Proteobacteria, *Bacteroidetes*, and *Actinobacteria* (Table 2.2). The seventh T-RF (length, 115 bp; T-115) was unidentified. Bacteria associated with T-115 (unidentified), T-202 (β-Proteobacteria), and T-350-360 (α-Proteobacteria) were stimulated by DMSP amendments as evidenced by greater relative abundance in the elevated-NA–larger-size populations. Bacteria associated with T-32 (*Actinobacteria*), T-57 (the *Roseobacter* lineage, α-Proteobacteria), T-98 (*Bacteroidetes*), and T-565 (γ-Proteobacteria) decreased in relative abundance in the higher bacterial cell NA-size class after DMSP amendments. In the Sept-02 experiment (Figure 2.4B), there were six responsive T-RFs. Bacteria represented by T-231 (γ-Proteobacteria) and T-518 (*Roseobacter* lineage, α-Proteobacteria) were relatively enriched in FCM populations with elevated NA content after DMSP amendments, while T-57 (*Roseobacter* lineage, α-Proteobacteria), T-98 (*Bacteroidetes*), T-202 (β-Proteobacteria), and T-370 (γ-Proteobacteria) reacted in the opposite manner.

Based on observed increases in cell counts (Figure 2.2), bacteria stimulated by DMSP amendments were likely present in the Jan-03 or May-03 samples. However, no significant change in either cell NA content or size was observed between FCM populations in control and DMSP-amended microcosms. Thus, DMSP-utilizing bacteria could not be sorted from the bulk population for these two samples.

T-RFLP analysis of small populations. Following optimization of the DNA extraction method, PCR amplification and T-RFLP analysis were carried out on bacterial population sizes as low as 2,000 cells from unpreserved samples and 20,000 cells from preserved samples. Bacterial populations containing 2×10^9, 2×10^6, and 2×10^3 unpreserved cells from the May-02 water sample were analyzed by T-RFLP. Regardless of the initial cell number, T-RFLP chromatograms typically contained the same major peaks in approximately the same relative abundances (Figure

2.5). Based on this finding, FCM populations differing in size by as much as 500,000 cells were compared in the T-RFLP analyses.

DISCUSSION

Determining the identity of bacterial taxa carrying out specific ecological activities, such as turnover of dissolved organic carbon in marine systems, is a central task of microbial ecology. In this study, we successfully employed a culture-independent method involving FCM and cell sorting to identify bacteria stimulated by DMSP amendments. Bacterial cell NA content is widely used as a proxy for bacterial cell activity. With few disagreements (19, 41), cells with higher NA content are considered to be more active than cells with lower NA content and responsible for a significant fraction of bulk activities (8, 9, 19, 20, 27). Based on this consensus, we assume that bacteria increasing their cell NA content above that in no-addition controls were actively responding to DMSP additions.

The taxa identified as responsive to DMSP amendments probably include those assimilating DMSP directly as well as those utilizing products of DMSP degradation. Taxonomic groups assimilating DMSP or its degradation products but maintaining the same NA content and size would be overlooked with our approach (as was the case for cells in the Jan-03 and May-03 samples). The high concentrations of DMSP in amended microcosms (20 μM) relative to natural concentrations (20 to 100 nM) (17) and the preincubation used to establish carbon-limited conditions might have introduced bias. Nonetheless, this approach allowed an initial identification of bacteria that respond to DMSP and that may be important under natural conditions.

T-RFLP analysis was used in this study because it provides an efficient way to view the community composition of many sorted FCM populations. However, like fluorescent in situ hybridization and denaturing gradient gel electrophoresis, this quick examination of community-level 16S rDNA heterogeneity includes a tradeoff with taxonomic resolution (7). To partially overcome this limitation, we also constructed 16S rDNA clone libraries for selected samples to access the within-T-RF variability (Table 2.3). The sequence similarity of clones within dominant T-RFs was highly variable and ranged from as low as 74.1% (for T-98, representing the *Bacteroidetes* group in May-02) to as high as 99.3% (for T-57, representing the *Roseobacter* lineage in Sept-02). High variability within *Bacteroidetes* T-RFs generated with commonly used restriction enzymes has been found previously (7).

The *Roseobacter* lineage (α-Proteobacteria) is ubiquitous in marine environments and is a major taxonomic group in many marine bacterioplankton communities (10). Recently, evidence from both lab and field studies have suggested that members of the *Roseobacter* lineage carry out DMSP transformations in marine systems (12, 22, 37). *Roseobacter* group isolates can degrade DMSP by two distinct pathways (cleavage and demethylation-demethiolation [1, 12, 15]) and may be important in regulating DMS formation in the surface ocean (31). During blooms of DMSP-producing algae in the open ocean, members of the *Roseobacter* lineage have been found in high abundance (13, 39). More recently, they have been directly linked to DMSP utilization by microautoradiography-fluorescent in situ hybridization methodology (24, 35). Several members of the *Roseobacter* lineages identified in the present study as responding positively to DMSP additions had high 16S rDNA sequence similarity to previously identified DMSP-utilizing bacteria, including *Ruegeria atlantica* and *Sulfitobacter pontiacus* (1, 11–13) (Table 2.1). However, the response of other members of the *Roseobacter* lineage indicated that they either

were uninvolved with DMSP metabolism or were outcompeted by other taxa (Table 2.2; Figure 2.6). This intragroup difference demonstrates the challenge of assigning biogeochemical functions to large and diverse marine bacterioplankton groups and clearly would be missed by the use of analysis methods with lower taxonomic resolution.

DMSP is not utilized exclusively by bacteria of the *Roseobacter* lineage (24, 25). In previous studies, DMSP-degrading β- and γ-Proteobacteria were readily isolated from estuarine water and surface sediment in a coastal marsh (1). Also, γ-Proteobacteria have been found to be abundant in the bacterioplankton community associated with algal blooms in the ocean (13, 40) and can assimilate DMSP in natural bacterial assemblages (35). In this study, bacteria related to *Brachymonas* and *Xenophilus* (β-Proteobacteria) in the May-02 sample and a novel γ-Proteobacterium in the Sept-02 sample were enriched in the FCM populations with elevated NA content after DMSP addition, indicating active linkage to DMSP utilization. In contrast, bacteria related to *Actinobacteria*, *Flexibacter* and *Flavobacterium* (*Bacteroidetes*), and *Neptunomonas* (γ-Proteobacteria) in the May-02 sample and *Cytophaga* (*Bacteroidetes*), *Brachymonas*, *Rhodoferax* (β-Proteobacteria), and *Pseudoalteromonas* and *Colwellia* (γ-Proteobacteria) in the Sept-02 sample were concentrated in the FCM populations with unchanged or low NA content (LNA1 and LNA2) (Table 2.2; Figure 2.2B). However, because not all DMSP-utilizing bacteria responded with an increase in bacterial cell NA content or size (Figure 2.2), we cannot assume that these taxa play no role in DMSP degradation.

Cluster analysis grouped bacterial FCM populations into four clusters coincident with sampling dates, indicating that temporal factors had a major influence on bacterial community composition. The bacterial assemblages in the control microcosms for the two samples in which an NA-size response to DMSP additions was evident (May-02 and Sept-02) were similar, with

relatively low contributions from β-Proteobacteria (Table 2.3). The bacterial assemblages in the control microcosms of the other two samples (with no NA-size response to DMSP; Jan-03 and May-03) were likewise similar to one another but distinct from the May-02 and Sept-02 samples, with a higher abundance of (β-Proteobacteria, and few γ-Proteobacteria or *Bacteroidetes* (Table 2.3). These differences in bacterial community composition may underlie the disparity in cell parameter response to DMSP amendments observed among samples, and they underscore the influence of temporal variability in the composition and physiology of the bacterioplankton that transform DMSP in coastal environments.

ACKNOWLEDGEMENTS

We thank J. Nelson for assistance with FCM and cell sorting; B. Binder for advice on data analysis; W. Ye, Edward Sheppard, and E. Biers for assistance with sampling at Sapelo Island; and N. Bano for advice on phylogenetic tree construction.

This study was supported by NSF grant 0084164 to the Sapelo Island Microbial Observatory (SIMO).

REFERENCES

1. **Ansede, J. H., R. Friedman, and D. C. Yoch.** 2001. Phylogenetic analysis of culturable dimethyl sulfide-producing bacteria from a *Spartina*-dominated salt marsh and estuarine water. *Appl. Environ. Microbiol.* **67:**1210-1217.

2. **Bihari, N., R. Batel, and R. K. Zahn.** 1999. Flow cytometry in marine environmental research. *Period. Biol.* 101:151-155.

3. **Collier, J. L., and L. Campbell.** 1999. Flow cytometry in molecular aquatic ecology. *Hydrobiologia* **401:**33-53.

4. **Dacey, J. W. H., G. M. King, and S. G. Wakeham.** 1987. Factors controlling emission of dimethylsulfide from salt marshes. *Nature* **330:**643-645.

5. **De Souza, M. P., and D. C. Yoch.** 1995. Comparative physiology of dimethyl sulfide production by dimethylsulfoniopropionate lyase in *Pseudomonas doudoroffii* and *Alcaligenes* sp. strain M3a. *Appl. Environ. Microbiol.* **61:**3986-3991.

6. **De Souza, M. P., and D. C. Yoch.** 1995. Purification and characterization of dimethylsulfoniopropionate lyase from an *Alcaligenes*-like dimethyl sulfide-producing marine isolate. *Appl. Environ. Microbiol.* **61:**21-26.

7. **Eiler, A., and S. Bertilsson.** 2004. Composition of freshwater bacterial communities associated with cyanobacterial blooms in four Swedish lakes. *Environ. Microbiol.* **6:**1228-1243.

8. **Gasol, J. M., and P. A. Del Giorgio.** 2000. Using flow cytometry for counting natural planktonic bacteria and understanding the structure of planktonic bacterial communities. *Sci. Mar.* **64:**197-224.

9. **Gasol, J. M., U. L. Zweifel, F. Peters, J. A. Fuhrman, and Å. Hagström.** 1999. Significance of size and nucleic acid content heterogeneity as measured by flow cytometry in natural planktonic bacteria. *Appl. Environ. Microbiol.* **65:**4475-4483.

10. **Giovannoni, S. J., and M. Rappé.** 2000. Evolution, diversity, and molecular ecology of marine prokaryotes, p. 47-84 *In* D. L. Kirchman (ed.), Microbial ecology of the oceans. John Wiley & Sons, New York, N.Y.

11. **González, J. M., J. S. Covert, W. B. Whitman, J. R. Henriksen, F. Mayer, B. Scharf, R. Schmitt, A. Buchan, J. A. Fuhrman, R. P. Kiene, and M. A. Moran.** 2003. *Silicibacter pomeroyi* sp. nov. and *Roseovarius nubinhibens* sp. nov., dimethylsulfoniopropionate-demethylating bacteria from marine environments. *Int. J. Syst. Evol. Microbiol.* **53:**1261-1269.

12. **González, J. M., R. P. Kiene, and M. A. Moran.** 1999. Transformation of sulfur compounds by an abundant lineage of marine bacteria in the alpha-subclass of the class *Proteobacteria*. *Appl. Environ. Microbiol.* **65:**3810-3819.

13. **González, J. M., R. Simó, R. Massana, J. S. Covert, E. O. Casamayor, C. Pedrós-Alió, and M. A. Moran.** 2000. Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. *Appl. Environ. Microbiol.* **66:**4237-4246.

14. **Kiene, R. P., and L. J. Linn.** 2000. Distribution and turnover of dissolved DMSP and its relationship with bacterial production and dimethylsulfide in the Gulf of Mexico. *Limnol. Oceanogr.* **45:**849-861.

15. **Kiene, R. P., L. J. Linn, and J. A. Bruton.** 2000. New and important roles for DMSP in marine microbial communities. *J. Sea Res.* **43:**209-224.

16. **Kiene, R. P., L. J. Linn, J. González, M. A. Moran, and J. A. Bruton.** 1999. Dimethylsulfoniopropionate and methanethiol are important precursors of methionine and protein-sulfur in marine bacterioplankton. *Appl. Environ. Microbiol.* **65:**4549-4558.

17. **Kiene, R. P., and S. K. Service.** 1991. Decomposition of dissolved DMSP and DMS in estuarine waters-dependence on temperature and substrate concentration. *Mar. Ecol. Prog. Ser.* **76:**1-11.

18. **Lebaron, P., N. Parthuisot, and P. Catala.** 1998. Comparison of blue nucleic acid dyes for flow cytometric enumeration of bacteria in aquatic systems. *Appl. Environ. Microbiol.* **64:**1725-1730.

19. **Lebaron, P., P. Servais, H. Agogué, C. Courties, and F. Joux.** 2001. Does the high nucleic acid content of individual bacterial cells allow us to discriminate between active cells and inactive cells in aquatic systems? *Appl. Environ. Microbiol.* **67:**1775-1782.

20. **Lebaron, P., P. Servais, A. C. Baudoux, M. Bourrain, C. Courties, and N. Parthuisot.** 2002. Variations of bacterial-specific activity with cell size and nucleic acid content assessed by flow cytometry. *Aquat. Microb. Ecol.* **28:**131-140.

21. **Ledyard, K. M., and J. W. H. Dacey.** 1994. Dimethylsulfide production from dimethylsulfoniopropionate by a marine bacterium. *Mar. Ecol. Prog. Ser.* **110:**95-103.

22. **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.

23. **Legendre, L., C. Courties, and M. Troussellier.** 2001. Flow cytometry in oceanography 1989-1999: environmental challenges and research trends. *Cytometry* **44:**164-172.

24. **Malmstrom, R. R., R. P. Kiene, and D. L. Kirchman.** 2004. Identification and enumeration of bacteria assimilating dimethylsulfoniopropionate (DMSP) in the North Atlantic and Gulf of Mexico. *Limnol. Oceanogr.* **49:**597-606.

25. **Mulholland, M. M., and M. L. Otte.** 2001. The effects of nitrogen supply and salinity on DMSP, glycine betaine and proline concentrations in leaves of *Spartina anglica*. *Aquat. Bot.* **71:**63-70.

26. **Pakulski, J. D., and R. P. Kiene.** 1992. Foliar release of dimethylsulfoniopropionate from *Spartina alterniflora*. *Mar. Ecol. Prog. Ser.* **81:**277-287.

27. **Servais, P., E. O. Casamayor, C. Courties, P. Catala, N. Parthuisot, and P. Lebaron.** 2003. Activity and diversity of bacterial cells with high and low nucleic acid content. *Aquat. Microb. Ecol.* **33:**41-51.

28. **Shapiro, H. M.** 2000. Microbial analysis at the single-cell level: tasks and techniques. *J. Microbiol. Methods* **42:**3-16.

29. **Simó, R., S. D. Archer, C. Pedrós-Alió, L. Gilpin, and C. E. Stelfox-Widdicombe.** 2002. Coupled dynamics of dimethylsulfoniopropionate and dimethylsulfide cycling and the microbial food web in surface waters of the North Atlantic. *Limnol. Oceanogr.* **47:**53-61.

30. **Simó, R., C. Pedrós-Alió, G. Malin, and J. O. Grimalt.** 2000. Biological turnover of DMS, DMSP and DMSO in contrasting open-sea waters. *Mar. Ecol. Prog. Ser.* **203:**1-11.

31. **Simó, R.** 2001. Production of atmospheric sulfur by oceanic plankton: biogeochemical, ecological and evolutionary links. *Trends Ecol. Evol.* **16:**287-294.

32. **Steen, H. B.** 2000. Flow cytometry of bacteria: glimpses from the past with a view to the future. J. Microbiol. Methods **42:**65-74.

33. **Stefels, J.** 2000. Physiological aspects of the production and conversion of DMSP in marine algae and higher plants. *J. Sea Res.* **43:**183-197.

34. **Stepanauskas, R., M. A. Moran, B. A. Bergamaschi, and J. T. Hollibaugh.** 2003. Covariance of bacterioplankton composition and environmental variables in a temperate delta system. *Aquat. Microb. Ecol.* **31:**85-98.

35. **Vila, M., R. Simó, R. P. Kiene, J. Pinhassi, J. M. González, M. A. Moran, and C. Pedrós-Alió.** 2004. Use of microautoradiography combined with fluorescence in situ hybridization to determine dimethylsulfoniopropionate incorporation by marine bacterioplankton taxa. *Appl. Environ. Microbiol.* **70:**4648-4657.

36. **Wallner, G., B. Fuchs, S. Spring, W. Beisker, and R. Amann.** 1997. Flow sorting of microorganisms for molecular analysis. *Appl. Environ. Microbiol.* **63:**4223-4231.

37. **Yoch, D. C.** 2002. Dimethylsulfoniopropionate: its sources, role in the marine food web, and biological degradation to dimethylsulfide. *Appl. Environ. Microbiol.* **68:**5804-5815.

38. **Yoch, D. C., J. H. Ansede, and K. S. Rabinowitz.** 1997. Evidence for intracellular and extracellular dimethylsulfoniopropionate (DMSP) lyases and DMSP uptake sites in two species of marine bacteria. *Appl. Environ. Microbiol.* **63:**3182-3188.

39. **Zubkov, M. V., B. M. Fuchs, S. D. Archer, R. P. Kiene, R. Amann, and P. H. Burkill.** 2001. Linking the composition of bacterioplankton to rapid turnover of dissolved dimethylsulphoniopropionate in an algal bloom in the North Sea. *Environ. Microbiol.* **3:**304-311.

40. **Zubkov, M. V., B. M. Fuchs, S. D. Archer, R. P. Kiene, R. Amann, and P. H. Burkill.** 2002. Rapid turnover of dissolved DMS and DMSP by defined bacterioplankton communities in the stratified euphotic zone of the North Sea. *Deep-Sea Res. Part II Top. Stud. Oceanogr.* **49:**3017-3038.

41. **Zubkov, M. V., B. M. Fuchs, P. H. Burkill, and R. Amann.** 2001. Comparison of cellular and biomass specific activities of dominant bacterioplankton groups in stratified waters of the Celtic Sea. *Appl. Environ. Microbiol.* **67:**5210-5218.

Table 2.1. Cell number and percentage of the total cell count for HNA and LNA subpopulations in four samples at 24 h incubation[a].

Sample	Population	Control (Avg ± SD)			DMSP (Avg ± SD)			Significant difference[c]
		# of cells ($\times 10^6$ ml^{-1})	% of total cells	# of cells sorted ($\times 10^3$)[b]	# of cells ($\times 10^6$ ml^{-1})	% of total cells	# of cells sorted ($\times 10^3$)[b]	
May-02	HNA	0.1 ± 0.1	2.1 ± 0.5	0.8 ± 0.1	5.2 ± 0.6	28.2 ± 1.7	**49.3 ± 1.8**	Yes
	LNA3	1.1 ± 0.3	16.8 ± 1.0	9.2± 0.3	6.4 ± 0.4	34.6 ± 1.0	**59.2 ± 1.5**	Yes
	LNA2	5.0 ± 0.2	76.8 ± 0.8	**27.2 ± 1.7**	5.4 ± 0.4	29.2 ± 0.8	**24.8 ± 0.9**	Yes
	LNA1	0.3 ± 0.0	4.4 ± 0.5	1.6 ± 0.1	1.5 ± 0.5	8.0 ± 1.8	6.6 ± 0.3	No
Sept-02	HNA	0.0 ± 0.0	2.1 ± 0.5	11.8 ± 1.9	5.8 ± 1.5	9.7 ± 2.6	**80.2 ± 25.0**	Yes
	LNA2	0.2 ± 0.0	11.2 ± 1.0	**62.9 ± 0.6**	40.7 ± 1.5	68.6 ± 2.5	**566.2 ± 79.8**	Yes
	LNA1	1.5 ± 0.0	80.5 ± 1.6	455.9 ± 50.0	9.1 ± 0.3	15.4 ± 0.5	126.4 ± 11.6	Yes
Jan-03	LNA2	2.6 ± 0.2	59.2 ± 1.9	**78.1 ± 2.5**	3.4 ± 0.2	56.1 ± 2.7	**96.0 ± 5.2**	No
	LNA1	1.8 ± 0.3	40.8 ± 5.1	**53.9 ± 7.1**	2.7 ± 0.2	43.9 ± 3.5	**75.1 ± 9.8**	No
May-03	LNA2	5.5 ± 0.8	64.6 ± 3.9	**99.5 ± 7.1**	11.4 ± 0.9	63.6 ± 5.0	**82.0 ± 8.2**	No
	LNA1	3.2 ± 0.2	36.7 ± 4.4	**56.5 ± 5.3**	6.5 ± 1.0	36.1 ± 4.7	**46.6 ± 7.6**	No

[a] Average numbers and standard deviations were generated based on independent analysis of 3 replicate incubations.

[b] Actual number of cells available for molecular analysis. Boldface indicates FCM populations that were able to provide DNA templates for subsequent molecular analysis.

[c] "Significant difference" indicates whether the number of cells differed between DMSP and control treatments for a given subpopulation at $p < 0.05$.

Table 2.2. Taxonomic identification of bacterial groups associated with peaks in T-RFLP profiles based on 16S rDNA clone library analysisa

Sample	TRFL (bp)	Clone (accession no.)	Group	Closest relative (accession no.)	Similarity (%)	Closest described relative (accession no.)	Similarity (%)
May-02	32	SIMO-1620 (AY476751)	*Actinobacteria*	Uncultured *Actinobacterium* MB11A03 (AY033296)	96	*Streptomyces africanus* (AY208912)	80
	57	SIMO-277 (AF547399)	*Roseobacter*	Uncultured α-proteobacterium (AY580445)	99	"*Citricella thiooxidans*" (AY639887)	95
		SIMO-293 (AF547406)	*Roseobacter*	Uncultured α-proteobacterium (AY038571)	97	*Roseovarius nubinhibens* (AF098495)	96
		SIMO-312 (AF547424)	*Sulfitobacter*	Arctic sea ice bacterium ARK9994 (AF468380)	97	*Sulfitobacter pontiacus* (Y13155)	97
		SIMO-313 (AF547425)	*Sulfitobacter*	Arctic sea ice bacterium ARK9994 (AF468380)	97	*Sulfitobacter pontiacus* (AY159887)	97
		SIMO-1603 (AY476739)	*Roseobacter*	Uncultured α-proteobacterium (AY580445)	100	*Ruegeria atlantica* (D88527)	95
		SIMO-1604 (AY476740)	*Sulfitobacter*	*Sulfitobacter pontiacus* (AY159887)	99	*Sulfitobacter pontiacus* (AY159887)	99
	98	SIMO-295 (AF547408)	*Bacteroidetes*	Uncultured *Bacteroidetes* (AY580726)	92	*Flexibacter canadensis* (AB078046)	76
		SIMO-296 (AF547409)	*Bacteroidetes*	Uncultured *Bacteroidetes* (AY580664)	94	*Flavobacterium columnare* (AB015480)	82
		SIMO-309 (AF547422)	*Bacteroidetes*	Uncultured *Bacteroidetes* (AY580666)	96	*Flavobacterium limicola* (AJ585236)	82
	115	Unidentified					
	202	SIMO-308 (AF547421)	β-Proteobacteria	Uncultured β-proteobacterium (AB113609)	93	*Xenophilus azovorans* (AF285414)	92
		SIMO-1616 (AY476750)	β-Proteobacteria	Uncultured bacterium (AF546929)	95	*Brachymonas petroleovorans* (AY275432)	93
	350-360	SIMO-292 (AF547405)	*Roseobacter*	*Roseovarius* sp. strain 2S5-2 (AB114422)	97	*Sulfitobacter dubius* (AY180102)	96
		SIMO-294 (AF547407)	*Roseobacter*	Bacterium K2-11 (AY345438)	98	*Loktanella vestfoldensis* (AJ582226)	95
		SIMO-300 (AF547413)	*Oceanicola*	*Oceanicola batsensis* (AY424898)	97	*Oceanicola batsensis* (AY424898)	97

		SIMO-301 (AF547414)	α-Proteobacteria	Uncultured bacterium (AJ459874)	95	*Caulobacter crescentus* (AE006011)	95
		SIMO-302 (AF547415)	*Roseobacter*	Bacterium K2-11 (AY345438)	99	*Loktanella vestfoldensis* (AJ582227)	95
		SIMO-1605 (AY476741)	*Roseobacter*	Bacterium K2-11 (AY345438)	98	*Loktanella vestfoldensis* (AJ582226)	96
	565	SIMO-1599 (AY476738)	γ-Proteobacteria	*Neptunomonas naphthovorans* (AF053734)	93	*Neptunomonas naphthovorans* (AF053734)	93
Sept-02	57	SIMO-335 (AY476763)	*Roseobacter*	Unidentified bacterium (Z88582)	99	*Ruegeria atlantica* (D88527)	95
		SIMO-341 (AY476769)	*Roseobacter*	Unidentified bacterium (Z88582)	99	*Ruegeria atlantica* (D88527)	95
		SIMO-342 (AY476770)	*Roseobacter*	Unidentified α-proteobacterium (AY580445)	99	*Ruegeria atlantica* (D88527)	94
		SIMO-343 (AY476771)	*Roseobacter*	Unidentified α-proteobacterium (AY580445)	99	*Ruegeria atlantica* (D88527)	96
		SIMO-346 (AY476774)	*Roseobacter*	Unidentified bacterium (Z88582)	98	"*Citricella thiooxidans*" (AY639887)	95
		SIMO-1606 (AY476802)	*Roseobacter*	Uncultured α-proteobacterium (AY580466)	99	*Roseovarius nubinhibens* (AF098495)	96
	98	SIMO-347 (AY476775)	*Bacteroidetes*	Uncultured *Bacteroidetes* (AY580722)	99	*Cytophaga fermentans* (M58766)	84
		SIMO-351 (AY476779)	*Bacteroidetes*	Uncultured *Bacteroidetes* (AY580722)	99	*Cytophaga fermentans* (M58766)	84
		SIMO-1600 (AY476799)	*Bacteroidetes*	Uncultured *Bacteroidetes* (AY580722)	99	*Cytophaga fermentans* (M58766)	88
		SIMO-1601 (AY476800)	*Bacteroidetes*	Uncultured prokaryote (AF477832)	99	*Cytophaga fermentans* (M58766)	82
	202	SIMO-1617 (AY476804)	*Rhodoferax*	*Rhodoferax antarcticus* (AY609198)	97	*Rhodoferax antarcticus* (AY609198)	97
		SIMO-1618 (AY476805)	β-Proteobacteria	Uncultured bacterium (AF546929)	96	*Brachymonas petroleovorans* (AY275432)	94
	231	SIMO-1602 (AY476801)	γ-Proteobacteria	Uncultured γ-proteobacterium (AY580764)	93	*Microbulbifer elongatus* (AF500006)	86
	370	SIMO-339 (AY476767)	γ-Proteobacteria	*Pseudoalteromonas* sp. strain Md 213 (AY461670)	99	"*Pseudoalteromonas porphyrae*" (AF475096)	95
		SIMO-348 (AY476776)	γ-Proteobacteria	*Pseudoalteromonas* sp. strain Md 213 (AY461670)	99	*Pseudoalteromonas prydzensis* (U85855)	95
		SIMO-350 (AY476778)	γ-Proteobacteria	Uncultured γ-proteobacterium (AY580367)	95	*Colwellia demingiae* (U85844)	95

	SIMO-352 (AY476780)	γ-Proteobacteria	*Pseudoalteromonas* sp. strain Md 213 (AY461670)	99	*Pseudoalteromonas mariniglutinosa* (AJ507251)	95
	SIMO-353 (AY476781)	γ-Proteobacteria	*Pseudoalteromonas* sp. strain Md 213 (AY461670)	98	*Pseudoalteromonas prydzensis* (U85855)	94
	SIMO-356 (AY476784)	γ-Proteobacteria	*Pseudoalteromonas* sp. strain A28 (AF227238)	92	"*Pseudoalteromonas porphyrae*" (AF475096)	88
	SIMO-357 (AY476785)	γ-Proteobacteria	*Pseudoalteromonas* sp. strain Md 213 (AY461670)	99	"*Pseudoalteromonas porphyrae*" (AF475096)	94
	SIMO-367 (AY476795)	γ-Proteobacteria	*Pseudoalteromonas* sp. strain FE1-03 (AJ784130)	94	"*Pseudoalteromonas porphyrae*" (AF475096)	93
518	SIMO-366 (AY476794)	*Roseobacter*	Marine α-proteobacterium AS-19 (AJ391181)	97	*Ketogulonicigenium vulgare* (AF136846)	93

[a]Group affiliations were determined using the RDP-II Sequence Match program (http://rdp.cme.msu.edu/html/analyses.html). Closest relative of each clone was determined using the BLASTN program of Genbank (http://www.ncbi.nlm.nih.gov/BLAST/). Only one representative was shown if more than one clone shared an identical sequence within the same sample.

Table 2.3. Contribution of major bacterial taxa to the total bacterioplankton assemblage in control microcosms after 24 h of incubation based on percentage of total area under the T-RFLP chromatogram

Group	T-RF(s)	Contribution to each sample (%)			
		May-02	Sept-02	Jan-03	May-03
α-Proteobacteria	T-57, T-350~360, T-518	32	19	37	34
β-Proteobacteria	T-202	0	4	25	30
γ-Proteobacteria	T-231, T-370, T-565	21	14	0	0
Bacteroidetes	T-98	6	35	0	0
Actinobacteria	T-32	3	0	0	0
Unidentified bacteria		38	28	38	36

Figure 2.1. Flow cytometric analysis of control (left panel) and DMSP treatment (right panel) samples collected after 24 h of incubation in May 2002, September 2002, January 2003, and May 2003. Beads and HNA and LNA cells are delimited by gates on each cytogram. Boldface letters are used to label subpopulations that were analyzed with T-RFLP. Asterisks are used to label FCM populations used for clone library construction.

