

**MICROBIAL PRODUCTION AND CONSUMPTION OF MARINE
DISSOLVED ORGANIC MATTER**

By

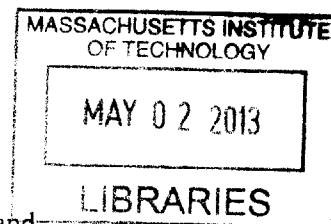
Jamie William Becker

B.S., University of North Carolina at Chapel Hill, 2004
Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
and the
WOODS HOLE OCEANOGRAPHIC INSTITUTION

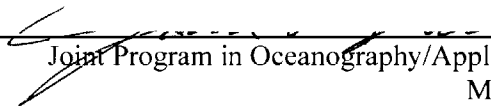
ARCHIVES



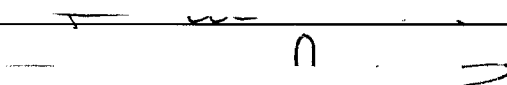
June 2013
© 2013 Jamie W. Becker
All rights reserved.

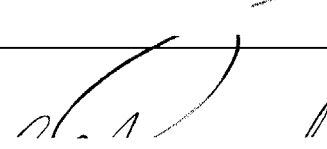
The author hereby grants to MIT and WHOI permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author

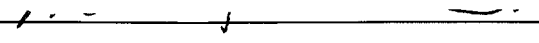

Joint Program in Oceanography/Applied Ocean Sciences and Engineering
Massachusetts Institute of Technology
and Woods Hole Oceanographic Institution
April 22, 2013

Certified by


Dr. Daniel J. Repeta
Thesis Supervisor


Prof. Edward F. DeLong
Thesis Supervisor

Accepted by


Dr. Michael Neubert
Chair, Joint Committee for Biological Oceanography
Woods Hole Oceanographic Institution

MICROBIAL PRODUCTION AND CONSUMPTION OF MARINE DISSOLVED ORGANIC MATTER

By Jamie William Becker

Submitted to the MIT-WHOI Joint Program in Oceanography / Applied Ocean Science and
Engineering on April 22, 2013 in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biological Oceanography

Marine phytoplankton are the principal producers of oceanic dissolved organic matter (DOM), the organic substrate responsible for secondary production by heterotrophic microbes in the sea. Despite the importance of DOM in marine food webs, details regarding how marine microbes cycle DOM are limited, and few definitive connections have been made between specific producers and consumers. Consumption is thought to depend on the source of the DOM as well as the identity of the consumer; however, it remains unclear how phytoplankton diversity and DOM composition are related, and the metabolic pathways involved in the turnover of DOM by different microbial taxa are largely unknown. The motivation for this thesis is to examine the role of microbial diversity in determining the composition, lability, and physiological consumption of marine DOM. The chemical composition of DOM produced by marine phytoplankton was investigated at the molecular level using mass spectrometry. Results demonstrate that individual phytoplankton strains release a unique suite of organic compounds. Connections between DOM composition and the phylogenetic identity of the producing organism were identified on multiple levels, revealing a direct relationship between phytoplankton diversity and DOM composition. Phytoplankton-derived DOM was also employed in growth assays with oligotrophic bacterioplankton strains to examine effects on heterotrophic growth dynamics. Reproducible responses ranged from suppressed to enhanced growth rates and cell yields, and depended both on the identity of the heterotroph and the source of the DOM. Novel relationships between specific bacterioplankton types and DOM from known biological sources were found, and targets for additional studies on reactive DOM components were identified. The physiology of DOM consumption by a marine *Oceanospirillales* strain was studied using a combined transcriptomic and untargeted metabolomic approach. The transcriptional response of this bacterium to *Prochlorococcus*-derived DOM revealed an increase in anabolic processes related to metabolism of carboxylic acids and glucosides, increased gene expression related to proteorhodopsin-based phototrophy, and decreased gene expression related to motility. Putative identification of compounds present in *Prochlorococcus*-derived DOM supported these responses. Collectively, these findings highlight the potential for linking detailed chemical analyses of labile DOM from a known biological source with bacterioplankton diversity and physiology.

Thesis Supervisors:

Dr. Daniel J. Repeta

Title: Senior Scientist, Marine Chemistry & Geochemistry Department, Woods Hole
Oceanographic Institution

Prof. Edward F. DeLong

Title: Morton and Claire Goulder Professor, Departments of Civil and Environmental
Engineering and Biological Engineering, Massachusetts Institute of Technology

Acknowledgments

Financial support for this work was provided by the National Science Foundation Center for Microbial Oceanography: Research and Education (award #EF0424599 and #DBI0424599) including generous contributions from the EDventures program, the Gordon and Betty Moore Foundation (grant #1711), and a student research grant from the Coastal Ocean Institute at the Woods Hole Oceanographic Institution.

Many individuals have helped me (both personally and professionally) over the last several years. So many in fact, that to properly thank them all would double the length of this thesis.

Nevertheless, big mahalos to:

My advisors, Dan & Ed, who despite their better judgment always asked for my opinion. You provided me with a mountain of support to complete an ambitious project and the freedom to find my own way throughout the process.

My committee members, John & Mike, for wisdom and coffee. Especially for the coffee.

Current and former members of the APO at WHOI, particularly Julia, Marsha, Valerie, Tricia, & Jim for always having my back.

C-MORE and Moore, for the money and for thinking big. You're changing the game in a good way.

My cohort, Elizabeth, Maya, Meredith, Ann, & Kathryn, for letting a guy join the party.

My officemates, Kathleen, Erin, Mar, Rene & Chris, for the much-needed distractions.

TJ, Nick, & Adrian for keeping it fun.

My parents and my brother, Bob, Ellen, & Brian, who taught me young to pursue whatever it is that makes me happy. You've not only stood by my side, but also kept me from falling on many occasions.

And finally...the menagerie. For love. For perspective. Misty, Will, Pac, & Lexi - I wouldn't trade a single crazy day.

All things are one thing and that one thing is all things – plankton, a shimmering phosphorescence on the sea and the spinning planets and an expanding universe, all bound together by the elastic string of time.

-John Steinbeck & Edward F. Ricketts

TABLE OF CONTENTS

ABSTRACT.....	3
ACKNOWLEDGMENTS.....	5
TABLE OF CONTENTS.....	7
CHAPTER ONE: Introduction.....	9
CHAPTER TWO: Phytoplankton-derived dissolved organic matter composition is influenced by producer phylogeny.....	29
CHAPTER THREE: The reactivity of marine dissolved organic matter depends on its biological source and sink.....	73
CHAPTER FOUR: Transcriptional profiling of a marine <i>Oceanospirillales</i> reveals increased expression of genes related to anabolism and proteorhodopsin-based phototrophy, and decreased expression of genes related to motility and chemotaxis in the presence of <i>Prochlorococcus</i> -derived dissolved organic matter.....	145
CHAPTER FIVE: Summary and future directions.....	225
APPENDICES: Dissolved organic matter consumption by mixed microbial assemblages.....	235
A. Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea.....	236
B. Distinct dissolved organic matter sources induce rapid transcriptional responses in sympatric populations of <i>Prochlorococcus</i> , <i>Pelagibacter</i> , and the OM60 clade.....	245

CHAPTER ONE

Introduction

Life in the sea, like most of our planet, is enabled by microbial activity. Half of global primary production is marine-based, conducted primarily by phytoplankton in euphotic waters (Longhurst et al. 1995, Field et al. 1998). Marine primary production is a major sink of atmospheric carbon dioxide (CO₂) and forms the basis of most marine food chains (Azam et al. 1983, Chavez et al. 2011). Although the direct consumption of phytoplankton by larger heterotrophic organisms transfers some nutrients and energy to higher trophic levels, about half of marine photosynthate ends up dissolved in seawater as nonliving material, commonly referred to as dissolved organic matter, or DOM (Hedges & OADES 1997, Hansell & Carlson 1998).

Marine DOM is a complex and dynamic reservoir of organic carbon and other nutrients, containing nearly as much carbon as the Earth's atmosphere and serving as the primary substrate for secondary production by heterotrophic microorganisms (Hedges 1992, Pomeroy et al. 2007, Azam & Malfatti 2007). The organic substrates that comprise DOM would remain unavailable to higher trophic levels if not for the activity of heterotrophic microbes, therefore this "microbial loop" of energy and materials is a key component of marine food webs (Figure 1). The recognition of microorganisms as an important component of marine food webs (Pomeroy 1974, Azam et al. 1983) has led to significant efforts in the field of microbial ecology to understand the functional roles that different microbes play in the cycling of marine DOM.

Extracellular release of fixed carbon from marine phytoplankton is a major source of aquatic DOM (Cole et al. 1982, Fogg 1983, Baines & Pace 1991). Previous studies have demonstrated that a diverse array of phytoplankton release between 5 and 50% of their photosynthate into their surrounding environment in laboratory culture as well as in the environment (Fogg 1977, Mague et al. 1980, Bjornsen 1988, Carlson 2002, Bertilsson & Jones 2003, Bertilsson et al. 2005). The underlying mechanisms of extracellular release are not clear, and the wide range in values may be due to environmental factors, growth conditions, physiological differences among the organisms surveyed, or variations in the methodology used. If different marine phytoplankton produce and release DOM of different composition, then DOM composition could be an important driver of microbial diversity and act as a link between autotrophic and heterotrophic community structure. Compositional differences in the DOM released by different phytoplankton taxa may alter the community structure of sympatric heterotrophic microbes. The extent to which DOM composition varies among different phytoplankton types and how these variations relate to phytoplankton diversity are largely unknown. It is currently unclear if phylogenetically related marine phytoplankton produce chemically similar DOM.

There are many possible fates for marine DOM, including: incorporation into heterotrophic biomass via anabolic metabolism, conversion back to inorganic nutrients and CO₂ via catabolic heterotrophic activity, persistence in the DOM pool, photodegradation from ultraviolet light, and aggregation with other DOM to form particulate organic matter (Carlson & Ducklow 1996, Chin et al. 1998, Covert & Moran

2001, Jiao et al. 2010, Loiselle et al. 2012). Due to the large reservoir size of marine DOM and the rate at which certain components turnover, the relative distribution of these various fates has a large impact on the global cycling of carbon, nitrogen, phosphorus and other biologically relevant elements. As microorganisms are the major consumers of marine DOM, studies that investigate how microbes degrade and alter DOM are essential for a more thorough understanding of marine DOM cycling and its impacts on biogeochemical cycles and microbial ecology. Any changes to the lability of marine DOM, rates of DOM consumption, or the partitioning of DOM between anabolic and catabolic pathways will have significant impacts on the cycling of carbon and other nutrients, including carbon export and storage in the ocean interior.

The microbial production and consumption of marine DOM involves both biological and chemical processes, and studies of microbial DOM cycling benefit by integrating techniques and approaches from these two fields. Decades of research by marine organic geochemists have revealed a wealth of information regarding the age and chemical characteristics of marine DOM; however, the identity and lability of individual DOM constituents remain largely unknown despite improvements in isolation and analytical technologies (Druffel et al. 1992, Benner et al. 1992, also reviews by Mopper et al. 2007, Nebbioso & Piccolo 2012). Marine microbiologists have been measuring rates of microbial production (both primary and secondary) and respiration for decades (Baines & Pace 1991, Williams & Purdie 1991, also reviews by Duarte et al. 2013, Williams et al. 2013) and have recently begun using molecular approaches to examine DOM effects on microbial community structure and activity (Poretsky et al. 2010, Vila-

Costa et al. 2010, McCarren et al. 2010, Tada et al. 2011, Trabelsi & Rassoulzadegan 2011, Mou et al. 2011, Rinta-Kanto et al. 2011, Romera-Castillo et al. 2011, Nelson & Carlson 2012, Teeling et al. 2012, Sarmiento & Gasol 2012, Gómez-Consarnau et al. 2012). These recent studies are revealing connections between various DOM sources and consumption by individual taxa, however it is not yet clear why these observed connections occur and the production of specific compounds by one microorganism has not yet been linked to their consumption by another.

DOM in the ocean arises from a wide variety of sources and processes. Bulk seawater analyses are not able to link specific components to particular biological sources, making it difficult to examine the relationship between phytoplankton diversity and DOM composition. In addition, highly labile DOM compounds are under constant degradation pressure from heterotrophic microbes and are therefore unlikely to accumulate, making it difficult to obtain and study this material. Most of what is currently known about the chemical composition of marine DOM may not be relevant to the preferred substrates of most marine microbes. Several recent reviews have highlighted the need to examine DOM heterogeneity at the molecular level and link DOM compounds to their sources in order to identify the roles of different microbial groups in DOM cycling (Nagata 2002, Hansell et al. 2009, Dittmar & Paeng 2009, Kujawinski 2011). Major broad questions regarding microbial cycling of marine DOM that have yet to be resolved include:

- How are microbial diversity and DOM composition related?
- How are DOM composition and lability related?

- How are DOM lability and microbial diversity related?
- What is the physiology of microbial DOM cycling?

Information regarding how the community structure of phytoplankton and heterotrophic bacterioplankton influence and are influenced by marine DOM will be vital in uncovering novel interactions that affect marine microbial diversity and the biogeochemical cycling of organic nutrients. Addressing these broad questions requires investigating both the chemistry of DOM and the biological processes involved in its production and consumption.

An interdisciplinary approach

With at least tens of thousands of organic compounds and a million microbial cells in every drop of seawater, determining the identity, origin, and metabolic fate of marine DOM is a formidable undertaking. To gain a full understanding of microbe-DOM interactions, one must study both the microbes and the DOM in tandem. Historically, this has not often transpired, likely due to disciplinary boundaries and the diverse expertise necessary to design successful experiments that address inherent complexities on both the biological and chemical sides of the story. In the rare cases in which such interdisciplinary studies have been performed, novel insight as to how microbes produce and consume marine DOM has been gained. For example, studies in which commercially available organic compounds are added to seawater to evaluate effects on microbial diversity have become commonplace and, although they unfortunately yield little insight into *how* microbes consume DOM (e.g. Cottrell & Kirchman 2000, Nelson

& Carlson 2012, Gómez-Consarnau et al. 2012), they have been instrumental in constructing a paradigm regarding trophic strategies among different heterotrophic bacterioplankton (Tang et al. 2012). The recent addition of gene and protein expression analyses to these experiments is providing a first glimpse at the physiology behind the consumption of specific DOM components (Vila-Costa et al. 2010, Mou et al. 2011), sources (Poretsky et al. 2010, Rinta-Kanto et al. 2011, Teeling et al. 2012), and size fractions (McCarren et al. 2010).

A handful of studies have employed cultured isolates of marine phytoplankton to evaluate the chemical composition of DOM from a known biological source, revealing information on DOM lability and production dynamics (Aluwihare et al. 1999, Bertilsson et al. 2005, Grossart & Simon 2007, Barofsky et al. 2009), and identifying novel metabolites (Baran et al. 2011). Three studies have looked specifically at interactions between DOM and SAR11, the most abundant marine bacterioplankton clade, and demonstrate the potential for identifying DOM source markers (Kujawinski et al. 2009), trophic strategies (Malmstrom et al. 2005), and environmental adaptations (Gómez-Pereira et al. 2012) of specific heterotrophic bacterioplankton taxa. The degradation of a few specific organic compounds have been tracked in-situ using molecular techniques, revealing succession patterns within individual heterotrophic functional groups and highlighting the extreme complexity of microbe-DOM interactions in nature (Lau et al. 2007, Watanabe et al. 2012).

Collectively, this body of research has identified many connections between bacterioplankton and the consumption of DOM from various sources and a few

connections between marine phytoplankton and DOM composition, however many questions remain. While the role of marine DOM as a link between autotrophic and heterotrophic microorganisms from a production standpoint is well established (Baines & Pace 1991, Duarte & Cebrian 1996, Carlson 2002), the role of marine DOM as a potential link between autotrophic and heterotrophic microbial community structure remains less understood. This is due in large part to a dearth of information regarding the composition of DOM produced by different types of marine phytoplankton. Although it is widely assumed that distinct marine phytoplankton groups produce DOM of varying composition, this has never been demonstrated at the molecular level and the role of producer phylogeny in determining DOM composition is currently unknown. Without this information, the extent to which DOM composition might influence the microbial community structure of heterotrophic assemblages remains a mystery (Figure 2). Direct connections between individual heterotrophic bacterioplankton taxa and consumption of DOM from a known biological source are rare and the physiology of DOM consumption is not well described for most oligotrophic taxa.

Finally, the research cited above is largely focused on high productivity regions (e.g. coastal environments, episodic blooms) and their indigenous organisms (e.g. diatoms, copiotrophic bacterioplankton), and is therefore not representative of systems found in pelagic, oligotrophic environments that dominate the globe (Longhurst et al. 1995). Additional studies are needed that provide more detail regarding how DOM composition and lability relate to the diversity and physiology of oligotrophic microbes (Figure 3). Robust, reproducible connections between specific microbes and naturally-

derived DOM that are worthy of additional investigation need to be identified in order to develop model systems of DOM production and consumption that can aid in modeling and understanding biogeochemical cycles in the sea.