Fig. 2.1

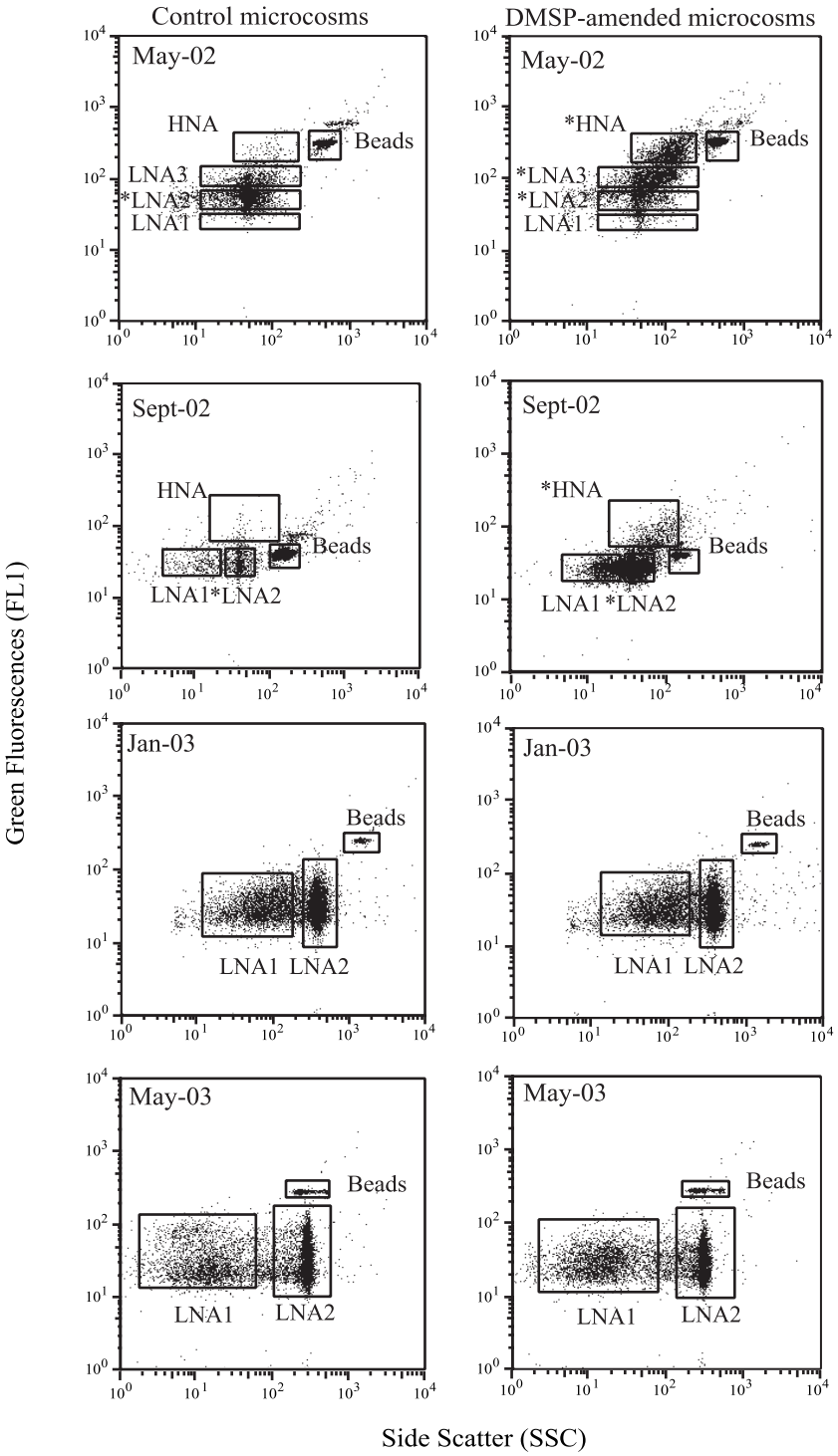

Control microcosms
DMSP-amended microcosms
May-02
HNA
Beads
LNA3
*LNA2
LNA1
*HNA
*LNA3
Sept-02
LNA1*LNA2
Jan-03
LNA1 LNA2
May-03
Green Fluorescences (FL1)
Side Scatter (SSC)

Figure 2.2. Total bacterial cell numbers (mean ± SD) counted by FCM in control and DMSP treatments over the 48 h incubation for each sample date.

Fig. 2.2

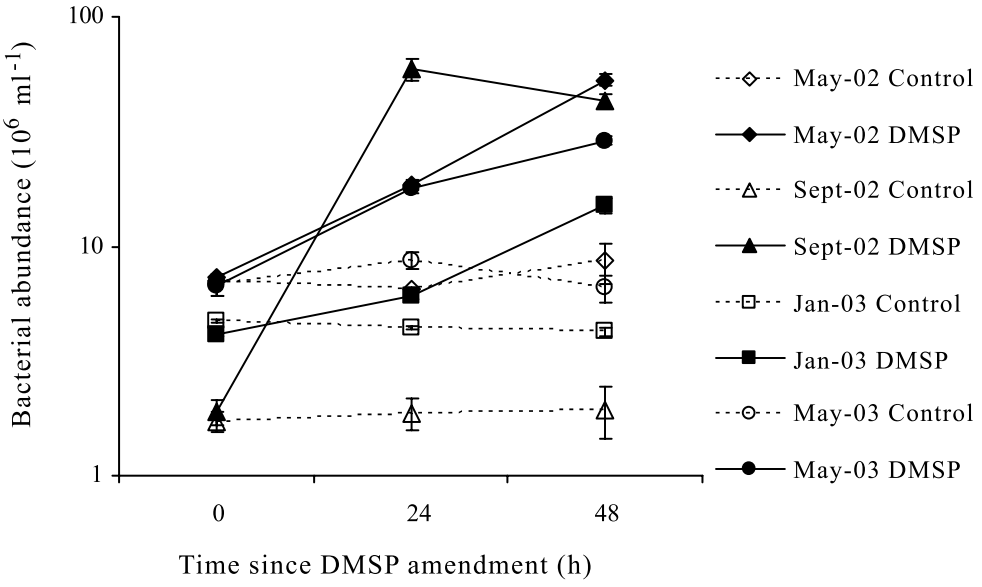

100
10
1
Bacterial abundance (10^6 ml^{-1})
0
24
48
Time since DMSP amendment (h)
May-02 Control
May-02 DMSP
Sept-02 Control
Sept-02 DMSP
Jan-03 Control
Jan-03 DMSP
May-03 Control
May-03 DMSP

Figure 2.3. Cluster analysis of T-RFLP analysis after 24 h of incubation. Similarity calculation was based on the Bray-Curtis dissimilarity on square-root-transformed relative peak area. FCM populations from triplicated incubations are shown, except for the May-02 sample, for which duplicates were available. Sample codes indicate treatment (D, DMSP amendment, C, Control), replicates (1, 2, or 3), and sorted subpopulation (LNA2, LNA3, or HNA).

Fig. 2.3

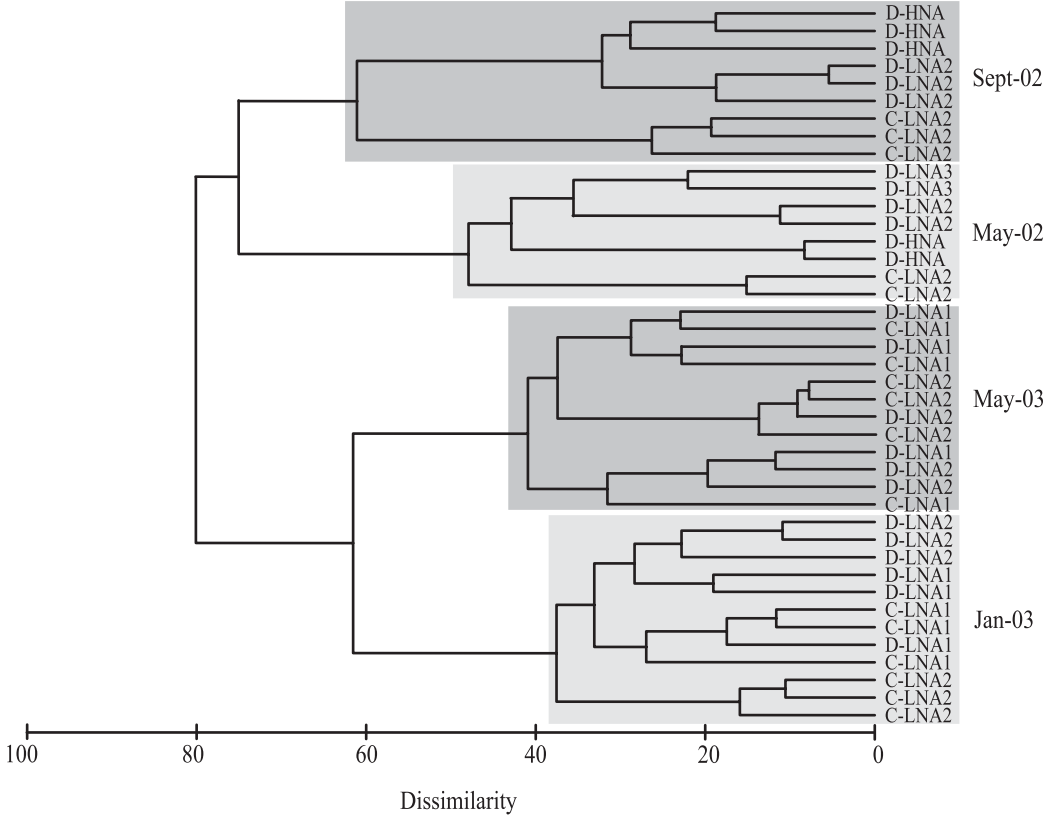

D-HNA
D-HNA
D-HNA
D-LNA2
D-LNA2
D-LNA2
C-LNA2
C-LNA2
C-LNA2
Sept-02
D-LNA3
D-LNA3
D-LNA2
D-LNA2
D-HNA
D-HNA
C-LNA2
C-LNA2
May-02
D-LNA1
C-LNA1
D-LNA1
C-LNA1
C-LNA2
C-LNA2
D-LNA2
C-LNA2
D-LNA1
D-LNA2
D-LNA2
C-LNA1
May-03
D-LNA2
D-LNA2
D-LNA2
D-LNA1
D-LNA1
C-LNA1
C-LNA1
D-LNA1
C-LNA1
C-LNA2
C-LNA2
C-LNA2
Jan-03
100
80
60
40
20
0
Dissimilarity

Figure 2.4. T-RFLP Chromatograms of the May-02 (A) and Sept-02 (B) samples. Empirically determined taxonomic identities of T-RFs and their lengths are indicated beneath the chromatogram. Alpha, α-Proteobacteria; beta, β-Proteobacteria; gamma, γ-Proteobacteria; unid, unidentified. Sample codes are as in Figure 2.3.

Fig. 2.4A

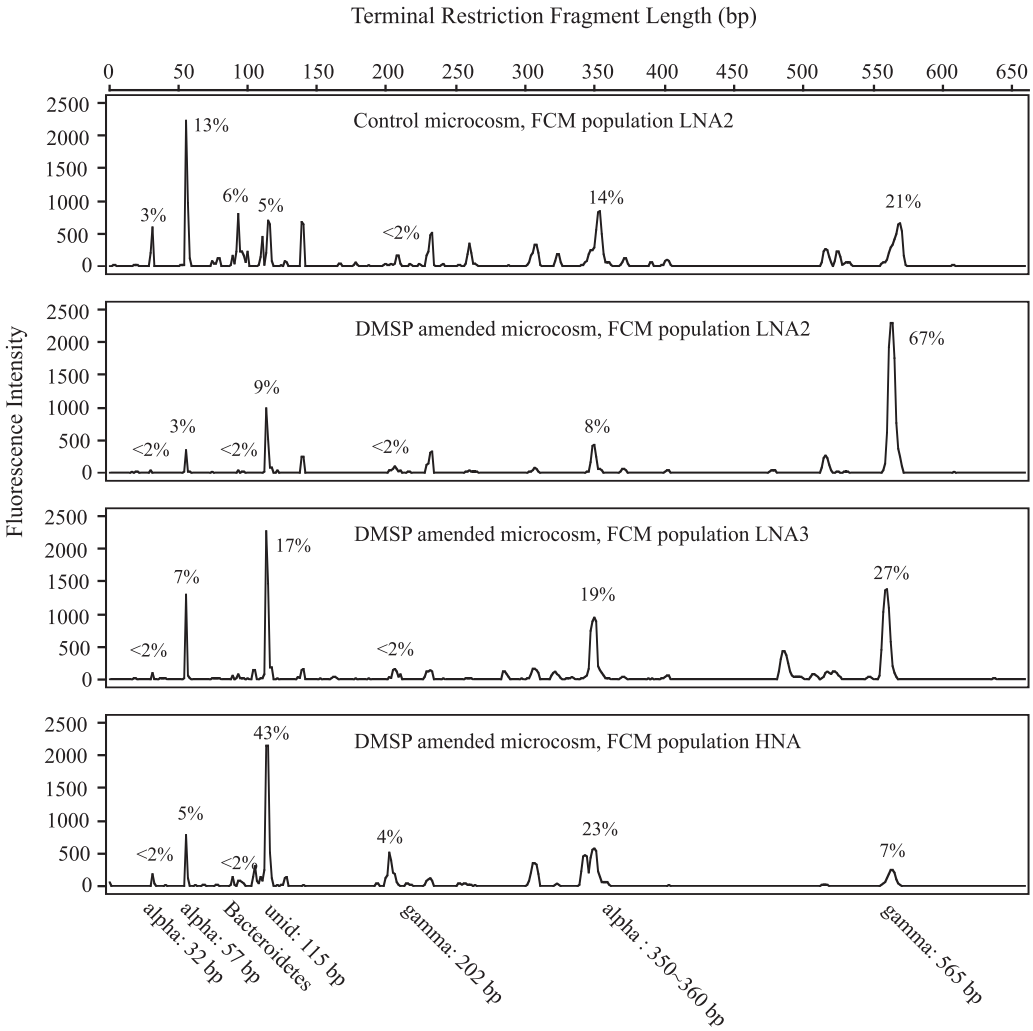

Terminal Restriction Fragment Length (bp)
0 50 100 150 200 250 300 350 400 450 500 550 600 650
Fluorescence Intensity
2500 2000 1500 1000 500 0
Control microcosm, FCM population LNA2
13%
6%
5%
3%
<2%
14%
21%
DMSP amended microcosm, FCM population LNA2
67%
9%
3%
<2%
<2%
<2%
8%
DMSP amended microcosm, FCM population LNA3
17%
7%
<2%
<2%
19%
27%
DMSP amended microcosm, FCM population HNA
43%
5%
<2%
<2%
4%
23%
7%
alpha: 32 bp
alpha: 57 bp
Bacteroidetes
unid: 115 bp
gamma: 202 bp
alpha : 350-360 bp
gamma: 565 bp

Fig. 2.4B

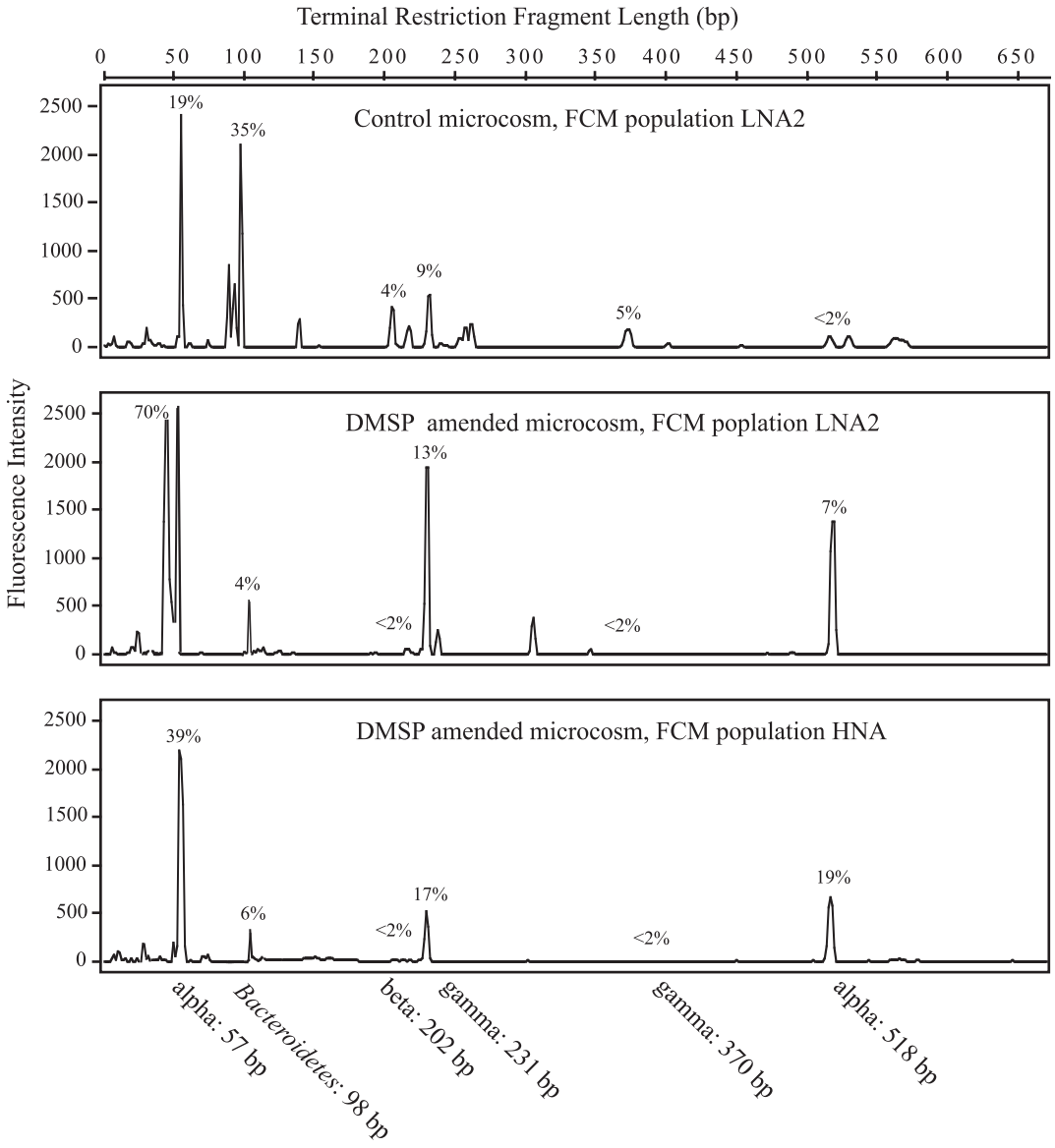

Terminal Restriction Fragment Length (bp)
0 50 100 150 200 250 300 350 400 450 500 550 600 650
Fluorescence Intensity
Control microcosm, FCM population LNA2
19%
35%
4%
9%
5%
<2%
DMSP amended microcosm, FCM poplation LNA2
70%
4%
<2%
13%
<2%
7%
DMSP amended microcosm, FCM population HNA
39%
6%
<2%
17%
<2%
19%
alpha: 57 bp
Bacteroidetes: 98 bp
beta: 202 bp
gamma: 231 bp
gamma: 370 bp
alpha: 518 bp

Figure 2.5. Community structures as revealed by T-RFLP chromatograms from different-sized bacterioplankton populations of 2×10^9 (A), 2×10^6 (B), and 2×10^3 (C) cells collected in May 2002.

Fig. 2.5

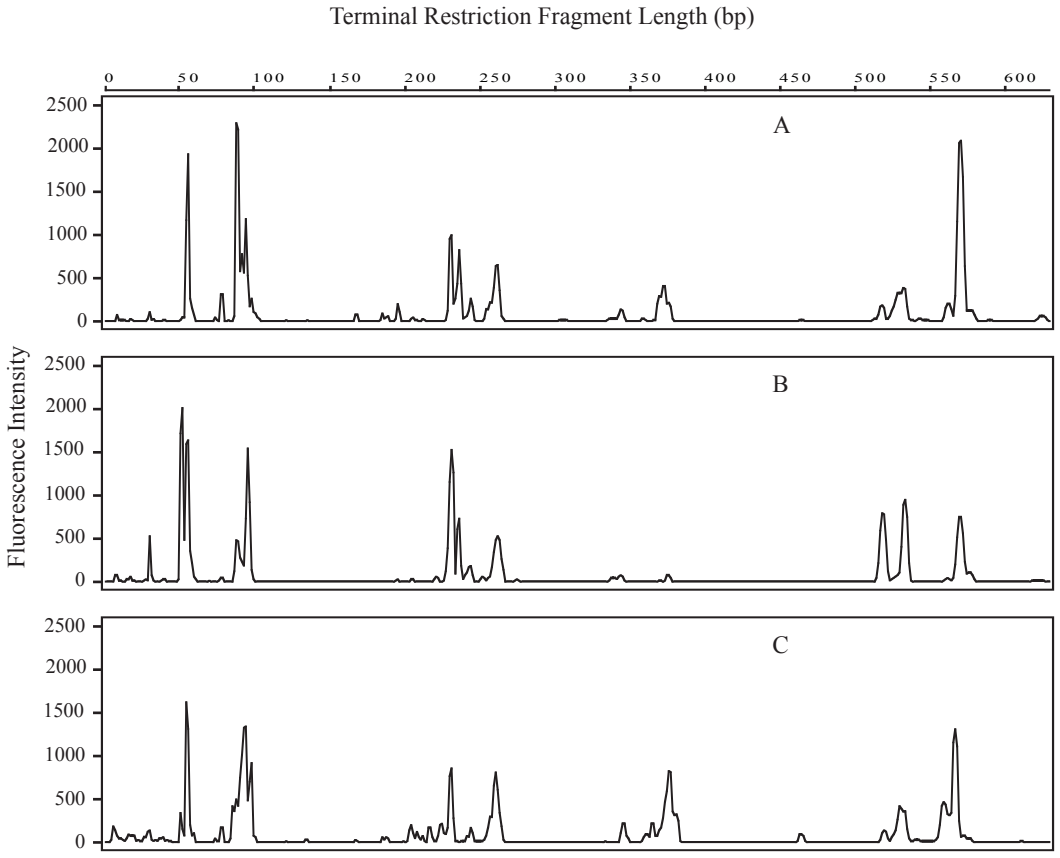
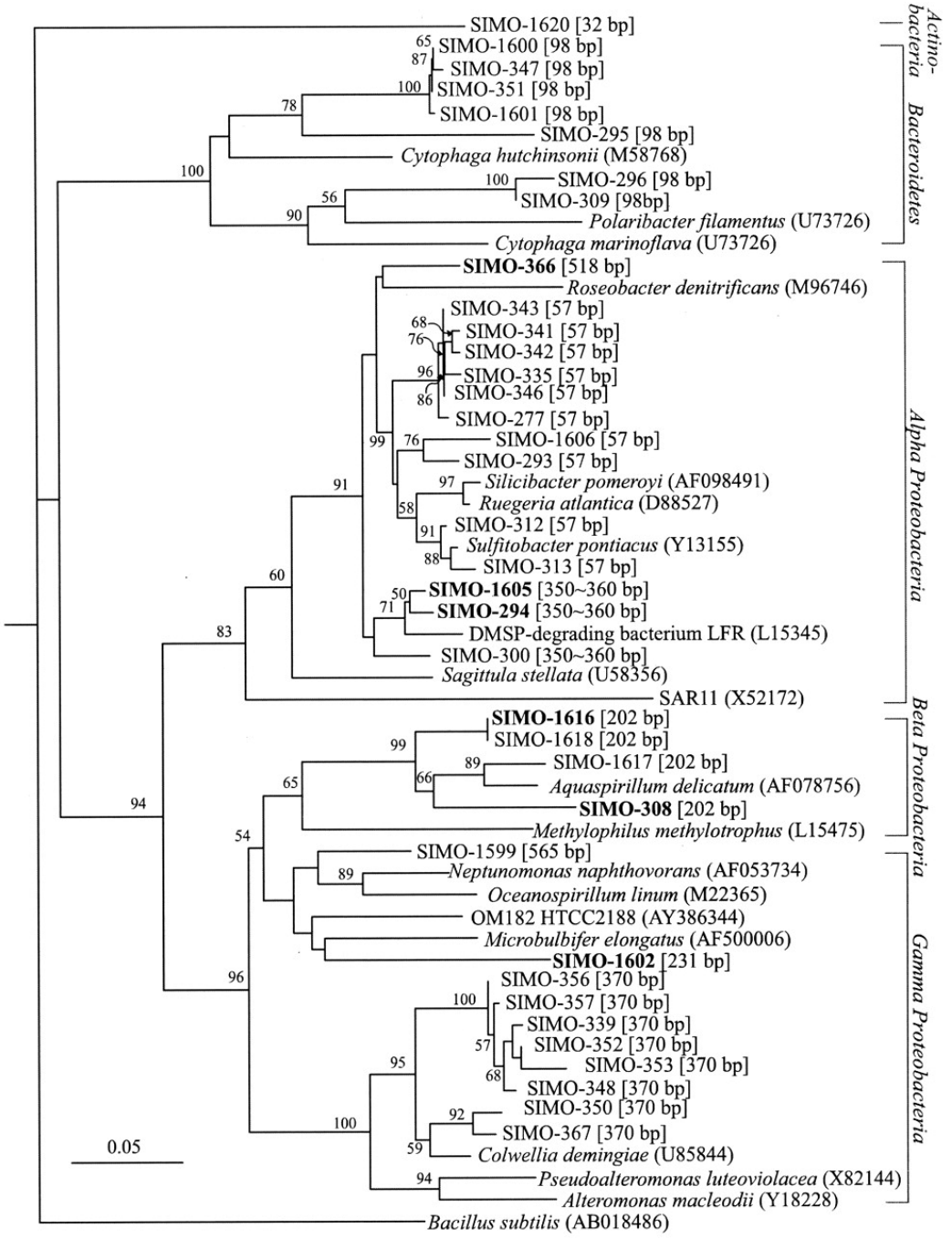

Terminal Restriction Fragment Length (bp)
0
50
100
150
200
250
300
350
400
450
500
550
600
Fluorescence Intensity
2500
2000
1500
1000
500
0
A
B
C

Figure 2.6. Phylogenetic tree based on partial sequences of 16S rDNA showing the relationship among clones from all subpopulations. Clones showed elevated NA content after DMSP addition are labeled by bold font. Clones showed depleted NA content in the DMSP and clones from control treatments are shown in regular font. The tree was constructed using Jukes-Cantor distance and neighbor-joining method. *Bacillus subtillis* was used as the outgroup. Bootstrap values of 50% are indicated at the branch nodes. The scale bar indicates the amount of genetic change in terms of the number of nucleotide substitutions per site. The Genbank accession numbers of the reference sequences are shown in parentheses. Empirically determined T-RF lengths are shown in brackets. Only one representative was shown if more than one clone shared an identical sequence in the same sample.

Fig. 2.6


SIMO-1620 [32 bp]
SIMO-1600 [98 bp]
SIMO-347 [98 bp]
SIMO-351 [98 bp]
SIMO-1601 [98 bp]
SIMO-295 [98 bp]
Cytophaga hutchinsonii (M58768)
SIMO-296 [98 bp]
SIMO-309 [98bp]
Polaribacter filamentus (U73726)
Cytophaga marinoflava (U73726)
SIMO-366 [518 bp]
Roseobacter denitrificans (M96746)
SIMO-343 [57 bp]
SIMO-341 [57 bp]
SIMO-342 [57 bp]
SIMO-335 [57 bp]
SIMO-346 [57 bp]
SIMO-277 [57 bp]
SIMO-1606 [57 bp]
SIMO-293 [57 bp]
Silicibacter pomeroyi (AF098491)
Ruegeria atlantica (D88527)
SIMO-312 [57 bp]
Sulfitobacter pontiacus (Y13155)
SIMO-313 [57 bp]
SIMO-1605 [350~360 bp]
SIMO-294 [350~360 bp]
DMSP-degrading bacterium LFR (L15345)
SIMO-300 [350~360 bp]
Sagittula stellata (U58356)
SAR11 (X52172)
SIMO-1616 [202 bp]
SIMO-1618 [202 bp]
SIMO-1617 [202 bp]
Aquaspirillum delicatum (AF078756)
SIMO-308 [202 bp]
Methylophilus methylotrophus (L15475)
SIMO-1599 [565 bp]
Neptunomonas naphthovorans (AF053734)
Oceanospirillum linum (M22365)
OM182 HTCC2188 (AY386344)
Microbulbifer elongatus (AF500006)
SIMO-1602 [231 bp]
SIMO-356 [370 bp]
SIMO-357 [370 bp]
SIMO-339 [370 bp]
SIMO-352 [370 bp]
SIMO-353 [370 bp]
SIMO-348 [370 bp]
SIMO-350 [370 bp]
SIMO-367 [370 bp]
Colwellia demingiae (U85844)
Pseudoalteromonas luteoviolacea (X82144)
Alteromonas macleodii (Y18228)
Bacillus subtilis (AB018486)
0.05
Actino-bacteria
Bacteroidetes
Alpha Proteobacteria
Beta Proteobacteria
Gamma Proteobacteria

CHAPTER 3

BACTERIOPLANKTON ASSEMBALGES TRANSFORMING SPECIFIC DISSOLOVED ORGANIC COMPOUNDS IN COASTAL SEAWATER[2]

[2] Mou, X., R. E. Hodson, and M. A. Moran. To be submitted to *Environmental Microbiology*.

SUMMARY

To characterize bacterioplankton functional assembalges that transform specific components of the coastal seawater dissolved organic carbon (DOC) pool, bromodeoxyuridine (BrdU) was used to label the cells that were active following addition of one of four model compounds representative of organic osmolytes [dimethylsulfoniopropionate (DMSP) and glycine betaine (GlyB)] and aromatic monomers [para-hydroxybenzoic acid (pHBA) and vanillic acid (VanA)] in a marsh bacterial community. Bacterioplankton populations were sorted based on activity level by fluorescence-activated cell sorting (FACS) and then characterized by16S rDNA-based terminal restriction fragment length polymorphisms (T-RFLP) and clone library analysis. Assemblages with high metabolic activities (HA) developed within 16 h incubation in all of four DOC amended samples. The HA groups in all four amendments were composed of bacteria from the same phyla/subphyla, but the relative abundance of each taxon differed between the samples receiving organic osmolytes and aromatic compounds. To look at the HA assemblages at more detailed taxonomic rankings (family and species levels), two high-resolution clone libraries (each contained ~ 200 sequenced clones) were constructed for GlyB and VanA amendments. *Actinobacteria*, *Bacteroidetes*, α-Proteobacteria (mainly members of the *Roseobacter* clade), β-Proteobacteria, γ-Proteobacteria (mainly members of *Altermonadaceae*, *Chromatiaceae*, *Oceanospirillaceae* and *Pseudomonadaceae*) were found in both libraries but the relative abundances of these groups differed. The GlyB library was significantly enriched with members affiliated with the genera *Marinomonas* and *Oceanospirillum* in *Oceanospirillaceae* family of the γ-Proteobacteria. The VanA library was significantly enriched with members affiliated with the genera *Silicibacter* and *Sagittula* in the *Roseobacter* clade of α-Proteobacteria, and an uncultured sulfur-oxidizing bacterium in the *Chromatiales* order of the γ-

Proteobacteria. The capture of members of OM60/241, OM185, SAR11, SAR86 and SAR116 into the actively growing population during FCM sorting indicated that members of these groups may be mediating GlyB or VanA, providing some of the first glimpses into heterotrophy by these ubiquitous but poorly described environmental clusters.

INTRODUCTION

Coastal marshes and estuaries serve as repositories of DOC produced internally or transported from terrestrial and oceanic sources. The transformation of coastal DOC is carried out by diverse communities of bacterioplankton whose rates and pathways of DOC processing ultimately determine the form and amount of carbon exported to the atmospheric and oceanic reservoirs (Moran and Hodson, 1994; Hullar et al., 1996; Moran et al., 1999; Raymond and Bauer, 2000). However, the composition of bacterioplankton functional assemblages (i.e., groups of cells that carry out a defined metabolic function) that are responsible for the turnover of individual components of the DOC pool is largely unknown. This is partly because widely-employed 16S rDNA based techniques, although providing taxonomic identification in a culture-independent manner, generally lack the ability to link bacterial species with their metabolic activities.

Techniques have been developed to label active cells within a bacterial community with bromodeoxyuridine (BrdU) (Urbach et al., 1999; Artursson and Jansson, 2003). As a halogenated analog for thymidine, BrdU is incorporated into newly synthesized DNA in active cells and readily immunodetected with high specificity and sensitivity (Borneman, 1999). Methods involving fluorescence-activated cell sorting (FACS) discriminate among populations within complex communities (Burton, 1996; Collier and Campbell, 1999), and the defined populations can be physically sorted and analyzed by subsequent molecular techniques (Mou et al., 2005). A successful combination of BrdU incorporation and FACS will permit species-level taxonomic identification of cells at defined levels of metabolic activity.

The goals of the present study were to characterize and compare the phylogenetic diversities of functional assemblages that are responsible for transforming two important coastal

DOC components, aromatic monomers and organic osmolytes. Individual compounds representing the two DOC classes were added at 100 nM concentrations to natural communities, and a protocol combining BrdU incorporation, FACS, and 16S rDNA analysis was employed to identify bacterioplankton with the strongest response to the enrichments at the species level. Our results indicate that the two DOC-transforming assemblages were both diverse and composed of the same major bacterial groups commonly found in coastal seawaters but with different relative abundance of some taxa. Our results also provide insights into DOC utilization by poorly characterized marine environmental clusters including OM60/241, OM185, SAR11, SAR86 and SAR116.

RESULTS

Development of active bacterial populations. Bacterioplankton communities in southeastern U. S. coastal waters are generally carbon and nitrogen limited (Pomeroy et al., 2000). Natural DOC concentrations average approximately 500 μM, but a significant fraction of this pool is biologically refractory (Moran et al., 1991). The addition of model compounds at 100 nM concentrations increased the availability of individual labile compounds representing aromatic monomers and organic osmolytes, but had relatively little impact on overall DOC levels.

As revealed by FCM analysis (Figure 3.1), the total number of bacterioplankton cells increased during the course of a 24-h incubation in the model compound amendments but not in the no-addition controls (Figure 3.2A). At the 24-h time point, the increase in cell numbers in amendments receiving organic osmolytes (DMSP and GlyB) was about 1.4-fold greater than controls, and for those receiving aromatic compounds (pHBA and VanA) was about 1.2-fold greater. Based on the apparent grouping pattern and level of BrdU incorporation, which is

positively related to cell metabolic activity, cells in DOC amendments and no-addition controls were categorized into 3 groups: high- activity population (HA), medium-activity population (MA) and low-activity population (LA; Figure 3.1). HA populations developed only in model compound amendments. After 16 h incubation, the HA populations reached their maximum abundance in the amended samples (i.e., 8.2% of the total cell counts in DMSP amendments, 10.1% in GlyB, 6.1% in pHBA and 7.6% in VanA). At this same time point, the number of HA cells in no-addition controls was significantly lower (only 0.5% of total cell counts; Kruskal-Wallis analysis $p < 0.05$; Figure 2.2b). To catch the point of maximum numbers of HA cells, flow sorting and subsequent molecular analysis were performed on samples taken at the 16-h time point.

16S rDNA T-RFLP of sorted populations. The same major terminal restriction fragments (T-RFs, typically 12 in total) were present in T-RFLP chromatograms for the HA assemblages and total communities, but their relative abundance varied among the treatments (Figure S3.1). Quantitative comparisons of T-RFLP data using non-metric multi-dimensional scaling (MDS; Figure 3.3) and one way analysis of similarity (ANOSIM) revealed that: 1) the composition of total bacterial community shifted in response to model compound additions; 2) bacterial populations in amendments receiving similar model compounds (i.e., aromatic monomers versus osmolytes) were more similar to each other than to those from a different type; and 3) the most active cells (HA) from different model compound amendments were grouped seperately from one another (Figure 3.3 and Table S3.1). The MDS analysis also indicated very good within-treatment reproducibility, as true replicates (incubated, sorted, and amplified independently) clustered closely together (Figure 3.3).

16S rDNA sequences of sorted populations. Bacterial diversity at the species level is generally inaccessible by 16S rDNA based T-RFLP analysis because different bacterial taxa may generate T-RFs with the same length (as in the present study, Figure S3.1). To assess the diversity at a finer scale, twelve clone libraries were constructed, including one for the original community, three for no-addition controls after 16-h incubation (Total, LA and MA), and four for each of GlyB and VanA amendments after 16-h incubation (total, LA, MA and HA), based on one replicate for each sample. The 40-48 sequences obtained for each library (534 total sequences with an average length of 650 bp) were distributed among five major taxa that are commonly found in coastal waters (Table 1). For each library, the sequences grouped into 26-33 operational taxonomic units (OTUs) at an evolutionary distance of 0.03 (roughly the species level) (Schloss and Handelsman, 2004), and most OTUs were singletons. Despite the low coverage of these small clone libraries (31% - 59% coverage based on Good's coverage calculation) (Good, 1953), several patterns emerged: 1) the community structure remained unchanged for the no-addition control during the course of the incubation; 2) the community in the GlyB and VanA amendments contained the same major groups but the relative abundance of these groups were different; 3) the relative abundance of γ-Proteobacteria increased in the GlyB amendments (from 61% to 81%); and 4) the relative abundance of α-Proteobacteria increased in the VanA amendments (from 3% to 11%; Table 3.1).

Comparisons between high activity populations. To more deeply sample the populations that responded most strongly to model DOC compounds, additional clones were sequenced from GlyB and VanA HA libraries. These two HA populations were most dissimilar to each other based on T-RFLP data (Figure 3.3 and Table S3.1), and represented one of each of the DOC classes. A total of 224 and 239 sequences were recovered from the two high-resolution clone

libraries, respectively (Table S3.2). Most of the sequences had highest similarities to environmental clones from the same study site [Sapelo Island Microbial Observatory (SIMO) microbial database, http://www.simo.marsci.uga.edu/public_db/] or other marine habitats including the North Sea (Wichels et al., 1998), open-ocean surface waters (Cho and Giovannoni, 2004), a hydrothermal vent (Lopez-Garcia et al., 2003), a cold seep (Heijs et al., 2005), and sea ice (Brinkmeyer et al., 2003).

Statistical analysis using ∫- libshuff (Schloss et al., 2004) indicated that the HA sequence libraries differed significantly between GlyB and VanA amendments [$\Delta C_{GlyB/VanA}$, $P = 0.0023$; $\Delta C_{VanA/GlyB}$, $P < 0.0001$), in agreement with the T-RFLP data. Within individual HA libraries, sequences grouped into 55 OTUs for the GlyB library (86% coverage at a 0.03 distance level) and 88 OTUs for the VanA library (80% coverage). OTUs from both samples were distributed among five major phylogenetic taxa, but with different patterns. The most apparent differences were for the *Oceanospirllaceae* (γ- Proteobacteria), the *Roseobacter* clades (α-Proteobacteria) and the β-Proteobacteria (Figure 3.4). Consistent with the low-resolution clone libraries, members of the γ- Proteobacteria were specifically enriched in the GlyB HA library (55% of total OTUs in GlyB HA library vs. 24% for the VanA HA library), while members of the α- and β-Proteobacteria were specifically enriched in the VanA library (19% and 7% for VanA vs. 4% and 0.4% for GlyB).

OTUs from the two high-resolution HA libraries were found to overlap in 10 cases. Overlapping OTUs were in γ-Proteobacteria groups *Oceanospirillaceae* (3 OTUs) and *Chromatiaceae* (2 OTUs), the α-Proteobacteria *Roseobacter* clade (2 OTUs), and two unclassified taxa (2 OTUs; Figure 3.5). These overlapping OTUs often contained a relatively high number of sequences, and therefore accounted for a larger percentage of the total sequences

(32% and 17% of GlyB and VanA sequences, respectively) than of the total OTUs (18% and 11% of GlyB and VanA OTUs, respectively). Assuming these OTUs represent the metabolic generalists that are able to rapidly respond to multiple substrates, they account for a considerable fraction of the HA populations but not much of the diversity.

OTU clusters that had higher than 90% sequence similarity and contained more than 10 sequences were identified and examined for potential enrichment towards specific DOC amendments using a Chi-square analysis (Figure 3.5A and 3.5B). Out of six identified OTU clusters, five contained significantly more sequences from one HA library than the other (Table S3.3). Clusters O and M were enriched with sequences from the GlyB HA library and these sequences were related to *Oceanospirillum beijerinckii* and *Marinomonas vaga*, respectively. In contrast, Clusters C, R1 and R2 were biased towards the VanA HA library. Cluster C fell within the *Chromatiaceae* group*;* members had no close marine relatives in culture but were distantly related to a sulfur-oxidizing bacterium ODIII6 isolated from a shallow water hydrothermal vent. Clusters R1 and R2 both fell within the *Roseobacter* clade and their members were closely related to *Silicibacter pomeroyi* and *Sagittula stellata*, respectively. Only Cluster N, whose members were related to *Neptunomonas napthovorans,* showed no significant difference in the distribution of HA sequences between model compound amendments.

DISCUSSION

Using an approach that couples BrdU-incorporation, FACS, and 16S rDNA analysis, we identified bacterioplankton that exhibited a substantial increase in metabolic activity in response to specific components of the DOC pool. Because this approach involves addition of specific compound at non-tracer levels (100 nM) and incubation times approaching 24 h, cells identified

as responding to these experiments may differ from those processing a given substrate under steady-state conditions in an undisturbed community. Nonetheless, this study identifies bacterioplankton that are capable of increasing their activity in response to specific compounds typical of coastal DOC pools at high taxonomic resolution and in a culture-independent manner. The most active bacteria (i.e. those captured in the HA populations) are assumed to belong to functional assemblages capable of transforming the added substrates or their derivatives.

The four model compounds studied here represent two classes of DOC, organic osmolytes and aromatic compounds, both of which are common and abundant (as high as 200 nM) in coastal waters (Moran et al., 1991; González et al., 2000) but are structurally and ecologically distinct. Osmolytes GlyB and DMSP are aliphatic compounds that provide organic carbon along with nitrogen (GlyB) or sulfur (DMSP) to heterotrophic marine bacterioplankton, or may even be stored undegraded in the cell to function as an osmolyte (Kiene et al., 2000). Aromatic monomers pHBA and VanA contain a stable ring structure that must be cleaved via specialized enzymes during the initial steps of degradation. All four compounds stimulated a subset of the bacterioplankton community that contained the same taxa but differed in the relative representation (Figure 3.3). For the two studied in more detail, GlyB induced a stronger metabolic response in γ-Proteobacteria, while VanA induced a stronger response in α- and β-Proteobacteria. A substantial fraction of the GlyB and VanA induced bacterioplankton overlapped, probably for taxa housing metabolic generalist. In that case, it suggests that possessing a variety of degradation pathways for commonly-occurring substrate may be a successful ecological strategy in coastal seawater.

These patterns were also fairly consistent at the family level, with two clusters in the *Oceanospirillaceae* group showing a more pronounced response to GlyB than VanA, and two in

the *Roseobacter* group showing a more pronounced response to VanA than GlyB. The differential growth of bacterioplankton groups on specific substrates has been shown in several previous studies at phylum/class level. For example, in a study using florescent in situ hybridization (FISH)-based analysis, α-Proteobacteria have been shown to better utilize low-molecular-weight DOC (such as amino acids and DMSP) than β-, γ-Proteobacteria and *Bacteroidetes* at relatively low DOC concentration (<10 nM) (Cottrell and Kirchman, 2000; Malmstrom et al., 2004). Another study showed that the variation of the contribution of bacterial taxa, including the *Roseobacter* clade (α-Proteobacteria), SAR86 (γ-Proteobacteria) and *Bacteroidetes*, to one specific DOC transforming process depends on the ambient concentration of that DOC compound (Alonso and Pernthaler, 2006). While identifying taxonomic patterns in metabolic ability for substrate utilization is an important step for understanding drivers of DOC turnover in marine ecosystems, low taxonomic resolution (phylum/subphylum, class, family) limits access to the biochemical and genetic basis for DOC utilization. Our finding that many neighboring OTUs (with a distance of 0.05, roughly at the genus level) (Schloss and Handelsman, 2004) belonged to different functional assemblages (GlyB or VanA-transforming) indicates a substantial variability in the ability to use specific DOC components utilization at the genus and species level (Figure 3.5), despite trends at broader taxon levels (Figure 3.4). This OTU-level variability may be overlooked by approaches targeting higher taxonomic groups, but is essential for understanding and predicting DOC processing based on taxonomic composition of the bacterial community.

The across-taxa differences in metabolic response to elevated DOC supply could be due to variation in genetic capabilities among bacterioplankton groups, such that only a subset of taxa possesses the genes necessary for metabolizing the substrate. Alternatively, the patterns

could be due to inequalities in physiological response time to increased substrate concentrations, such that some taxa outcompete others that are also capable of utilizing the substrate. We examined these possibilities by screening marine bacterial genomes for evidence of GlyB and VanA degradation pathways using 46 genomes available in draft form from the Moore Marine Microbial Genome Sequencing Initiative. Most of these genomes are from heterotrophic surface water bacterioplankton in the α- and γ-Proteobacteria (17 genomes each) and *Bacteriodetes* (8 genomes). Homologs of two genes shown to be involved in degradation of GlyB (dimethyldiglycine dehydrogenase and sarcosine oxidase α subunit), and VanA (protocatechuate dioxygenase α subunit and vanillate O-demethylase oxidoreductase) were used as an index of genetic ability to metabolize the compounds (Table 3.2). Homologs of at least one of the two genes selected to represent the GlyB degradation pathway were present in 11 (65%) of the α-Proteobacteria and 3 (18%) of the γ-Proteobacteria. Homologs of at least one of the two VanA degradation pathway genes were present in 8 (47%) of α-Proteobacteria and 2 (12%) of γ-Proteobacteria. Both metabolic capabilities are therefore inferred to be represented in the two proteobacterial groups, as predicted by the checkered OTU patterns in Figure 3.5. The strong response by γ-Proteobacteria to addition of GlyB does not correlate well with the overall abundance of GlyB degrading enzymes in the currently available genomes, but the response by α-proteobacteria, and particularly those in the *Roseobacter* cluster, to VanA agrees with VanA-related degradative genes in seven of nine *Roseobacter* genomes (Table 3.2). Two caveats to this analysis are that alternative pathways of GlyB and VanA degradation may exist that are not being considered, and exchange of degradation pathway intermediates or cellular metabolites between cells could result in an enhanced BrdU incorporation by cells unable to metabolize the model compound directly.

Overall, we found that the ability to degrade GlyB and VanA is widespread across several major heterotrophic groups and bacterial generalists were accounted for a substantial fraction of the functional assemblages. Taxonomic patterns in functional assemblages at the phylum and sub-phylum levels (Figure 3.4) are superimposed on considerable variability at the genus and species levels (Figure 3.5). If these model compounds are representative of other DOC components, predicting rates and efficiencies of DOC degradation based on dynamics of taxonomic markers may be a challenge.

Members of OM60/OM241, OM182, SAR11, SAR116 and SAR86 were captured in the HA populations, indicating that they are possibly able to transport and metabolize the model compounds (Figure 3.5). The ubiquitous OMG (oligotrophic marine gamma-Proteobacteria) and SAR clusters are commonly present in this coastal site at low abundance (each averages < 3% of 16S rDNA SIMO clone libraries; simo.marsci.uga.edu/public_db). However, they can be numerically significant in surface waters of the open ocean, and thus are of considerable importance in the ecology of the marine environment (Giovannoni and Rappé, 2000). Little is yet known of their carbon acquisition strategies, but our results suggest heterotrophic assimilation of GlyB by members of the SAR116 cluster, and assimilation of VanA by members of the SAR 86, SAR116 OM60/241 and OM182 clusters. The presence of some of these taxa in the HA libraries might alternatively result from use of metabolites of the added compound.

Our approach provides species-level identification of bacterioplankton that metabolize individual DOC components and is sensitive enough to assign potential ecological functions to minor members of the bacterial community. However, the method will miss active bacterial cells that are insensitive to BrdU labeling (Pernthaler et al., 2002) and may produce false positives

from non-degraders that are indirectly stimulated by the added compound (e.g. those that use the degradation product of the added compounds or other metabolites of the stimulated bacteria).

Our experiment was conducted in the late autumn, a season characterized by a high number of γ-Proteobacteria (36.5 % in average of the total community in three October samples) and low number of α-Proteobacteria (10.8%), in southeastern U.S. coastal waters. The annual averages for these two bacterial taxa at this site are 18.6 % and 16.2 %, respectively (simo.marsci.uga.edu/public_db). Whether seasonal variation exists in the response of bacterial taxa to temporally increased DOC supply is not yet known. To fully understand the composition and dynamics of the functional assemblages of bacterioplankton that are responsible for turnover of the coastal DOC pool, better temporal resolution of the taxonomy and genetic capabilities of the bacterial functional assemblages is needed. Toward that end, we used an approach that couples BrdU-incorporation and FACS to allow us to identify members of the bacterioplankton that respond to specific DOC compounds at high taxonomic resolution. This provided insights into bacterially-mediated transformations of DOC under natural conditions.

EXPERIMENTAL PROCEDURES

Sample collection, processing and BrdU labeling. Surface water was taken from Dean Creek (a salt marsh tidal creek on Sapelo Island, GA) in acid-washed Nalgene carboys in October 2004. Once collected, water was sequentially filtered though 3.0- and 1.0-μm pore-size spun polypropylene filter cartridges (J B Systems) to obtain a bacterioplankton size fraction and exclude large particles and bacteriovores. Manipulative experiments were initiated immediately by adding 800 ml of 1.0-μm filtrate to ashed glass media bottles containing 10 μM BrdU (Roche) and one of four model compounds at 100 nM concentration: GlyB or DMSP, representing

organic osmolytes (Kiene et al., 2000); or pHBA or VanA, representing aromatic monomers (Moran and Hodson, 1989, 1994). No-addition controls were also established that received no model compounds. Samples (100 ml) were collected after 0, 8, 16, 24 h of incubation and preserved with 1% paraformaldehyde at room temperature for 2 h. Preserved cells were then collected on 0.22-μm pore-size white polycarbonate filters (Poretics Products) and washed with PBS three times before storage at 4°C. All incubations were carried out in triplicate and performed in the dark at in situ temperature with occasional agitation.

Immunodetection. Each filter was sequentially treated with lysozyme (10 mg ml^{-1}, 30 min at room temperature), proteinase K (2 μg ml^{-1}, 30 min at room temperature) and exonuclease III (50 U ml^{-1} in a buffer containing 5 mM $MgCl_2$ and 50 mM Tris HCl, 30 min at 37 °C). Immunodetection was carried out using an in situ cell proliferation kit (FLUOS; Roche) following a protocol modified from the manufacturer's instructions. First, filters were incubated with incubation buffer (0.5% BSA and 0.1% Tween 20 in PBS) for 10 min at room temperature. Then an anti-BrdU-FLUOS working solution (FITC conjugated antibody) was added to the surface of the filters and incubated for 3 h at 37°C in a humid chamber. Bacterial cells were counterstained with 4', 6-diamidino-2-phenylindole (DAPI, 1 μg ml^{-1}). Three gentle PBS washes were included between each step. Bacterial cells were brought back into suspension by incubating filters in 5 ml PBS at room temperature with shaking at 200 rpm for 10 min and votexing at high speed for 10 s before and after incubation. Over 80% cells were recovered for the FACS analysis, based on microscopic cell counts of samples before and after immunodection.

Fluorescence-activated flow cytometric cell sorting. FACS was performed with a MoFlo flow cytometer (Dakocytomation). A sterilized PBS solution (Puraflo; Dakocytomation) served as the sheath fluid. BrdU labeling was detected by green fluorescence (FL1) emitted from FITC and

used as a proxy for bacterial activity. The total bacterial cells were detected by UV fluorescence (FL6) emitted from DAPI. Single stained (DAPI or FITC) no-addition controls served to compensate for the spectral overlap between FITC and DAPI fluorescence within the Summit program (Dakocytomation). An internal standard (1-μm diameter yellow-green fluorescent beads; Fluoresbrite YG Microspheres; Polysciences) was mixed with the samples to estimate the cell density, but was excluded during the sorting step. Data acquisition was triggered by FL6. FCM populations were initially determined on the FCM cytograms of one no-addition replicate based on apparent grouping pattern after 16 h of incubation, and the gating frames were kept fixed through all other measurements (Fig 3.1). Typically, there were 3 apparent groups in one no-addition control replicate. The one that showed the lowest values of FL1 (BrdU-incorporation) and FL6 (DAPI stain) was designated the “particles or cell debris group’ (P). The other two groups were designated the “low activity group” (LA) and “medium activity group” (MA), based on the value of FL1, which is positively relatd to the metabolic activity of cells. The “high activity groups” (HAs) were defined to include cells that were only observed in the model compound amendments and had greatest FL1 values. Samples taken after 16 h incubations exhibited the biggest difference in HA abundance between DOC amendments and no-addition controls and were selected for sorting. FCM populations were sorted in the "purify 1 drop" mode and collected in sterile 1.5 ml polypropylene test tubes containing 1 ml PBS solution. Sorting was terminated when the number of sorted cells was >500,000 counts for each FCM population or at 30 min.

A Kruskal-Wallis ANOVA was performed to compare total cell number as well as relative abundance of HAs between control and nutrient amendments. Significant differences were reported at $p < 0.05$.

DNA extraction. DNA extraction from sorted cells was performed using a procedure described previously (Mou et al., 2005). Before extraction, the cells were incubated with 1 M NaCl and washed with PBS three times. DNA from the total bacterial community (unsorted) was obtained by using the UltraClean soil DNA extraction kit (MoBio).

T-RFLP analysis. For the total bacterial community, i.e., unsorted cells, three replicates were used for PCR amplification. For the FCM populations, one replicate was sacrificed to determine optimal PCR conditions and two were used for PCR amplification. Bacterial 16S rDNAs were amplified with a FailSafe PCR Premix selection kit (Epicentre) using 0.4 μM concentrations of 5' end FAM (carboxyfluorescein)-labeled 27F and unlabelled 1492R primers (Delong et al., 1989). PCR amplification was performed using a touch-down program with annealing temperature sequentially decreasing from 62 to 52°C per cycle, followed by 15 cycles at 52°C. Each cycle included denaturing (at 96°C for 50s), annealing (at 62-52°C for 50s), and extension (at 72°C for 60s). An initial 4-min denaturation and final 7-min extension step were also included in the PCR program. PCR amplification was confirmed by electrophoresis on ethidium bromide-stained 1% agarose gels. PCR amplicons were recovered from the gels with the QIAquick gel extraction kit (Qiagen), and digested with CfoI restriction enzyme (Roche) at 37 °C for 4.5 h, after which an ethanol precipitation was performed. The restricted amplicons were resuspended in 12 μl of deionized formamide plus 0.7 μl of DNA fragment length standard (Gene-Scan-2500 TAMRA; Applied Biosystems). The lengths of T-RFs were determined on an ABI PRISM 310 genetic analyzer (Applied Biosystems) in GeneScan mode.

T-RFLP output data (length and relative area of T-RFs) were standardized as described previously (Mou et al., 2005), and then quantitatively compared by MDS and ANOSIM based on a square-root-transformed Bray-Curtis dissimilarity matrix using Primer v5 (Primer-E LTD,

Plymouth, UK). The length of the T-RF of particular OTUs was later predicted based on in silico matching of T-RF peaks with cloned partial 16S rDNA sequences using a Visual Basic program for Microsoft Excel (González et al., 2000).

Clone library construction and sequencing analysis. One replicate of the total and sorted populations from each treatment was further analyzed by cloning and sequencing. Bacterial 16S rDNAs were amplified and purified using the protocol described above, with the exception that the 27F primer was not fluorescently labeled. PCR amplicons were cloned using a TA cloning kit (Invitrogen) following the manufacturer's instructions. 16S rDNA were sequenced in a 96-well format by SeqWright DNA Technology Services (Houston, TX) using the 27F primer. Sequences were assembled using Sequencher 4.1 (Gene Codes Corporation) and were checked for chimeras using the CHECK_CHIMERA program of the Ribosomal Database Project II (Cole et al., 2003). Bacterial taxonomic identities were automatically assigned using the SIMO RDP Agent (http://www.simo.marsci.uga.edu/public_db/bioinformatics.htm).

Phylogenetic analysis. The percent coverage of the clone libraries was calculated using the formula [1– (n/N)] x 100, where n is the number of sequence singletons and N is the total number of sequences in the corresponding library (Good, 1953).

The distance matrix of sequences was prepared based on Kimura's two-parameter calculations using the MEGA 3.1 package (Kumar et al., 2004). OTUs for each clone library were assigned using the DOTUR program based on the furthest-neighbor clustering algorithm and were reported at the 0.03 distance level (http://www.plantpath.wisc.edu/fac/joh/DOTUR.html) (Schloss and Handelsman, 2005). The Chao1 (Chao, 1984) and Shannon-Weaver estimations (Shannon and Weaver, 1963) of OTU diversity were also calculated using DOTUR.

The distance matrix of all OTUs (with each OTU represented by only one sequence) was prepared using the same method described above. Bootstrapping with 500 replicates was used to assign confidence levels to the nodes in the trees. Taxonomic affiliation of the OTUs was determined using Smith-Waterman alignments to representative sequences of common taxa in marine ecosystems (Moran et al. 2004).

Total clone library composition was statistically compared using ∫-libshuff with the default settings (http://www.plantpath.wisc.edu/fac/joh/S-LIBSHUFF.html) (Schloss et al., 2004). ∫-libshuff compares homologous and heterologous coverage curves for multiple pairwise reciprocal comparisons among a set of clone libraries based on the integral form of the Cramér-von Mises statistic. Significant difference between libraries was reported when P value was lower than the calculated standard error.

OTU clusters that contained sequences with higher than 90% similarity and more than 10 sequences were identified. The random distribution of members of the OTUs between the two DOC treatments (GlyB and VanA) was assessed using Chi-square statistic. A significant difference was reported when $p < 0.05$.

Bioinformatic analysis. Genomic sequences of marine bacterial isolates from the Moore microbial genome database were searched for two key genes of each from the GlyB and VanA pathways. Selected query sequences (given in Table 3.2) were analyzed against the genomes using BLASTP with a cutoff expected value $E < 10^{-30}$.

Nucleotide sequence accession numbers. The GenBank accession numbers for 16S rRNA gene sequences determined in this study are DQ189509-DQ189987, DQ295922-DQ295597 and DQ421418-DQ421778.

ACKNOWLEDGEMENTS

We thank J. Nelson for assistance with FCM and cell sorting; W. Sheldon, J. González and S. Sun for bioinformatics advice; E. Biers for assistance with sampling. This study was supported by NSF (MCB 0084164 to the Sapelo Island Microbial Observatory) and the Gordon and Betty Moore Foundation.

REFERENCES

Alonso, C., and Pernthaler, J. (2006) Concentration-dependent patterns of leucine incorporation by coastal picoplankton. *Appl Environ Microbiol* **72**: 2141-2147.

Artursson, V., and Jansson, J.K. (2003) Use of Bromodeoxyuridine immunocapture to identify active bacteria associated with *Arbuscular Mycorrhizal* Hyphae. *Appl Environ Microbiol* **69**: 6208-6215.

Borneman, J. (1999) Culture-independent identification of microorganisms that respond to specified stimuli. *Appl Environ Microbiol* **65**: 3398-3400.

Brinkmeyer, R., Knittel, K., Jurgens, J., Weyland, H., Amann, R., and Helmke, E. (2003) Diversity and structure of bacterial communities in arctic versus antarctic pack ice. *Appl Environ Microbiol* **69**: 6610-6619.

Burton, R.S. (1996) Molecular tools in marine ecology. *J Exp Mar Biol Ecol* **200**: 85-101.

Chao, Q (1984) Non-parametric estimation of the number of classes in a population. *Scand J Stat* **11**: 265-270.

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

Cole, J.R., Chai, B., Marsh, T.L., Farris, R.J., Wang, Q., Kulam, S.A. et al. (2003) The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucl Acids Res* **31**: 442-443.

Collier, J.L., and Campbell, L. (1999) Flow cytometry in molecular aquatic ecology. *Hydrobiologia* **401**: 33-53.

Cottrell, M.T., and Kirchman, D.L. (2000) Natural assemblages of marine proteobacteria and members of the *Cytophaga-Flavobacter* cluster consuming low- and high- molecular-weight dissolved organic matter. *Appl Environ Microbiol* **66**: 1692-1697.

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

Giovannoni, S.J., and Rappé, M. (2000) Evolution, diversity, and molecular ecology of marine prokaryotes. In *Microbial Ecology of the Oceans*. Kirchman, D.L. (eds). Wiley-Liss, Inc.,pp. 47-84.

González, J.M., Simó, R., Massana, R., Covert, J.S., Casamayor, E.O., Pedros-Alio, C., and Moran, M.A. (2000) Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. *Appl Environ Microbiol* **66**: 4237-4246.

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

Heijs, S.K., Damste, J.S.S., and Forney, L.J. (2005) Characterization of a deep-sea microbial mat from an active cold seep at the Milano mud volcano in the Eastern Mediterranean Sea. *FEMS Microbiol Ecol* **54**: 47-56.

Hullar, M., Fry, B., Peterson, B., and Wright, R. (1996) Microbial utilization of estuarine dissolved organic carbon: a stable isotope tracer approach tested by mass balance. *Appl Environ Microbiol* **62**: 2489-2493.

Kiene, R.P., Linn, L.J., and Bruton, J.A. (2000) New and important roles for DMSP in marine microbial communities. *J of Sea Res* **43**: 209-224.

Kumar, S., Tamura, K., and Nei, M. (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* **5**: 150-163.

Lopez-Garcia, P., Duperron, S., Philippot, P., Foriel, J., Susini, J., and Moreira, D. (2003) Bacterial diversity in hydrothermal sediment and epsilon proteobacterial dominance in experimental microcolonizers at the Mid-Atlantic Ridge. *Environ Microbio* **5**: 961-976.

Malmstrom, R.R., Kiene, R.P., and Kirchman, D.L. (2004) Identification and enumeration of bacteria assimilating dimethylsulfoniopropionate (DMSP) in the North Atlantic and Gulf of Mexico. *Limnol Oceanogr* **49**: 597-606.

Moran, M.A., and Hodson, R.E. (1989) Formation and bacterial utilization of dissolved organic carbon derived from detrital lignocellulose. *Limnol Oceanogr* **34**: 1034-1047.

Moran, M.A., Pomeroy, L.R., Sheppard, E.S., Atkinson, L.P., and Hodson, R.E. (1991) Distribution of terrestrially derived dissolved organic matter on the southeastern United States Continental Shelf. *Limnol Oceanogr* **36**: 1134-1149.

Moran, M.A., and Hodson, R.E. (1994) Dissolved humic substances of vascular plant origin in a coastal marine environment. *Limnol Oceanogr* **39**: 762-771.

Moran, M.A., Sheldon, W.M., and Sheldon, J.E. (1999) Biodegradation of riverine dissolved organic carbon in five estuaries of the southeastern United States. *Estuaries* **22**: 55-64.

Mou, X., Moran, M.A., Stepanauskas, R., González, J.M., and Hodson, R.E. (2005) Flow-cytometric cell sorting and subsequent molecular analyses for culture-independent identification of bacterioplankton involved in dimethylsulfoniopropionate transformations. *Appl Environ Microbiol* **71**: 1405-1416.

Pernthaler, A., Pernthaler, J., Schattenhofer, M., and Amann, R. (2002) Identification of DNA-synthesizing bacterial cells in coastal North Sea plankton. *Appl Environ Microbiol* **68**: 5728-5736.

Pomeroy, L.R., Sheldon, J.E., Sheldon, W.M., Blanton, J.O., Amft, J., and Peters, F. (2000) Seasonal changes in microbial processes in estuarine and continental shelf waters of the south-eastern USA. *Estuar Coast Shelf S* **51**: 415-428.

Raymond, P.A., and Bauer, J.E. (2000) Bacterial consumption of DOC during transport through a temperate estuary. *Aquatic Microbial Ecology* **22**: 1-12.

Schloss, P.D., and Handelsman, J. (2004) Status of the microbial census. *Microbiol Mol Biol R* **68**: 686-691.

Schloss, P.D., and Handelsman, J. (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* **71**: 1501-1506.

Schloss, P.D., Larget, B.R., and Handelsman, J. (2004) Integration of microbial ecology and statistics: a test to compare gene libraries. *Appl Environ Microbiol* **70**: 5485-5492.