Advantages of a cultivation-based approach

The use of pure cultures in studies of DOM production and consumption is an excellent approach for acquiring the detailed, robust connections described above. Techniques for the isolation and cultivation of marine phytoplankton and heterotrophic bacterioplankton from oligotrophic environments are now well established (Becker et al. 2007, Andersen 2005, Stingl et al. 2007) and have produced isolates of some of the most ubiquitous microbes in the sea (e.g. Rappé et al. 2002, Moore et al. 2007). These isolates are excellent candidates for studies regarding DOM cycling in oligotrophic environments. Pure cultures allow for the direct analysis of DOM that is produced and/or consumed by a single organism, providing unequivocal evidence of microbial sources and sinks of marine DOM. DOM derived from pure phytoplankton cultures can be obtained and analyzed in the absence of any heterotrophic activity, so even the most labile material can be studied. Analyzing how pure cultures of heterotrophic bacterioplankton alter their transcriptome when supplied with DOM has the potential to provide more detail on the microbial physiology of DOM consumption. The goal of this thesis is to use cultivation-based approaches to identify direct connections between individual marine phytoplankton and the DOM they produce, and to uncover details regarding the consumption of this DOM by heterotrophic bacterioplankton. Fundamental knowledge regarding the origin,

composition, and degradation of marine DOM was acquired en route to achieving this goal. The data chapters are outlined below along with an appendix containing two relevant in-situ studies on DOM cycling conducted in the oligotrophic North Pacific Subtropical Gyre.

Chapter 2: Phytoplankton-derived dissolved organic matter composition is influenced by producer phylogeny.

The composition of DOM produced by eight pure strains of marine phytoplankton was analyzed using mass spectrometry to identify relationships between DOM composition and phytoplankton diversity. This is the first look at the molecular composition of DOM from several model marine phytoplankton, including *Prochlorococcus*, the dominant phytoplankton in oligotrophic regions.

Chapter 3: The reactivity of marine dissolved organic matter depends on its biological source and sink.

Batch cultures of axenic heterotrophic bacterioplankton were supplemented with bulk DOM derived from seawater and DOM from individual phytoplankton strains (Chapter 2) in order to examine effects on bacterial growth rate and total cell yield. Reactive substrates for individual heterotrophic taxa were identified and assays with DOM fractions revealed a general connection between the polarity and reactivity of *Prochlorococcus*-derived DOM. This chapter demonstrates a direct carbon cycling

connection between *Prochlorococcus* and SAR11, likely the most abundant autotrophic and heterotrophic organisms on Earth.

Chapter 4: Transcriptional profiling of a marine *Oceanospirillales* reveals increased expression of genes related to anabolism and proteorhodopsin-based phototrophy, and decreased expression of genes related to motility and chemotaxis in the presence of *Prochlorococcus*-derived dissolved organic matter.

The transcriptional response of a marine gammaproteobacterium from the *Oceanospirillales* order to DOM derived from *Prochlorococcus* was monitored at several phases of growth. This application of next generation RNA-sequencing technology to a pure culture of marine bacteria degrading DOM provided an unprecedented level of detail regarding the physiological response of a heterotrophic bacterium to organic substrate from a known biological source.

Appendices: Dissolved organic matter consumption by mixed microbial assemblages.

Microcosm experiments were conducted in the North Pacific Subtropical Gyre in the fall of 2007 (Appendix A) and spring of 2010 (Appendix B) to investigate the metabolic response of oligotrophic microbial communities to elevated concentrations of DOM from specific chemical and biological sources using a metatranscriptomic approach.

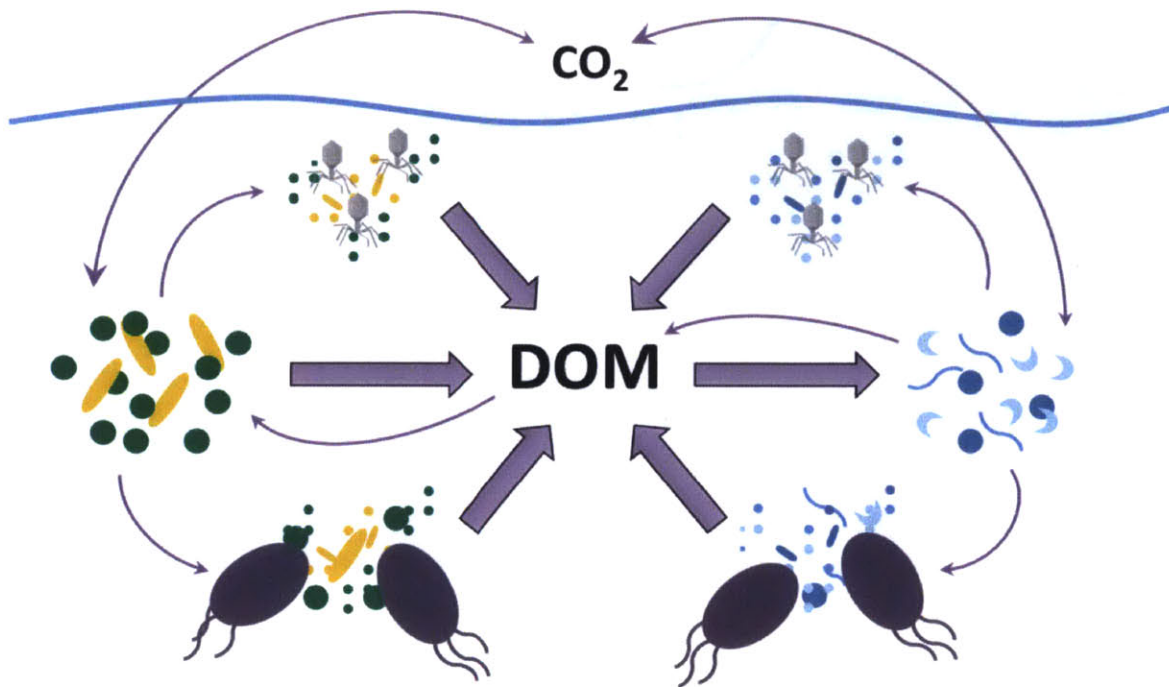


Figure 1. A simplified microbial loop highlighting processes discussed throughout this thesis. The conversion of inorganic carbon and nutrients to dissolved organic matter (DOM) by marine phytoplankton (left; green/orange cells) occurs through a variety of mechanisms including viral lysis (top), sloppy feeding by grazers (bottom) and direct release (center). Heterotrophic bacterioplankton (right; blue cells) are the primary consumers of marine DOM, converting much of this organic material back into inorganic forms and altering the DOM pool in a variety of ways.

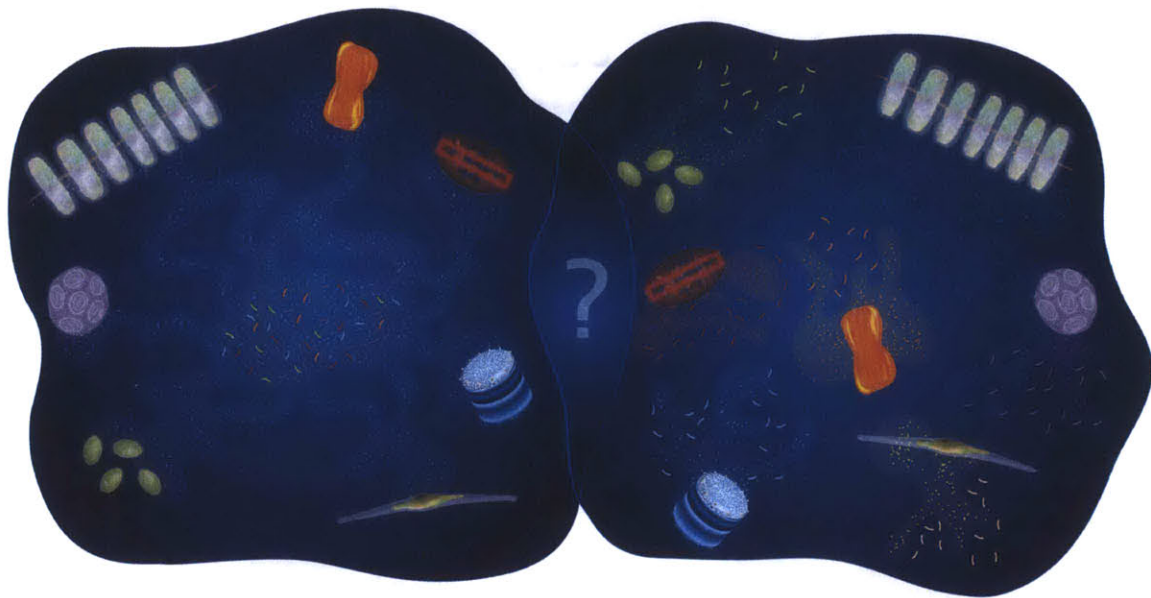


Figure 2. Hypothetical scenarios of DOM-mediated microbial interactions. In the scenario on the left, various marine phytoplankton taxa produce a highly similar suite of dissolved organic matter (DOM) as indicated by the blue smoke/dots, having little effect on the community composition of surrounding heterotrophic bacterioplankton (shown in the middle). In the scenario on the right, each phytoplankton type is producing and releasing unique “flavors” into the DOM pool, some of which are preferentially utilized by specific heterotrophic taxa. These opposing scenarios represent extremes, however if reality is more akin to the scenario on the right, then autotrophic diversity would directly influence heterotrophic diversity via a DOM link. Image by Amy Caracappa-Qubeck, WHOI graphics department.

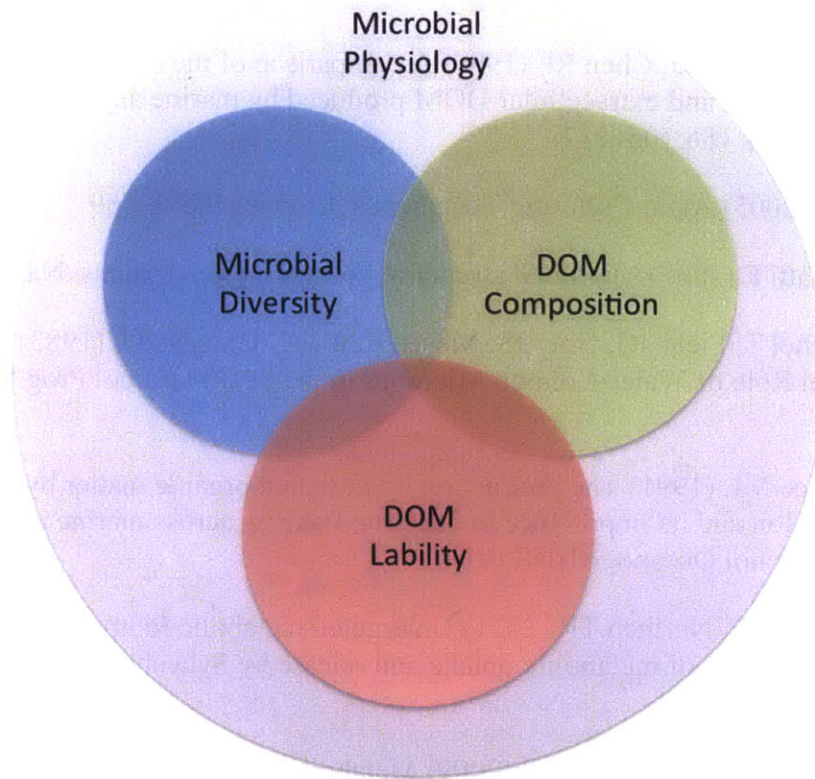


Figure 3. Research areas addressed in this thesis. Relationships between the diversity of marine microbes, the composition of marine dissolved organic matter, and the lability of this complex substrate were examined using both cultivation-based (chapters 2, 3, and 4) and cultivation-independent approaches (appendices A and B). Particular focus was placed on uncovering the underlying microbial physiology pertinent to these relationships in chapter 4.

References

- Aluwihare LI, Repeta DJ, Chen RF (1999) A comparison of the chemical characteristics of oceanic DOM and extracellular DOM produced by marine algae. *Marine Ecology Progress Series* 186:105–117
- Andersen RA (2005) *Algal Culturing Techniques*. Elsevier Inc. 1–589
- Azam F, Malfatti F (2007) Microbial structuring of marine ecosystems. *Nature*:782–791
- Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA, Thingstad F (1983) The Ecological Role of Water-Column Microbes in the Sea. *Mar Ecol Prog Ser* 10:257–263
- Baines SB, Pace ML (1991) The production of dissolved organic matter by phytoplankton and its importance to bacteria: Patterns across marine and freshwater systems. *Limnol Oceanogr*:1078–1090
- Baran R, Bowen BP, Northen TR (2011) Untargeted metabolic footprinting reveals a surprising breadth of metabolite uptake and release by *Synechococcus* sp. PCC 7002. *Mol BioSyst* 7:3200
- Barofsky A, Vidoudez C, Pohnert G (2009) Metabolic profiling reveals growth stage variability in diatom exudates. *Limnol Oceanogr-Meth* 7:382–390
- Becker JW, Brandon ML, Rappé MS (2007) Cultivating microorganisms from dilute aquatic environments: melding traditional methodology with new cultivation techniques and molecular methods. In: Hurst CJ (ed) *Manual of Environmental Microbiology*, 3rd edn. ASM Press, p 399–406
- Benner R, Pakulski JD, McCarthy M, Hedges JI, Hatcher PG (1992) Bulk chemical characteristics of dissolved organic matter in the ocean. *Science* 255:1561–1564
- Bertilsson S, Jones JB Jr (2003) 1 - Supply of Dissolved Organic Matter to Aquatic Ecosystems: Autochthonous Sources. In: Findlay SEG, Sinsabaugh RL (eds) *Aquatic Ecosystems*. Academic Press, Burlington, p 3–24
- Bertilsson S, Berglund O, Pullin MJ, Chisholm SW (2005) Release of dissolved organic matter by *Prochlorococcus*. *Vie et Milieu* 55:225–232
- Bjornsen PK (1988) Phytoplankton exudation of organic matter: Why do healthy cells do it? *Limnol Oceanogr*:151–154
- Carlson CA (2002) Production and removal processes. In: Hansell DA, Carlson CA (eds) *Biogeochemistry of Marine Dissolved Organic Matter*. Academic Press, p 91–151

- Carlson CA, Ducklow HW (1996) Growth of bacterioplankton and consumption of dissolved organic carbon in the Sargasso Sea. *Aquat Microb Ecol* 10:69–85
- Chavez FP, Messié M, Pennington JT (2011) Marine Primary Production in Relation to Climate Variability and Change. *Annu Rev Marine Sci* 3:227–260
- Chin WC, Orellana MV, Verdugo P (1998) Spontaneous assembly of marine dissolved organic matter into polymer gels. *Nature* 391:568–572
- Cole JJ, Likens GE, Strayer DL (1982) Photosynthetically produced dissolved organic carbon: An important carbon source for planktonic bacteria. *Limnol Oceanogr*:1080–1090
- Cottrell MT, Kirchman DL (2000) Natural assemblages of marine proteobacteria and members of the Cytophaga-Flavobacter cluster consuming low-and high-molecular-weight dissolved organic matter. *Applied and Env. Microbiology* 66:1692–97
- Covert JS, Moran MA (2001) Molecular characterization of estuarine bacterial communities that use high-and low-molecular weight fractions of dissolved organic carbon. *Aquat Microb Ecol* 25:127–139
- Dittmar T, Paeng J (2009) A heat-induced molecular signature in marine dissolved organic matter. *Nature Geosci* 2:175–179
- Druffel E, Williams P, Bauer J, Ertel J (1992) Cycling of dissolved and particulate organic matter in the open ocean. *J Geophys Res* 97:15–639–15–659
- Duarte CM, Cebrian J (1996) The fate of marine autotrophic production. *Limnol Oceanogr* 41:1758–1766
- Duarte CM, Regaudie-de-Gioux A, Arrieta JM, Delgado-Huertas A, Agustí S (2013) The Oligotrophic Ocean Is Heterotrophic. *Annu Rev Marine Sci* 5:551–569
- Fogg GE (1983) The Ecological Significance of Extracellular Products of Phytoplankton Photosynthesis. *Botanica Marina* 26:3–14
- Field CB, Behrenfeld MJ, Randerson JT, Falkowski P (1998) Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science* 281:237–240
- Fogg GE (1977) Excretion of organic matter by phytoplankton. *Limnol Oceanogr*; (United States) 22
- Gómez-Consarnau L, Lindh MV, Gasol JM, Pinhassi J (2012) Structuring of bacterioplankton communities by specific dissolved organic carbon compounds. *Environmental Microbiology* 14:2361–2378