Shannon, C.E., and Weaver, W. The mathematical theory of communication. Urbana, Univerisity of Illinois Press; 1963.

Urbach, E., Vergin, K.L., and Giovannoni, S.J. (1999) Immunochemical detection and isolation of DNA from metabolically active bacteria. *Appl Environ Microbiol* **65**: 1207-1213.

Wichels, A., Biel, S.S., Gelderblom, H.R., Brinkhoff, T., Muyzer, G., and Schutt, C. (1998) Bacteriophage diversity in the North Sea. *Appl Environ Microbiol* **64**: 4128-4133.

Table 3.1. Distribution of sequences from low-resolution clone libraries at the phylum and sub-phylum level for initial and 16-h samples, expressed as a percent of each library. Population abbreviations are the same as in Figure 3.1. Boldface letters indicate FCM populations that were also used for high-resolution clone library construction.

Treatments	**Populations**	***Actinobacteria* (%)**	***Bacteroidetes* (%)**	**Proteobacteria (%)**			**Other (%)**
				α	β	γ	
Initial community	Total	20	12	5	0	61	2
Control 16-h	Total	16	8	3	3	67	3
	LA	9	16	4	9	53	9
	MA	7	21	12	5	55	0
GlyB 16-h	Total	12	2	5	0	81	0
	LA	5	17	0	12	66	0
	MA	2	5	5	5	81	2
	HA	8	5	5	0	79	3
VanA 16-h	Total	16	7	11	0	64	2
	LA	4	8	12	4	72	0
	MA	13	7	11	2	65	2
	HA	3	7	20	10	53	7

Table 3.2. The presence of key genes of GlyB and VanA degradation in all available marine bacterioplankton genome sequences from the Moore Genome Sequencing Initiative (https://research.venterinstitute.org/moore/). A gene was considered present if the E value from BLASTP analysis was < 10-30. The query sequence of selected genes were obtained from NCBI's Genbank: dimethylglycine dehydrogenase (AAZ22057, *Pelagibacter ubique* HTCC1062), sarcosine oxidase alpha chain (*soxA*, AAZ21871, *Pelagibacter ubique* HTCC1062), vanillate O-demethylase oxidoreductase (*vanB*, AAN69333, *Pseudomonas putida* KT2440), protocatechuate 3,4-dioxygenase beta subunit (*pcaH*, AAV97184, *Silicibacter pomeroyi* DSS-3). Using query sequences from bacteria in different phyla or classes did not substantially change the outcome of the analysis.

Phylum/class	**Bacterial species**	**GlyB degradation**		**VanA degradation**	
		Dimethylglycine dehydrogenase	*soxA*	*vanB*	*pcaH*
Actinobacteria	*Janibacter* sp. HTCC2649				
	Marine actinobacterium PHSC20C1				
Bacterioidetes	*Robiginitalea biformata* HTCC2501	+			
	Cellulophaga sp. MED134				
	Flavobacteria bacterium BBFL7	+	+		
	Flavobacteriales bacterium HTCC2170				
	Flavobacterium sp. MED217			+	
	Polaribacter irgensii 23-P				
	Psychroflexus torquis ATCC 700755	+	+		
	Tenacibaculum sp. MED152	+			
Firmicutes	*Bacillus* sp. NRRL B-14911	+			
Planctomycetes	*Blastopirellula marina* DSM 3645				
α-Proteobacteria *Roseobacter* clade		+	+		+
	Loktanella vestfoldensis SKA53				
	Oceanicola batsensis HTCC2597	+	+		+
	Oceanicola granulosus HTCC2516		+	+	
	Rhodobacterales bacterium HTCC2654	+	+	+	
	Roseobacter sp. MED193	+	+	+	+
	Roseovarius nubinhibens ISM	+	+	+	
	Roseovarius sp. 217	+	+		+
	Sulfitobacter sp. EE-36	+	+		
	Sulfitobacter sp. NAS-14.1	+	+		
α-Proteobacteria Other clades	*Aurantimonas* sp. SI85-9A1				

	Erythrobacter litoralis HTCC2594				
	Erythrobacter sp. NAP1				
	Nitrobacter sp. Nb311A				
	Oceanicaulis alexandrii HTCC2633				
	Parvularcula bermudensis HTCC2503				
	Sphingomonas sp. SKA58		+		
	Pelagibacter ubique HTCC1002	+	++		+
γ-Proteobacteria				+	
Oceanospirillaceae	*Marinomonas* sp. MED121				
	Oceanospirillum sp. MED92		+		
	Oceanobacter sp. RED65		+		
γ-Proteobacteria other clades	*Alteromonas macleodii*				
	Gamma proteobacterium KT71				
	Gamma proteobacterium sp. HTCC2207	+	+		
	Idiomarina baltica OS145			+	
	Nitrococcus mobilis Nb-231				
	Photobacterium profundum 3TCK				
	Photobacterium sp. SKA34				
	Pseudoalteromonas tunicata D2				
	Psychromonas sp. CNPT3				
	Reinekea sp. MED297				
	Vibrio alginolyticus 12G01				
	Vibrio angustum S14				
	Vibrio sp. MED222				
	Vibrio splendidus 12B01				

Figure 3.1. Flow-cytometric analysis of control and model compound-amended samples collected after 16 h of incubation. Gate notation was based on BrdU incorporation, which is positively related to cell metabolic activity (HA, high-activity; MA, medium-activity; LA, low-activity; and P, nonliving particles and cell debris). Boldface letters are used to indicate FCM populations that were analyzed with T-RFLP. All three replicates of the DMSP HA populations failed to provide enough template for PCR amplification and therefore were excluded from T-RFLP and sequence analysis. Italic letters are used to indicate FCM populations that were used for the high-resolution (>200 clones) clone library construction.

Fig. 3.1

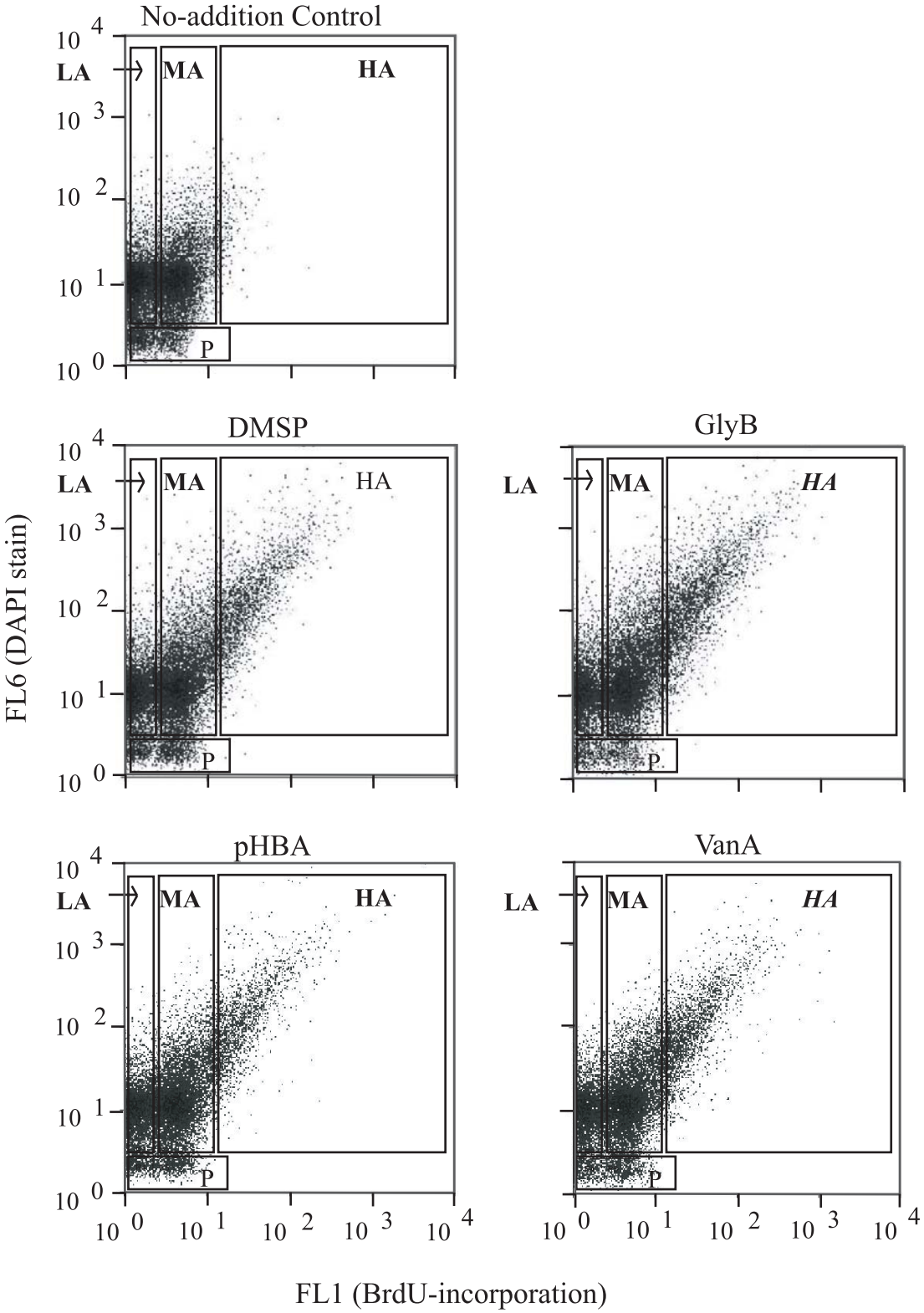

No-addition Control
DMSP
GlyB
pHBA
VanA
LA
MA
HA
P
FL6 (DAPI stain)
FL1 (BrdU-incorporation)

Figure 3.2. Total numbers of bacterial cells (A) and abundances of HA cells (B) (avg.± std., n=3) in the no-addition controls and model compound amendments over a 24-h incubatioin based on FCM analysis. Sample symbols in Figure 3.2A indicate treatments: solid square, DMSP; solid triangle, GlyB; open square, pHBA; open triangle, VanA; solid circle, no-addition control.

Fig. 3.2

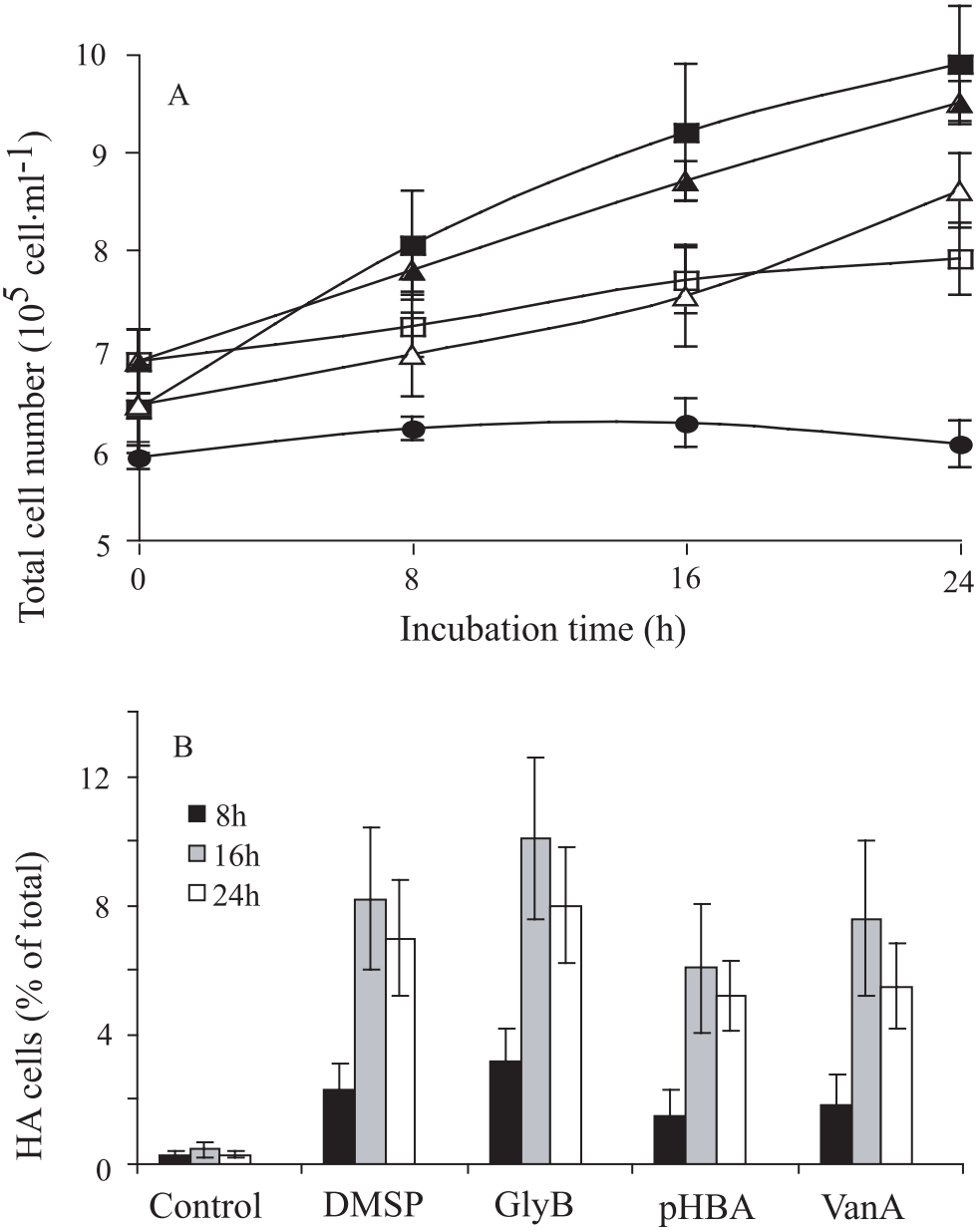

A
Total cell number (10^5 cell·ml^{-1})
10
9
8
7
6
5
0
8
16
24
Incubation time (h)
B
8h
16h
24h
HA cells (% of total)
12
8
4
0
Control
DMSP
GlyB
pHBA
VanA

Figure 3.3. MDS of T-RFLP data after 16 h of incubation. FCM populations from replicate incubations are shown, except for DMSP HA communities for which no 16S rRNA genes were successfully amplified. Sample symbols: circle, control; square, DMSP; star, GlyB; diamond, VanA; hexagon, pHBA. FCM population codes: white, total; light grey, LA populations; dark grey, MA populations; black, HA populations. Stress = 0.1

Fig. 3.3

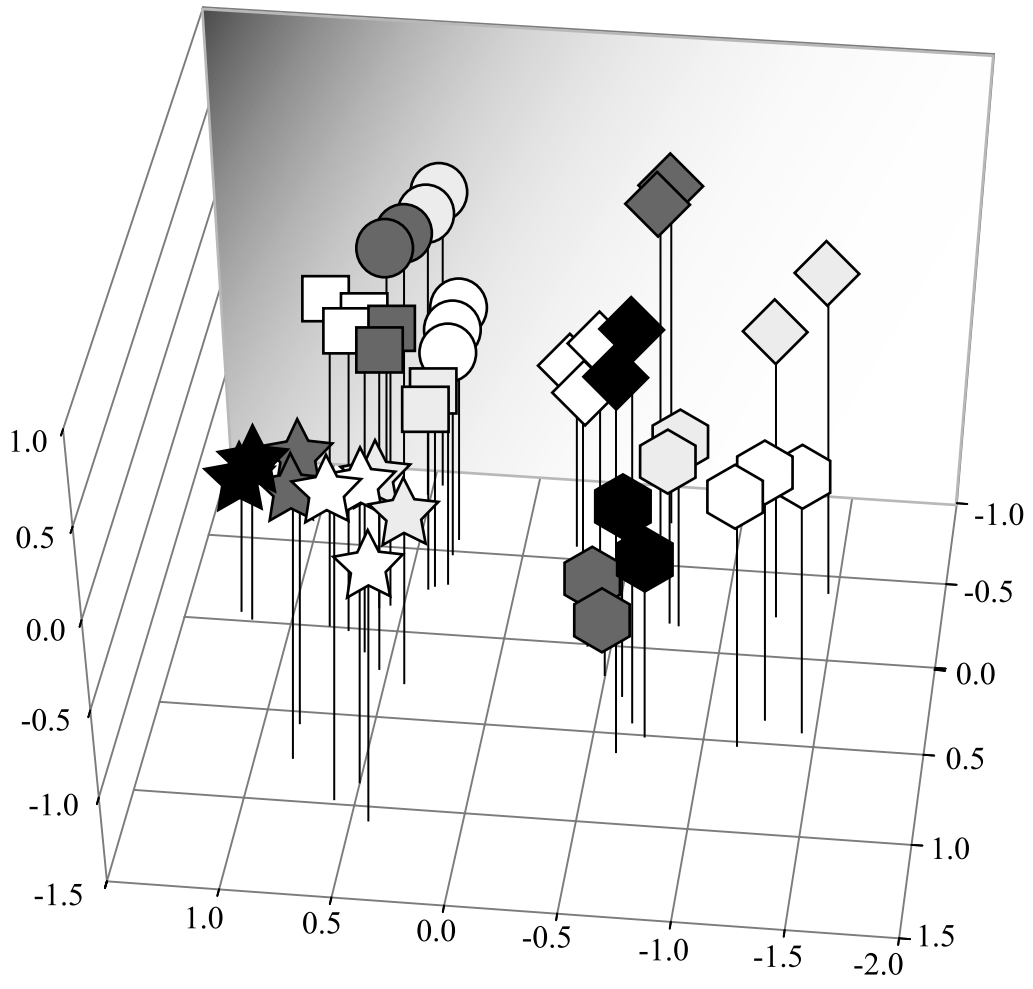


Figure 3.4. Comparison of bacterioplankton OTU distributions among major phylogenetic groups in the GlyB and VanA HA populations based on the high resolution 16S rDNA clone libraries.

Fig. 3.4

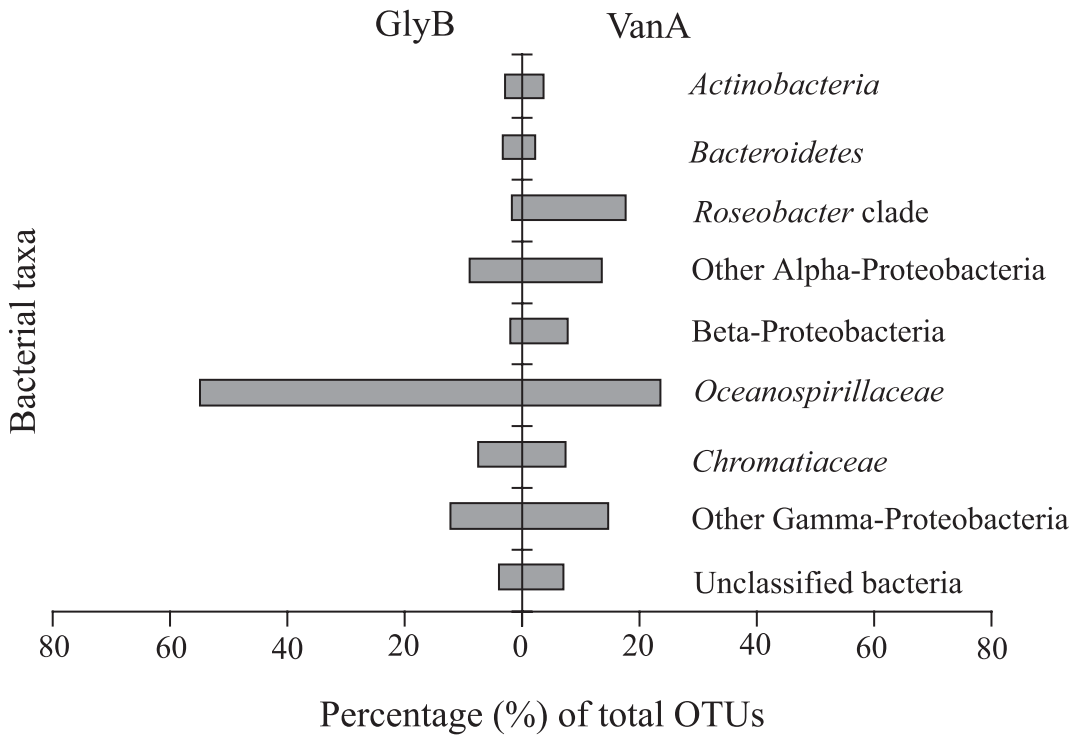

GlyB
VanA
Actinobacteria
Bacteroidetes
Roseobacter clade
Other Alpha-Proteobacteria
Beta-Proteobacteria
Oceanospirillaceae
Chromatiaceae
Other Gamma-Proteobacteria
Unclassified bacteria
Bacterial taxa
80
60
40
20
0
20
40
60
80
Percentage (%) of total OTUs

Figure 3.5. Phylogenetic tree of OTUs (0.03 distance) based on partial sequences of 16S rDNAs recovered from the GlyB and VanA HA libraries, showing γ-Proteobacteria (A), and other phylogenetic groups (B). Bootstrap values higher than 50% are indicated at the branch nodes. The scale bar indicates the amount of genetic change as the number of nucleotide substitutions per site. Solid circles denote OTUs from the GlyB HA library, while open circles denote OTUs from the VanA HA library. OTUs found in both HA libraries that are identical at the species level are highlighted by black arrows. The number of clones in each OTU is shown in brackets. The GenBank accession numbers of the reference sequences are shown in parentheses.

Fig. 3.5A

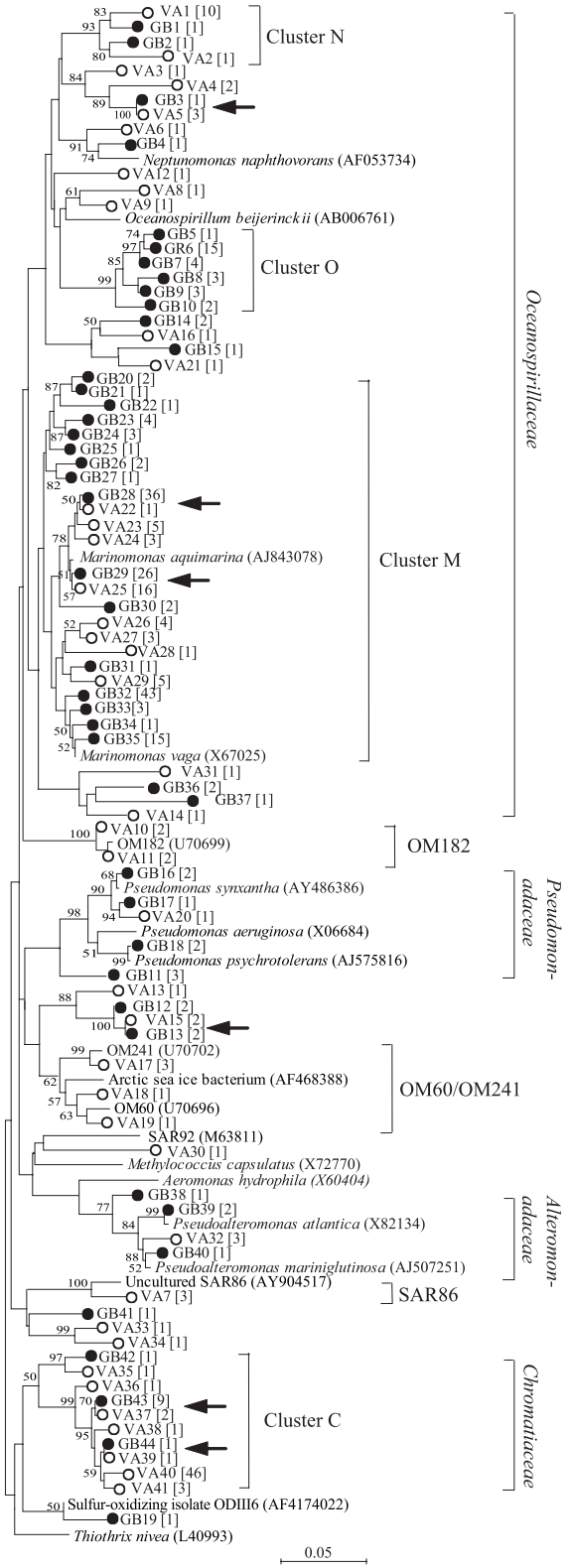

VA1 [10]
GB1 [1]
GB2 [1]
VA2 [1]
Cluster N
VA3 [1]
VA4 [2]
GB3 [1]
VA5 [3]
VA6 [1]
GB4 [1]
Neptunomonas naphthovorans (AF053734)
VA12 [1]
VA8 [1]
VA9 [1]
Oceanospirillum beijerinckii (AB006761)
GB5 [1]
GR6 [15]
GB7 [4]
GB8 [3]
GB9 [3]
GB10 [2]
Cluster O
GB14 [2]
VA16 [1]
GB15 [1]
VA21 [1]
GB20 [2]
GB21 [1]
GB22 [1]
GB23 [4]
GB24 [3]
GB25 [1]
GB26 [2]
GB27 [1]
GB28 [36]
VA22 [1]
VA23 [5]
VA24 [3]
Marinomonas aquimarina (AJ843078)
GB29 [26]
VA25 [16]
GB30 [2]
VA26 [4]
VA27 [3]
VA28 [1]
GB31 [1]
VA29 [5]
GB32 [43]
GB33[3]
GB34 [1]
GB35 [15]
Marinomonas vaga (X67025)
Cluster M
VA31 [1]
GB36 [2]
GB37 [1]
VA14 [1]
Oceanospirillaceae
VA10 [2]
OM182 (U70699)
VA11 [2]
OM182
GB16 [2]
Pseudomonas synxantha (AY486386)
GB17 [1]
VA20 [1]
Pseudomonas aeruginosa (X06684)
GB18 [2]
Pseudomonas psychrotolerans (AJ575816)
GB11 [3]
Pseudomon-adaceae
VA13 [1]
GB12 [2]
VA15 [2]
GB13 [2]
OM241 (U70702)
VA17 [3]
Arctic sea ice bacterium (AF468388)
VA18 [1]
OM60 (U70696)
VA19 [1]
OM60/OM241
SAR92 (M63811)
VA30 [1]
Methylococcus capsulatus (X72770)
Aeromonas hydrophila (X60404)
GB38 [1]
GB39 [2]
Pseudoalteromonas atlantica (X82134)
VA32 [3]
GB40 [1]
Pseudoalteromonas mariniglutinosa (AJ507251)
Alteromon-adaceae
Uncultured SAR86 (AY904517)
VA7 [3]
SAR86
GB41 [1]
VA33 [1]
VA34 [1]
GB42 [1]
VA35 [1]
VA36 [1]
GB43 [9]
VA37 [2]
VA38 [1]
GB44 [1]
VA39 [1]
VA40 [46]
VA41 [3]
Cluster C
Chromatiaceae
Sulfur-oxidizing isolate ODIII6 (AF4174022)
GB19 [1]
Thiothrix nivea (L40993)
0.05

Fig. 3.5B

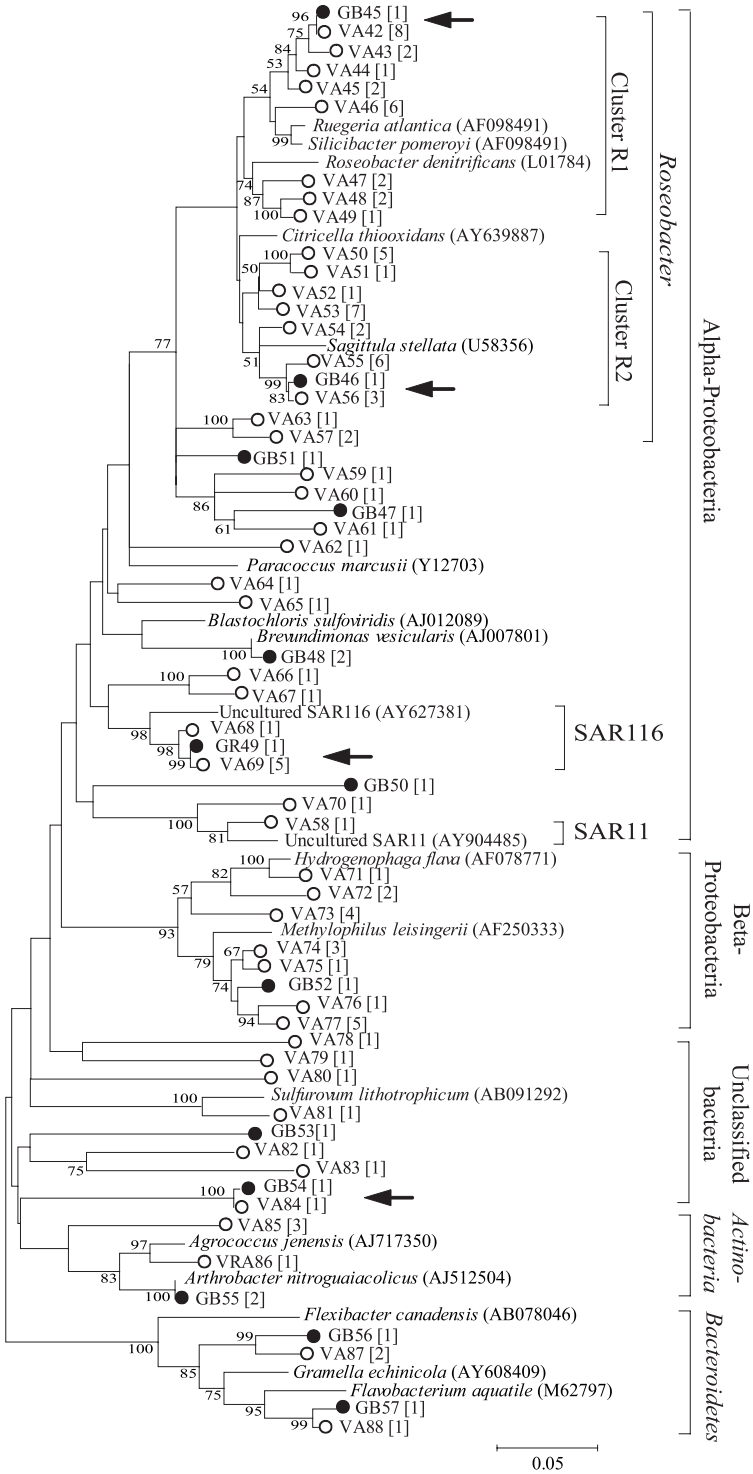

GB45 [1]
VA42 [8]
VA43 [2]
VA44 [1]
VA45 [2]
VA46 [6]
Ruegeria atlantica (AF098491)
Silicibacter pomeroyi (AF098491)
Roseobacter denitrificans (L01784)
VA47 [2]
VA48 [2]
VA49 [1]
Citricella thiooxidans (AY639887)
VA50 [5]
VA51 [1]
VA52 [1]
VA53 [7]
VA54 [2]
Sagittula stellata (U58356)
VA55 [6]
GB46 [1]
VA56 [3]
VA63 [1]
VA57 [2]
GB51 [1]
VA59 [1]
VA60 [1]
GB47 [1]
VA61 [1]
VA62 [1]
Paracoccus marcusii (Y12703)
VA64 [1]
VA65 [1]
Blastochloris sulfoviridis (AJ012089)
Brevundimonas vesicularis (AJ007801)
GB48 [2]
VA66 [1]
VA67 [1]
Uncultured SAR116 (AY627381)
VA68 [1]
GR49 [1]
VA69 [5]
GB50 [1]
VA70 [1]
VA58 [1]
Uncultured SAR11 (AY904485)
Hydrogenophaga flava (AF078771)
VA71 [1]
VA72 [2]
VA73 [4]
Methylophilus leisingerii (AF250333)
VA74 [3]
VA75 [1]
GB52 [1]
VA76 [1]
VA77 [5]
VA78 [1]
VA79 [1]
VA80 [1]
Sulfurovum lithotrophicum (AB091292)
VA81 [1]
GB53[1]
VA82 [1]
VA83 [1]
GB54 [1]
VA84 [1]
VA85 [3]
Agrococcus jenensis (AJ717350)
VRA86 [1]
Arthrobacter nitroguaiacolicus (AJ512504)
GB55 [2]
Flexibacter canadensis (AB078046)
GB56 [1]
VA87 [2]
Gramella echinicola (AY608409)
Flavobacterium aquatile (M62797)
GB57 [1]
VA88 [1]
0.05
Cluster R1
Cluster R2
Roseobacter
SAR116
SAR11
Alpha-Proteobacteria
Beta-Proteobacteria
Unclassified bacteria
Actino-bacteria
Bacteroidetes

Table S3.1. ANOSIM statistics for comparisons of bacterial populations between the no-addition control and model compound amendments based on T-RFLP data generated for samples taken after 16 h of incubation. The T-RFLP data from the same treatments were pooled to generate enough replicates for this analysis. Global R = 0.66, significance level of sample statistic is 0.1%. We interpreted R-values > 0.75 as statistically distinct; R > 0.5 as different but overlapping.

Comparison	Sample statistic R	Level of sample statistic (%)
Control vs. DMSP	0.618	0.1
Control vs. GlyB	0.688	0.1
Control vs. *p*HBA	0.583	0.2
Control vs. VanA	0.942	0.1
DMSP vs. GlyB	0.415	0.3
DMSP vs. *p*HBA	0.717	0.1
DMSP vs. VanA	0.913	0.1
GlyB vs. *p*HBA	0.835	0.1
GlyB vs. VanA	0.913	0.1
*p*HBA vs. VanA	0.335	0.2

Table S3.2. Description of clone libraries constructed for the GlyB and VanA HA populations

Measurements	GlyB HA library	VanA HA library
Sequences #	224	239
OTU #	55[a]	88
OTU singletons #	31	51
Good's Coverage (%)	86	80
Chao1 species richness (OTUs)	99 (75, 155)[b]	178 (137, 297)
Shannon's diversity index (H')	3.09 (2.90, 3.28) [b]	3.86 (3.68, 3.95)

a. Numbers of OTUs were 78 and 88 for the GlyB and VanA HA libraries, respectively, at a 0.01 distance level.

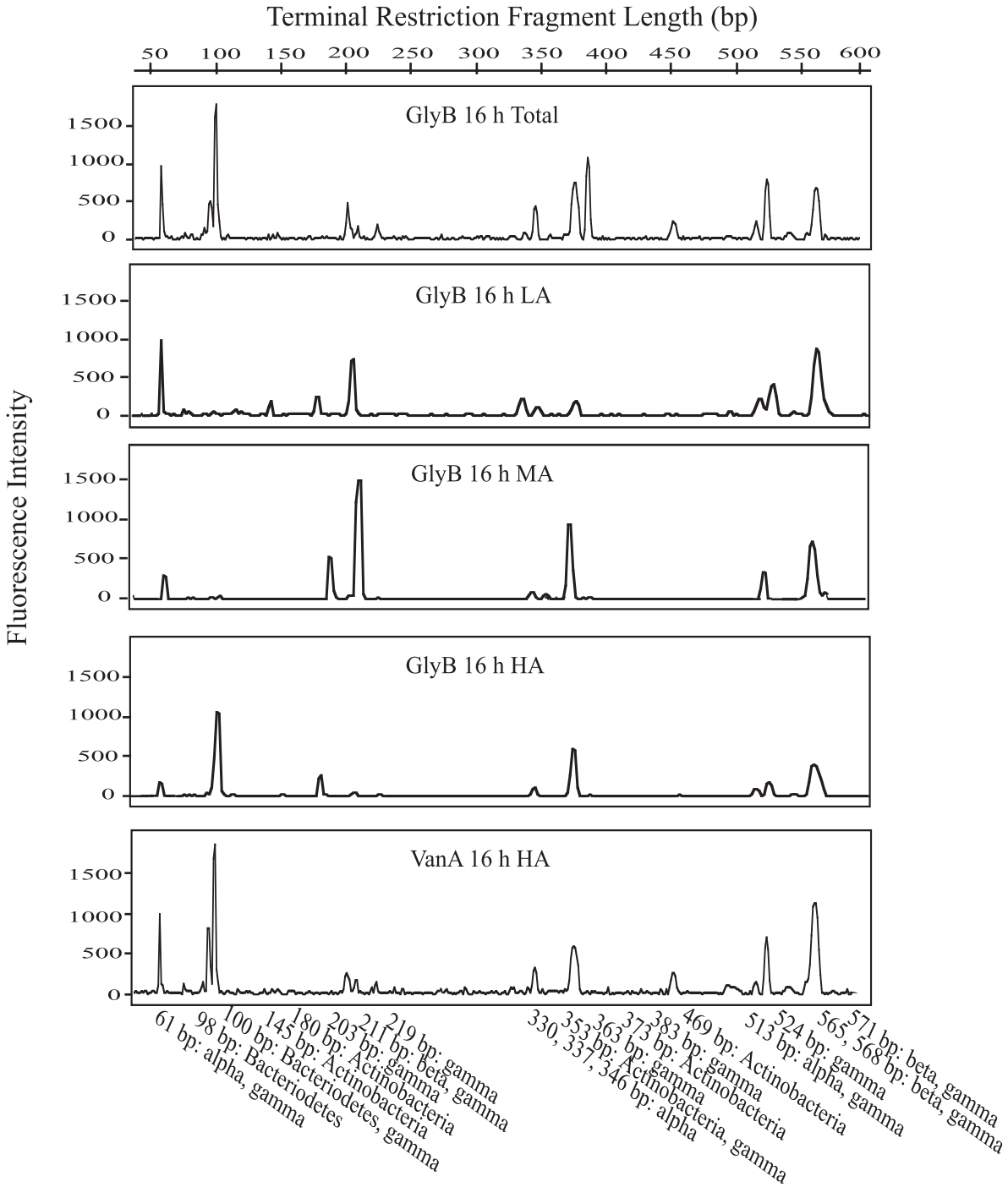
b. Numbers in parenthesis indicate the high and low values of 95% confidence intervals.

Table S3.3. OTU clusters of HA clone sequences containing more than 10 sequences with >90% similarity for the GlyB and VanA amendments. Significant enrichment in one of the model compound amendments was tested using a Chi-square analysis ($p < 0.05$) with expected values based on the ratio of GlyB: VanA in the total libraries.

Taxa	Clusters	Observed # of Seq		P_{χ^2}	Significant enrichment
		GlyB	VanA		
Oceanospirillaceae	Cluster N	2	11	0.1	No
Oceanospirillaceae	Cluster O	28	0	0.001	Yes
Oceanospirillaceae	Cluster M	142	38	0.001	Yes
Chromatiaceae	Cluster C	11	55	0.001	Yes
Roseobacter Clade	Cluster R1	1	24	0.001	Yes
Roseobacter Clade	Cluster R2	1	25	0.001	Yes

Figure S3.1: T-RFLP chromatograms of total community and all sorted FCM populations in one GlyB replicate, and the HA population in one VanA replicate. Putative taxonomic identities of T-RFs are indicated along the bottom axis. Alpha, α-Proteobacteria; beta, β-Proteobacteria; gamma, γ-Proteobacteria.

Fig. S3.1


Terminal Restriction Fragment Length (bp)
50 100 150 200 250 300 350 400 450 500 550 600
Fluorescence Intensity
GlyB 16 h Total
GlyB 16 h LA
GlyB 16 h MA
GlyB 16 h HA
VanA 16 h HA
1500
1000
500
0
61 bp: alpha, gamma
98 bp: Bacteriodetes
100 bp: Bacteriodetes, gamma
145 bp: Actinobacteria
180 bp: Actinobacteria
203 bp: gamma
217 bp: beta, gamma
219 bp: gamma
330, 337, 346 bp: alpha
353 bp: Actinobacteria, gamma
363 bp: gamma
373 bp: Actinobacteria
383 bp: gamma
469 bp: Actinobacteria
513 bp: alpha, gamma
524 bp: gamma
565, 568 bp: beta, gamma
571 bp: beta, gamma

CHAPTER 4

COMPARATIVE METAGENOMICS OF BACTERIOPLANKOTN ASSEMBLAGES TRANSFORMING DISSOLVED ORGANIC CARBON IN COASTAL SEAWATER[3]

[3] Mou, X., S. Sun, R. Edwards, R. E. Hodson, and M. A. Moran. To be submitted to *Environmental Microbiology*.

SUMMARY

Bacterioplankton play a central role in transforming dissolved organic carbon (DOC) in marine systems, yet how closely these genetic capabilities track environmental changes in an ecological time frame is poorly understood. Here we report metagenomic analysis of bacterioplankton enriched with 100 nM dimethylsulfoniopropionate (DMSP) or vanillic acid (VanA), two ecologically important and structurally distinct DOC compounds in coastal seawater. A total of 105,161 and 46,219 random sequence reads with an average length of 97 bp were obtained from the DMSP and VanA metagenomic libraries by pyrosequencing, respectively. Prior to sequence annotation, the reliability of gene prediction from short sequences was evaluated by *in silico* analysis of random fragments of known 16S rDNA and functional genes. Using the criteria generated by this analysis, the metagenomic sequences were annotated by BLASTN against the RDP II database for taxonomic identification and by BLASTX against the NCBI's RefSeq, COG and KEGG databases for functional gene assignments. The taxonomic markers in the metagenomic libraries were generally consistent with those revealed by PCR-based 16S rDNA clone libraries generated from the same sample templates. The enriched bacterioplankton in the DMSP and VanA amendments were primarily from *Altermomonadales*, *Oceanospirillales* (γ-Proteobacteria), *Burkhoderiales* (β-Proteobacteria) and the *Roseobacter* clade (α-Proteobacteria). No significant differences were identified in the protein repertoires of the DMSP and VanA enriched bacterioplankton metagenomes. Compared with the average content of previously sequenced marine microbial genomes, the enriched bacterioplankton metagenomes included more genes for the metabolism of carbohydrates, amino acids, glycans and lipids. The enriched metagenomes were also high in genes for the synthesis of the nucleotides, cofactors and vitamins. The similar taxonomic and functional structure of DMSP

and VanA enriched bacterioplankton indicates that bacterial generalists dominated the transformation of these common DOC components. Our results provide insights into strategies for bacterial responses to changing environmental conditions within an ecological time frame.

INTRODUCTION

Dimethylsulfoniopropionate (DMSP) and vanillic acid (VanA) are typical components of the dissolved organic carbon (DOC) pool in coastal seawater (concentrations can be as high as 200 nM) (Moran et al., 1991; González et al., 2000) and are particularly important in southeastern U.S. marsh-dominated ecosystems (Moran and Hodson, 1990; Otte et al., 2004). DMSP is an organic osmolyte synthesized by marine algae and coastal vascular plants to help organisms survive under high salt conditions. Its degradation product dimethylsulfide (DMS) is a sulfur gas with important implications for global climate change (Kiene et al., 2000). VanA is one of the major intermediates of biological degradation of lignin (Crawford and Olson, 1978), a highly refractory structural material produced by vascular plants and the second-most abundant organic compound in the biosphere (after cellulose).

In marine environments, dissolved DMSP and VanA are transformed by the bacterial community. A number of studies have identified the bacteria responding to an elevated supply of DMSP or VanA under *in situ* or manipulated conditions (Buchan et al., 2000; González et al., 2000; Ansede et al., 2001; Mou et al., 2005). However, taxonomy alone cannot provide insights into the pathways and regulations of degradation. To obtain information on these two aspects, new studies are needed to directly address the genes harbored by the DOC-degrading bacterioplankton

The study of mixed microbial genomes from natural environments (i.e. metagenomics or environmental genomics) can be an ideal approach to link biogeochemical processes with the underlying genetic capabilities of the bacterial community (Handelsman, 2004; Riesenfeld et al., 2004). By comparing unassembled shotgun sequence libraries, clear differences in functional gene inventories were shown across microbial communities from disparate environments (Tringe

et al., 2005) and at different vertical locations in the ocean (DeLong et al., 2006). For example, genes for cellulose degradation and antibiotic production were more enriched in a farm soil, genes for photosynthesis and organic osmolyte transporter systems were more enriched in surface seawater, and genes for type IV secretion systems and glycan degradation were more enriched in the deep sea whale falls (Tringe et al., 2005). These findings suggest that the relative abundance of functional genes in natural microbial communities is fine-tuned to represent those functions that are most valuable in that environment. However, the extent of temporal variability in community gene pools within a given habitat or, in other words, how closely the metagenome tracks the ever-changing environmental conditions, is poorly understood.

As powerful as metagenomics can be, however, its wide application is hindered by the high cost associated with DNA sequencing. Recently, a relatively faster and more affordable sequencing method has become available based on massively parallel pyrosequencing (Margulies et al., 2005). Different from the classic Sanger sequencing (di-deoxynucleotide based chain termination) method, pyrosequencing is based on dectection of pyrophosphate releases during synthesizing the complemenary stand along a single strand of DNA (Ronaghi et al., 1996). The tradeoff, however, is that the average sequence length obtained from this method is about 100 bp. While the short-read DNA sequences typically generated by pyrosequncing present no particular problem for pure-culture genetic analysis, they present a challenge for use in metagenomic studies where database coverage is generally low and sequences are often novel (Handelsman, 2004).

In this study, we tracked the short-term shift in a bacterioplankton metagenome in response to increased availability of single components of the DOC pool. Newly synthesized genes of bacterioplankton enriched with either DMSP or vanillic acid (100 nM, final

concentration) were extracted by immunocapture of bromodeoxyuridine (BrdU)-labeled DNA and sequenced by pyrosequencing. We hypothesized that the bacterioplankton responding to these different (although common) DOC components were metabolic specialists and thus the newly synthesized genomic DNA would contain genes that are highly substrate-specific. In this case, there would be significant difference in the metagenomic sequences synthesized in response to the DMSP and VanA additions. To make best use of the data, we systematically examined the reliability of phylogenetic and functional predictions from short sequences, similar to those generated by pryrosequncing, by *in silico* analysis of random fragments of known genes.

RESULTS

Cell growth in amended samples. Compared with the cell numbers in the initial community ($1.29 \pm 0.08 \times 10^6$ cells ml^{-1}), the cell number in the two replicate DMSP amendments increased by 36% ($1.84 \pm 0.14 \times 10^6$ cells ml^{-1}) and in the two VanA replicates by 19% ($1.6 \pm 0.08 \times 10^6$ cells ml^{-1}) after a 12-h incubation. The cell number did not change in the no-addition control ($1.35 \pm 0.09 \times 10^6$ cells ml^{-1}; t-test, $p < 0.05$). The observed increases therefore indicated that cells in the DMSP and VanA amendments were growing in response to the added substrates. The newly synthesized genomic DNA of bacterioplankton in amended samples was labeled with BrdU, a thymidine analog supplied exogenously during the incubation. The BrdU-labeled newly-synthesized DNA of these bacterioplankton was separated from the remainder of community DNA using an immunochemical approach then used as the template for directly genomic pyrosequencing and PCR-based analysis of 16S rDNA compoents of the samples.

Metagenomic data validation. Because of concern over the reliability of gene annotations from short-read sequences generated by pyrosequencing, we systematically investigated the gene

predictions of random short fragments from known 16S rDNA and functional genes. This *in silico* analysis was used to establish criteria for gene predictions from the pyrosequences by BLAST analysis. One hundred near full-length 16S rDNAs were selected from the Ribosomal Database Project (RDP) II database to cover major groups of marine bacteria. The genes were fragmented randomly into 9 lengths ranging from 20 bp to 500 bp and analyzed by BLASTN against the RDP II database. To mimic the scenario of annotating environmental sequences, self hits were excluded and the taxonomic identities of the best not-to-self (BNTS) hits were assigned to the corresponding query sequence fragments.

As the length of the 16S rDNA fragments decreased, the expected value (E value), which is a statistical measure of the probability that query and target sequences match by chance alone, increased (Figure 4.1A). For example, all query sequences of 500 bp matched a 16S rDNA sequence with an E value $\leq 10^{-50}$, whereas most query sequences of 20 bp matched a 16S rDNA sequence with an E value $\leq 10^{-2}$. Sequence fragments longer than 35 bp [roughly the shortest sequences obtained from 454 pyrosequencing (Figure 4.1B)] all had the BNTS hits of 16S rDNA with E value $\leq 10^{-5}$. This value was thus used as the cutoff E value for identifying 16S rDNA sequences within the metagenomic libraries.

Once 16S rDNA sequences were identified, the validity of taxonomic assignment was tested with the simulated random 16S rDNA fragments using the SIMO RDP Agent (http://www.simo.marsci.uga.edu/public_db/taxonomy.htm#rdpagent) and local Smith-Waterman alignment to reference sequences of marine bacteria (Moran et al., 2004). The two programs retrieve classification information from the near full-length type sequences in the RDP II database and a local marine reference sequence dataset, respectively, and perform pairwise Smith-Waterman alignments to assign taxonomic ranks based on similarity-weighted cutoff

limits (Lasher et al., submitted). The majority (>90%) of simulated 16S rDNA sequence fragments were assigned to the correct order when the query sequences were at least 65 bp long and were at least 80% similar to the corresponding BNTS hits from the database(Figure 4.2). Because the pyrosequences may contain partial sequences from two adjacent genes, the overlapping length instead of the sequence length was used as the cutoff limit. Based on these analyses, we set up two criteria for accepting a taxonomic assignment of pyrosequences at the order level: overlapping length ≥65 bp and similarity ≥80% of the best hit to a 16S rDNA database sequence. To further assign the 16S rDNA at a finer taxonomic resolution (i.e., to assign to typical marine environmental clusters), an additional cutoff limit was set at a best hit similarity ≥90% in local Smith-Waterman alignment to reference sequences of marine bacteria (Moran et al., 2004).

In a similar *in silico* prediction exercise for protein-encoding genes, 100 known functional genes were arbitrarily selected based on the best hits for the first 100 pyrosequences in the DMSP metagenomic library. BLASTX analysis of the 100 pyrosequences against the National Center for Biotechnology Information (NCBI)'s Clusters of Orthologous Groups (COG) database (Tatusov et al., 2001) with a cutoff E value <0.1 were used to generate the *in silico* gene pool. Random sequence fragments generated from protein sequences in this gene pool were analyzed by BLASTX against the NCBI's curated non-redundant reference sequence (RefSeq) (Pruitt et al., 2005) and COG databases to determine if the BNTS hit belonged to the same gene or COG group. Most BNTS hits belonged to the same role category as the query fragments' parent sequences. However, the shorter the sequence fragment, the greater the chance that no hit was returned with the default cutoff limits of these programs (E value <10; Figure 4.3A and 4.4). For sequence fragments longer than 65 bp and with an E value <0.01 and a similarity >40% to

their BNTS hits in the RefSeq database, over 70% were assigned to the correct functional roles (Figure 4.3). At these cutoff limits, the phylogenetic affiliation derived from the taxonomic bins of the BNTS hit protein homolog had over 90% accuracy at the phylum/subphylum level. A similar trend was found for the COG analysis, except that the cutoff limit for the E value was decreased to 0.1 (Figure 4.4). Based on these analyses, the cutoff criteria for protein prediction from pyrosequences using BLASTX analysis against the RefSeq database were established as E-value <0.01, similarity >40%, and overlapping length >65 bp to the corresponding best hit. The cutoff criteria for functional protein identification based on orthlogous groups using BLASTX analysis against the COG database were established as: E value <0.1; similarity >40%; and overlapping length > 65 bp to the corresponding best hit. The Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa and Goto, 2000) is close to the COG database in size (and smaller than the RefSeq database). Therefore, the COG cutoff criteria were applied to BLASTX analysis results for KEGG pathway prediction.

The simulated fragments of known genes were also analyzed by BLASTN against the NCBI's NR database (16S rDNA sequences and protein-encoding sequences) and RDP II database (protein-encoding sequences). None returned a cross-category hit with the program's default setting (E value <10). This indicates that 16S rDNA sequences and protein-encoding sequences were distinct enough from each other that the chance of cross category mis-prediction from short sequence fragments is quite low.

Taxonomic structure of enriched bacterioplankton: metagenomic analysis. A total of 54,848 and 50,313 pyrosequences were generated from the enriched bacterioplankton metagenomes in the two DMSP replicates and 12,446 and 33,773 pyrosequences from the two VanA replicates. These sequences had an average length of 97 bp (Table 4.1). Before gene prediction, sequences

from the replicates were pooled, resulting in two libraries containing approximately 10 Mbp (DMSP) and 4 Mbp (VanA) of total sequence data (Table 4.1). Using the cutoff criteria established as described above, the metagenomic sequences were queried against several public databases for 16S rDNA and functional gene predictions at different resolution levels.

Approximately 0.2% of the sequences were predicted to be 16S rDNA (176 and 59 sequences in the DMSP and VanA metagenomic libraries, respectively) by BLASTN analysis against the RDP II database (Table 4.1). This value is consistent with the expected fraction of 16S rDNA in marine genomes, assuming two rRNA operons and 2 Mb of DNA per bacterial genome (Klappenbach et al., 2001), and indicates that sequences were randomly captured into the metagenomic libraries. When the predicted 16S rDNA sequences were used to query the NCBI's NR database, none returned a hit from a protein sequence with the default cutoff limit of the program (E value <10).

Generally, the taxonomic structure of the two enriched bacterioplankton metagenomes was similar at the order level (Figure 4.5). Members of γ-Proteobacteria, primarily in the orders *Alteromonadales* and *Oceanospirillales*, comprised the largest fraction of both enriched metagenomes, accounting for 62% and 56% of 16S rDNA in the DMSP and VanA libraries, respectively. Members of the α- and β-Proteobacteria were also abundant, accounting for 17% and 10% of the 16S rDNA in the DMSP library and 7% and 22% in the VanA library. Members of the α-Proteobacteria were mainly affiliated with the *Roseobacter* clade, while members of the β-Proteobacteria were mainly affiliated with the *Burkholderiales*. Smaller numbers of predicted 16S rDNAs were affiliated with *Bacteroidetes* and *Cyanobacteria*, *Firmicutes*, δ- and ε-Proteobacteria, and *Verrucomicrobia* (Figure 4.5) in both metagenomic libraries. Besides these common coastal bacterioplankton taxa, a few 16S rDNA sequences were affiliated with marine

environmental clusters typical of oligotrophic oceans, including SAR116 (α- Proteobacteria) in the DMSP metagenomic library and SAR86 (γ-Proteobacteria) in the VanA metagenomic library. SAR432 (*Actinobacteria*), SAR11 (α- Proteobacteria) and OM60/241 (γ-Proteobacteria) were found in both metagenomic libraries. Each of these environmental taxa accounted for 1- 4% of detected 16S rDNA.

The remaining metagenomic sequences that were not identified as 16S rDNA were queried against the NCBI's RefSeq database using BLASTX analysis. The taxonomic bins of bacterial protein homologs returned in those queries reflected a similar taxonomic distribution pattern of the enriched DMSP and VanA metagenomes at the phylum/subphylum level, although resolution was much less than for 16S rDNA sequences. As found for the 16S rDNA analysis, the proteins in the DMSP and VanA libraries were dominated by sequences with best hits to γ-Proteobacteria followed by α- and β-Proteobacteria (Figure 4.6A). Sequences with best hits to archaeal, eukaryotic and viral proteins were minor members in both metagenomic libraries, accounting for approximately 0.5%, 3% and 2% of the communities, respectively (Figure 4.6B).

Taxonomic structure of enriched bacterioplankton: PCR-based 16S rDNA analysis. The taxonomic composition of the enriched bacterioplankton metagenomes was also determined by PCR-based 16S rDNA clone library analysis, and was compared to the composition of total community in the no-addition controls. Analysis of the 16S rDNA clone libraries using the Libshuff program (Singleton et al., 2001) revealed a generally similar taxonomic composition of the DMSP and VanA enriched metagenomes at evolutionary distance >0.17 (roughly the level of class or higher),, although significant difference was found at lower phylogenetic levels (Figure 4.7). The classification of the 16S rDNA clone sequences was consistent with this result, indicating similar distribution patterns among the major bacterial groups as was found for the

metagenomic 16S rDNA, including γ-Proteobacteria, α-Proteobacteria, β-Proteobacteria, *Bacteroidetes* and *Actinobacteria* (Figure 4.6B). Differences between the PCR-based clone libraries and the metagenomic libraries were mainly found in the relative abundances of *Bacteroidetes* and α-Proteobacteria datasets. For the DMSP amendment, the PCR-based clone libraries had more members of *Bacteroidetes* (19%) than did the metagenomic library (2%), particularly for sequences affiliated with *Sphingobacteriales*. For the VanA amendment, the PCR-based clone libraries had more members of the α-Proteobacteria in the *Roseobacter* clade (22%) than did the metagenomic library (2%). A greater number of members of environmental clusters, including SAR11, SAR116 (α-Proteobacteria), SAR86 (γ-Proteobacteria) and SAR432 (*Actinobacteria*) were represented in metagenomic libraries. This cannot be simply explained by the difference in library size, given that the VanA metagenomic library contained a smaller number of 16S rDNA than the clone library (59 vs. 81 sequences). Difference in taxon diversity and richness between PCR-based 16S rDNA clone libraries and metagenomic libraries has been found previously (Cottrell et al., 2005). It is currently thought that PCR-based 16S rDNA clone library analysis is less representative of the actual community because of bias in the PCR and cloning steps (Edwards et al., 2006).

Metabolic capabilities of enriched bacterioplankton. To predict the functional roles of the enriched bacterioplankton, metagenomic sequences were compared against the NCBI's RefSeq, COG and KEGG databases separately using BLASTX analysis based on the cutoff criteria established by the *in silico* fragmentation analysis of known genes.

A total of 24,896 and 10,911 proteins were predicted by BLASTX analysis against the RefSeq database in the DMSP and VanA metagenomic libraries, respectively. These predictions accounted for ~24% of the sequences in each metagenomic library (Table 4.1). The remaining

sequences were not similar enough to entries in the RefSeq database to be definitively identified as a protein. This may be because the fragment contains short partial sequences of two adjacent genes, or the fragment is targeting a novel gene, or the fragment is from a poorly conserved region or an intergenic region of a known gene.

The resulting protein repertoires were similar between the DMSP and VanA metagenomic libraries and both contained a wide range of functional genes. A large number of transporters were predicted in both metagenomic libraries for a diversity of substrates, including carbohydrates (e.g. ABC-type sugar transporters and maltosaccharide transporters), protein derivatives (e.g. ABC-type peptide transporters and branched amino acid transporters), organic osmolytes (glycine betaine transporters), inorganic compounds (ABC-type nitrate/sulfonate/bicarbonate transporters and phosphate transporters), organic acids (TRAP C4-dicarboxylate transport systems), and secondary metabolites (TRAP-type mannitol/chloroaromatic compound transport systems). These indicate that the bacterioplankton in coastal seawater responding to DOC amendments can fulfill their carbon demands through a wide variety of heterotrophic processes.

Different types of aerobic carbon monoxide dehydrogenases (*coxCFGILS*; 31 copies in DMSP and 10 copies in VanA libraries) and sulfur oxidation proteins (*soxABXY*; 6 copies in DMSP and 4 copies in VanA libraries) were identified in both libraries, consistent with the presence of these mechanisms for lithotrophic energy generation in the genome of a coastal bacterium, *Silicibacter pomeroyi* (Moran et al., 2004). A recently-identified gene that is specifically involved in DMSP transformation by demethylating DMSP to methylmecaptopropionate (*dmdA*; Howard et al., submitted) appeared with comparable representation in DMSP (6 copies) and VanA (4 copies) libraries. Genes specifically involved in

VanA transformation such as *vanAB* (which converts vanillic acid into protocatechuate) and *pcaH* (which converts protocatechuate into succinate and acetyl-CoA via the β-ketoadipate pathway) likewise appeared in both the DMSP and VanA libraries with comparable representation (4 and 3 copies, repectively).

A total of 9,900 and 6,150 proteins were predicted by BLASTX analysis against the COG database in the DMSP and VanA metagenomic libraries, respectively. These proteins fell in 1,769 and 1,461 COG groups (Table 4.1). The sequences classified based on the COG groups were subsets of those annotated by analysis against the RefSeq database (i.e., only ~50% of sequences classified by the RefSeq analysis were also classified by the COG analysis; Table 4.1). The distribution of COG categories was similar between the metagenomes of DMSP and VanA enriched bacterioplankton (Figure 4.10). Protein orthologs involved in DNA replication appeared overrepresented in both DMSP and VanA enriched metagenomes, compared with the average content of four previously sequenced marine microbial genomes representing two major taxa, α- and γ-Proteobacteria in coastal marine systems (Figure 4.9 and Table S4.1). These COG groups included F0F1-type ATP synthase, Methionine, superfamily II DNA and RNA helicases, pyrovate dehyrogenase and ribonucleotide reductase. Such an overrepresentation of these essential genes may reflect a smaller average genome size in the enriched bacterioplankton than in the four cultured marine bacteria used for comparison. Other COG groups with considerably higher representation in the enriched metagenomes included those for amino acid and protein synthesis (methionine synthase I, glutamate synthase, leucyl-tRNA synthetase, and translation elongation factors), and energy production and conversion (pyruvate/2-oxoglutarate dehydrogenase complex and F0F1-type ATP synthase) (Figure 4.9).

A total of 9,339 and 7,003 proteins were predicted by BLASTX analysis against the KEGG database in the DMSP and VanA metagenomic libraries, respectively. These proteins fell in 148 and 125 KEGG pathways (Table 4.1). Consistent with the COG analysis, the KEGG pathways also suggested a similar distribution of genes between the DMSP and VanA enriched metagenomes (Figure 4.8). KEGG pathways for metabolizing carbohydrates (pyruvate metabolism and the citrate cycle), amino acids (phenylanlanine, tyrosine and tryptophan biosynthesis), lipids (fatty acid metabolism) and nucleotides (purine and pyrimidine metabolism) had higher representation in the DMSP and VanA enriched metagenomes than in the average marine bacterial genome, as did proteins involved in energy metabolism in the cell (oxidative phosphorylation and reduced carboxylate cycle), and cofactor and vitamin biosythesis (ubiquinone and folate biosynthesis) (Figure 4.10 and Table S4.2).

Statistical comparison of the protein repertoires between the metagenomes of the DMSP and VanA enriched bacterioplankton were performed using a bootstrap resampling method (Rodriguez-Brito et al., 2006). Only 13 COG groups and 4 KEGG pathways were found to be statistically overrepresented in one sample versus the other at a confidence level of 95% ($p < 0.05$; Table 4.2 and 4.3), and the numbers were zero when the confidence level was increased to 98% ($p < 0.02$). Given that 2,011 COG groups and 124 KEGG pathways were compared, some of the identified differences at $p < 0.05$ are likely to be by chance alone. In two relevant studies using the same statistical approach, one which compared metagenomic libraries from an aerobic and an anaerobic iron mine (Edwards et al., 2006), and one which compared metagenomic libraries from various depths at a single site in the North Pacific Ocean (DeLong et al., 2006), many more differences were found. Using the SEED subsystem, a functional category roughly equivalent to a KEGG pathway, about 60 (Edwards et al., 2006) and 100 (DeLong et al., 2006) subsystems

were significantly overrepresented in one habitat relative to the other. The small numbers of significant COG and KEGG categories found in our study indicates that the metagenomes of the DMSP and VanA enriched bacterioplankton had quite similar protein repertoires.