- Gómez-Pereira PR, Hartmann M, Grob C, Tarran GA, Martin AP, Fuchs BM, Scanlan DJ, Zubkov MV (2012) Comparable light stimulation of organic nutrient uptake by SAR11 and *Prochlorococcus* in the North Atlantic subtropical gyre. *The ISME Journal*
- Grossart HP, Simon M (2007) Interactions of planktonic algae and bacteria: effects on algal growth and organic matter dynamics. *Aquat Microb Ecol* 47:163
- Hansell DA, Carlson CA (1998) Deep-ocean gradients in the concentration of dissolved organic carbon. *Nature* 395:263–266
- Hansell DA, Carlson CA, Repeta DJ, Schlitzer R (2009) Dissolved organic matter in the ocean: a controversy stimulates new insights.
- Hedges JJ (1992) Global biogeochemical cycles: progress and problems. *Marine Chemistry* 39:67–93
- Hedges JJ, Oades JM (1997) Comparative organic geochemistries of soils and marine sediments. *Organic Geochemistry* 27:319–361
- Jiao N, Herndl GJ, Hansell DA, Benner R, Kattner G, Wilhelm SW, Kirchman DL, Weinbauer MG, Luo T, Chen F, Azam F (2010) Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. *Nature Publishing Group* 8:593–599
- Kujawinski EB (2011) The Impact of Microbial Metabolism on Marine Dissolved Organic Matter. *Annu Rev Marine Sci* 3:567–599
- Kujawinski EB, Longnecker K, Blough NV, Del Vecchio R, Finlay L, Kitner JB, Giovannoni SJ (2009) Identification of possible source markers in marine dissolved organic matter using ultrahigh resolution mass spectrometry. *Geochimica et Cosmochimica Acta* 73:4384–4399
- Lau WWY, Keil RG, Armbrust EV (2007) Succession and Diel Transcriptional Response of the Glycolate-Utilizing Component of the Bacterial Community during a Spring Phytoplankton Bloom. *Applied and Environmental Microbiology* 73:2440–2450
- Loiselle S, Vione D, Minero C, Maurino V, Tognazzi A, Dattilo AM, Rossi C, Bracchini L (2012) Chemical and optical phototransformation of dissolved organic matter. *Water Research* 46:3197–3207
- Longhurst A, Sathyendranath S, Platt T, Caverhill C (1995) An estimate of global primary production in the ocean from satellite radiometer data. *Journal of Plankton Research* 17:1245–1271

- Mague T, Friberg E, Hughes D, Morris I (1980) Extracellular release of carbon by marine phytoplankton; a physiological approach. *Limnol Oceanogr*:262–279
- Malmstrom RR, Cottrell MT, Elifantz H, Kirchman DL (2005) Biomass production and assimilation of dissolved organic matter by SAR11 bacteria in the Northwest Atlantic Ocean. *Applied and Environmental Microbiology* 71:2979–2986
- McCarren J, Becker JW, Repeta DJ, Shi Y, Young CR, Malmstrom RR, Chisholm SW, DeLong EF (2010) Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. *Proc Natl Acad Sci USA* 107:16420–16427
- Moore LR, Coe A, Zinser ER, Saito MA, Sullivan MB, Lindell D, Frois-Moniz K, Waterbury J, Chisholm SW (2007) Culturing the marine cyanobacterium *Prochlorococcus*. *Limnol Oceanogr-Meth* 5:353–362
- Mopper K, Stubbins A, Ritchie JD, Bialk HM, Hatcher PG (2007) Advanced Instrumental Approaches for Characterization of Marine Dissolved Organic Matter: Extraction Techniques, Mass Spectrometry, and Nuclear Magnetic Resonance Spectroscopy. *Chem Rev* 107:419–442
- Mou X, Vila-Costa M, Sun S, Zhao W, Sharma S, Moran MA (2011) Metatranscriptomic signature of exogenous polyamine utilization by coastal bacterioplankton. *Environmental Microbiology Reports* 3:798–806
- Nagata T (2002) Organic matter-bacteria interactions in seawater (DL Kirchman, Ed.). Academic Press
- Nebbio A, Piccolo A (2012) Molecular characterization of dissolved organic matter (DOM): a critical review. *Anal Bioanal Chem*
- Nelson CE, Carlson CA (2012) Tracking differential incorporation of dissolved organic carbon types among diverse lineages of Sargasso Sea bacterioplankton. *Environmental Microbiology* 14:1500–1516
- Pomeroy LR (1974) The ocean's food web, a changing paradigm. *Bioscience*:499–504
- Pomeroy LR, Williams PJL, Azam F, Hobbie JE (2007) The Microbial Loop. *Oceanography* 20:28–33
- Poretsky RS, Sun S, Mou X, Moran MA (2010) Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon. *Environmental Microbiology* 12:616–627

- Rappé MS, Connon SA, Vergin KL, Giovannoni SJ (2002) Cultivation of the ubiquitous SAR 11 marine bacterioplankton clade. *Nature* 418:630–633
- Rinta-Kanto JM, Sun S, Sharma S, Kiene RP, Moran MA (2011) Bacterial community transcription patterns during a marine phytoplankton bloom. *Environmental Microbiology* 14:228–239
- Romera-Castillo C, Sarmiento H, Alvarez-Salgado XA, Gasol JM, Marrase C (2011) Net Production and Consumption of Fluorescent Colored Dissolved Organic Matter by Natural Bacterial Assemblages Growing on Marine Phytoplankton Exudates. *Applied and Environmental Microbiology* 77:7490–7498
- Sarmiento H, Gasol JM (2012) Use of phytoplankton-derived dissolved organic carbon by different types of bacterioplankton. *Environmental Microbiology* 14:2348–2360
- Stingl U, Tripp HJ, Giovannoni SJ (2007) Improvements of high-throughput culturing yielded novel SAR11 strains and other abundant marine bacteria from the Oregon coast and the Bermuda Atlantic Time Series study site. *The ISME Journal*
- Tada Y, Taniguchi A, Nagao I, Miki T, Uematsu M, Tsuda A, Hamasaki K (2011) Differing Growth Responses of Major Phylogenetic Groups of Marine Bacteria to Natural Phytoplankton Blooms in the Western North Pacific Ocean. *Applied and Environmental Microbiology* 77:4055–4065
- Tang K, Jiao N, Liu K, Zhang Y, Li S (2012) Distribution and Functions of TonB-Dependent Transporters in Marine Bacteria and Environments: Implications for Dissolved Organic Matter Utilization. *PLoS ONE* 7:e41204
- Teeling H, Fuchs BM, Becher D, Klockow C, Gardebrecht A, Bennke CM, Kassabgy M, Huang S, Mann AJ, Waldmann J, Weber M, Klindworth A, Otto A, Lange J, Bernhardt J, Reinsch C, Hecker M, Peplies J, Bockelmann FD, Callies U, Gerdt G, Wichels A, Wiltshire KH, Glockner FO, Schweder T, Amann R (2012) Substrate-Controlled Succession of Marine Bacterioplankton Populations Induced by a Phytoplankton Bloom. *Science* 336:608–11
- Trabelsi A, Rassoulzadegan F (2011) Effect of bacterial community dynamics on DOC seasonal changes in the north-western Mediterranean Sea. *Journal of Plankton Research* 33:1249–1262
- Vila-Costa M, Rinta-Kanto JM, Sun S, Sharma S, Poretsky R, Moran MA (2010) Transcriptomic analysis of a marine bacterial community enriched with dimethylsulfoniopropionate. *The ISME Journal* 4:1410–1420

- Watanabe K, Komatsu N, Kitamura T, Ishii Y, Park H-D, Miyata R, Noda N, Sekiguchi Y, Satou T, Watanabe M, Yamamura S, Imai A, Hayashi S (2012) Ecological niche separation in the Polynucleobacter subclusters linked to quality of dissolved organic matter: a demonstration using a high sensitivity cultivation-based approach. *Environmental Microbiology* 14:2511–2525
- Williams PJ, Purdie DA (1991) In vitro and in situ derived rates of gross production, net community production and respiration of oxygen in the oligotrophic subtropical gyre of the North Pacific Ocean. *Deep Sea Research Part A Oceanographic Research Papers* 38:891–910
- Williams PJLB, Quay PD, Westberry TK, Behrenfeld MJ (2013) The Oligotrophic Ocean Is Autotrophic. *Annu Rev Marine Sci* 5:535–549

CHAPTER TWO:

Phytoplankton-derived dissolved organic matter composition is influenced by producer phylogeny

Abstract

The production of dissolved organic matter (DOM) by marine phytoplankton supplies the majority of substrate consumed by heterotrophic bacterioplankton in the sea. This production and subsequent consumption converts a vast quantity of carbon, nitrogen and phosphorus between organic and inorganic forms, directly impacting global cycles of these biologically important elements. Details regarding the chemical composition of DOM produced by different groups of marine phytoplankton are sparse, and while often assumed, it is not currently known if phylogenetically distinct groups of marine phytoplankton release characteristic suites of DOM. To investigate the relationship between phytoplankton diversity and DOM composition, hydrophobic dissolved organic matter released by eight axenic strains of marine phytoplankton (DOM_P) was characterized and compared using high-performance liquid chromatography coupled to mass spectrometry. We found the composition of DOM_P to be complex and highly strain dependent. Connections between DOM_P composition (both quality and quantity) and phylogenetic relatedness were identified on multiple levels, indicating that marine phytoplankton produce different DOM suites that tend to follow their phylogenetic relationships. Chemical information regarding the size and polarity of metabolites from known biological sources was also obtained. Our findings reveal DOM_P composition to be influenced by the phylogeny of its producer and further implicate marine DOM as a potential factor impacting microbial diversity in the sea by acting as a link between autotrophic and heterotrophic microbial community structure.

Introduction

The extracellular release of dissolved organic matter (DOM) by marine phytoplankton fuels secondary production in the sea (Pomeroy 1974, Mague et al. 1980, Baines & Pace 1991). As much as 50% of the total carbon fixed by photoautotrophs may be released into seawater as a by-product of normal metabolism or through active processes for waste removal, substrate acquisition, defense, or communication (Bjornsen 1988, Carlson 2002, Bertilsson & Jones 2003). The amount of extracellular DOM released and its composition depend on the organism and its physiological state (Mykkestad 1995, Meon & Kirchman 2001, Wetz & Wheeler 2007, Romera-Castillo et al. 2010) and additional factors including: temperature, light, growth phase, availability of inorganic nutrients and the presence of other organisms (Hirt et al. 1971, Obernosterer & Herndl 1995, Grossart & Simon 2007, Barofsky et al. 2009, Engel et al. 2011). Factors affecting DOM uptake are less clear, but recent evidence suggests that DOM resource partitioning within microbial communities may impact the ecological success and community structure of heterotrophic bacterioplankton (Poretsky et al. 2010, McCarren et al. 2010, Sarmiento & Gasol 2012). Substrate specificity may be of particular importance for heterotrophs that have evolved streamlined genomes and are therefore fine-tuned to particular resources (Giovannoni et al. 2005, 2008).

Genomic and physiological differences among marine phytoplankton likely influence the types of organic matter they produce and consequently, what substrates are available to sympatric heterotrophic communities. If different phytoplankton groups release DOM of varying composition and quality, then DOM could provide a direct link

between autotrophic and heterotrophic microbial diversity and community structure. For example, *Prochlorococcus* is frequently the dominant photoautotroph in oligotrophic surface waters and is an essential source of organic material for heterotrophic bacterioplankton in this environment (Partensky et al. 1999, Bertilsson et al. 2005). It has been suggested that DOM released by *Prochlorococcus* could be responsible for supporting 12-41% of bacterial production in tropical and subtropical regions (Bertilsson et al. 2005). If DOM released by *Prochlorococcus* is distinct in composition, metabolic specialization within associated heterotrophic populations to utilize *Prochlorococcus*-derived DOM might influence heterotrophic bacterioplankton community structure in these regions.

Axenic laboratory cultures of well-characterized marine phytoplankton provide an excellent tool for examining the biological origin and composition of even the most labile marine DOM. In this study, we examined the chemical composition of DOM released by eight marine phytoplankton strains to investigate the relationship between phylogenetic relatedness and the DOM produced by each strain. Dissolved organic matter recovered by solid phase extraction (SPE) from spent culture media (hereafter referred to as phytoplankton DOM, or DOM_p) was analyzed by high performance liquid chromatography–electrospray ionization mass spectrometry (HPLC-ESI-MS). Similar untargeted metabolomics approaches have been successfully employed to study many aspects of microbial interactions with DOM (Kujawinski et al. 2009, Barofsky et al. 2010, Baran et al. 2011), but none have yet examined the relationship between DOM composition and the taxonomy of the source organisms. We obtained and analyzed a

complex data set comprised of hundreds to thousands of metabolites from each organism tested to compare features produced by different phytoplankton strains. Here, we report on how phytoplankton phylogeny affects DOM composition at the molecular level.

Materials and Methods

Phytoplankton cultures

Eight species of marine phytoplankton were selected for this study to investigate variability in the composition of organic matter produced and released by unicellular photoautotrophs at various taxonomic levels. Individual strains were chosen on the criteria that a published sequenced genome is or will soon be available and that axenic strains were capable of growth to high cell densities (ca. 10^7 cells/ml) in natural seawater-based media. All media was produced by the National Center for Marine Algae and Microbiota (NCMA) using filtered, autoclaved oceanic surface water collected from the Sargasso Sea. Specific media amendments were added aseptically after sterilization. Triplicate 1 L cultures of each strain were prepared alongside triplicate controls containing all media amendments but no cell additions. Cells were pre-conditioned in 60 to 400 ml of the appropriate medium prior to 1 L inoculation. *Thalassiosira pseudonana* (CCMP str. 1335), *Thalassiosira rotula* (CCMP str. 1647), *Phaeodactylum tricornerutum* (CCMP str. 632), and both *Synechococcus* strains (WH8102/CCMP2370 and WH7803/CCMP1334) were grown at the NCMA in L1 medium prepared according to existing protocols (Guillard & Hargraves 1993). Diatom strains were grown at 20 °C and *Synechococcus* strains at 24 °C. All five strains were grown on a 13/11 light/dark cycle

with gentle periodic mixing. *Synechococcus* str. WH7803 was grown under lower light conditions at 10-20 $\mu\text{E}/\text{m}^2/\text{s}$, while the rest were grown at 100-120 $\mu\text{E}/\text{m}^2/\text{s}$. Diatom cells were monitored for growth using a Palmer-Maloney counting chamber at 40X magnification on a Zeiss light microscope. *Synechococcus* cells were monitored for growth via epifluorescence microscopy. Three *Prochlorococcus* strains (MED4, MIT9301, and MIT9313) were grown in Pro99 medium prepared by the NCMA according to existing protocols (Moore et al. 2002, 2007) with the addition of 5 mM sodium bicarbonate at 22 °C. All three strains were grown on a 14/10 light/dark cycle. *Prochlorococcus* str. MIT9313 was grown under lower light conditions at ca. 20 $\mu\text{E}/\text{m}^2/\text{s}$, while MIT9301 and MED4 were grown at ca. 40 $\mu\text{E}/\text{m}^2/\text{s}$. Cultures were monitored for growth using total fluorescence and flow cytometry and for changes in pH, which were mitigated by additions of 2 mM HEPES buffer and 5 mM sodium bicarbonate 6 days after inoculation. Samples for direct cell counts by flow cytometry were fixed with 0.125% final concentration of grade I glutaraldehyde (Sigma) and stored at -80°C prior to analysis. Fixed samples were diluted in 0.2 μm filtered seawater and *Prochlorococcus* were enumerated using an Influx Cell Sorter (BD Biosciences) as previously described (Olson et al. 1985, Cavender-Bares et al. 1999).

Quantification of organic matter production

Samples for cell abundance and particulate and dissolved organic carbon (POC and DOC respectively) were taken at the onset of stationary phase growth to evaluate the organic carbon production of each strain. All organic carbon samples were processed

using combusted (450 °C for 8 h) glassware. Duplicate samples for POC analysis were taken by vacuum filtering 25 ml of sample onto a combusted 25 mm 0.7 µm glass fiber filter (Whatman GF/F). Filters were then placed inside a combusted glass petri dish, wrapped in foil, and immediately frozen. A blank filter was also prepared at each sampling point. Filters were later thawed and dried (60 °C; overnight) before encapsulation into 9x10 mm tin capsules. Measurements were performed by the University of California Davis Stable Isotope Facility. POC filtrates were transferred into combusted glass vials, acidified with 150 µl of a 25% phosphoric acid solution and measured for DOC using high temperature catalytic oxidation on a Shimadzu TOC-V_{CSH} with platinumized aluminum catalyst. Sample concentrations were determined alongside potassium hydrogen phthalate standards and consensus reference materials provided by the DOC-CRM program (www.rsmas.miami.edu/groups/biogeochem/CRM.html).