DISCUSSION

Recent metagenomic studies have supported the hypothesis that the gene distributions within microbial assemblages reflect the conditions unique to each environment (Tringe et al., 2005; DeLong et al., 2006; Edwards et al., 2006). However, how would the community gene pool respond to temporary environmental shifts, such as seasonal or event-driven changes in the supply of specific classes of DOC? One possibility is that evolution has packaged important functional genes independently into bacterial cells, such that each gene can be selected independently of most other genes encoding other key environmental functions. In this case, the community's functional genes would be harbored within specialist taxa, and changes in abundance of specialists in response to environmental shifts would therefore allow the community to respond dynamically and specifically. An alternative possibility is that evolution has packaged many functional genes together in bacterial cells, linking the genes that are co-beneficial in a specific environment. In this case, the community's functional genes co-occur in generalist bacteria and changes in abundance of generalists in response to environmental shifts would therefore result in parallel increases of all physically linked genes, regardless of whether additional copies of all these genes are helpful in processing the substrate. Results of our study support the second scenario. After a short-term incubation, the recently synthesized genomic DNA (16S rDNA and functional genes) by coastal bacterioplankton in response to structurally different types of DOC, i.e. DMSP and VanA, were quite similar.

Our results, however, do not rule out the possibility that specialists for DMSP or VanA use were present in the bacterial community, but suggest that they were minor members at the time of sampling. The relatively short incubation time (e.g. 12 h in this study) is not likely to be sufficient to allow growth of rare members of the community to the point where their genome signal would be detected. In addition, the tested compounds, DMSP and VanA, are consistently present in coastal seawater (Moran and Hodson, 1990; Kiene et al. 2000), and thus it is not surprising that successful taxa may be capable of utilizing both compounds.

Our studies have specifically emphasized only a small component of the multidimensional niches of bacteria in this complex coastal environment. Thus, we might expect different results if the experiments were run over a longer term (days, months or seasons), or run with different substrates, or run with the same substrates but at lower concentrations.

Some of the bacterial genes captured in the enriched metagenomes may have been responding to metabolites of the added compound produced by other bacteria, rather than to the compound itself. However, this would not affect our general conclusions because high numbers of bacterial specialists involved in secondary utilization of added DOC compounds would nonetheless be evident in the enriched metagenomes and result in statistical differences between the metagenomic libraries.

Given the above caveats, our study provides insights into how bacterioplankton communities cope with short term changes in environmental conditions in a particular coastal ecosystem, and has implications for understanding the role of evolution in functional gene packaging among bacterial taxa. Together with the knowledge of the distribution pattern of bacterial taxa and functional inventories over temporally and spatially broader scales (Tringe et

al., 2005; DeLong et al., 2006; Edwards et al., 2006), our study allows a more thorough view of the interaction between environments and the microorganisms that inhabit them.

Two of the recognized obstacles to metagenomic studies are the high cost associated with Sanger sequencing and, as a consequent the low coverage of most metagenomic libraries (Handelsman, 2004). In this study we enriched the metagenome to focus our sequencing efforts on the bacterioplankton genomes of just those responding to a specific manipulation, in this case, an increase in the availability of a defined component of the DOC pool. This was done by using an approach involved BrdU-labeling and immunochemical purification of newly synthesized DNA from the bulk bacterial community DNA. Sequencing the newly synthesized DNA by pyrosequencing provided advantages in terms of cost and time. The tradeoff, however, was a short sequence length for the metagenomic library (average of 97 bp in this study). Nonetheless, for pyrosequences with longer than 65 bp overlap with known genes, there is sufficient sequence data for reliable predictions of taxonomy of 16S rDNAs and roles of functional genes. Pyrosequencing can therefore serve as an alternative sequencing method for the promising technology of metagenomics.

EXPERIMENTAL PROCEDURES

Sample collection, procession and BrdU labeling. Surface water was collected in acid-washed Nalgene carboys in Nov 2005 from Dean Creek (a salt marsh tidal creek on Sapelo Island, GA, which serves as one of the sampling sites for Sapelo Island Microbial Observatory, SIMO). Afterwards, water was immediately filtered through 293-mm-diameter, 3.0-μm-pore size polycarbonate filters (Poretics Products) to exclude most eukaryotes and big particles. Six microcosms were established in polycarbonate carboys containing 20 L of filtrate and 10 μM

(final concentration) of BrdU (Roche Applied Science). DMSP or VanA (100 nM, final concentration) was added to two of the microcosms, and the remaining two microcosms served as no-addition controls. Total time from water collection to beginning the incubations was < 2 h. The bacterial communities in the initial sample and the microcosms after a 12-h incubation (in the dark at room temperature with occasional agitations) were collected on 0.2-μm-pore size polycarbonate filters (Poretics Products) by filtration. Subsamples (10 ml) were taken from each microcosm at the beginning and the end of the incubation and preserved in 4% paraformaldehyde (final concentration) for 2 h at room temperature and stored at 4°C before counted by epifluorescence microscopy. The cell counts in treatments after incubation were compared to those in the initial sample, and significant differences were reported at $p < 0.05$ in the t-test.

Community DNA extraction and BrdU-labeled DNA purification. Filters with collected bacterioplankton (from approximately 20 L of water) were cut into pieces and processed using a PowerMax soil mega prep DNA isolation kit (MoBio) to obtain community genomic DNA. BrdU-labeled DNA fraction was immunochemically purified using a method modified from a previous study (Urbach et al., 1999). Briefly, herring sperm DNA [0.63 mg ml^{-1} in phosphate buffered saline (PBS); Promega] was heated at 98 °C for 5 min then immediately put in ethanol-dry ice (5 min) before being mixed with monoclonal anti-BrdU antibodies (Zymed Laboratories) at a 9:1 ratio. The denatured mixture was then incubated for 45 min at room temperature. A 25 μl of extracted community genomic DNA was supplemented with 10 μl of PBS and denatured as described above then mixed with 30 μl of the pre-incubated herring sperm DNA-antibody mixture and incubated for another 1 h in the dark at room temperature with constant agitation. Dynabeads coated with goat anti-mouse immunoglobulin G (Dynal Biotech ASA) were washed 5 times with 1 mg ml^{-1} acetylated BSA in PBS buffer (PBS-BSA) using a magnetic particle

concentrator (Dynal) and resuspended in PBS-BSA to the initial concentration. The 10 μl of washed dynabeads were mixed with 65 μl of the genomic DNA-herring sperm DNA-antibody mixture and incubated for additional 1 h. Beads were subsequently washed three times with 0.5 ml of PBS-BSA. To elute the BrdU-labeled DNA fraction, 100 μl of 1.7 mM BrdU (in PBS-BSA) were added to the washed dynabeads and incubated for 1 h in the dark at room temperature with constant agitation. The BrdU-labeled DNA solution was finally separated from the beads with a magnetic particle concentrator and purified by ethanol precipitation.

Validation of gene prediction from pyrosequences. The nucleotide sequences of 100 known near full length (> 1200 nt) bacterial 16S rRNA genes that covered major phyla and classes of marine bacteria were arbitrarily selected and downloaded from the RDP II database. The sequences included 2 members from the *Acidobacteria*, 5 from the *Bacteroidetes*, 34 from the α-Proteobacteria, 19 from the β-Proteobacteria, 13 from the γ-proteobacteria, 15from the δ-Proteobacteria and 12 from the ε-Proteobacteria. Nine sequence fragments with lengths gradually decreasing from 500 to 20 bp (500, 300, 200, 100, 80, 65, 50, 35, 20 bp) were randomly generated from each 16S rDNA sequence and used in BLASTN analysis against the RDP II database to determine the cutoff limits for reliable prediction 16S rDNAs at the domain level. The taxonomy of predicted 16S rDNAs at more detailed phylogenetic levels were then determined using the SIMO RDP Agent and local Smith-Waterman similarity comparisons to marine reference sequences. Compared to the non-dynamic BLAST algorithm, the Smith-Waterman approximation was slower but more sensitive (Brenner et al., 1998). The taxonomy of the NTSB hit was assigned to each query sequence fragment and the accuracy of the assignment was evaluated at all the phylogenetic levels except for the species and strain levels to establish suitable cutoff limits for reliable taxonomic prediction.

Likewise, the random sequence fragments of 100 randomly picked functional genes were analyzed by BLASTX against the NCBI's RefSeq and COG databases. The reliability of the functional predictions based on the NTSB hits was then evaluated to determine the suitable cutoff limits for reliable functional prediction using BLAST analysis.

Genomic DNA pyrosequencing and analysis. BrdU-labeled genomic DNA from duplicated DMSP and VanA amendments were sequenced by 454 Life Sciences using pyrosequencing technology. Each sample contained approximately 3 μg DNA. Prior to analysis, potential sequence artifacts were removed from the sequence libraries and the duplicated sequence libraries were pooled to obtain a bigger dataset.

The unassembled sequences were analyzed by BLASTN against the RDP II database to screen for 16S rDNA. The identified 16S rDNAs were further assigned to the order level and to major marine groups using SIMO RDP Agent and local Smith-Waterman analysis with reference marine sequences. The functions of the sequences were predicted with analysis by BLASTX against the NCBI's RefSeq, COG and KEEG databases separately. Functional identities of the best hits were assigned to the corresponding query sequences if the established cutoff criteria for pyrosequences were met. Otherwise, the sequences were not considered further.

To determine if the DOC additions led to enrichment of specific gene categories, the relative abundance of each COG and KEGG category in the DMSP and VanA enriched bacterioplankton genomes was calculated and compared with the average gene content of four previously sequenced marine bacterial genomes representing two major taxa, α- and γ-Proteobacteria, in the DOC enriched marine genomes [*Silicibacter pomeroyi* DSS3, *Pelagibacter ubique* HTCC1062 (α-Proteobacteria), *Idiomarina ioihiensis*, and *Colwellia psychreythraea* (γ-Proteobacteria)].

To determine if the addition of DMSP and VanA led to the divergence of gene repertories in the enriched bacterial communities, a bootstrapping analysis was performed using a program developed by Edwards et al. (Edwards et al., 2006). The program was run with a sample size of 7,000 protein orthologous groups, and 20,000 repeated samplings were carried out. The significance level was set at $p < 0.02$ and $p < 0.05$. The bootstrapping analysis was repeated at pathway level (KEGG) using a sample size of 9,000. Typically, bootstrapping sample sizes should match that of the actual sample.

16S rDNA clone libraries. Bacterial 16S rDNA of total community and BrdU-lableled cells after 12 h incubation were PCR amplified using FailSafe PCR Premix selection kit (Epicentre) with 0.4 μM concentrations of 27 forward and 1492 reverse primers (DeLong et al., 1989). A touchdown PCR program was performed with the annealing temperature sequentially decreasing from 62 to 52°C by 1°C per cycle, followed by 15 cycles at 52°C. In each cycle, denaturing (at 95°C), annealing (at 62 to 52°C), and extension (at 72°C) steps were of 40-s duration. An initial 4 m denaturation hot start and final 7 m extension step were also included in the PCR program. PCR amplification was confirmed by electrophoresis on ethidium bromide-stained 1% agarose gels and cleaned with the QIAquick gel extraction kit (Qiagen). The cleaned 16S rDNA amplicons were cloned using a TA cloning kit (Invitrogen) with pCR 2.1 vector following manufacturer's instructions. Cloned 16SrDNA were sequenced in a 96-well format on an ABI Prism 3100 genetic analyzer (Applied Biosystems) in sequencing mode using the 27F primer.

Sequences were assembled using Sequencher 4.1 (Gene Codes Corporation) and were checked for chimeras and potential vector contamination using the CHECK_CHIMERA program of the RDP II (Cole et al., 2003) and vector search of EMBL-EBI (http://www.ebi.ac.uk/blastall/vectors.html), respectively. Bacterial taxonomic identities were

automatically assigned using the SIMO RDP Agent and Smith-Waterman alignment with marine reference sequences. The clone libraries of the two enriched metagenomes were quantitatively compared using the Libshuff program (Singleton et al., 2001) based on the distance matrix calculated by the Kimura's two-parameter method in the MEGA 3.1 package (Kumar et al., 2004)..

Nucleotide sequence accession numbers. The GenBank accession numbers for 16S rRNA gene sequences determined in this study are DQ880941-DQ881441.

ACKNOWLEDGMENTS

We thank W. Sheldon and J. González for assistance with bioinformatics and E. Biers for assistance with sampling. This study was supported by NSF (grant 0084164 to the Sapelo Island Microbial Observatory) and the Gordon and Betty Moore Foundation.

REFERENCES

Ansede, J.H., Friedman, R., and Yoch, D.C. (2001) Phylogenetic analysis of culturable dimethyl sulfide-producing bacteria from a *Spartina*-dominated salt marsh and estuarine water. *Appl Environ Microbiol* **67**: 1210-1217.

Brenner, S.E., Chothia, C., and Hubbard, T.J. (1998) Assessing sequence comparison methods with reliable structurally identified distant evolutionary relationships. *Proc Natl Acad Sci* **95**: 6073-6078.

Buchan, A., Collier, L.S., Neidle, E.L., and Moran, M.A. (2000) Key aromatic-ring-cleaving enzyme, protocatechuate 3,4- dioxygenase, in the ecologically important marine *Roseobacter* lineage. *Appl Environ Microbiol* **66**: 4662-4672.

Cole, J.R., Chai, B., Marsh, T.L., Farris, R.J., Wang, Q., Kulam, S.A. et al. (2003) The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucl Acids Res.* **31**: 442-443.

Cottrell, M.T., Waidner, L.A., Yu, L.Y., and Kirchman, D.L. (2005) Bacterial diversity of metagenomic and PCR libraries from the Delaware River. *Environ Microbiol* **7**: 1883-1895.

Crawford, R.L., and Olson, P.P. (1978) Microbial catabolism of vanillate: decarboxylation to guaiacol. *Appl Environ Microbiol* **36**: 539-543.

DeLong, E.F., Wickham, G.S., and Pace, N.R. (1989) Phylogenetic Stains - Ribosomal Rna-Based Probes for the Identification of Single Cells. *Science* **243**: 1360-1363.

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

Edwards, R.A., Rodriguez-Brito, B., Wegly, L., Haynes, M., Breitbart, M., Peterson, D.M. et al. (2006) Using pyrosequencing to shed light on deep mine microbial ecology under extreme hydrogenologic conditions. *BMC Genomics* **7**: 57.

González, J.M., Simó, R., Massana, R., Covert, J.S., Casamayor, E.O., Pedros-Alio, C., and Moran, M.A. (2000) Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. *Appl Environ Microbiol* **66**: 4237-4246.

Handelsman, J. (2004) Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* **68:** 669-685.

Howard, E., Henriken, J. R., Buchan, A., Reisch, C. R., Bürgmann, H., Welsh, R. et al. Bacterial taxa limiting sulfur flux from the ocean. *Submitted*.

Kanehisa, M., and Goto, S. (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucl Acids Res.* **28**: 27-30.

Kiene, R.P., Linn, L.J., and Bruton, J.A. (2000) New and important roles for DMSP in marine microbial communities. *J Sea Res* **43**: 209-224.

Klappenbach, J.A., Saxman, P.R., Cole, J.R., and Schmidt, T.M. (2001) Rrndb: the ribosomal RNA operon copy number database. *Nucl Acids Res* **29**: 181-184.

Kumar, S., Tamura, K., and Nei, M. (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* 5: 150-163.

Lasher, C., Dyszynski, G., Everett, K., Edmonds, J., Ye, W., Seldon, W. et. al. (2006) The diverse bacterial community in intertidal, anaerobic sediments at Sapelo Island, Georgia. *Submitted.*

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A. et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376-380.

Moran, M.A., and Hodson, R.E. (1990) Contributions of degrading *Spartina alterniflora* lignocellulose to the dissolved organic carbon pool of a salt marsh. *Mar Ecol-Prog Ser* **62**: 161-168.

Moran, M.A., Pomeroy, L.R., Sheppard, E.S., Atkinson, L.P., and Hodson, R.E. (1991) Distribution of terrestrially derived dissolved organic matter on the southeastern United States Continental Shelf. *Limnol Oceanogr* **36**: 1134-1149.

Moran, M.A., Buchan, A., González, J.M., Heidelberg, J.F., Whitman, W.B., Kiene, R.P. et al. (2004) Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature* **432**: 910-913.

Mou, X., Moran, M.A., Stepanauskas, R., González, J.M., and Hodson, R.E. (2005) Flow-cytometric cell sorting and subsequent molecular analyses for culture-independent identification of bacterioplankton involved in dimethylsulfoniopropionate transformations. *Appl Environ Microbiol* **71**: 1405-1416.

Otte, M.L., Wilson, G., Morris, J.T., and Moran, B.M. (2004) Dimethylsulphoniopropionate (DMSP) and related compounds in higher plants. *J Exp Bot* **55**: 1919-1925.

Pruitt, K.D., Tatusova, T., and Maglott, D.R. (2005) NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucl Acids Res* **33**: 501-504.

Riesenfeld, C.S., Schloss, P.D., and Handelsman, J. (2004) Metagenomics: Genomic analysis of microbial communities. *Annu Rev Genet* **38**: 525-552.

Rodriguez-Brito, B., Rohwer, F., and Edwards, R. (2006) An application of statistics to comparative metagenomics. *BMC Bioinformatics* **7**: 162.

Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M., and Nyren, P. (1996) Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem* **242**: 84-89.

Singleton, D.R., Furlong, M.A., Rathbun, S.L., and Whitman, W.B. (2001) Quantitative comparisons of 16S rRNA gene sequence libraries from environmental samples. *Appl Environ Microbiol* **67**: 4374-4376.

Tatusov, R.L., Natale, D.A., Garkavtsev, I.V., Tatusova, T.A., Shankavaram, U.T., Rao, B.S. et al. (2001) The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucl Acids Res* **29**: 22-28.

Tringe, S.G., von Mering, C., Kobayashi, A., Salamov, A.A., Chen, K., Chang, H.W. et al. (2005) Comparative metagenomics of microbial communities. *Science* **308**: 554-557.

Urbach, E., Vergin, K.L., and Giovannoni, S.J. (1999) Immunochemical detection and isolation of DNA from metabolically active bacteria. *Appl Environ Microbiol* **65**: 1207-1213.

Zubkov, M.V., Fuchs, B.M., Archer, S.D., Kiene, R.P., Amann, R., and Burkill, P.H. (2001) Linking the composition of bacterioplankton to rapid turnover of dissolved dimethylsulphoniopropionate in an algal bloom in the North Sea. *Environ Microbiol* **3**: 304-311.

Table 4.1. Summary of metagenomic libraries for the enriched bacterioplankton in the DMSP and VanA amendments after 12 h incubation. The data are based on the combined metagenomic sequences of the two replicates for each amendment.

Measurements	DMSP Amendments	VanA Amendments
Size of library (bp)	10,091,785	4,460,308
Number of sequences	105,161	46,219
Average sequence length (bp)	96.6 ± 16.8	96.5 ± 16.3
Average quality score[a]	24.9 ± 0.6	24.5 ± 0.7
Number (percentage) of predicted 16S rDNAs	176 (0.17 %)	59 (0.13%)
Number (percentage) of predicted proteins categorized into functional genes[b]	24,896 (23.6 %)	10,911 (23.6 %)
Number (percentage) of predicted proteins categorized into COG groups[b]	9,900 (9.4%)	6,150 (13.3 %)
Number (percentage) of predicted proteins categorized into KEGG pathways[b]	9,339 (8.9 %)	7,003 (15.2 %)

[a]The quality score of each base was provided by 454 Life sequences, and is analogous to the Phred score of Sanger Sequencing methods. The value cited here is the average of the mean quality score per sequence. Typically, bases with a Phred score higher than 20 is predicted to have an accuracy of 99% or higher.

[b]Functional gene and category predictions were based on separate BLASTX analyses against the NCBI's RefSeq, COG and KEGG databases.

Table 4.2 COG groups that were statistically overrepresented in the metagenomes of enriched bacterioplankton in either the DMSP or VanA amendments. Statistical analysis is based on a bootstrapping method with a sample size of 7000 proteins, 20,000 repeated samples, and a confidence interval of 95% ($p < 0.05$).

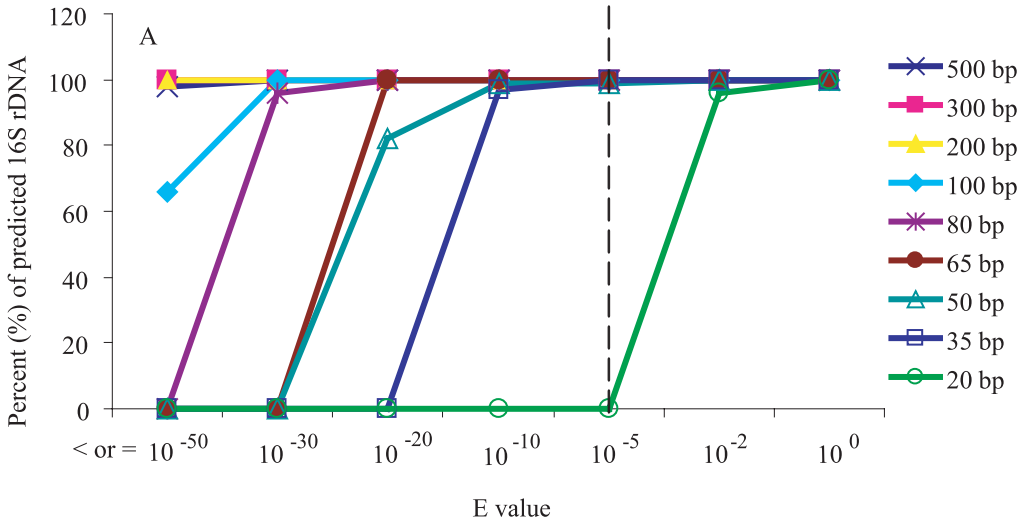
DMSP amendment		**VanA amendment**	
COG groups (COG #)	**Functional category**	**COG groups (COG #)**	**Functional category**
Isocitrate/isopropylmalate dehydrogenase (COG0473)	Energy production and amino acid metabolism	Protein chain release factor A (COG0216)	Translation ribosome structure and biosynthesis
DNA polymerase III, alpha subunit (COG0587)	Replication, recombination and repair	Folylpolyglutamate synthase (COG0285)	Coenzyme transport and metabolism
ABC-type cobalt transport system (COG1122)	Inorganic ion transport	Molecular chaperone (COG0443)	Post-translational modification, protein turnover, chaperones
ABC-type multidrug transport system (COG1132)	Defense mechanism	UDP-N-acetyl-D-mannosaminuronate dehydrogenase (COG0677)	Cell wall/membrane/envelope biogenesis
RTX toxins and related Ca^{2+}-binding proteins (COG2931)	Secondary metabolites biosynthesis	5'-nucleotidase/2',3'-cyclic phosphodiesterase (COG0737)	Nucleotide transport and metabolism
		Nitroreductase (COG0778)	Energy production and conversion
		Uroporphyrinogen-III synthase (COG1587)	Coenzyme transport and metabolism
		Phospholipase, lecithinase, hemolysin (COG3240)	Lipid transport and metabolism

Table 4.3. KEGG pathways (KEGG pathway numbers) that were statistically overrepresented in the metagenomes of enriched bacterioplankton in either the DMSP or VanA amendments. The statistical analysis is based on a bootstrapping method with a sample sized of 9,000 proteins, 20,000 repeated samples and confidence interval of 95% ($p < 0.05$). Because a given protein could participate in more than one pathway, the distribution table for KEGG hits was generated by recording the occurrence of each pathway number (at the third KEGG level).

DMSP amendment	**VanA amendment**
DNA polymerase (path: ko03030)	The citrate cycle (path: ko00020)
	Aminosugars metabolism (path: ko00530)
	Glyoxylate and dicarboxylate metabolism (path: ko00630)
	One carbon pool by folate (path: ko00670)
	Folate biosynthesis (path: ko00790)

Figure 4.1. (A) The percent of correct 16S rDNA predictions (at the domain level) from the random sequence fragments of 100 known 16S rDNA based on the E values in BLASTN analysis against the RDP II database. The dotted line shows the cutoff E value of the best hits for reliable 16S rDNA predictions. (B) Length distribution of the sequences in the enriched bacterioplankton metagenomes for the DMSP and VanA amendments. The dotted line shows the cutoff sequence length of gene prediction from pyrosequences and the percent of pyrosequences above that cutoff limit.

Fig. 4.1


A
120
100
80
60
40
20
0
Percent (%) of predicted 16S rDNA
< or = 10 -50
10 -30
10 -20
10 -10
10 -5
10 -2
10 0
E value
500 bp
300 bp
200 bp
100 bp
80 bp
65 bp
50 bp
35 bp
20 bp

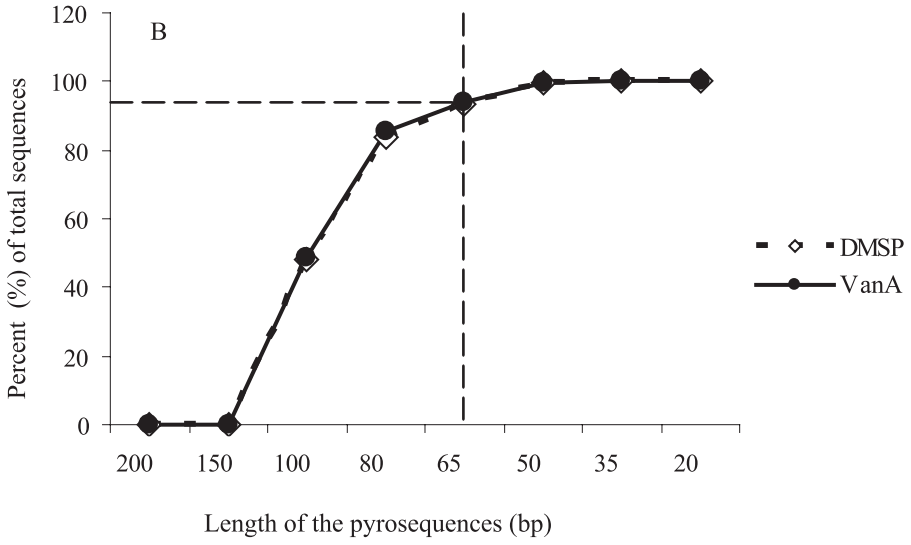

B
120
100
80
60
40
20
0
Percent (%) of total sequences
200
150
100
80
65
50
35
20
Length of the pyrosequences (bp)
DMSP
VanA

Figure 4.2. (A) The percent of correct taxonomic assignments from the random sequence fragments of 100 known 16S rDNA sequences based on the SIMO RDP Agent at different taxonomic ranks. The dotted line shows the reliability of taxonomic assignment at the order level for sequences longer than 65 bp. (B) The distribution of the random sequence fragments based on similarity corrected at the order level. The dotted line shows the cutoff sequence similarity of best hits for the order level of taxonomic classification using the SIMO RDP Agent.

Fig. 4.2

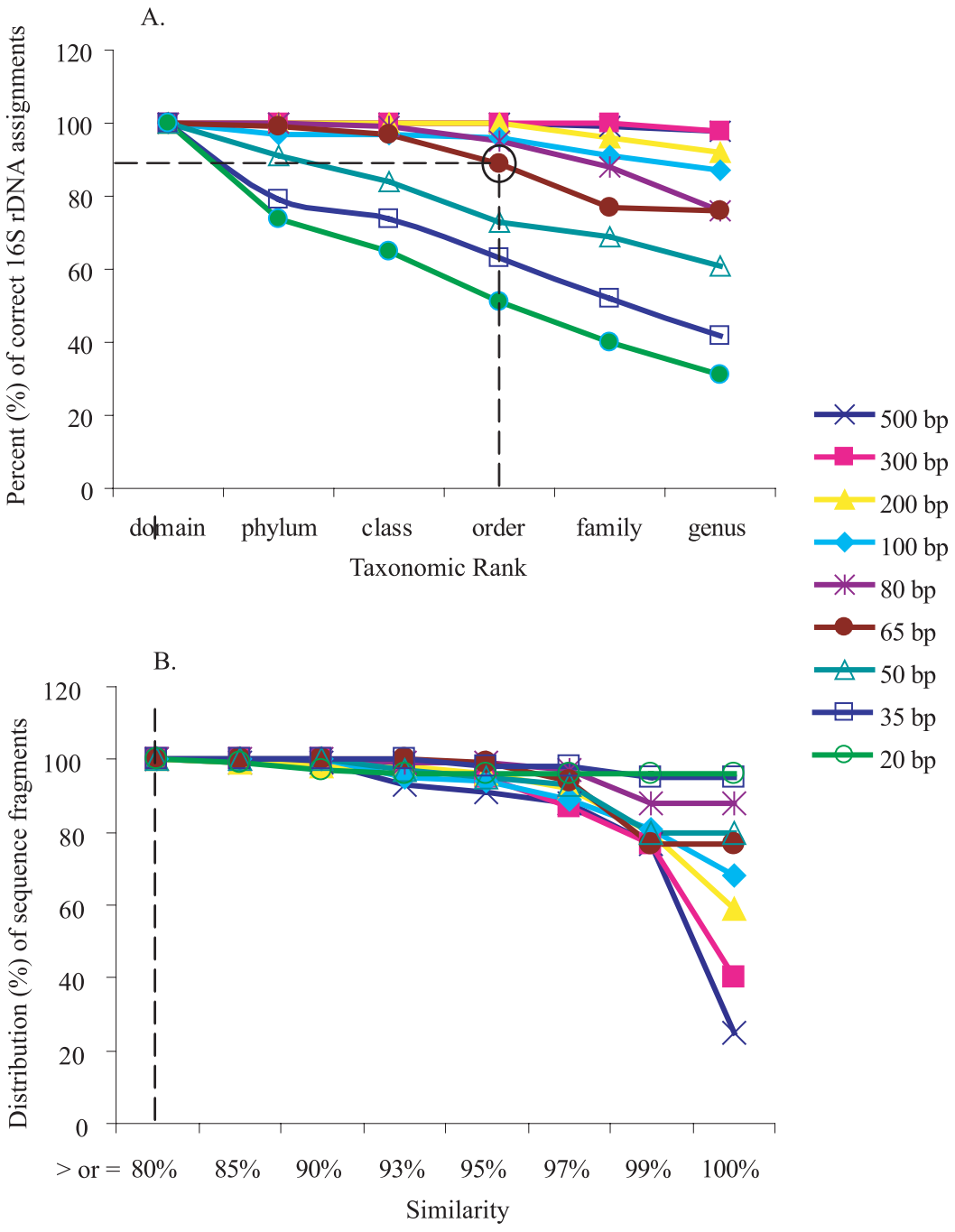

A.
Percent (%) of correct 16S rDNA assignments
120
100
80
60
40
20
0
domain
phylum
class
order
family
genus
Taxonomic Rank
500 bp
300 bp
200 bp
100 bp
80 bp
65 bp
50 bp
35 bp
20 bp
B.
Distribution (%) of sequence fragments
120
100
80
60
40
20
0
> or = 80%
85%
90%
93%
95%
97%
99%
100%
Similarity

Figure 4.3. The percent of correct protein predictions from the random sequence fragments of 100 known functional genes based on BLASTX analysis against the NCBI's RefSeq database at different ranges of E values (A) and sequence similarity (B). The dotted line shows the cutoff E value (A) or sequence similarity (B) to a query's best hit for functional gene prediction using BLASTX analysis against the NCBI's RefSeq database.