Extraction and analysis of DOM_p

Upon entering stationary phase, cells were removed by centrifugation at either 878 RCF (diatoms) or 10,751 RCF (cyanobacteria) for 15 min, followed by gentle filtration (0.1 µm; Whatman Polycap 36 TC capsule filter). Filtrates were then stored briefly in the dark at 4 °C until SPE. Media controls were processed and stored alongside the culture samples. Filtrates were acidified to pH 2-3 by adding trace metal grade hydrochloric acid and organic matter was extracted onto ISOLUTE C18(EC) SPE columns (0.5g, Biotage) at a rate of 1 ml/min. SPE columns were preconditioned with 5 ml HPLC-grad methanol followed by 10 ml ultrapure water. After sampling, mineral

salts were washed from the columns with acidified ultrapure water (pH 2-3) at a flow rate of 1 ml/min. Organic matter was recovered with gravity elution using 10 column volumes of acidified HPLC-grade methanol (pH 2-3). Samples were concentrated to a small volume by rotary evaporation, and then taken to dryness under filtered, high purity nitrogen. Samples were resuspended in 156 μ l of methanol, and stored briefly in combusted amber vials at 4 °C in the dark prior to chemical analysis. A 3 μ l subsample of each was placed onto a combusted 25 mm, 0.7 μ m glass fiber filter (Whatman GF/F) and submitted for POC analysis to quantify the carbon recovered by SPE.

Chromatographic and spectrometric analyses were performed using an Agilent 1200 series liquid chromatograph coupled to an Agilent 6130 (single quadrupole) mass spectrometer. Organic extracts were separated on a ZORBAX SB-C18 column (Agilent; 3.5 μ m 4.6x150 mm) eluted at 0.8 ml min⁻¹ using a linear gradient (% solvent A, % solvent B, minutes): 100, 0, 0; 20, 80, 31.25; 0, 100, 43.75; 0, 100, 64, where solvent A is aqueous formic acid (0.1%) and solvent B is methanolic formic acid (0.1%). Mass spectrometry was performed using an atmospheric electrospray ionization source. Drying gas was set at 11.5 l min⁻¹ and 300 °C, the nebulizer was at 60 psig, and capillary voltage was set to 4000 V. Data was obtained in the positive ion mode from 100-2000 *m/z* with a 4.0 fragmentor, 150 threshold, and 0.1 step size.

Feature identification and comparison

Positive mode mass spectral data was analyzed using MZmine 2 molecular profiling software (Pluskal et al. 2010). Ions with signal intensity at least 5-fold greater

than the maximum noise level were detected using a centroid mass detector.

Chromatograms were then built from the raw data using a retention time tolerance of ± 5 s, and a mass tolerance of ± 0.3 m/z . Alignments were made to correct for small shifts in retention time between samples and individual peaks were identified by setting a minimum acceptable intensity (5-fold greater than the noise level) and duration (5 s) to remove noise within each chromatogram and also by searching for local minima within each chromatogram. Sample feature lists were then aligned and gaps were filled in using a secondary threshold of 3-fold greater than the noise to identify common peaks that fell just below the initial strict threshold. A feature was defined as a unique m/z at a specific retention time, thus multiple features could share a specific m/z or retention time, but never both.

Biological triplicates of each strain were first compared against triplicate media controls to distinguish material produced by the organism from any background material, including DOM present in the seawater medium used to cultivate each strain. Features were considered to be derived from a particular organism only if they were present in all 3 biological replicates of that strain and absent in all 3 replicate controls of the appropriate media type. The removal of features present in sterile media controls reduces the possibility of correlation effects due to variations in growth media (i.e. Pro99 medium vs. L1 medium). Features were also removed if they were present in any of the blank samples tested, which included triplicate instrument blank injections of pure methanol and triplicate processing blanks that were created by rinsing and eluting pre-cleaned resins without any prior sample loading. Feature lists for each strain were then aligned to

identify both common and unique features among the 8 strains tested. Aligned feature lists were also used to investigate intensity level differences among features common to multiple samples.

To provide a rigorous comparison of the similarity between all cultures, the aligned feature list was converted into a binary matrix where each column represented an individual culture and each row a feature present in at least one culture. Multiplying this matrix by its transpose created a second matrix where each element was the number of features common to each pairwise culture comparison. The matrix was then normalized by the total number of features present in each pairwise comparison. This normalized similarity matrix was used to quantify the similarity of feature lists and their parent strains and to generate cluster diagrams as a visual representation of DOM_p similarity among strains.

Results

All eight of the phytoplankton strains tested yielded hundreds of extracellular hydrophobic metabolites recovered by SPE (Table 1). The majority of these metabolites were consistently associated with only one of the strains tested, while a relatively small subset appeared to be produced by a diverse range of organisms. Of the 2,426 culture-derived features found in this study ranging from 101-1,481 *m/z*, 1,972 (81%) were unique to a particular strain while only 6 (0.25%) were produced by all eight strains.

Metabolite composition also varied among biological replicates (Table 1). Variation among replicates was found in all strains tested and was potentially due to

uncontrollable differences in growth conditions and/or slight variations in growth stage at the time of processing. On average, 41% of the features for a given set of triplicates were found in all three replicates, while 24% were found in 2 out of the 3 replicates and the remaining 35% were unique to one replicate. On average, 60% of the features in a given replicate were also present in the other two replicates for that organism. Features common to all replicates for a given strain were compared to other strains to stringently identify both common and unique features. Features that were present in any replicate sample were also used to generate a similarity matrix for running all possible pairwise comparisons to conduct a parallel investigation of DOM composition using a distinct approach.

Prochlorococcus (clade-level)

Metabolites derived from three strains of *Prochlorococcus* (MIT9313, MIT9301 and MED4) representing different clades were each analyzed in triplicate. MIT9313, MIT9301 and MED4 reached an average maximum cell yield of 6.64×10^7 cells/ml, 2.71×10^8 cells/ml and 1.55×10^8 cells/ml respectively over a growth period of 12 days and yielded between 0.6 – 1.2 mM POC (Table 2).

HPLC-MS analysis of triplicate MIT9313 cultures revealed the production and release of 693 features when compared to sterile controls. 148 (21%) features were common to all three replicates and accounted for 39 - 42% of the features found in each replicate (Fig. 1a). HPLC-MS analysis of triplicate MIT9301 cultures revealed the production and release of 765 features when compared to sterile controls. 253 (33%)

features were common to all three replicates and accounted for 48 - 53% of the features found in each replicate. HPLC-MS analysis of duplicate MED4 cultures revealed the production and release of 525 features when compared to sterile controls. Purity broth tests revealed subsequent heterotrophic contamination in one of the MED4 replicate cultures; therefore this sample was removed from all analyses. 155 (29%) features were common to the two remaining replicates and accounted for 40 - 53% of the features found in each replicate. *Prochlorococcus*-derived features ranged in mass from 101-1,377 *m/z*, with the majority falling between 150-400 *m/z*. Most features were detected between 20-50 min when the mobile phase composition was between 51 and 100% methanol (Fig. 2).

Of the total *Prochlorococcus*-derived features detected in this study, 54 (15%) were found associated with all three strains. High-light adapted strains MIT9301 and MED4 were the most similar strains, sharing 105 (35%) of their features, followed by MIT9313 (low-light adapted) and MED4 sharing 62 (26%) features and finally MIT9313 and MIT9301 sharing 75 (23%) features (Fig. 3a). All three strains exhibited unique features that were not found in the other two strains; MIT9301 having the most with 127 unique features, or 50.2% of its total. Correlation analysis revealed that metabolite variation was generally greater between the strains (0.13-0.39) than among triplicates within any of the three strains tested (0.27-0.65), although four *Prochlorococcus* samples (MED4B, MED4C, 9301B and 9313B) had greater correlations with a *Prochlorococcus* replicate from another strain than with one of their own replicates (Table S1).

Synechococcus (clade-level)

Metabolites derived from two strains of *Synechococcus* representing different clades (WH8102/CCMP2370 and WH7803/CCMP1334) were analyzed in triplicate. WH8102 and WH7803 reached an average maximum cell yield of 1.14×10^8 cells/ml and 3.66×10^7 cells/ml respectively over a growth period of 22 days. WH8102 produced 1.4 – 1.6 mM POC and 437 - 512 μ M extracellular DOC, while WH7803 produced 0.6 – 0.7 mM POC and 168 - 180 μ M extracellular DOC (Table 2). On average, the two *Synechococcus* strains released 22% of the total fixed carbon as DOC, or 4.5 fmol per cell. Of the total DOC released, an average of 4% was recovered by SPE.

HPLC-MS analysis of triplicate WH8102 cultures revealed the production and release of 893 features when compared to sterile controls. 449 (50%) features were common to all three replicates and accounted for 62 - 75% of the features found in each replicate. HPLC-MS analysis of triplicate WH7803 cultures revealed the production and release of 504 features when compared to sterile controls. 186 (37%) features were common to all three replicates and accounted for 50 – 64% of the features found in each replicate (Fig. 1b). *Synechococcus*-derived features ranged from 107-1,377 m/z , with the majority falling between 150-500 m/z . Most features were detected between 12-50 min when the mobile phase was between 31 and 100% methanol (Fig. 2).

Of the total *Synechococcus*-derived features detected in this study, 77 (14%) were found associated with both strains. Both strains produced unique features that were not found in the other strain; WH8102 had 372 unique features, or 83% of its total and WH7803 had 109 unique features, or 59% of its total (Fig. 3b). Correlation analysis

shows that metabolite variation was greater between the strains (0.13-0.21) than among triplicates within either of the strains tested (0.42-0.72) (Table S1).

Cyanobacteria comparison (genus-level)

Of the 54 and 77 total features consistently produced and released by *Prochlorococcus* and *Synechococcus* respectively, 17 (15%) were found associated with all five cyanobacteria strains. The most similar strains between the two groups were *Prochlorococcus* str. MIT9301 and *Synechococcus* str. WH8102, sharing 114, or 16% of their features. All *Prochlorococcus* strains shared more features with each other than with either *Synechococcus* strain. *Synechococcus* str. WH7803 shared more features with the other *Synechococcus* strain tested (WH8102) than with any of the *Prochlorococcus* strains; however the reverse was not true. *Synechococcus* str. WH8102 had the most overlap with *Prochlorococcus* str. MIT9301, sharing 16%, followed by MIT9313 (13%) and then WH7803 (12%). Low-light adapted *Synechococcus* str. WH7803 was more similar to the low-light adapted *Prochlorococcus* str. MIT9313 than it was to either of the high-light adapted *Prochlorococcus* strains. Percent similarities are summarized in Table S2.

Correlation analysis of DOM compositions confirmed that *Synechococcus* str. WH8102 and *Prochlorococcus* str. MIT9301 were the most related *Synechococcus* and *Prochlorococcus* strains (0.15-0.21) with respect to the DOM features analyzed, and also revealed that MIT9301 and WH7803 were the least related cyanobacteria tested (0.07-0.13). In general, all of the *Prochlorococcus* strains were more similar to WH8102 (0.10-

0.21) than WH7803 (0.05-0.13) and WH8102 was less like MED4 (0.10-0.17) than the other two *Prochlorococcus* strains (0.13-0.21) (Table S1).

Thalassiosira sp. (species-level)

Metabolites derived from two centric diatoms of the *Thalassiosira* genus were analyzed in triplicate. *T. pseudonana* (CCMP1335) and *T. rotula* (CCMP1647) were chosen to represent different species of the *Thalassiosira* genus. *T. pseudonana* and *T. rotula* reached an average maximum cell yield of 1.39×10^7 cells/ml and 1.72×10^7 cells/ml respectively over a growth period of 19 days. *T. pseudonana* produced 2.5 – 2.6 mM POC and 240 – 325 μ M extracellular DOC, while *T. rotula* produced 1.5 – 1.8 mM POC and 358 – 512 μ M extracellular DOC (Table 2). On average, the two *Thalassiosira* strains released 15% of the total fixed carbon as DOC, or 22.9 fmol of DOC per cell. Of the total DOC released, an average of 9% was obtained via SPE.

HPLC-MS analysis of triplicate *T. pseudonana* cultures revealed the production and release of 1,226 features when compared to sterile controls. 665 (54%) features were common to all three replicates and accounted for 63 - 82% of the features found in each replicate. HPLC-MS analysis of triplicate *T. rotula* cultures revealed the production and release of 1,269 features when compared to sterile controls. 484 (38%) features were common to all three replicates and accounted for 43 – 79% of the features found in each replicate. *Thalassiosira*-derived features ranged from 104-1,417 m/z , with the majority falling between 150-700 m/z . Most features were detected between 20-45 min when the mobile phase went from 51 to 100% methanol (Fig. 2).

Of the total *Thalassiosira*-derived features detected in this study, 108 (10%) were found associated with both strains. Both strains produced many unique features that were not found in the other strain; *T. pseudonana* had 557 unique features, or 84% of its total and *T. rotula* had 376 unique features, or 78% of its total. Correlation analysis revealed that metabolite variation was greater between the strains (0.11-0.14) than among triplicates within either of the strains tested (0.44-0.70) (Table S1).

Phaeodactylum tricornutum

Metabolites derived from the pennate diatom *Phaeodactylum tricornutum* (CCMP632) were analyzed in triplicate. *P. tricornutum* reached an average maximum cell yield of 9.32×10^6 cells/ml over a growth period of 19 days and produced 2.8 mM POC and 147 – 189 μ M extracellular DOC (Table 2). On average, *P. tricornutum* released 6% of the total fixed carbon as DOC, or 18.5 fmol of DOC per cell. Of the total DOC released, an average of 23% was obtained via SPE.

HPLC-MS analysis of triplicate *P. tricornutum* cultures revealed the production and release of 1,443 features when compared to sterile controls. 845 (59%) features were common to all three replicates and accounted for 63 - 79% of the features found in each replicate (Fig. 1c). *Phaeodactylum*-derived features ranged from 135-1,481 *m/z*, with the majority falling between 150-1,000 *m/z*. Most features were detected between 12-45 min when the mobile phase went from 31 to 100% methanol (Fig. 2). Correlation analysis revealed that *P. tricornutum* triplicates were more similar to each other (0.66-0.73) than to any other sample (0.02-0.09).

Diatom comparison (order-level)

In general, DOM_P composition was more diverse among the diatom strains than among the cyanobacteria. Of the 1,782 total features consistently produced and released by the diatom strains tested, 32 (2%) were found associated with all three strains. *T. pseudonana* and *T. rotula* were the most similar diatoms, sharing 108 (10%) of their features, followed by *T. pseudonana* and *P. tricornutum* sharing 75 (5%) features and finally *T. rotula* and *P. tricornutum* sharing 61 (5%) features (Fig. 3c). All three strains produced a large quantity of unique features that were not found in the other two strains; *P. tricornutum* having the most with 741 unique features, or 88% of its total. *T. pseudonana* and *T. rotula* were both more similar to each other than to any of the cyanobacteria strains tested and *P. tricornutum* was more similar to both *Thalassiosira* diatoms than to any cyanobacteria strain. Correlation analysis revealed that metabolite variation was greater between the three diatom strains (0.06-0.14) than among triplicates within any of the three strains tested (0.44-0.73) and confirmed that *T. pseudonana* and *T. rotula* are the most similar diatoms (0.11-0.14), followed by *T. pseudonana* and *P. tricornutum* (0.07-0.09) and finally *T. rotula* and *P. tricornutum* (0.06-0.08) (Table S1).

Domain-level comparisons

The 2,426 DOM_P features detected in this study were organized into 3 groups: produced by all *Prochlorococcus* strains, produced by all *Synechococcus* strains, and produced by all diatom strains. Of the 138 features that can be classified this way, only 6 (4%) were common to all 3 groups. *Prochlorococcus* and *Synechococcus* were the most

similar groups, sharing 17 (15%) of their features, followed by diatoms and *Prochlorococcus* sharing 8 (10%) features and finally diatoms and *Synechococcus* sharing 6 (6%) features (Fig. 3d). All 3 groups produced a large quantity of features unique to that group; the *Synechococcus* group having the most with 60 unique features, or 78% of its total. Combining the *Prochlorococcus* and *Synechococcus* groups into 1 cyanobacteria group and comparing this with the diatom group reduced the number of total features to 43. The same 6 features (14%) mentioned above were common to both groups, while 11 features (26%) were unique to the cyanobacteria group and 26 features (60%) were unique to the diatom group.

Correlation analysis of individual cyanobacteria and diatom strains revealed that all five cyanobacteria strains were more similar to one another (0.07-0.39) than to any diatom strain tested (0.02-0.12). Comparisons between cyanobacteria and diatom strains showed *Prochlorococcus* str. MIT9301 and *T. rotula* to be the most similar, sharing 50 (7%) features, followed closely by *Synechococcus* str. WH8102 and *T. rotula*, which shared 62 (7%). *Prochlorococcus* str. MED4 and *P. tricornutum* were the least similar strains, sharing only 12 (1%) features. Correlations between cyanobacteria and diatoms confirmed that all five cyanobacteria strains were most similar to *T. rotula* (0.03-0.12), followed by *T. pseudonana* (0.03-0.08) and finally *P. tricornutum* (0.02-0.06). *T. pseudonana* and *P. tricornutum* were both more similar to all diatom strains than to any cyanobacteria tested. While *T. rotula* was most similar to *T. pseudonana* (10%), it had more features in common with *Prochlorococcus* str. MIT9301 (7%) and *Synechococcus* str. WH8102 (7%) than with *P. tricornutum* (5%). All three diatom strains were more

correlated to *Synechococcus* str. WH8102 (0.03-0.12) than to any other cyanobacteria strain (0.02-0.10) and more correlated to *Prochlorococcus* str. MIT9301 (0.03-0.10) than any other *Prochlorococcus* strain (0.02-0.07). All three diatoms strains were least correlated to *Prochlorococcus* strains MED4 (0.02-0.06) and MIT9313 (0.02-0.07). Correlations are summarized in Table S1, while percent similarities and similarity rankings are shown in Tables S2 & S3. All eight phytoplankton strains produced unique features that were not found in any other strain. *P. triornutum* was the most distinct strain with 727 unique features (86% of its total), while *Prochlorococcus* str. MED4 was the least distinct with 34 unique features (22% of its total). Metabolite composition diversity was greatest at the domain-level (1- 7% common features), and lowest at the clade-level (14 - 35% common features).

Abundance of common features

Although strictly quantitative comparisons of features within a sample are not feasible due to uncertainties in ionization efficiency in ESI, if we assume that matrix effects are minimal and that abundance scales linearly with signal intensity, then the intensity level of features detected in multiple samples can provide semi-quantitative results. Intensity level comparisons of features found in multiple DOM_p samples reveals semi-quantitative information regarding common features produced by different phytoplankton. Comparing a particular strain with its respective media control can identify features present in the seawater-based medium that were also produced by the

organism as well as highlight features that were consumed by different phytoplankton cells during growth.

Scatter plots of signal intensity provide a visual representation of semi-quantitative differences among features common to multiple samples (Fig. 4). For each panel in Figure 4, common features are indicated by blue dots and their position in space is determined by their intensity in the two samples being compared. The axes are intensity ranges for two representative DOM_p samples and dots along the black 1:1 line indicate common features of similar intensity. Dots that fall outside of the two red 4:1 lines indicate features that are >4-fold as intense in one sample. 4:1 lines were chosen as the majority of common features (84% on average) among all replicate samples and blanks were within a 4-fold intensity difference (Fig. 4a). A representative comparison of common features found in a culture sample (*Prochlorococcus* str. 9301 replicate C) and its respective media control (Pro99 replicate C) demonstrates that while the majority of features (88% on average) were within a 4-fold intensity difference, there were metabolites present in the medium that were both produced (dots above the upper 4:1 line) and consumed (dots below the lower 4:1 line) during phytoplankton growth (Fig. 4b). Pairwise comparisons of different strains exhibited greater intensity differences than replicates (average of 69% within a 4-fold intensity difference) and the degree of intensity differences between strains varied widely. On average, 87% of common features among different *Prochlorococcus* strains were within a 4-fold intensity difference (Fig. 4c). An average of 75% of common features among the different *Synechococcus* strains were within a 4-fold intensity difference and 72% of common

features were within a 4-fold intensity difference among the different diatom strains. On average, 75% of features common to *Prochlorococcus* and *Synechococcus* were within a 4-fold intensity difference, while an average of 66% and 59% of common features were within a 4-fold intensity difference when comparing *Prochlorococcus* to diatoms and *Synechococcus* to diatoms, respectively (Fig. 4d-f).

Discussion

The composition of DOM produced by axenic cultures of eight marine phytoplankton strains chosen to represent various levels of phylogenetic relatedness was characterized in an effort to investigate how the chemical diversity of DOM relates to the phylogenetic identity of its source organism. A variable portion (5 – 25%) of the organic carbon produced by each phytoplankton strain was released as extracellular DOM where it might be accessible to heterotrophic microbes in a natural setting (Table 2). Solid-phase extraction of this DOM from spent culture media allowed for a direct comparison of DOM_P composition from a variety of biological sources under controlled conditions. Mass spectral analysis revealed DOM_P to be complex chemical mixtures containing between 293 and 1,237 different features using stringent signal-to-noise criteria (Table 1). A large proportion of these features (81%) were unique to a particular strain, indicating that DOM_P composition is highly variable.

The phytoplankton strains employed in this study can be grouped in several ways based on phylogenetic relationships. The three diatom strains (2 centric strains from the same genus and 1 pennate strain) and five cyanobacteria strains (2 *Synechococcus* strains

and 3 *Prochlorococcus* strains) chosen permit comparisons at multiple levels of phylogenetic relatedness. Cluster analysis of these strains based on their DOM_p composition displayed a striking similarity between biological and chemical diversity and illustrates that DOM_p composition is closely related to the phylogeny of its microbial source (Fig. 5). Comparing features present in all three replicates of each strain and correlation analysis of all pairwise comparisons (in which replicates were treated as individual samples) both revealed two major clusters based on DOM_p composition. These clusters match the phylogenetic relationship of these eight strains at the domain level. One group contains all three diatom (eukarya) strains tested, while the other contains all five cyanobacteria (bacteria) strains tested. Deeper investigation into the diatom group revealed another division into two clusters based on DOM_p composition that follows phylogenetic variation at the order level: one containing both centric diatoms from the same genus (*T. rotula* and *T. pseudonana*) and the other containing the lone pennate diatom tested, *P. tricorutum*. Deeper investigation into the cyanobacteria group revealed a division into two clusters based on DOM_p composition that follows phylogenetic variation of these strains at the genus level: one containing both *Synechococcus* strains and the other containing all three *Prochlorococcus* strains. Further investigation into the *Prochlorococcus* group revealed a final division into two clusters based on DOM_p composition that follows phylogenetic variation at the clade level: one containing both high-light adapted *Prochlorococcus* strains (MIT9301 and MED4) and the other containing the lone low-light adapted *Prochlorococcus* strain tested, MIT9313. Overall, these results demonstrate that similarities in DOM_p composition reflect the

phylogenetic relationship of its producers to a certain extent, a potential consequence of connections between the genomes and exometabolomes (defined here as DOM_P) of these marine phytoplankton. A relationship between microbial phylogeny and the presence of specific functional traits has recently been demonstrated (Martiny et al. 2012) and our findings suggest that this relationship may also extend beyond the phytoplankton cell, with influence on the chemical composition of the surrounding DOM pool as well.

DOM_P produced by these eight strains was also compared by investigating the signal intensities of features common to multiple strains. The degree of variation in intensity among common features also parallels phylogenetic variation, indicating that genetically related strains not only produce DOM of similar composition, but also produce common metabolites in similar relative abundance (Fig. 4). These comparisons indicate that not only the quality, but also the quantity of specific metabolites within DOM_P are indicative of phylogenetic origin and gives further support to the hypothesis that DOM composition and phytoplankton diversity are connected in the marine environment. Semi-quantitative analysis in this fashion also provides an opportunity to examine growth dynamics and mixotrophic activity of marine phytoplankton. Many marine phytoplankton have been shown to exhibit mixotrophic tendencies (Ceron Gargia et al. 2006, Bronk et al. 2007, Baran et al. 2011, Gomez-Pereira et al. 2012) and features found in greater abundance in media samples compared to DOM_P samples provide targets for further chemical identification of organic compounds consumed by marine phytoplankton during growth in laboratory culture (Fig. 4b).

Molecular weight and polarity distributions of DOM_P were analyzed for additional connections between DOM_P composition and phytoplankton diversity. While all eight strains produced low-molecular-weight ($\leq 1,481$ *m/z*) material over a broad polarity spectrum, *Prochlorococcus*-derived DOM_P was comprised of smaller, less polar metabolites when compared to diatom-derived DOM_P (Fig. 2). *Synechococcus*-derived DOM_P tended to fall between these groups in terms of both size and polarity. *Phaeodactylum tricornerutum* cultures in particular produced large polar metabolites, while *Thalassiosira rotula* produced smaller material that more closely resembled cyanobacterial DOM profiles. The low-light adapted *Synechococcus* str. WH7803 produced slightly larger material, resembling centric diatom profiles. The intermediate polarity of *Synechococcus*-derived DOM_P also exhibited two maxima in distinct polarity regions (Fig. 2b). Overall, these chemical trends indicate that larger eukaryotic phytoplankton cells may contribute a greater amount of slightly higher molecular weight metabolites over a broad polarity range when compared to smaller cyanobacterial cells and that both the size and polarity of DOM_P components vary according to biological origin. These trends warrant additional investigation and could lead to a better understanding of lability differences in DOM from various phytoplankton sources.

Connections between the composition of DOM_P and factors other than phylogeny were also identified. Similarity of conserved gene sequences does not equate to functional similarity and many environmental and physiological factors likely influence DOM_P composition. Among the diatoms tested, the DOM_P composition of *P. tricornerutum* was found to be more similar to that of *T. pseudonana*, than to that of *T.*

rotula. Both *P. tricornutum* and *T. pseudonana* were isolated from the North Atlantic and under lower temperature conditions than *T. rotula*, which was isolated from the Mediterranean Sea (ncma.bigelow.org). These variations in temperature and isolation location may have contributed to variations in DOM_P composition found in our study. We also found that DOM_P derived from *Prochlorococcus* str. MED4 was more compositionally similar to that of *T. rotula* (the only other strain isolated from the Mediterranean Sea) than to either of the other diatom strains tested. Additionally, DOM_P derived from the low-light adapted *Prochlorococcus* str. MIT9313 as well both of the *Synechococcus* strains tested was more similar to that of the high-light adapted *Prochlorococcus* str. MIT9301 than to that of the high-light adapted *Prochlorococcus* str. MED4. MIT9301, MIT9313 and both *Synechococcus* strains were all isolated from the Sargasso Sea and both MIT9313 and MIT9301 were also isolated from deeper depths (135m and 90m respectively) than MED4 (5m) (ncma.bigelow.org). *Prochlorococcus* str. MIT9301 and *Synechococcus* str. WH8102 had the highest correlation in DOM_P composition between the two cyanobacteria groups. These two strains were originally isolated from the warmest and most pelagic locations of all the strains tested in this study. Although many of the strains employed here have been in laboratory culture for many years, these potential connections between environmental niche and DOM_P composition are worthy of additional investigation in the field.

Although differences in DOM_P composition were generally much greater between strains than among biological replicates of the same strain, a surprising amount of variation was also found to exist among biological replicates (Fig. 1). While some of this

variation may be attributed to uncontrolled variations in sample processing, this finding more likely indicates that DOM_P composition depends not only on the source organism, but also on additional factors such as growth conditions, cell density and phase of growth. All replicates were grown in the same batch of media under identical growth conditions and harvested at the same time, therefore minor variations in parameters such as light and nutrient levels and exact growth phase at time of harvest likely resulted in physiological differences that impacted DOM_P composition. Growth phase in particular has been shown to greatly affect the composition of DOM released by several diatom strains in culture (Barofsky et al. 2009). Our findings support this conclusion and suggest that even subtle shifts in growth phase and other parameters may significantly impact the composition of phytoplankton-derived DOM. For instance, growth rate differences in our study could have affected DOM_P composition as cell densities at time of harvest varied among biological replicates (Table 2). The five cyanobacteria strains tested in this study all displayed more DOM_P variation among replicates than the diatom strains, suggesting that cyanobacteria may have been more susceptible to these additional factors. The high degree of variation in DOM_P composition produced by replicate cultures has implications for designing experiments that use DOM derived from phytoplankton cultures to examine DOM cycling. While the application of DOM derived from a single phytoplankton culture can be useful in identifying possible connections between heterotrophic bacteria and a known DOM source, experiments utilizing DOM from biological replicate phytoplankton cultures is likely needed to ensure that identified connections are robust under a variety of conditions.

The protocols described here allow for the creation of a spectral database on DOM_P composition and for identification of target metabolites of particular interest for additional chemical analyses such as nuclear magnetic resonance spectroscopy to examine compound structure and high-resolution mass spectrometry to determine elemental formulae (Aluwihare et al. 1999, Soule et al. 2010). These types of chemical data can be funneled into existing metabolomic databases for compound identification (Tautenhahn et al. 2012). With each phytoplankton strain tested here producing and releasing a complex suite of DOM_P containing hundreds to thousands of hydrophobic metabolites, it is useful to identify ways in which to reduce this vast collection down to more manageable lists of targets worthy of additional attention, as the effort required to definitively identify a given feature is not trivial (Bowen & Northen 2010). One way to accomplish this task is to group metabolites into categories based off their presence or absence in particular samples according to phylogenetic grouping. For instance, only six features were found to be common in all eight strains tested and likely represent metabolites that are more universally produced by a diverse array of phytoplankton and therefore worthy of further chemical characterization. The experimental design employed here highlights the potential for characterizing metabolites of interest in greater detail, identifying biologically relevant targets for environmental metabolomics studies, and using representative metabolites as biomarkers for particular phytoplankton groups in natural seawater samples. Data from additional strains of marine phytoplankton will be essential for identification of metabolites that are truly diagnostic of a particular phytoplankton group.

In conclusion, we present here a molecular comparison of DOM derived from multiple strains of marine phytoplankton using a method that can be applied to any cultivable organism. Analysis of phytoplankton exudate suites by HPLC/ESI-MS identified hundreds to thousands of metabolites from known biological sources, in the absence of any heterotrophic activity. Results from this initial investigation of compositional differences in DOM_p indicate that marine phytoplankton release highly unique suites of organic material. Qualitative and semi-quantitative connections between DOM_p composition and phylogeny were identified on multiple levels, indicating a previously undemonstrated relationship between phytoplankton diversity and DOM composition. If most or all marine phytoplankton produce and release different suites of DOM that mirror their genetic similarities, then spatial and temporal variations in phytoplankton community structure (over smaller scales such as seasonal phytoplankton blooms or larger scales including gradual shifts due to global climate conditions) will induce concurrent variations in the composition of marine DOM. Compositional differences in DOM have been shown to affect both the relative abundance and activity of particular heterotrophic bacteria groups, suggesting a relationship between DOM composition and heterotrophic bacterioplankton diversity (Cottrell & Kirchman 2000, McCarren et al. 2010, Romera-Castillo et al. 2011, Nelson & Carlson 2012, Sarmiento & Gasol 2012). If DOM composition and producer phylogeny are indeed related as our findings indicate, then phytoplankton diversity could directly impact heterotrophic bacterioplankton diversity via a marine DOM link.

Acknowledgements

Thanks to P. Berube and S. Chisholm for providing *Prochlorococcus* cultures and valuable discussions on the manuscript regarding phytoplankton phylogeny along with J. Waterbury. Thanks to C. Follett for assistance running correlations of mass spectral data. I also gratefully acknowledge the National Center for Marine Algae and Microbiota, particularly C. Schuman and T. Riggins, for providing phytoplankton strains and media. I'd also like to thank V. Starczak, K. Longnecker and M. Bhatia for insightful discussions regarding data analysis, M. Rappé for helpful comments on the manuscript, and C. Johnson for technical assistance with mass spectrometry method development. This work was funded by a grant to D.J.R. from the Gordon and Betty Moore Foundation (grant #1711) and grants to D.J.R., E.F.D., and S.W.C. from the National Science Foundation Science and Technology Center Award #EF0424599 and #DBI0424599.

Table 1. The number of DOM_p features detected in each sample using HPLC-ESI-MS. Features are defined as a unique m/z value at a given retention time with a signal intensity at least 5-fold above the noise level. Features that were also detected in any replicate media control or blank were not included. Feature totals for individual biological replicates of each phytoplankton strain are shown along with the sum of all features for each strain. The number of features unique to an individual replicate as well the number of features common to all replicates for a given strain are also shown. Total and common features are also displayed for each organism group. Identical features were defined using a retention time tolerance of +/-5 s after normalization to correct for retention time shifts and a mass tolerance of +/-0.3 m/z.