Fig. 4.3

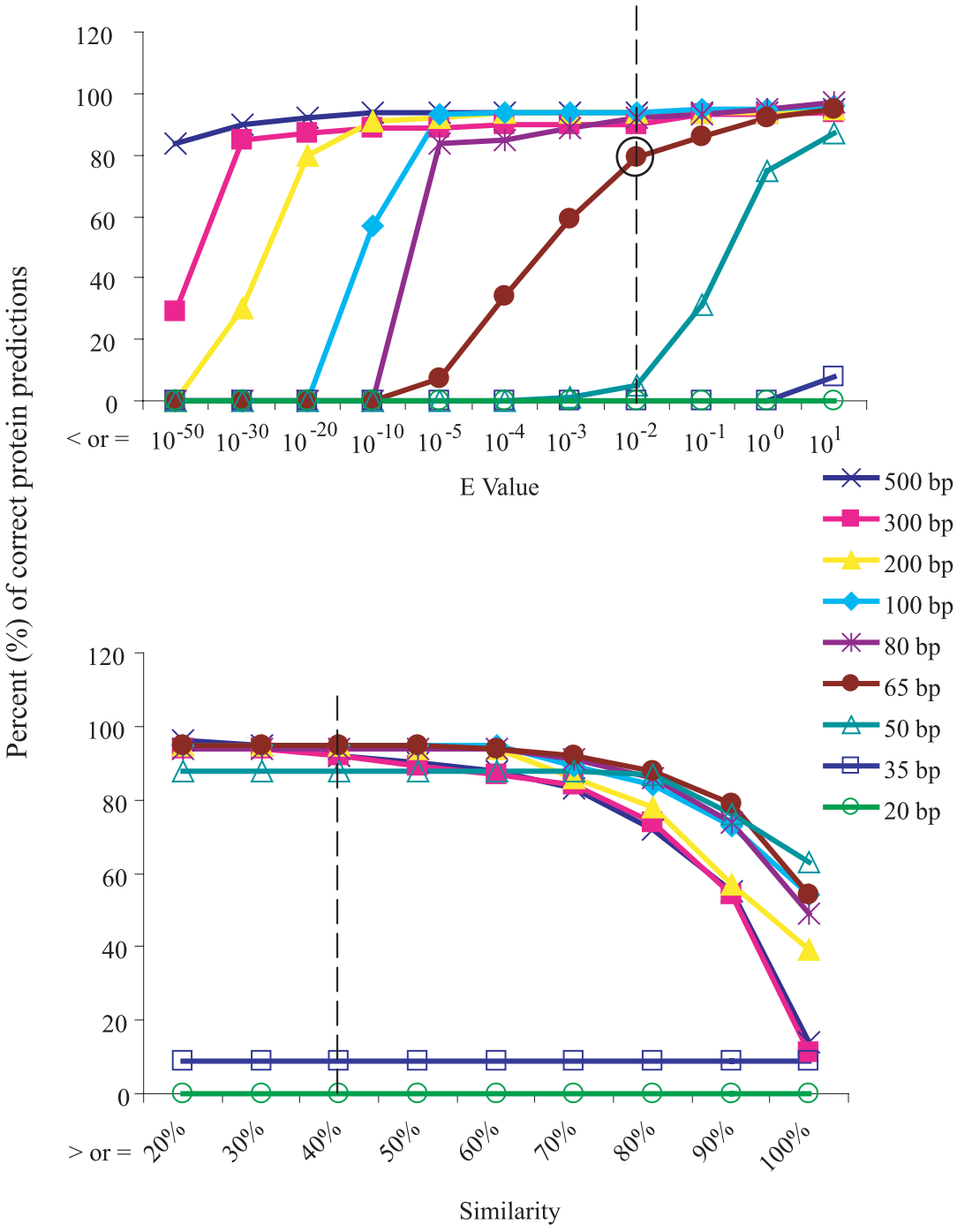

Percent (%) of correct protein predictions
120
100
80
60
40
20
0
< or =
10-50
10-30
10-20
10-10
10-5
10-4
10-3
10-2
10-1
100
101
E Value
500 bp
300 bp
200 bp
100 bp
80 bp
65 bp
50 bp
35 bp
20 bp
120
100
80
60
40
20
0
> or =
20%
30%
40%
50%
60%
70%
80%
90%
100%
Similarity

Figure 4.4. The percent of correct COG predictions from the random sequence fragments of 100 known functional genes based on BLASTX analysis against the NCBI's COG database at different ranges of the E values (A) and similarity levels (B). The dotted line shows the cutoff expected value (A) or sequence similarity (B) to a query's best hit for functional gene prediction using BLASTX analysis against the NCBI's COG database.

Fig. 4.4

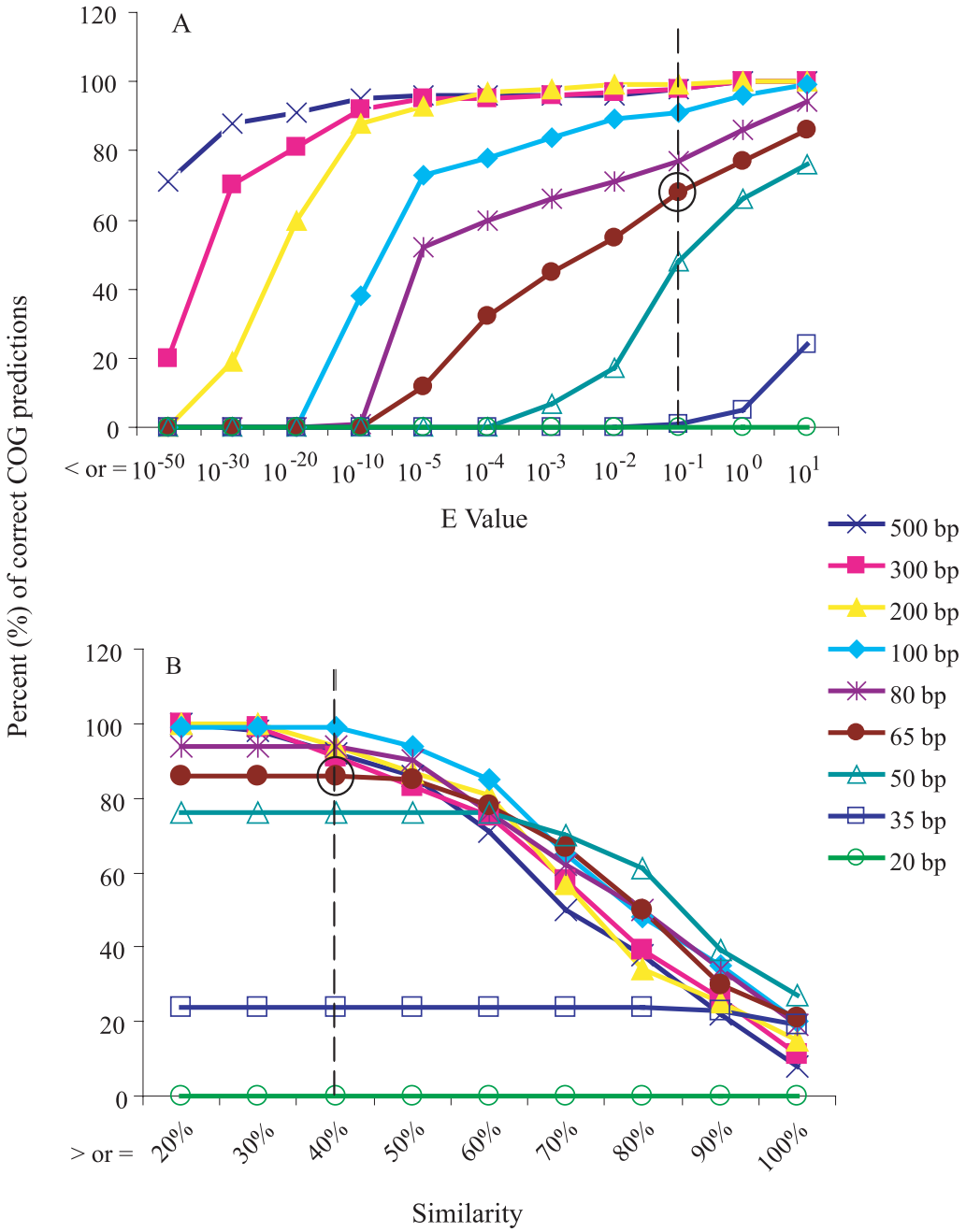

A
120
100
80
60
40
20
0
< or = 10-50 10-30 10-20 10-10 10-5 10-4 10-3 10-2 10-1 100 101
E Value
Percent (%) of correct COG predictions
B
120
100
80
60
40
20
0
> or = 20% 30% 40% 50% 60% 70% 80% 90% 100%
Similarity
500 bp
300 bp
200 bp
100 bp
80 bp
65 bp
50 bp
35 bp
20 bp

Figure 4.5. Bacterial taxon distributions of the enriched metagenome following DMSP and VanA additions as derived from 16S rDNA prediction of the metagenomic libraries (A) and PCR-based 16S rDNA clone libraries (B). The 16S rDNA prediction for both libraries was based on Smith-Waterman alignment against the RDP II database (SIMO RDP Agent) and a marine reference sequence dataset.

Fig. 4.5

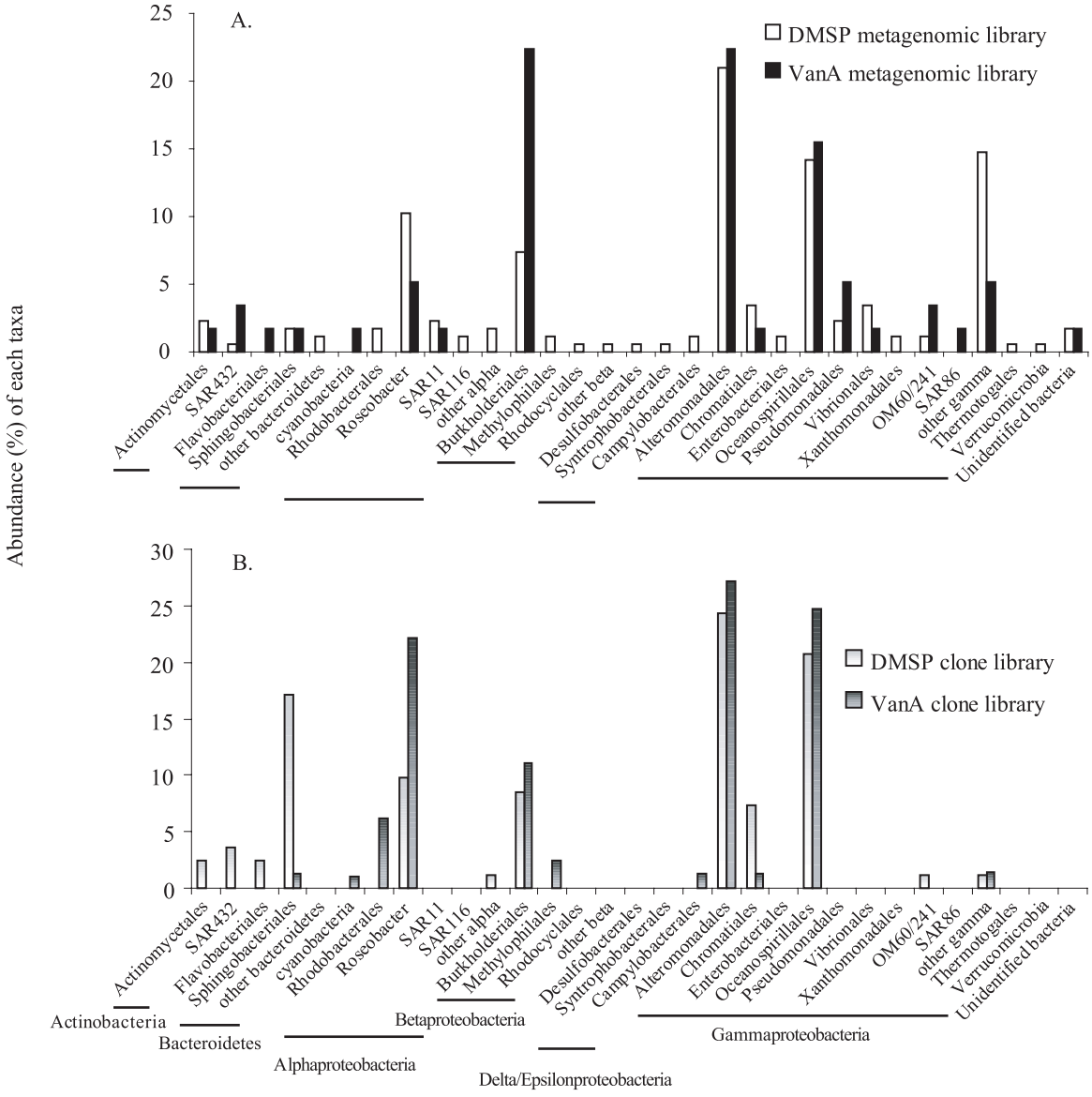

A.
DMSP metagenomic library
VanA metagenomic library
Abundance (%) of each taxa
25
20
15
10
5
0
Actinomycetales
SAR432
Flavobacteriales
Sphingobacteriales
other bacteroidetes
cyanobacteria
Rhodobacterales
Roseobacter
SAR11
SAR116
other alpha
Burkholderiales
Methylophilales
Rhodocyclales
other beta
Desulfobacterales
Syntrophobacterales
Campylobacterales
Alteromonadales
Chromatiales
Enterobacteriales
Oceanospirillales
Pseudomonadales
Vibrionales
Xanthomonadales
OM60/241
SAR86
other gamma
Thermotogales
Verrucomicrobia
Unidentified bacteria
B.
DMSP clone library
VanA clone library
30
25
20
15
10
5
0
Actinobacteria
Bacteroidetes
Alphaproteobacteria
Betaproteobacteria
Delta/Epsilonproteobacteria
Gammaproteobacteria

Figure 4.6. The metagenomic 16S rDNA (A) and protein sequences (B) revealed similar taxonomic distribution patterns between the enriched bacterioplankton in the DMSP and VanA amendments. The protein prediction was based on BLASTX analysis against the NCBI's RefSeq database. Alpha, α-Proteobacteria; Beta, β-Proteobacteria; Gamma, γ-Proteobacteria; Delta/Epsilon, δ- and ε-Proteobacteria; Unclassified, unclassified bacteria.

Fig. 4.6

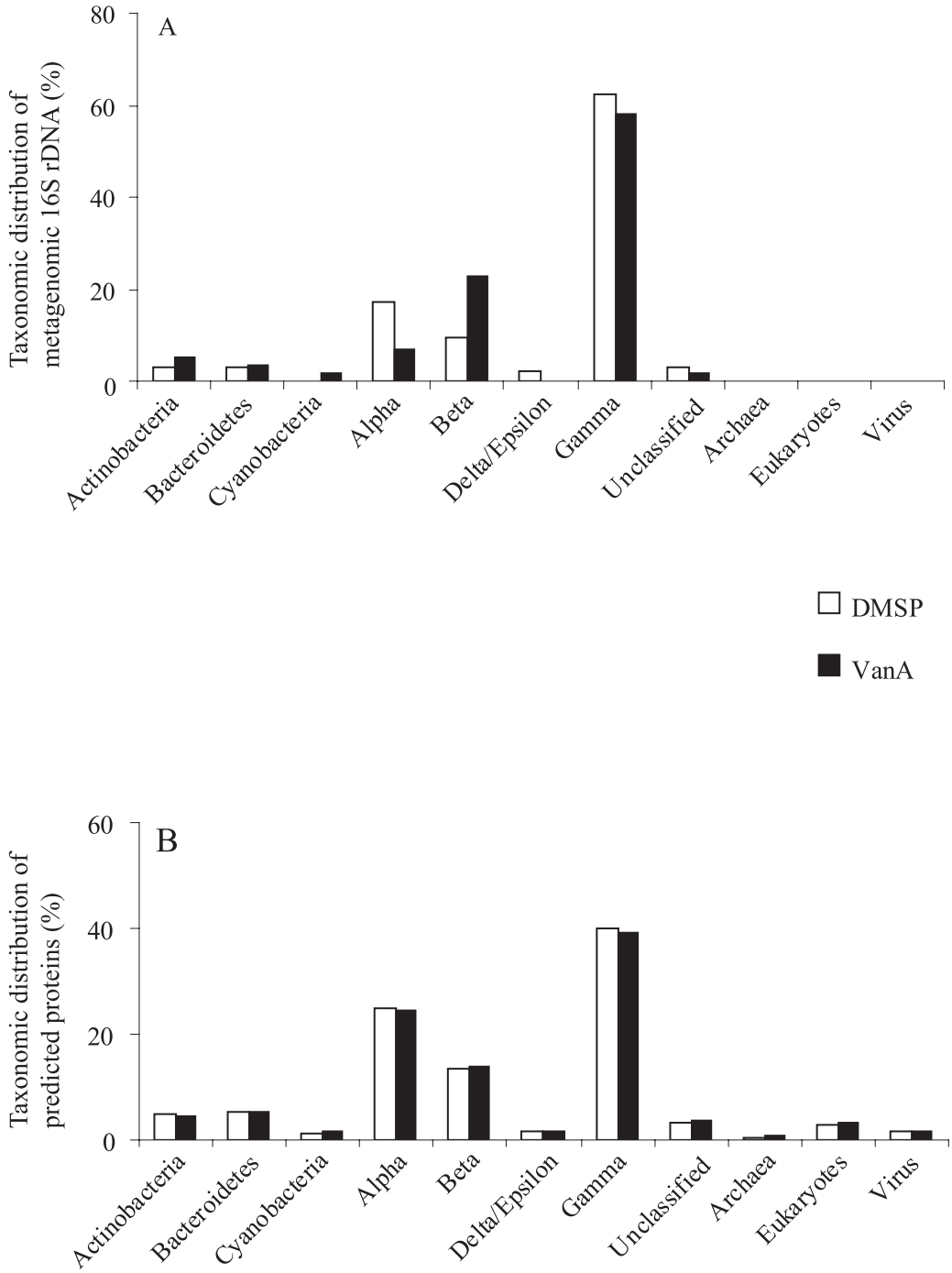

A
Taxonomic distribution of metagenomic 16S rDNA (%)
80
60
40
20
0
Actinobacteria
Bacteroidetes
Cyanobacteria
Alpha
Beta
Delta/Epsilon
Gamma
Unclassified
Archaea
Eukaryotes
Virus
DMSP
VanA
B
Taxonomic distribution of predicted proteins (%)
60
40
20
0
Actinobacteria
Bacteroidetes
Cyanobacteria
Alpha
Beta
Delta/Epsilon
Gamma
Unclassified
Archaea
Eukaryotes
Virus

Figure 4.7. Libshuff comparisons of the 16S rDNA clone libraries for the enriched bacterioplankton in the DOC amendments. DMSP library (X-axis) compared with VanA library (Y-axis) (A). VanA library (X-axis) compared with DMSP library (Y-axis) (B). Coverage (C) curves were generated for both homologous (open circle) and heterologous (dark circle) libraries. Solid lines indicate the value of $(C_X\text{-}C_{XY})^2$ for the original samples at each value of evolutionary distance, which is calculated the by MEGA3 program based on the Kimura's two-parameter method. Dashed lines indicate the value of $(C_X\text{-}C_{XY})^2$ for the randomized samples at $p = 0.05$. The X library is significantly different from the Y library when the value of $(C_X\text{-}C_{XY})^2$ for the randomized samples (dashed line) is lower than that for the original samples (solid line). The p value for the overall comparison of DMSP clone library vs. VanA clone library = 0.001. The p value for the comparison of VanA clone library vs. DMSP clone library = 0.053.

Fig. 4.7

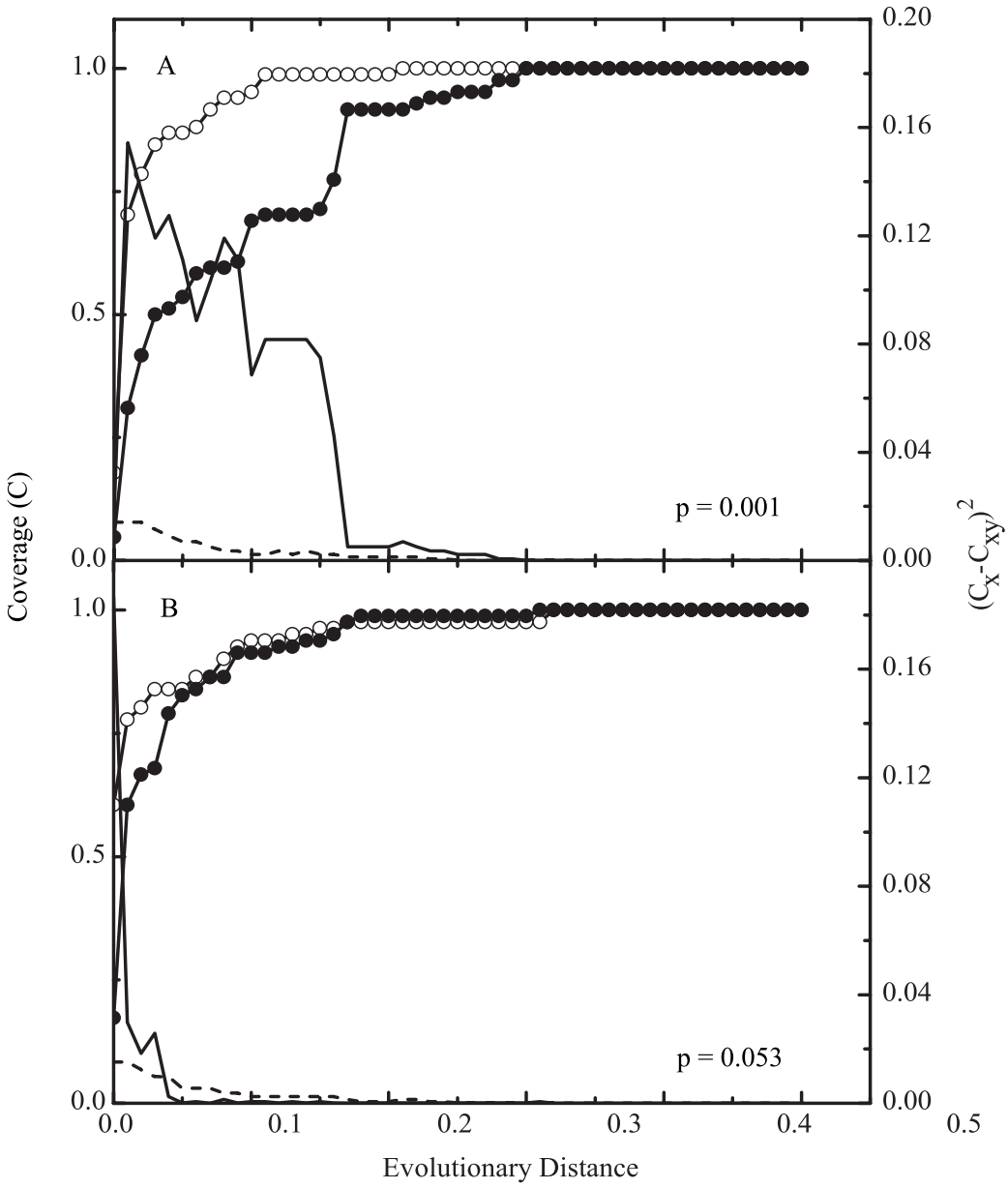

A
B
p = 0.001
p = 0.053
Coverage (C)
$(C_X-C_{XY})^2$
Evolutionary Distance
1.0
0.5
0.0
0.20
0.16
0.12
0.08
0.04
0.00
0.1
0.2
0.3
0.4
0.5

Figure 4.8. The distribution of predicted COG groups (A) and KEGG pathways (B) in the DMSP (above the origin) and VanA (below the origin) enriched metagenomes.

Fig. 4.8

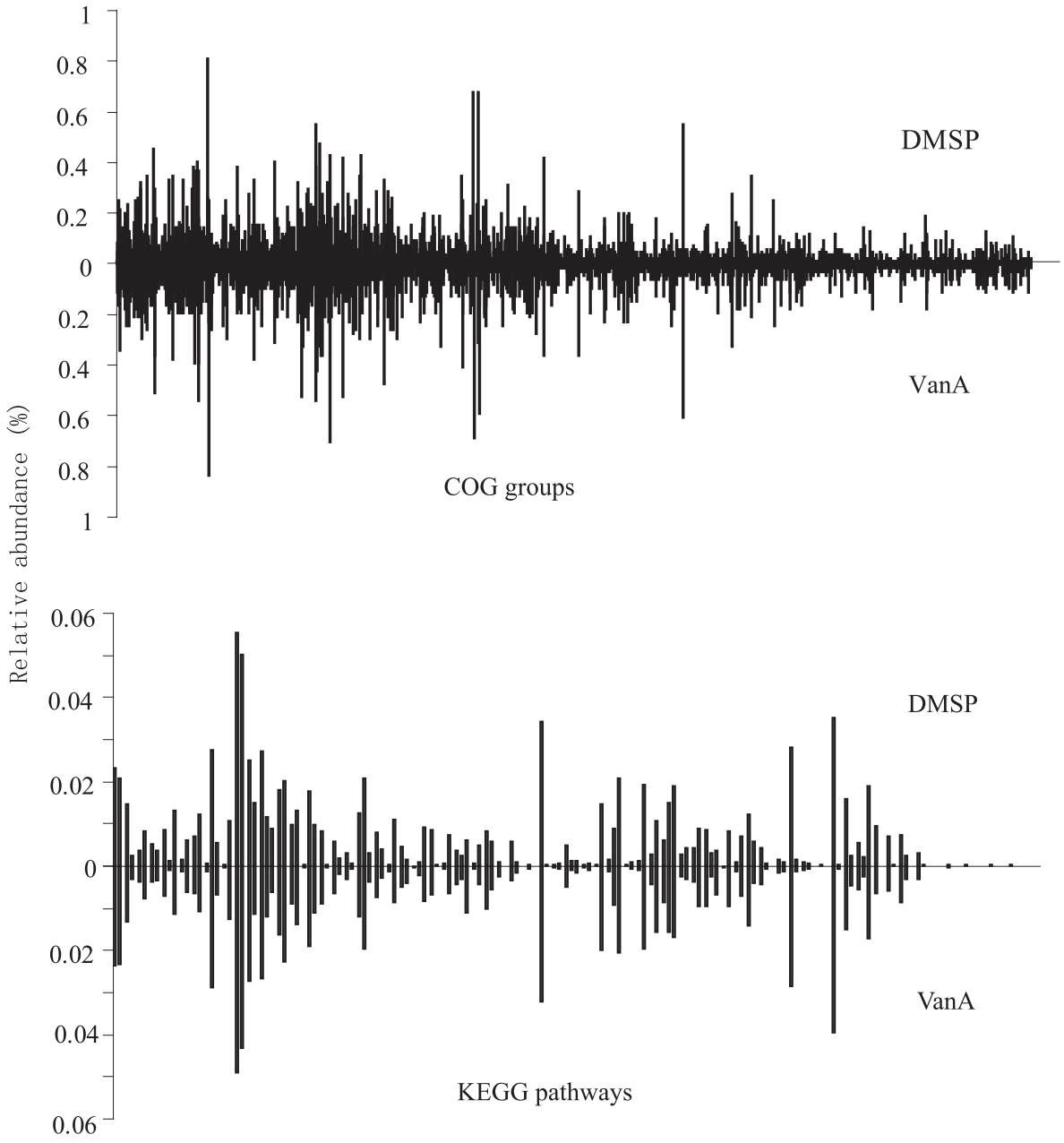

Relative abundance (%)
1
0.8
0.6
0.4
0.2
0
0.2
0.4
0.6
0.8
1
DMSP
VanA
COG groups
0.06
0.04
0.02
0
0.02
0.04
0.06
DMSP
VanA
KEGG pathways

Figure 4.9. The overrepresented COG groups in the DMSP and VanA enriched metagenomes based on the comparisons with the average protein content of 4 selected marine bacterioplankton genomes. For display purposes, COG groups with < 0.2% abundance in the enriched metagenomes are not shown.

Fig. 4.9

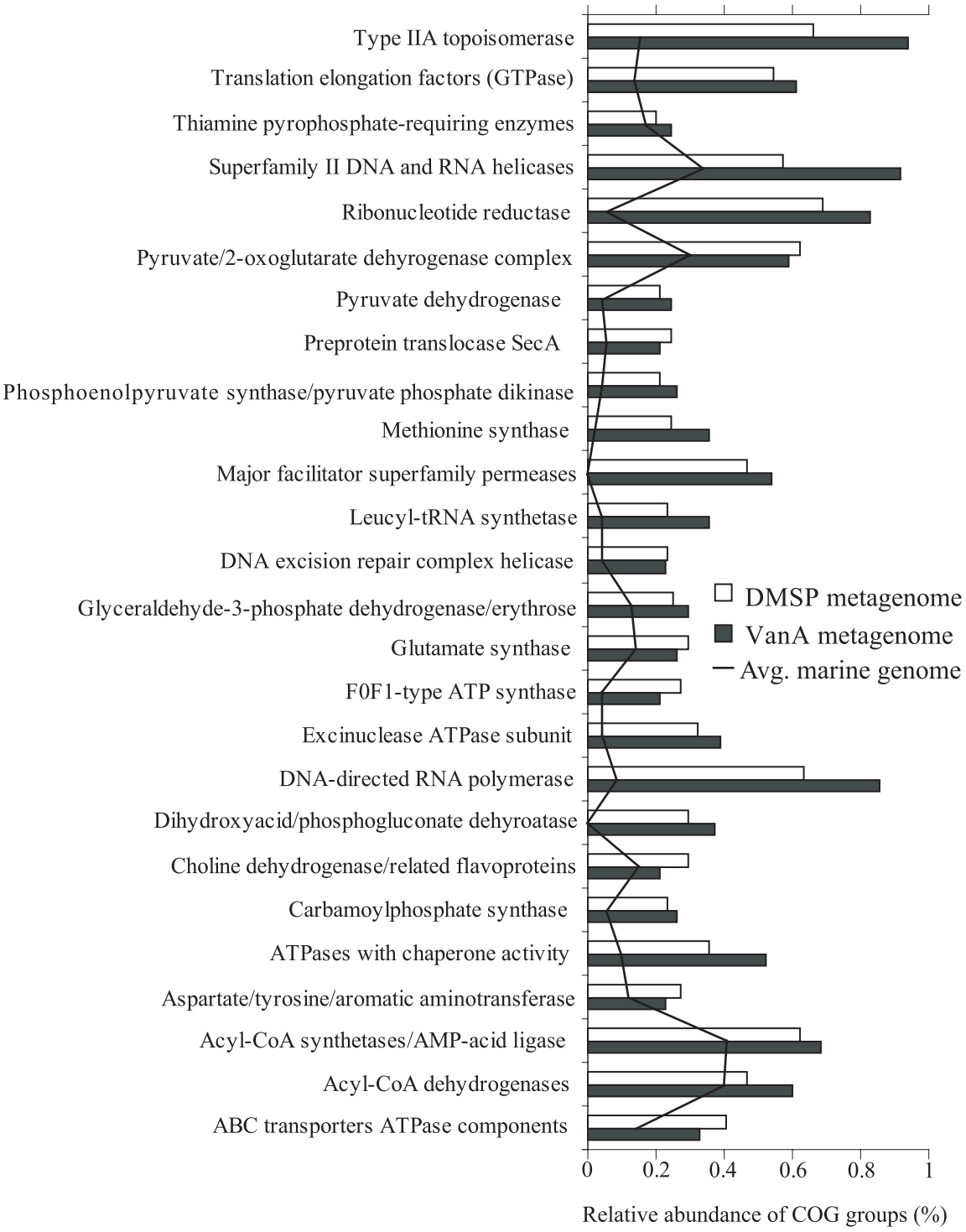

Type IIA topoisomerase
Translation elongation factors (GTPase)
Thiamine pyrophosphate-requiring enzymes
Superfamily II DNA and RNA helicases
Ribonucleotide reductase
Pyruvate/2-oxoglutarate dehyrogenase complex
Pyruvate dehydrogenase
Preprotein translocase SecA
Phosphoenolpyruvate synthase/pyruvate phosphate dikinase
Methionine synthase
Major facilitator superfamily permeases
Leucyl-tRNA synthetase
DNA excision repair complex helicase
Glyceraldehyde-3-phosphate dehydrogenase/erythrose
Glutamate synthase
F0F1-type ATP synthase
Excinuclease ATPase subunit
DNA-directed RNA polymerase
Dihydroxyacid/phosphogluconate dehyroatase
Choline dehydrogenase/related flavoproteins
Carbamoylphosphate synthase
ATPases with chaperone activity
Aspartate/tyrosine/aromatic aminotransferase
Acyl-CoA synthetases/AMP-acid ligase
Acyl-CoA dehydrogenases
ABC transporters ATPase components
DMSP metagenome
VanA metagenome
Avg. marine genome
0
0.2
0.4
0.6
0.8
1
Relative abundance of COG groups (%)

Figure 4.10. The overrepresented KEGG pathways in the DMSP and VanA enriched metagenomes based on comparisons with the average protein content of 4 selected cultured marine bacterioplankton genomes. For display purposes, KEGG pathways with < 1% abundance in the enriched metagenomes are not shown.