Strain	Total Features			Unique Features			Total Features	Common Features
	<i>Replicate A</i>	<i>Replicate B</i>	<i>Replicate C</i>	<i>Replicate A</i>	<i>Replicate B</i>	<i>Replicate C</i>		
<i>Prochlorococcus</i>							368	54
MIT9313	379	351	378	119	154	153	693	148
MIT9301	502	477	525	83	142	54	765	253
MED4	NA	387	293	NA	232	138	525	155
<i>Synechococcus</i>							558	77
WH8102	730	679	597	84	69	76	893	449
WH7803	348	373	293	54	51	75	504	186
<i>Diatoms</i>							1782	32
CCMP632	1065	1118	1237	74	53	184	1443	845
CCMP1647	799	1137	614	51	375	46	1269	484
CCMP1335	1062	916	813	207	62	57	1226	665

Table 2. Biomass and organic carbon production by all phytoplankton cultures in this study. Organic carbon data is background subtracted using media controls. Measurement were taken as described in ‘Materials and Methods’ and used to calculate the quantity of DOC and POC produced per cell as well as the % of the total organic carbon produced (or net primary production) that was released as extracellular DOC at the time of biomass removal. The % of that DOC that was recovered by solid phase extraction was estimated with POC measurements as described in ‘Materials and Methods’. Organic carbon data for the *Prochlorococcus* strains was not acquired due to the additions of HEPES buffer during growth to mitigate pH changes.

Sample	Growth Period (Days)	Maximum Cell Density (cells/ml)	DOC produced (μM)	POC produced (mM)	DOC produced per cell (fmol)	POC produced per cell (fmol)	% of total OC released as DOC	%DOC recovered via SPE
MIT9313A	12	6.38E+07	NA	1.0	NA	15.1	NA	NA
MIT9313B	12	6.81E+07	NA	1.0	NA	14.7	NA	NA
MIT9313C	12	6.74E+07	NA	1.0	NA	15.0	NA	NA
MIT9301A	12	2.55E+08	NA	1.1	NA	4.1	NA	NA
MIT9301B	12	3.01E+08	NA	1.2	NA	4.0	NA	NA
MIT9301C	12	2.57E+08	NA	1.1	NA	4.2	NA	NA
MED4B	12	1.63E+08	NA	0.6	NA	3.8	NA	NA
MED4C	12	1.61E+08	NA	0.6	NA	3.7	NA	NA
WH8102A	22	1.06E+08	512.1	1.6	4.8	14.8	25	3
WH8102B	22	1.27E+08	436.7	1.4	3.4	11.2	24	2
WH8102C	22	1.10E+08	499.0	1.6	4.5	14.2	24	3
WH7803A	22	3.14E+07	180.5	0.6	5.7	19.7	23	4
WH7803B	22	4.24E+07	167.8	0.6	4.0	15.1	21	4
WH7803C	22	3.61E+07	172.6	0.7	4.8	18.3	21	6
CCMP1335A	19	1.74E+07	325.3	2.6	18.7	150.6	11	7
CCMP1335B	19	1.06E+07	261.3	2.5	24.7	233.8	10	7
CCMP1335C	19	1.38E+07	239.7	2.5	17.4	182.4	9	9
CCMP1647A	19	1.89E+07	511.5	1.7	27.1	90.9	23	11
CCMP1647B	19	1.95E+07	426.7	1.8	21.9	93.3	19	10
CCMP1647C	19	1.31E+07	358.4	1.5	27.4	117.1	19	11
CCMP632A	19	1.07E+07	172.3	2.8	16.1	264.3	6	22
CCMP632B	19	9.28E+06	147.4	2.8	15.9	302.4	5	24
CCMP632C	19	7.98E+06	188.5	2.8	23.6	356.7	6	24

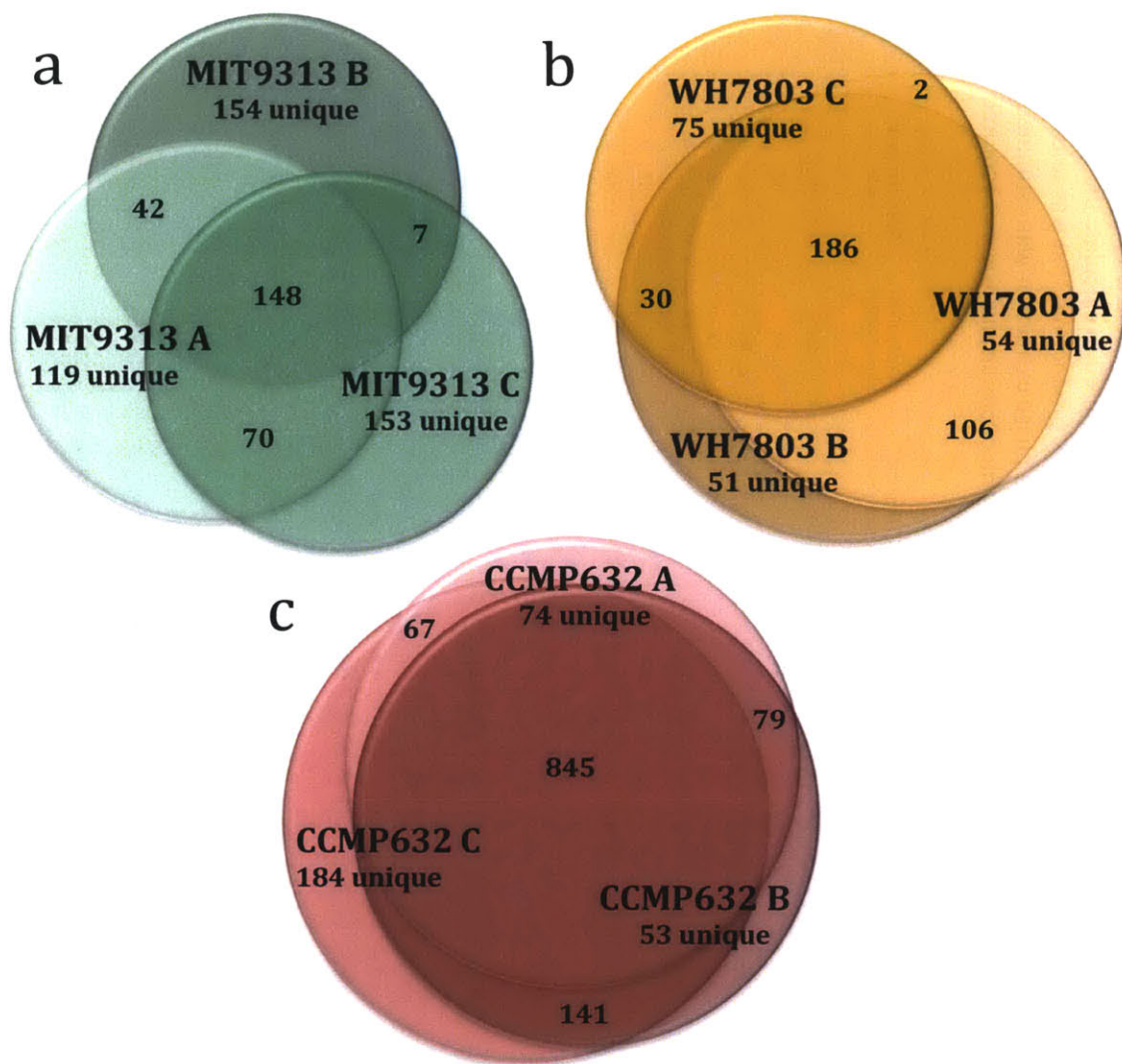


Fig. 1. Venn diagrams showing the degree to which triplicate cultures of representative phytoplankton strains share DOM_p features. Each culture is represented by a circle, the area of which is proportional to the total number of features identified in that sample. The degree of overlap between circles is proportional to their number of shared features. Low-light adapted *Prochlorococcus* str. MIT9313 replicates had the most variation of any strain tested (a), while *P. tricornutum* (CCMP632) replicates had the least variation (b). Low-light adapted *Synechococcus* str. WH7803 replicates exhibited an average amount of variation indicative of most strains tested in this study (c). For more detailed information regarding replicate correlations, see Table S1.

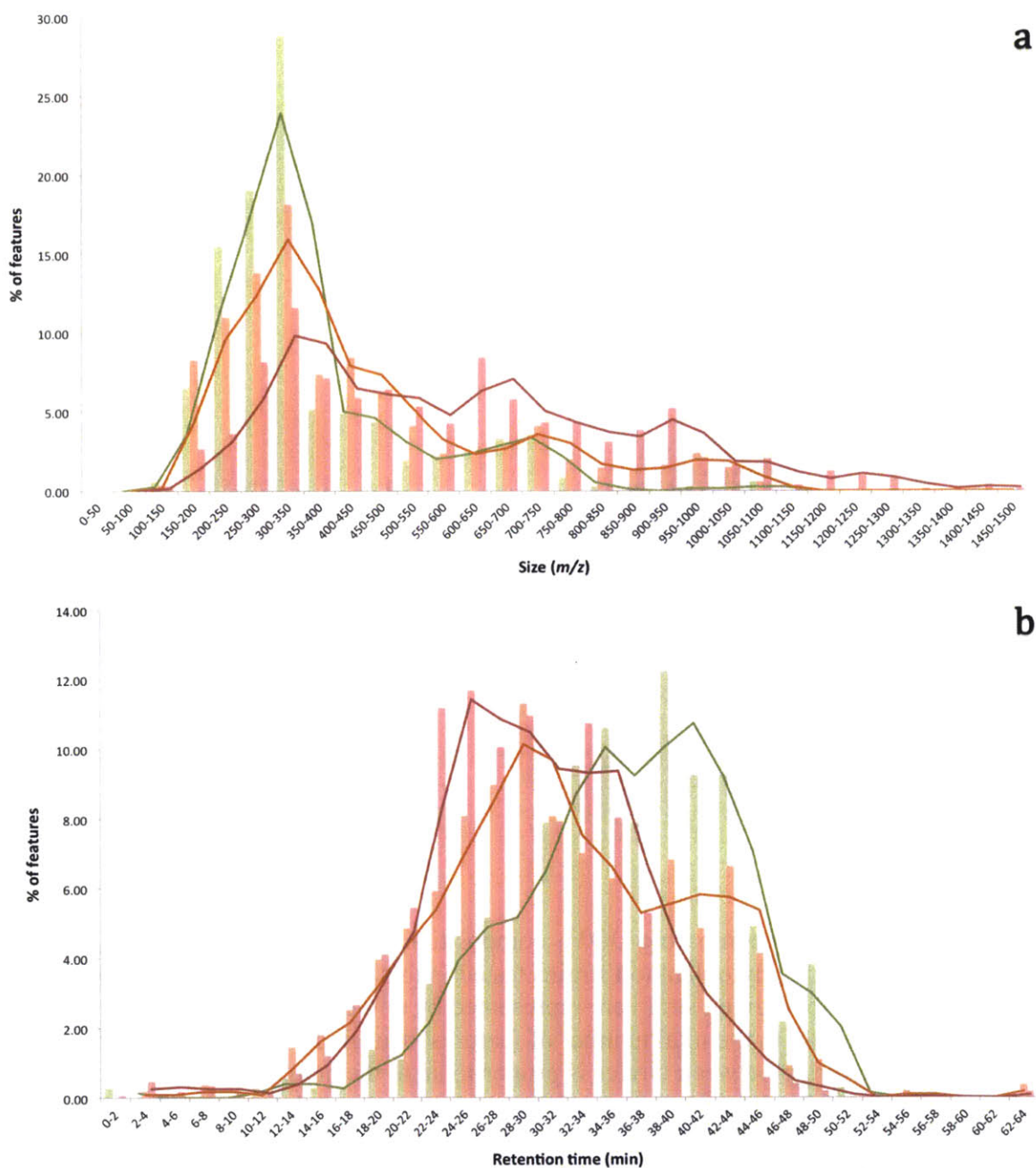


Fig. 2. Histograms comparing the molecular weight (a) and polarity (b) distributions of DOM_p derived from representative strains from each major phytoplankton group; *Prochlorococcus* (green), *Synechococcus* (orange), and diatoms (red). Bars indicate the % of features that fall within each bin labeled on the X-axes and moving average lines are displayed to highlight trends. In general, diatom-derived DOM_p had a more even size distribution and a greater abundance of larger, more polar metabolites than cyanobacteria-derived DOM_p. DOM_p suites derived from *Prochlorococcus* were comprised of the smallest, least polar material when compared to other strains and DOM_p derived from *Synechococcus* tended to fall between *Prochlorococcus* and diatom DOM_p suites in terms of both size and polarity.

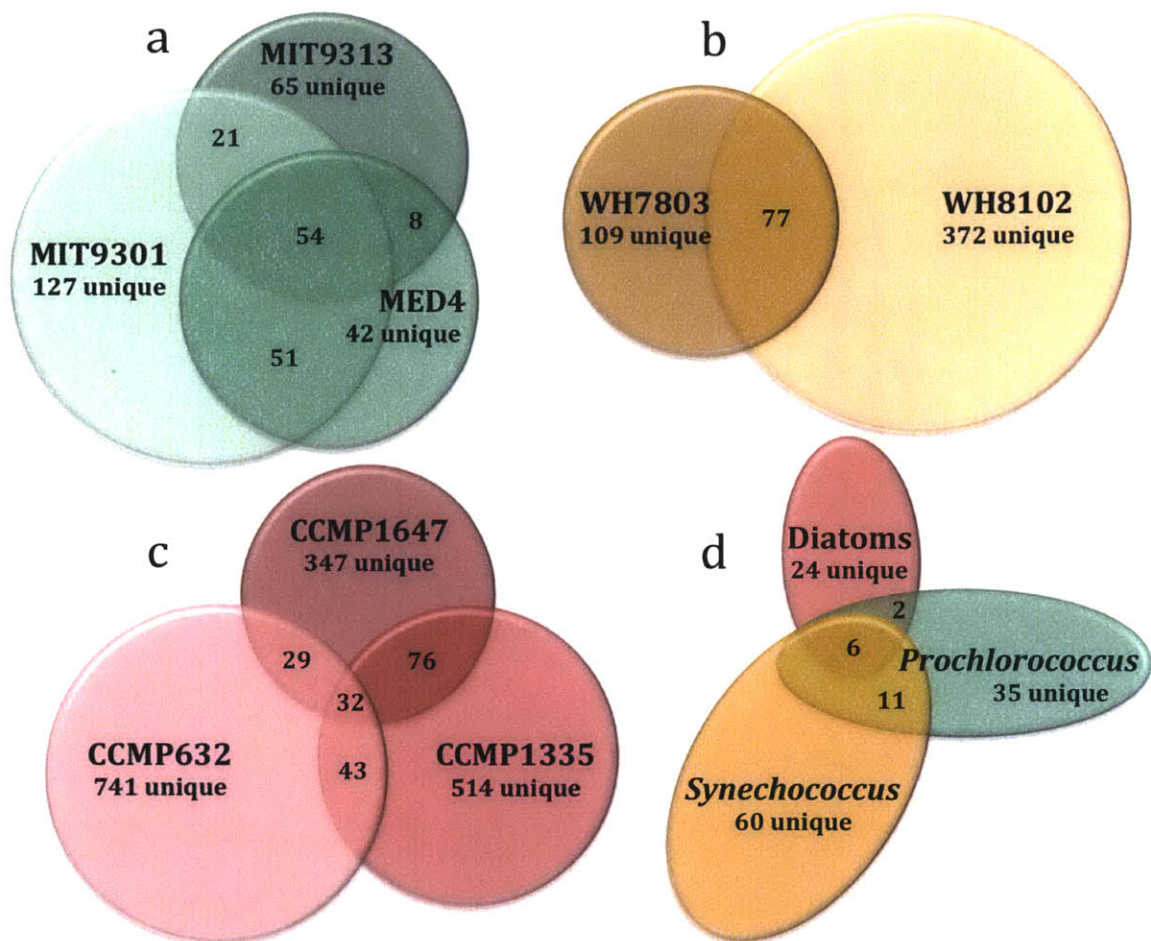


Fig. 3. Venn diagrams comparing DOM_p composition on several levels of phylogenetic variation for all eight phytoplankton strains tested in this study. Features common to all replicate cultures for a given strain are represented by circles or ellipses, the area of which is proportional to the total number of features identified as common that strain. The degree of overlap between circles is proportional to the number of shared features between strains of *Prochlorococcus* (a), *Synechococcus* (b), and diatoms (c). Features common to each of these 3 broad phytoplankton groupings were also compared (d) to examine variation at the genus and domain level. For more detailed information regarding percentages of shared DOM_p features among different strains and their similarity rankings, see Tables S2 & S3.

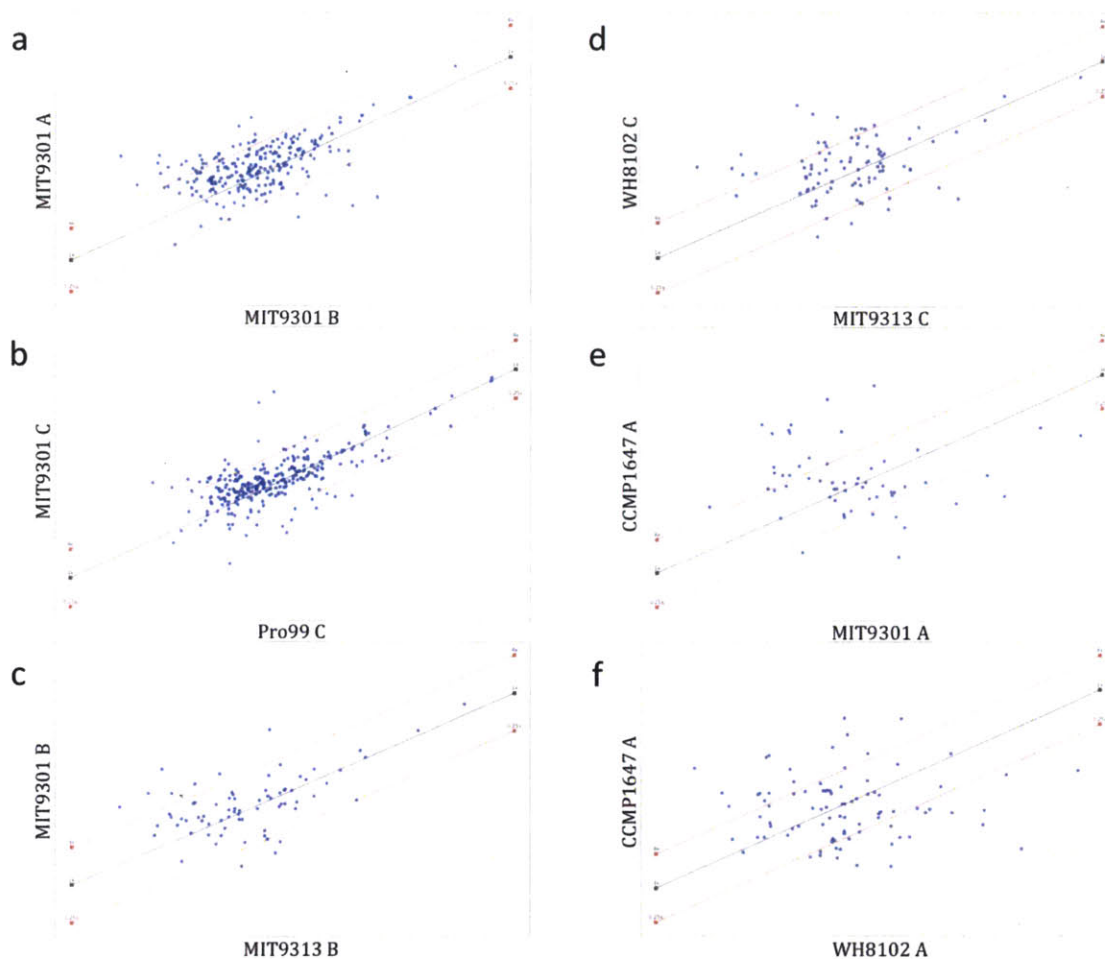


Fig. 4. Scatter plots illustrating variations in signal intensity among features common to multiple samples. Signal intensity is used as a semi-quantitative proxy for feature abundance. Features are indicated by dots and their location is a function of their signal intensity in the 2 samples labeled on the axes for each plot. Black lines indicate 1:1 and red lines indicate 4:1. Variations in common feature intensities are shown for biological replicates (a) indicating some abundance variation among replicate samples. Comparing a culture sample to its respective media control (b) indicates some features present in the medium that are both produced (above the upper red line) and consumed (below the lower red line) by the phytoplankton strain. Variations in common feature intensities can also be compared at multiple levels of phylogenetic variation including at the clade-level (c), genus-level (d) and domain level (e,f), indicating that the quantity of different features produced by multiple marine phytoplankton is also related to producer phylogeny.

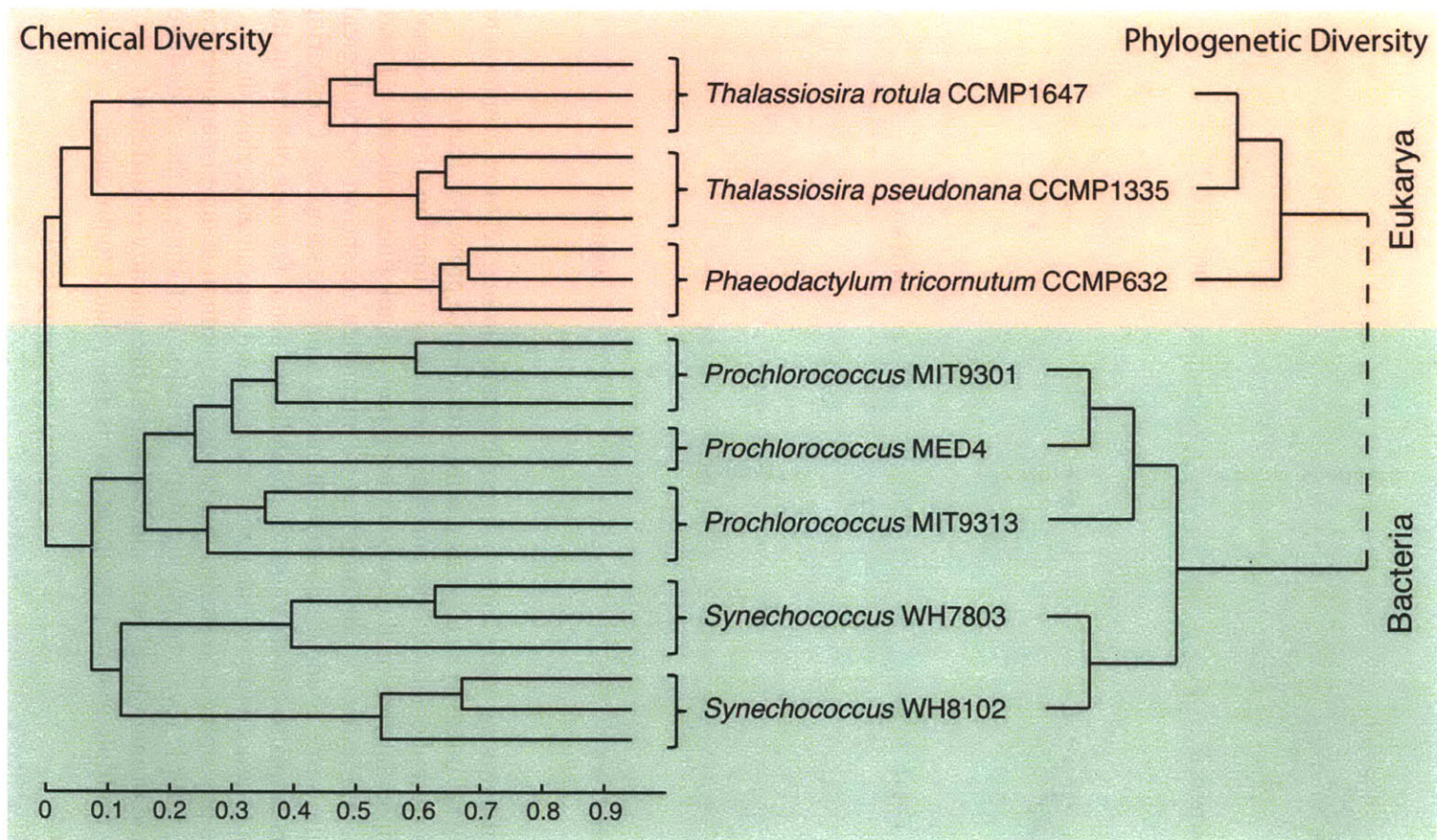


Fig. 5. Phylogeny of phytoplankton and the chemical similarity of the DOM they produce show similar relationship patterns. The suites of compounds produced by the eight model phytoplankton tested here is consistent with their phylogenetic relationship. At left is shown a dendrogram created from the normalized similarity matrix of all DOM_p features identified in each culture replicate after media subtraction using the unweighted pair group method average (UPGMA). At right is a schematic representation of the phylogenetic relationships between the 8 organisms based on (Rocap et al. 2002), (Kettler et al. 2007) and (Sorhannus 2004). Branch lengths do not correspond to phylogenetic distances. The scale bar corresponds to correlation values of DOM composition for all pairwise comparisons given in Table S1.

References

- Aluwihare LI, Repeta DJ, Chen RF (1999) A comparison of the chemical characteristics of oceanic DOM and extracellular DOM produced by marine algae. *Marine Ecology Progress Series* 186:105–117
- Baines SB, Pace ML (1991) The production of dissolved organic matter by phytoplankton and its importance to bacteria: Patterns across marine and freshwater systems. *Limnol Oceanogr*:1078–1090
- Baran R, Bowen BP, Northen TR (2011) Untargeted metabolic footprinting reveals a surprising breadth of metabolite uptake and release by *Synechococcus* sp. PCC 7002. *Mol BioSyst* 7:3200–3206
- Barofsky A, Simonelli P, Vidoudez C, Troedsson C, Nejstgaard JC, Jakobsen HH, Pohnert G (2010) Growth phase of the diatom *Skeletonema marinoi* influences the metabolic profile of the cells and the selective feeding of the copepod *Calanus* spp. *Journal of Plankton Research* 32:263–272
- Barofsky A, Vidoudez C, Pohnert G (2009) Metabolic profiling reveals growth stage variability in diatom exudates. *Limnol Oceanogr-Meth* 7:382–390
- Bertilsson S, Jones JB Jr (2003) 1 - Supply of Dissolved Organic Matter to Aquatic Ecosystems: Autochthonous Sources. In: Findlay SEG, Sinsabaugh RL (eds) *Aquatic Ecosystems*. Academic Press, Burlington, p 3–24
- Bertilsson S, Berglund O, Pullin MJ, Chisholm SW (2005) Release of dissolved organic matter by *Prochlorococcus*. *Vie et Milieu* 55:225–232
- Bjornsen PK (1988) Phytoplankton exudation of organic matter: Why do healthy cells do it? *Limnol Oceanogr*:151–154
- Bowen BP, Northen TR (2010) Dealing with the Unknown: Metabolomics and Metabolite Atlases. *J Am Soc Mass Spectrom* 21:1471–1476
- Carlson CA (2002) Production and removal processes. In: Hansell DA, Carlson CA (eds) *Biogeochemistry of Marine Dissolved Organic Matter*. Academic Press, p 91–151
- Cavender-Bares KK, Mann EL, Chisholm SW, Ondrusek ME, Bidigare RR (1999) Differential response of equatorial Pacific phytoplankton to iron fertilization. *Limnol Oceanogr*:237–246
- Cottrell MT, Kirchman DL (2000) Natural assemblages of marine proteobacteria and members of the Cytophaga-Flavobacter cluster consuming low- and high-molecular-weight dissolved organic matter. *Applied and Env. Microbiology* 66:1692–97

- Engel A, Handel N, Wohlers J, Lunau M, Grossart HP, Sommer U, Riebesell U (2011) Effects of sea surface warming on the production and composition of dissolved organic matter during phytoplankton blooms: results from a mesocosm study. *Journal of Plankton Research* 33:357–372
- Giovannoni SJ, Hayakawa DH, Tripp HJ, Stingl U, Givan SA, Cho J-C, Oh H-M, Kitner JB, Vergin KL, Rappé MS (2008) The small genome of an abundant coastal ocean methylotroph. *Environmental Microbiology* 10:1771–1782
- Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D, Bibbs L, Eads J, Richardson TH, Noordewier M, Rappé MS, Short JM, Carrington JC, Mathur EJ (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* 309:1242–1245
- Grossart HP, Simon M (2007) Interactions of planktonic algae and bacteria: effects on algal growth and organic matter dynamics. *Aquat Microb Ecol* 47:163–176
- Guillard RRL, Hargraves PE (1993) *Stichochrysis immobilis* is a diatom, not a chrysophyte. *Phycologia* 32:234–236
- Hirt G, Tanner W, Kandler O (1971) Effect of Light on the Rate of Glycolysis in *Scenedesmus obliquus*. *Plant Physiol* 47:841–843
- Kettler GC, Martiny AC, Huang K, Zucker J, Coleman ML, Rodrigue S, Chen F, Lapidus A, Ferreira S, Johnson J, Steglich C, Church GM, Richardson P, Chisholm SW (2007) Patterns and Implications of Gene Gain and Loss in the Evolution of *Prochlorococcus*. *PLoS Genet* 3:2515–2528
- Kujawinski EB, Longnecker K, Blough NV, Del Vecchio R, Finlay L, Kitner JB, Giovannoni SJ (2009) Identification of possible source markers in marine dissolved organic matter using ultrahigh resolution mass spectrometry. *Geochimica et Cosmochimica Acta* 73:4384–4399
- Mague T, Friberg E, Hughes D, Morris I (1980) Extracellular release of carbon by marine phytoplankton; a physiological approach. *Limnol Oceanogr*:262–279
- Martiny AC, Treseder K, Pusch G (2012) Phylogenetic conservatism of functional traits in microorganisms. *The ISME Journal* 7:830–838
- McCarren J, Becker JW, Repeta DJ, Shi Y, Young CR, Malmstrom RR, Chisholm SW, DeLong EF (2010) Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. *Proc Natl Acad Sci USA* 107:16420–16427

- Meon B, Kirchman DL (2001) Dynamics and molecular composition of dissolved organic material during experimental phytoplankton blooms. *Marine Chemistry* 75:185–199
- Moore LR, Coe A, Zinser ER, Saito MA, Sullivan MB, Lindell D, Frois-Moniz K, Waterbury J, Chisholm SW (2007) Culturing the marine cyanobacterium *Prochlorococcus*. *Limnol Oceanogr-Meth* 5:353–362
- Moore LR, Post AF, Rocap G, Chisholm SW (2002) Utilization of different nitrogen sources by the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. *Limnol Oceanogr*:989–996
- Myklestad SM (1995) Release of extracellular products by phytoplankton with special emphasis on polysaccharides. *Science of the total Environment* 165:155–164
- Nelson CE, Carlson CA (2012) Tracking differential incorporation of dissolved organic carbon types among diverse lineages of Sargasso Sea bacterioplankton. *Environmental Microbiology* 14:1500–1516
- Obernosterer I, Herndl GJ (1995) Phytoplankton extracellular release and bacterial growth: dependence on the inorganic N: P ratio. *Marine Ecology - Progress Series* 116:247–257
- Olson RJ, Vaultot D, Chisholm SW (1985) Marine phytoplankton distributions measured using shipboard flow cytometry. *Deep Sea Research Part A Oceanographic Research Papers* 32:1273–1280
- Partensky F, Hess WR, Vaultot D (1999) *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiology and Molecular Biology Reviews* 63:106–127
- Pluskal T, Castillo S, Villar-Briones A, Orešič M (2010) MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 11:395–406
- Pomeroy LR (1974) The ocean's food web, a changing paradigm. *Bioscience*:499–504
- Poretsky RS, Sun S, Mou X, Moran MA (2010) Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon. *Environmental Microbiology* 12:616–627
- Rocap G, Distel DL, Waterbury JB, Chisholm SW (2002) Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S-23S ribosomal DNA internal transcribed spacer sequences. *Applied and Environmental Microbiology* 68:1180–1191

- Romera-Castillo C, Sarmiento H, Alvarez-Salgado XA, Gasol JM, Marrase C (2010) Production of chromophoric dissolved organic matter by marine phytoplankton. *Limnol Oceanogr* 55:446–454
- Romera-Castillo C, Sarmiento H, Alvarez-Salgado XA, Gasol JM, Marrase C (2011) Net Production and Consumption of Fluorescent Colored Dissolved Organic Matter by Natural Bacterial Assemblages Growing on Marine Phytoplankton Exudates. *Applied and Environmental Microbiology* 77:7490–7498
- Sarmiento H, Gasol JM (2012) Use of phytoplankton-derived dissolved organic carbon by different types of bacterioplankton. *Environmental Microbiology* 14:2348–2360
- Sorhannus U (2004) Diatom phylogenetics inferred based on direct optimization of nuclear-encoded SSU rRNA sequences. *Cladistics* 20:487–497
- Soule MCK, Longnecker K, Giovannoni SJ, Kujawinski EB (2010) Impact of instrument and experiment parameters on reproducibility of ultrahigh resolution ESI FT-ICR mass spectra of natural organic matter. *Organic Geochemistry* 41:725–733
- Tautenhahn R, Cho K, Uritboonthai W, Zhu Z, Patti GJ, Siuzdak G (2012) An accelerated workflow for untargeted metabolomics using the METLIN database. *Nature Biotechnology* 30:826–828
- Wetz MS, Wheeler PA (2007) Release of dissolved organic matter by coastal diatoms. *Limnol Oceanogr* 52:798–807

Table S1. Correlation values for all pairwise comparisons. All features present in any replicate media control or blank sample were removed prior to correlation analysis. For details on creation of the normalized similarity matrix used to obtain correlation values, see ‘Materials and Methods’.