Fig. 4.10

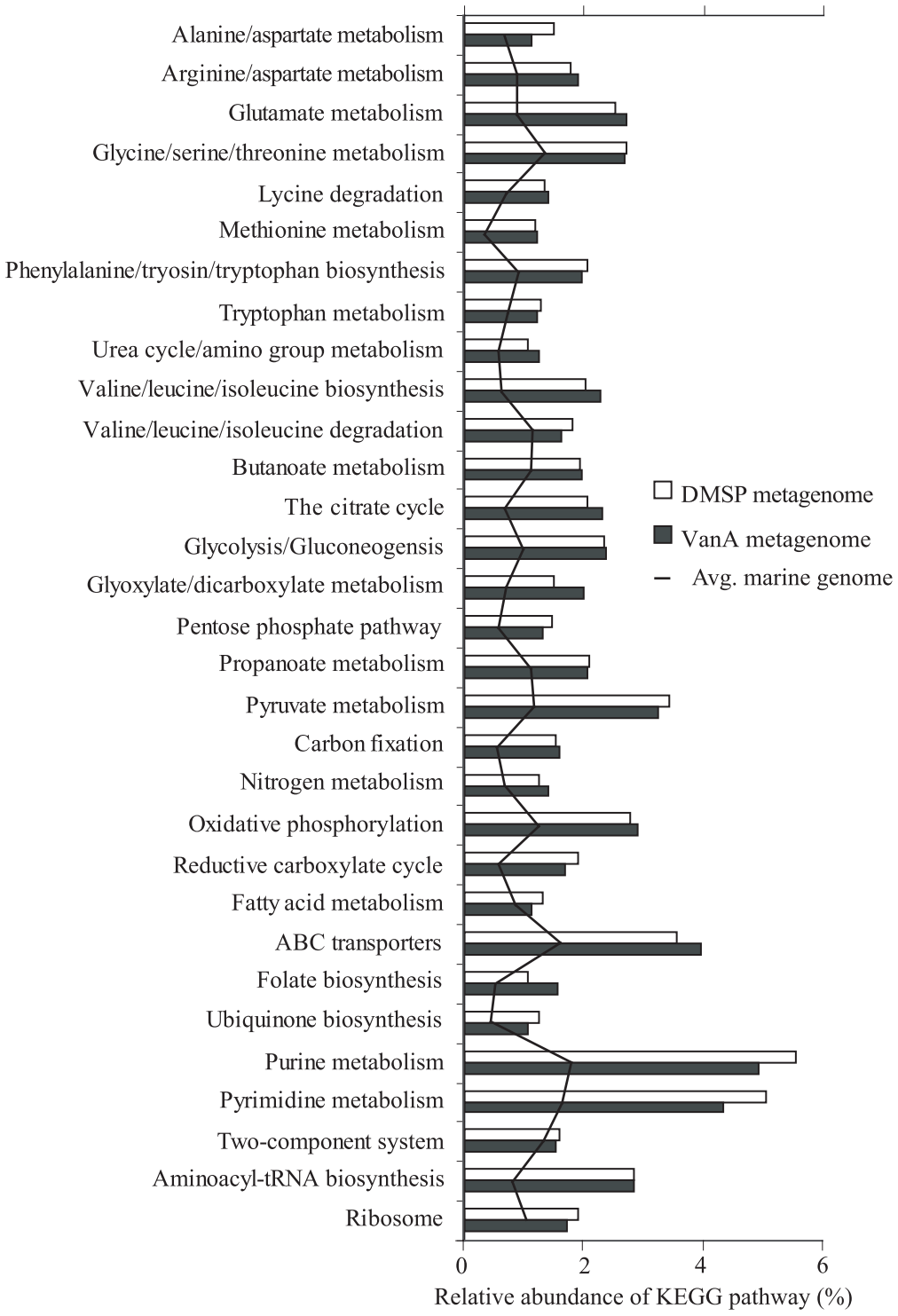

Alanine/aspartate metabolism
Arginine/aspartate metabolism
Glutamate metabolism
Glycine/serine/threonine metabolism
Lycine degradation
Methionine metabolism
Phenylalanine/tryosin/tryptophan biosynthesis
Tryptophan metabolism
Urea cycle/amino group metabolism
Valine/leucine/isoleucine biosynthesis
Valine/leucine/isoleucine degradation
Butanoate metabolism
The citrate cycle
Glycolysis/Gluconeogensis
Glyoxylate/dicarboxylate metabolism
Pentose phosphate pathway
Propanoate metabolism
Pyruvate metabolism
Carbon fixation
Nitrogen metabolism
Oxidative phosphorylation
Reductive carboxylate cycle
Fatty acid metabolism
ABC transporters
Folate biosynthesis
Ubiquinone biosynthesis
Purine metabolism
Pyrimidine metabolism
Two-component system
Aminoacyl-tRNA biosynthesis
Ribosome
DMSP metagenome
VanA metagenome
Avg. marine genome
0
2
4
6
Relative abundance of KEGG pathway (%)

Table S4.1. The COG groups that were overrepresented in the DMSP and VanA enriched metagenomes compared to the average of 4 selected cultured marine bacterioplankton genomes. COG groups that had abundance < 0.1% are not shown. Notation n.a. means not applicable. A ratio of the relative abundance for either one of the DOC enriched metagenomes to the average of 4 selected marine bacterioplankton genomes (AMG) > 1 indicates overrepresentation in the DOC-enriched metagenomes.

COG groups	Functional role	Relative abundance			Ratio	
		DMSP	VanA	Avg. marine genome (AMG)	DMSP vs. AMG	VanA vs. AMG
COG0004	Ammonia permease	0.0021	0.0016	0	n.a.	n.a.
COG0008	Glutamyl- and glutaminyl-tRNA synthetases	0.0018	0.0034	0.0011	1.6749	3.146
COG0021	Transketolase	0.0017	0.0024	0.0004	3.9546	5.617
COG0028	Thiamine pyrophosphate-requiring enzymes	0.0020	0.0024	0.0017	1.1631	1.404
COG0042	tRNA-dihydrouridine synthase	0.0014	0.0021	0.0009	1.6284	2.434
COG0046	Phosphoribosylformylglycinamidine (FGAM) synthase,	0.0021	0.0018	0.0004	4.8852	4.119
COG0049	Ribosomal protein S7	0.0011	0.0011	0.0004	2.5589	2.621
COG0050	GTPases - translation elongation factors	0.0022	0.0021	0.0008	2.9244	2.782
COG0056	F0F1-type ATP synthase, alpha subunit	0.0027	0.0021	0.0004	6.2809	4.868
COG0057	Glyceraldehyde-3-phosphate dehydrogenase/erythrose	0.0025	0.0029	0.0013	1.9386	2.247
COG0060	Isoleucyl-tRNA synthetase	0.0011	0.002	0.0004	2.5589	4.494
COG0065	3-isopropylmalate dehydratase large subunit	0.001	0.0023	0	n.a.	n.a.
COG0069	Glutamate synthase domain 2	0.0029	0.0026	0.0014	2.0757	1.844
COG0082	Chorismate synthase	0.0013	0.0016	0	n.a.	n.a.
COG0085	DNA-directed RNA polymerase, beta subunit/140 kD	0.0038	0.0036	0.0004	8.8398	8.238
COG0086	DNA-directed RNA polymerase, beta' subunit/160 kD	0.0025	0.005	0.0004	5.8157	11.61
COG0090	Ribosomal protein L2	0.0015	0.0011	0.0004	3.4894	2.621
COG0112	Glycine/serine hydroxymethyltransferase	0.0013	0.0016	0.001	1.3441	1.664
COG0119	Isopropylmalate/homocitrate/citramalate synthases	0.0028	0.002	0.0009	3.2568	2.247
COG0129	Dihydroxyacid/phosphogluconate dehydratase	0.0029	0.0037	0	n.a.	n.a.
COG0137	Argininosuccinate synthase	0.0012	0.0015	0.0004	2.7915	3.37
COG0138	AICAR transformylase/IMP cyclohydrolase	0.0012	0.0015	0	n.a.	n.a.
COG0141	Histidinol dehydrogenase	0.0011	0.002	0	n.a.	n.a.
COG0151	Phosphoribosylamine-glycine ligase	0.0013	0.0018	0.0004	3.0241	4.119
COG0154	Asp-tRNAAsn/Glu-tRNAGln amidotransferase A subunit	0.0028	0.002	0	n.a.	n.a.
COG0156	7-keto-8-aminopelargonate synthetase	0.001	0.0013	0.0008	1.3293	1.712
COG0166	Glucose-6-phosphate isomerase	0.0012	0.002	0.0005	2.2332	3.595
COG0173	Aspartyl-tRNA synthetase	0.0011	0.0013	0.0004	2.5589	2.996
COG0174	Glutamine synthetase	0.0025	0.002	0	n.a.	n.a.
COG0178	Excinuclease ATPase subunit	0.0032	0.0039	0.0004	7.444	8.987
COG0187	Type IIA topoisomerase (DNA gyrase/topo II)	0.0034	0.0039	0.0008	4.5196	5.136
COG0188	Type IIA topoisomerase (DNA gyrase/topo II)	0.0031	0.0054	0.0008	4.1208	7.061

COG0189	Glutathione synthase/Ribosomal protein S6 modific	0.0015	0.0013	0.0008	1.9939	1.712
COG0192	S-adenosylmethionine synthetase	0.0012	0.002	0.0004	2.7915	4.494
COG0195	Transcription elongation factor	0.0011	0.0015	0.0004	2.5589	3.37
COG0208	Ribonucleotide reductase, beta subunit	0.0022	0.0013	0.0003	6.8237	3.994
COG0209	Ribonucleotide reductase, alpha subunit	0.0069	0.0083	0.0005	12.655	15.28
COG0210	Superfamily I DNA and RNA helicases	0.0021	0.0023	0.001	2.1712	2.33
COG0218	Predicted GTPase	0.001	0.0011	0.0003	3.1017	3.495
COG0243	Anaerobic dehydrogenases, typically selenocysteine	0.0016	0.0024	0.001	1.6542	2.496
COG0247	Fe-S oxidoreductase	0.0021	0.0016	0.0009	2.4426	1.872
COG0249	Mismatch repair ATPase (MutS family)	0.0016	0.0029	0.0003	4.9627	8.987
COG0277	FAD/FMN-containing dehydrogenases	0.0032	0.002	0.0017	1.861	1.123
COG0281	Malic enzyme	0.0016	0.0018	0.001	1.6542	1.831
COG0304	3-oxoacyl-(acyl-carrier-protein) synthase	0.0023	0.002	0.0011	2.1402	1.797
COG0318	Acyl-CoA synthetases (AMP-forming)/AMP-acid ligase	0.0028	0.0037	0.0025	1.1328	1.498
COG0323	DNA mismatch repair enzyme (predicted ATPase)	0.0013	0.0015	0.0004	3.0241	3.37
COG0365	Acyl-coenzyme A synthetases/AMP-(fatty) acid ligase	0.0034	0.0031	0.0016	2.1091	1.897
COG0372	Citrate synthase	0.0015	0.0021	0.0007	2.3263	3.245
COG0436	Aspartate/tyrosine/aromatic aminotransferase	0.0027	0.0023	0.0012	2.284	1.906
COG0439	Biotin carboxylase	0.0015	0.0018	0.0004	3.4894	4.119
COG0443	Molecular chaperone	0.0018	0.0052	0.0009	2.0936	5.992
COG0445	NAD/FAD-utilizing enzyme apparently involved in ce	0.0017	0.0023	0.0004	3.9546	5.243
COG0449	Glucosamine 6-phosphate synthetase	0.001	0.0015	0.0004	2.3263	3.37
COG0458	Carbamoylphosphate synthase large subunit	0.0023	0.0026	0.0005	4.2803	4.793
COG0459	Chaperonin GroEL (HSP60 family)	0.0025	0.0018	0.0004	5.8157	4.119
COG0462	Phosphoribosylpyrophosphate synthetase	0.0014	0.0013	0.0004	3.2568	2.996
COG0465	ATP-dependent Zn proteases	0.0012	0.0015	0.0007	1.861	2.247
COG0466	ATP-dependent Lon protease	0.002	0.0016	0.0005	3.722	2.996
COG0477	77 Permeases of the major facilitator superfamily	0.0046	0.0054	0	n.a.	n.a.
COG0480	Translation elongation factors (GTPases)	0.0032	0.0042	0.0007	4.9627	6.491
COG0488	ATPase components of ABC transporters with duplica	0.004	0.0033	0.0014	2.8631	2.304
COG0493	NADPH-dependent glutamate synthase beta chain and	0.0012	0.0021	0.0007	1.861	3.245
COG0495	Leucyl-tRNA synthetase	0.0023	0.0036	0.0004	5.3504	8.238
COG0504	CTP synthase (UTP-ammonia lyase)	0.0013	0.0011	0.0004	3.0241	2.621
COG0508	Pyruvate/2-oxoglutarate dehydrogenase complex	0.0027	0.0023	0.0014	1.9326	1.613
COG0513	3 Superfamily II DNA and RNA helicases	0.0036	0.007	0.0025	1.4564	2.8
COG0516	IMP dehydrogenase/GMP reductase	0.0015	0.0021	0.0004	3.4894	4.868
COG0525	Valyl-tRNA synthetase	0.0019	0.0011	0.0004	4.4199	2.621
COG0532	Translation initiation factor 2 (IF-2; GTPase)	0.002	0.002	0.0004	4.6525	4.494
COG0541	Signal recognition particle GTPase	0.0015	0.0011	0.0004	3.4894	2.621
COG0542	ATPases with chaperone activity	0.0035	0.0052	0.001	3.6186	5.326
COG0548	Acetylglutamate kinase	0.001	0.0015	0.0005	1.861	2.696
COG0556	Helicase subunit of the DNA excision repair comple	0.0023	0.0023	0.0004	5.3504	5.243
COG0557	Exoribonuclease R	0.0019	0.0011	0.0003	5.8932	3.495
COG0567	2-oxoglutarate dehydrogenase complex, dehydrogenas	0.0012	0.0028	0.0004	2.7915	6.366
COG0568	DNA-directed RNA polymerase, sigma subunit	0.0016	0.0018	0.0012	1.3535	1.498
COG0574	Phosphoenolpyruvate synthase/pyruvate phosphate dikinase	0.0021	0.0026	0.0004	4.8852	5.992
COG0587	DNA polymerase III, alpha subunit	0.0036	0.002	0.0007	5.583	2.996

COG0621	2-methylthioadenine synthetase	0.0024	0.0011	0.0008	3.1903	1.498
COG0626	Cystathionine beta-lyases/cystathionine gamma-synthases	0.0013	0.0018	0.0012	1.0997	1.498
COG0653	Preprotein translocase subunit SecA	0.0024	0.0021	0.0005	4.4664	3.895
COG0659	Sulfate permease and related transporters	0.0018	0.0026	0.0008	2.3927	3.424
COG0682	Prolipoprotein diacylglyceryltransferase	0.0012	0.0011	0	n.a.	n.a.
COG0749	DNA polymerase I - 3'-5' exonuclease and polymeras	0.0017	0.0013	0.0004	3.9546	2.996
COG0768	Cell division protein FtsI/penicillin-binding protein	0.0016	0.0011	0.0009	1.861	1.311
COG0772	Bacterial cell division membrane protein	0.001	0.0011	0	n.a.	n.a.
COG0773	UDP-N-acetylmuramate-alanine ligase	0.0012	0.002	0.0007	1.861	2.996
COG0784	FOG: CheY-like receiver	0.0016	0.0033	0	n.a.	n.a.
COG0825	Acetyl-CoA carboxylase alpha subunit	0.001	0.0011	0.0004	2.3263	2.621
COG0843	Heme/copper-type cytochrome/quinol oxidases	0.0021	0.002	0.0007	3.2568	2.996
COG1003	Glycine cleavage system protein P	0.0016	0.002	0.0004	3.722	4.494
COG1032	Fe-S oxidoreductase	0.0011	0.0016	0.0003	3.4119	4.993
COG1048	Aconitase A	0.0019	0.0021	0.0005	3.5359	3.895
COG1053	Succinate dehydrogenase/fumarate reductase	0.0021	0.0016	0.0007	3.2568	2.496
COG1088	dTDP-D-glucose 4,6-dehydratase	0.001	0.0011	0.0002	4.6525	5.243
COG1109	Phosphomannomutase	0.0017	0.0024	0.0008	2.2598	3.21
COG1138	Cytochrome c biogenesis factor	0.0012	0.0013	0.0005	2.2332	2.397
COG1154	Deoxyxylulose-5-phosphate synthase	0.0012	0.0018	0.0005	2.2332	3.295
COG1160	Predicted GTPases	0.0012	0.0013	0.0005	2.2332	2.397
COG1171	Threonine dehydratase	0.0015	0.0013	0.0012	1.2689	1.089
COG1185	Polyribonucleotide nucleotidyltransferase	0.0019	0.0015	0.0004	4.4199	3.37
COG1196	Chromosome segregation ATPases	0.0012	0.0018	0.0007	1.861	2.746
COG1197	Transcription-repair coupling factor	0.001	0.0021	0.0003	3.1017	6.491
COG1200	RecG-like helicase	0.0011	0.0018	0.0004	2.5589	4.119
COG1207	N-acetylglucosamine-1-phosphate uridyltransferase	0.0011	0.002	0.0004	2.5589	4.494
COG1217	Predicted membrane GTPase involved in stress response	0.0012	0.0028	0.0004	2.7915	6.366
COG1219	ATP-dependent protease Clp	0.0015	0.0013	0.0004	3.4894	2.996
COG1249	Pyruvate/2-oxoglutarate dehydrogenase complex	0.0035	0.0036	0.0016	2.1712	2.197
COG1290	Cytochrome b subunit of the bc complex	0.0011	0.0011	0.0004	2.5589	2.621
COG1410	Methionine synthase I, cobalamin-binding domain	0.0024	0.0036	0.0002	11.166	16.48
COG1529	Aerobic-type carbon monoxide dehydrogenase, large	0.0015	0.0023	0.0009	1.7447	2.621
COG1530	Ribonucleases G and E	0.0012	0.0018	0.0008	1.5952	2.354
COG1643	HrpA-like helicases	0.0016	0.0023	0.0005	2.9776	4.194
COG1653	ABC-type sugar transport system, periplasmic compo	0.0017	0.0013	0.0008	2.2598	1.712
COG1674	DNA segregation ATPase FtsK/SpoIIIE and related pr	0.0013	0.0015	0.0004	3.0241	3.37
COG1804	Predicted acyl-CoA transferases/carnitine dehydrat	0.0015	0.0013	0.0012	1.2689	1.089
COG1894	NADH:ubiquinone oxidoreductase, NADH-binding	0.001	0.0024	0.0008	1.3293	3.21
COG1960	Acyl-CoA dehydrogenases	0.0046	0.006	0.004	1.1568	1.498
COG2072	Predicted flavoprotein involved in K+ transport	0.0012	0.0011	0.0008	1.5952	1.498
COG2217	Cation transport ATPase	0.0014	0.0018	0.001	1.4475	1.831
COG2225	Malate synthase	0.0012	0.0018	0.0005	2.2332	3.295
COG2303	Choline dehydrogenase and related flavoproteins	0.0029	0.0021	0.0015	1.9275	1.391
COG2609	Pyruvate dehydrogenase complex, dehydrogenase (E1)	0.0021	0.0024	0.0004	4.8852	5.617
COG2873	O-acetylhomoserine sulfhydrylase	0.001	0.0015	0.0004	2.3263	3.37
COG3333	Uncharacterized protein conserved in bacteria	0.0011	0.0011	0.0007	1.7059	1.748

COG3839	ABC-type sugar transport systems, ATPase component	0.0016	0.0018	0.0011	1.4888	1.648

Table S4.2. The KEGG pathways that were overrepresented in the DMSP and VanA enriched metagenomes compared to the average of 4 selected cultured marine bacterioplankton genomes. Notation n.a. means not applicable. A ratio of the relative abundance for either one of the DOC enriched metagenomes to the average of 4 selected marine bacterioplankton genomes (AMG) > 1 indicates overrepresentation in the DOC-enriched metagenomes.

KEGG groups	Pathway	Relative abundance			Ratio	
		DMSP	VanA	Avg. marine genome	DMSP	VanA
path00251	Glutamate metabolism	0.0223	0.0424	0.0088	2.5269	4.793
path00252	Alanine and aspartate metabolism	0.0132	0.0176	0.0068	1.9545	2.606
path00260	Glycine, serine and threonine metabolism	0.0241	0.0417	0.0136	1.7744	3.065
path00271	Methionine metabolism	0.0104	0.0187	0.0033	3.146	5.688
path00280	Valine, leucine and isoleucine degradation	0.0161	0.0254	0.0114	1.4042	2.222
path00290	Valine, leucine and isoleucine biosynthesis	0.0180	0.0355	0.0062	2.9196	5.762
path00310	Arginine and proline metabolism	0.0118	0.0216	0.0071	1.6585	3.044
path00330	Histidine metabolism	0.0157	0.0297	0.0088	1.7742	3.355
path00380	Tryptophan metabolism	0.0112	0.0187	0.0073	1.5407	2.573
path00400	Phenylalanine, tyrosine and tryptophan biosynt	0.0183	0.0308	0.0090	2.0354	3.414
path00010	Glycolysis / Gluconeogenesis	0.0199	0.0350	0.0100	1.9933	3.513
path00020	Citrate cycle (TCA cycle)	0.0183	0.0361	0.0068	2.7138	5.344
path00030	Pentose phosphate pathway	0.0130	0.0205	0.0058	2.2427	3.533
path00620	Pyruvate metabolism	0.0304	0.0504	0.0118	2.5807	4.276
path00630	Glyoxylate and dicarboxylate metabolism	0.0131	0.0310	0.0071	1.8458	4.361
path00640	Propanoate metabolism	0.0185	0.0319	0.0111	1.6709	2.874
path00650	Butanoate metabolism	0.0171	0.0306	0.0113	1.5186	2.711
path00190	Oxidative phosphorylation	0.0244	0.045	0.0126	1.9439	3.584
path00710	Carbon fixation	0.0134	0.0245	0.0054	2.4943	4.565
path00720	Reductive carboxylate cycle (CO2 fixation)	0.0167	0.0263	0.0057	2.9247	4.6
path00910	Nitrogen metabolism	0.0109	0.0221	0.0068	1.5966	3.224
path00071	Fatty acid metabolism	0.0117	0.0176	0.0086	1.3627	2.053
path02010	ABC transporters	0.0314	0.0618	0.0161	1.9459	3.832
path00130	Ubiquinone biosynthesis	0.0109	0.0167	0.0043	2.5226	3.859
path00230	Purine metabolism	0.0491	0.0767	0.0180	2.7261	4.255
path00240	Pyrimidine metabolism	0.0447	0.0676	0.0165	2.7131	4.103
path02020	Two-component system	0.0142	0.0236	0.0133	1.0681	1.783
path00970	Aminoacyl-tRNA biosynthesis	0.0251	0.0444	0.0080	3.1473	5.565
path03010	Ribosome	0.0167	0.0268	0.0104	1.6086	2.573

CHAPTER 5

SUMMARY

The intent of the studies described herein was to identify and characterize the taxonomy and metabolic capabilities of bacterioplankton involved in the degradation of important components of the coastal DOC pool, aromatic monomers and organic osmolytes in particular, and to reveal the genetic mechanisms underlying these processes using modern culture-independent molecular techniques.

Aromatic monomers are abundant in coastal marine systems, where a considerable fraction of them is derived from lignin, a structural material of higher plants and one of the most abundant organic polymers in the biosphere (Peng et al., 1999). Due to the stable nature of the aromatic rings, lignin-related aromatic compounds are refractory to most coastal marsh dwellers and have to be transformed to aliphatic structures by fungi and bacteria before utilization by other organisms in the coastal marsh food webs (Benner et al., 1984). The cleavage of the aromatic rings in lignin-related aromatic monomers [e.g. vanillic acid (VanA) and parahyroxybenzoic acid (pHBA)] is the last step necessary to generate aliphatic compounds, and thus is essential to coastal food webs and global carbon cycles.

Aliphatic organic osmolytes dimethylsulfoniopropionate (DMSP) and its nitrogen analog glycine betaine (GlyB) are ubiquitous in marine systems and accumulate in marine organisms at high cytoplasmic concentrations to stabilize intracellular macromolecules under high osmolality conditions (Kiene et al., 1998). DMSP is responsible for the generation of over 90% of oceanic

dimethylsulfide (DMS), which accounts for most of the global biogenic sulfur flux to the atmosphere (Yoch, 2002). In addition, DMSP and GlyB are ubiquitous sources of labile organic matter in marine surface waters, where they provide reduced carbon, sulfur (DMSP) and nitrogen (GlyB) to heterotrophic bacterial communities (Kiene et al., 1999; Roberts, 2005). DMSP alone can provide up to 15% of the carbon requirement of heterotrophic bacteria and fulfill most of the bacterial cellular sulfur demand (Kiene et al., 2000).

It is well accepted that in the coastal systems, the transformation of aromatic monomers and organic osmolytes are mainly mediated by bacteria (Benner et al., 1984; Yoch, 2002). Heterotrophic bacteria in the coastal seawater ultimately determine the form and amount of DOC exported to the oceanic and atmospheric reservoirs. However, the identities of these functional assemblages at the taxonomic resolution of genus and below, and the genetic bases for their functional capabilities, have rarely been determined.

An investigation of the taxonomy of bacterioplankton responding to DMSP addition (20 μM) with elevated metabolic activity supported the hypothesis that the DMSP-transforming assemblage of bacterioplankton is a subset of the total bacterial community. As revealed by this study, the major taxa in the DMSP-transforming assemblage were the same over the four seasons tested, including the *Actinobacteria*, *Bacteriodetes*, α-Proteobacteria (mainly the *Roseobacter* clade), β-Proteobacteria and γ-Proteobacteria, but their relative abundance varied seasonally (Figure 2.4, Figure 2.6 and Table 2.2).

In a similar experiment but with model DOC compounds representing both organic osmolytes and aromatic monomers indicated that functional assemblages transforming common DOC compounds overlap in taxonomic composition. The same major taxa (phylum/subphylum level) were found to dominate the bacterioplankton assemblages responding to the addition of

organic osmolytes and to aromatic monomers (Figure 3.4). The overlap in membership between the GlyB- and VanA-induced bacterioplankton indicated that bacteria that are capable of transforming both DOC compounds (i.e. bacterial generalists) may play an important role in transforming GlyB and VanA (Figure 3.4 and Figure 3.5). Possessing a variety of degradation pathways for commonly-occurring substrate may be a successful ecological strategy in coastal seawater.

The third experiment that used a metagenomic approach to investigate DOC transformations also revealed similar taxonomic structure between bacterioplankton enriched by the addition of DMSP and VanA (Figure 4.1 and Figure 4.2). The functional gene repertoires of DMSP- and VanA-enriched bacterioplankton were also similarly distributed (Figure 4.3 and Figure 4.4). These results suggest that functional genes that are co-beneficial in specific environments are simultaneously harbored in genomes of bacterial generalists. In this case, the stimulation of the bacterial generalists by a specific substrate may result in parallel increases of physically linked genes, and resulting in no obvious enrichment of genes specifically involved in the degradation of a single substrate.

The members of OMG (oligotrophic marine gamma-Proteobacteria, e.g., OM60/241 and OM185) and SAR (e.g., SAR11, SAR86 and SAR116) clusters were repeatedly found in the bacterioplankton assemblage responding to the addition of aromatic compounds or organic osmolytes in the second and third experiments of this study, indicating that these bacteria may be able to transport and metabolize the model compounds used here (Figure 3.5 and Figure 4.1). These clusters are minor members in coastal seawater but can be numerically significant in open ocean surface water, and thus are of considerable importance in the ecology of the marine

environment (Giovannoni and Rappe, 2000). Our study provides some of the first glimpses into heterotrophy by these ubiquitous but poorly described environmental clusters.

REFERENCES

Benner, R., Newell, S.Y., Maccubbin, A.E., and Hodson, R.E. (1984) Relative contributions of bacteria and fungi to rates of degradation of lignocellulosic detritus in salt marsh sediments. *Appl Environl Microbiol* **48**: 36-40.

Giovannoni, S.J., and Rappe, M. (2000) Evolution, diversity, and molecular ecology of marine prokaryotes. *In* D. L. Kirchman [ed.], Microbial Ecology of the Oceans. Wiley-Liss, Inc.: 47-84.

Kiene, R.P., Willams, L.P.H., and Walker, J.E. (1998) Seawater microorganisms have a high affinity glycine betaine uptake system which also recognizes dimethylsulfoniopropionate. *Aquat Microb Ecol* **15**: 39-51.

Kiene, R.P., Linn, L.J., and Bruton, J.A. (2000) New and important roles for DMSP in marine microbial communities. *J Sea Res* **43**: 209-224.

Kiene, R.P., Linn, L.J., Gonzalez, J., Moran, M.A., and Bruton, J.A. (1999) Dimethylsulfoniopropionate and methanethiol are important precursors of methionine and protein-sulfur in marine bacterioplankton. *Appl Environ Microbiol* **65**: 4549-4558.

Peng, X., Masai, E., Katayama, Y., and Fukuda, M. (1999) Characterization of the meta-cleavage compound hydrolase gene involved in degradation of the lignin-related biphenyl structure by *Sphingomonas paucimobilis* SYK-6. *Appl Environ Microbiol* **65**: 2789-2793.

Roberts, M.F. (2005) Organic compatible solutes of halotolerant and halophilic microorganisms. *Saline Syst* **1:5**: doi: 10.1186/1746-1448-1181-1185. .

Yoch, D.C. (2002) Dimethylsulfoniopropionate: Its sources, role in the marine food web, and bioal degradation to dimethylsulfide. *Appl Environ Microbiol* **68**: 5804-5815.