	MIT9313 C	MIT9313 B	MIT9313 A	MIT9301 C	MIT9301 B	MIT9301 A	MED4 C	MED4 B	WH8102 C	WH8102 B	WH8102 A	WH7803 A	WH7803 C	WH7803 B	CCMP632 B	CCMP632 C	CCMP632 A	CCMP1647 C	CCMP1647 A	CCMP1647 B	CCMP1335 C	CCMP1335 B	CCMP1335 A
MIT9313 C	1.0000	0.2700	0.4045	0.1503	0.1430	0.1399	0.2068	0.1250	0.1285	0.1293	0.1329	0.0885	0.0651	0.0759	0.0198	0.0202	0.0205	0.0323	0.0379	0.0470	0.0312	0.0303	0.0315
MIT9313 B		1.0000	0.3519	0.2321	0.2938	0.2116	0.2361	0.2768	0.1533	0.1319	0.1391	0.0888	0.0662	0.0855	0.0251	0.0278	0.0291	0.0535	0.0512	0.0538	0.0374	0.0326	0.0329
MIT9313 A			1.0000	0.2538	0.2006	0.2550	0.2491	0.1589	0.1975	0.1861	0.1950	0.1116	0.0701	0.0946	0.0289	0.0273	0.0278	0.0575	0.0641	0.0684	0.0401	0.0444	0.0465
MIT9301 C				1.0000	0.4692	0.6485	0.3679	0.2459	0.2143	0.1968	0.2149	0.1294	0.0777	0.1073	0.0366	0.0401	0.0331	0.0868	0.0844	0.0956	0.0535	0.0518	0.0587
MIT9301 B					1.0000	0.3769	0.3391	0.3935	0.1712	0.1514	0.1595	0.0927	0.0665	0.0801	0.0324	0.0325	0.0273	0.0717	0.0669	0.0717	0.0454	0.0403	0.0406
MIT9301 A						1.0000	0.3475	0.2279	0.2130	0.1763	0.2008	0.1126	0.0700	0.0870	0.0391	0.0382	0.0309	0.0814	0.0833	0.0890	0.0428	0.0404	0.0511
MED4 C							1.0000	0.2952	0.1680	0.1422	0.1469	0.1206	0.0772	0.0936	0.0254	0.0234	0.0249	0.0522	0.0551	0.0538	0.0290	0.0289	0.0328
MED4 B								1.0000	0.1182	0.1012	0.1038	0.0637	0.0526	0.0541	0.0245	0.0253	0.0254	0.0460	0.0496	0.0481	0.0300	0.0284	0.0255
WH8102 C									1.0000	0.5773	0.6104	0.1842	0.1323	0.1858	0.0438	0.0468	0.0375	0.1151	0.0966	0.1003	0.0846	0.0746	0.0601
WH8102 B										1.0000	0.7246	0.2082	0.1289	0.1968	0.0442	0.0470	0.0344	0.0847	0.0796	0.0894	0.0619	0.0741	0.0609
WH8102 A											1.0000	0.1938	0.1254	0.1835	0.0542	0.0547	0.0448	0.1025	0.0976	0.1093	0.0686	0.0737	0.0692
WH7803 A												1.0000	0.4150	0.6807	0.0427	0.0434	0.0367	0.0583	0.0494	0.0532	0.0469	0.0464	0.0437
WH7803 C													1.0000	0.4800	0.0292	0.0380	0.0560	0.0461	0.0390	0.0377	0.0356	0.0342	0.0328
WH7803 B														1.0000	0.0347	0.0427	0.0345	0.0534	0.0455	0.0457	0.0431	0.0471	0.0376
CCMP632 B															1.0000	0.7202	0.7339	0.0711	0.0770	0.0821	0.0776	0.0819	0.0787
CCMP632 C																1.0000	0.6561	0.0718	0.0750	0.0811	0.0841	0.0901	0.0844
CCMP632 A																	1.0000	0.0613	0.0633	0.0700	0.0683	0.0714	0.0688
CCMP1647 C																		1.0000	0.5805	0.4376	0.1281	0.1308	0.1129
CCMP1647 A																			1.0000	0.5830	0.1226	0.1365	0.1361
CCMP1647 B																				1.0000	0.1092	0.1194	0.1283
CCMP1335 C																					1.0000	0.6968	0.6108
CCMP1335 B																						1.0000	0.6920
CCMP1335 A																							1.0000

Table S2. Percentage of features shared among all pairwise strain comparisons (top) alongside sequence distances based on 16S-18S rRNA gene sequences (bottom in bold) for the eight phytoplankton strains tested in this study. Features used in percentage calculations were considered to be produced by a strain only if present in all replicates for that strain and absent in all replicate media controls and blank samples. Sequences for distance calculations were taken from the SILVA SSURef database release 108, 19.08.2011 and aligned using Phylip version 3.6.8. Phylogeny was constructed from 1,349 sites using the unweighted LogDet algorithm. Comparisons between *Prochlorococcus* strains are highlighted in green, while comparisons between *Synechococcus* strains are highlighted in yellow and between diatom strains in red. See Rocap et al. 2002, Kettler et al. 2007 and Sorhannus et al. 2004 for more information regarding phylogenetic relationships within these groups.

	MIT9301	MED4	MIT9313	WH8102	WH7803	CCMP1647	CCMP1335	CCMP632
MIT9301	x	34.7	23.0	19.4	8.1	7.3	3.3	1.9
MED4	0.0081	x	25.7	12.1	8.3	3.9	1.9	1.2
MIT9313	0.0223	0.0240	x	15.3	9.2	3.6	2.3	1.7
WH8102	0.0322	0.0322	0.0232	x	13.8	7.1	4.4	2.9
WH7803	0.0320	0.0319	0.0176	0.012246	x	2.8	2.3	1.8
CCMP1647	0.7499	0.7510	0.7566	0.7350	0.7489	x	10.4	4.8
CCMP1335	0.7631	0.7641	0.7719	0.7462	0.7602	0.0366	x	5.2
CCMP632	0.7754	0.7771	0.7812	0.7678	0.7826	0.0916	0.0883	x

Table S3. Similarity ranking of strains based off the percentage data in Table S2. Rankings should be read by row, not column (i.e. WH8102 is most similar to MIT9301, not WH7803).

	MIT9301	MED4	MIT9313	WH8102	WH7803	CCMP1647	CCMP1335	CCMP632
MIT9301	X	1	2	3	4	5	6	7
MED4	1	X	2	3	4	5	6	7
MIT9313	2	1	X	3	4	5	6	7
WH8102	1	4	2	X	3	5	6	7
WH7803	4	3	2	1	X	5	6	7
CCMP1647	2	6	5	3	7	X	1	4
CCMP1335	4	7	6	3	5	1	X	2
CCMP632	4	7	6	3	5	2	1	X

CHAPTER THREE:

The reactivity of marine dissolved organic matter depends on its biological source and sink

Abstract

Marine phytoplankton are the paramount source of organic substrate for secondary production by heterotrophic microbes in the sea. Despite the importance of this secondary production for marine food webs and carbon cycling, details regarding the role of microbial diversity in the uptake and utilization of dissolved organic matter (DOM) are limited, and few connections have been made between specific producers and consumers of marine DOM. We amended fifteen axenic strains of heterotrophic marine bacteria, including members of the SAR11, *Rhodobacterales*, and *Oceanospirillales* lineages, with DOM from both natural seawater and that produced by axenic cultures of five marine phytoplankton to evaluate the effects of various DOM sources on heterotrophic growth characteristics. Phytoplankton-derived DOM was isolated by solid-phase extraction, permitting concurrent chemical analysis of the amendments by mass spectrometry. The dose specific effects on growth rate and cell yield were monitored with high-resolution cell enumeration in replicate cultures. Responses to DOM amendments ranged from suppressed to enhanced growth rate and yield, and depended on both the identity of the heterotroph and the source of the DOM. We demonstrate that cultivated strains of bacterioplankton are capable of utilizing DOM compounds derived from various sources and support the hypothesis that bacterioplankton taxa have differing organic carbon substrate requirements, providing a link between microbial diversity and DOM composition. The fractionation of *Prochlorococcus*-derived DOM refined the identification of reactive material to specific components within the DOM pool, revealing an unexpected connection between reactivity and polarity. Our results highlight the

potential for linking detailed chemical analyses of DOM from a defined source with controls on heterotrophic microbial diversity.

Introduction

Marine microorganisms consume the vast majority of reactive dissolved organic matter (DOM) found in seawater. Estimates indicate that gigatonnes (10^{15} g) of fixed carbon are transferred from marine phytoplankton to heterotrophic microbes every year in what is often referred to as the microbial loop (reviews by Hedges 2002, Pomeroy et al. 2007, Azam & Malfatti 2007). Low-molecular weight (LMW) organic compounds are transported directly through cell membranes, while larger compounds are first hydrolyzed to LMW compounds by ecto-cellular hydrolytic enzymes. After consuming marine DOM, microbes act as a trophic link by converting dissolved organic carbon (DOC) to bacterial biomass via anabolic reactions. This biomass can then be consumed by micrograzers and subsequently transferred to higher trophic levels. Marine microbes are also important for regenerated production processes, respiring DOC back to inorganic carbon via catabolic reactions. Any changes to the lability of marine DOM, the rate of DOM consumption by heterotrophic bacteria, or the partitioning of carbon between anabolic and catabolic pathways will have significant impacts on the oceanic carbon cycle including carbon export and storage in the ocean's interior.

Despite the importance of microbial secondary production for marine food webs and carbon cycle dynamics, details regarding the consumption of organic carbon and other organic nutrients are limited, and few definitive connections have been made

between specific producers and consumers of DOM in the sea. In large part, this is due to both the chemical complexity of marine DOM and the diversity of bacterioplankton communities. In every milliliter of surface seawater, millions of cells are interacting with organic compounds in a myriad of microenvironments, making detailed observations of individual processes within the microbial loop a daunting task. To date, the substrate requirements for many of the most abundant heterotrophic marine bacteria are not well characterized; therefore, one major constraint on our understanding of microbial metabolism is our limited knowledge of the organic compounds that fuel secondary production. Overall, researchers have yet to link the production of a substrate by a particular phytoplankton to its uptake and utilization by a specific heterotrophic bacterium. The lability of any given DOM molecule, or how fast a molecule is acquired and broken down by microorganisms, likely depends on the identity of the heterotrophs that encounter it, as well as abiotic factors such as temperature and the presence of other organic and inorganic material. New ways of identifying and investigating connections between individual microbes and specific DOM components are needed to make these detailed observations and uncover how and why variations in DOM composition affect bacterioplankton diversity.

Taxon-specific responses of bacterioplankton to DOM derived from phytoplankton cultures have been demonstrated for decades (Bell et al. 1974). However, the recognition of the uncultured microbial majority (Rappé & Giovannoni 2003) and widespread use of molecular tools to track microbial community structure dynamics has shifted research away from strain-specific approaches and toward studies focused on

whole community dynamics (e.g. Moreira et al. 2011, Romera-Castillo et al. 2011, Nelson & Carlson 2012, Sarmiento & Gasol 2012). A significant amount of work has also been performed in order to examine the response of bacterioplankton communities to other sources of DOM, including specific compounds, particular chemical and size fractions, and *in-situ* phytoplankton blooms (e.g. Cottrell & Kirchman 2000, Rink et al. 2007, McCarren et al. 2010, Tada et al. 2011, Rinta-Kanto et al. 2011, Teeling et al. 2012, Gómez-Consarnau et al. 2012). While these studies provide strong evidence for the structuring of bacterioplankton communities through variations in DOM composition and highlight the importance of both syntrophic and antagonistic interactions in DOM consumption, their inherent complexity precludes the identification of specific interactions between a single taxa and DOM component without *a priori* knowledge of a component's lability. A handful of recent studies have demonstrated how phytoplankton bloom dynamics and DOM quality can affect heterotrophic diversity on a fine phylogenetic scale (i.e. within a specific genus or functional group) (Lau et al. 2007, Watanabe et al. 2012). These findings highlight the need to conduct DOM uptake studies on individual bacterioplankton strains in addition to whole communities, and caution against the generalization of a given DOM response to entire taxa. Research that moves beyond the question of which bacterioplankton taxa respond to which DOM source and into questions regarding why these observed variations occur is now needed, and feasible. Simple, yet representative model systems for DOM consumption would be a valuable laboratory tool with which to examine microbial utilization of marine DOM in sufficient detail.

Advances in culturing methods, including high-throughput dilution to extinction isolation techniques, have led to the acquisition of many previously “unculturable” marine bacteria isolates (Connon & Giovannoni 2002, Rappé et al. 2002). Representative isolates from several phylogenetic lineages common in the marine environment, including genetically diverse subgroups of the ubiquitous SAR11 clade, SAR116 clade, OM43 clade, *Pseudoalteromonas*, and *Rhodobacterales*, offer a unique opportunity to test DOM enrichments on an array of microbial isolates that are only recently domesticated. Axenic isolates provide a direct means to study the consumption of various DOM sources by individual heterotrophic taxa in the absence of compounding factors such as grazing and competition. While these factors are important aspects of natural marine systems, removing them provides a means to unequivocally determine if specific taxa are able to utilize a particular DOM source and further manipulate these relationships in a controlled setting.

Here we present a screening assay based on cultivated bacterial isolates in order to identify model systems for DOM consumption. Representative marine bacteria from diverse phylogenetic lineages were grown in small volume replicate axenic cultures on an extensive matrix of organic substrates to investigate microbial consumption of marine DOM. To obtain high throughput measurements of cell abundance, a Guava EasyCyte Plus flow cytometer (Millipore) was employed. This instrument allowed for the automated determination of absolute cell enumeration of bacterioplankton by direct measure of sample volume without the use of any secondary measurements. Measurements were made on small sample volumes in a 96-well plate format, allowing

for rapid analysis of the thousands of individual culture-substrate combinations necessary to test replicate amendments at various doses alongside extensive controls. DOM sources included various concentrations of high-molecular weight (HMW) DOM from distinct productivity regimes in the North and South Pacific Oceans, and hydrophobic DOM derived from axenic cultures of marine phytoplankton. A range of DOM concentrations were tested as the growth potential of heterotrophic bacteria in seawater has been shown to peak at intermediate levels of phytoplankton abundance (Tada et al. 2011). DOM exudate from phytoplankton cultures was chosen over lysate because it is likely more representative of what bacterioplankton encounter *in-situ*, and it has recently been shown to be the more reactive DOM fraction when supplied to natural microbial populations (Nelson & Carlson 2012). Phytoplankton-derived DOM are known to induce unique changes in bacterioplankton community structure when compared to more traditional organic substrates used to assess secondary production (i.e. leucine, glucose, etc.), demonstrating the importance of learning more about how heterotrophic bacteria respond to these environmentally relevant substrate types (Nelson & Carlson 2012, Sarmiento & Gasol 2012).

In this study, the use of phytoplankton-derived DOM obtained by solid-phase extraction (SPE) provided a unique opportunity to examine uptake of labile DOM from a known biological source that could also be chemical interrogated by mass spectrometry. Here we define labile DOM as material that reproducibly enhanced the growth of an axenic bacterioplankton strain over the course of several days in batch culture. Highly labile DOM compounds in nature are in such great demand that their concentration in the

marine environment is likely kept extremely low, making it difficult to study the consumption of this material in detail. DOM derived from pure cultures of marine phytoplankton has not been processed by exogenous heterotrophic activity; therefore even the most labile compounds can be isolated for use in degradation studies.

Compositional analyses of these same DOM sources via mass spectrometry allowed for a rare glimpse into the chemical nature (compound size and polarity range) of biologically relevant DOM components.

The development of a rapid screening assay to monitor the growth response of bacterial isolates amended with various DOM substrates is useful to investigate DOM utilization among a diverse group of heterotrophic bacteria and can also be used to identify specific bacterium-substrate combinations for detailed metabolic and chemical analyses. Using this approach, substrates that enhance the growth of particular bacterioplankton taxa can also be potentially identified for use in future isolation and cultivation-based studies. An additional screening assay using individual chemical fractions from a single phytoplankton-derived DOM source was employed to resolve higher order variations in the uptake and utilization of specific compounds by individual bacterioplankton taxa. Dividing DOM sources into separate fractions can be used to reduce the pool of potentially reactive components in a given DOM sample, making the chemical identification of reactive compounds from a known biological source a realistic goal.

Materials and Methods

Control medium preparation

The base medium used for all screens was sterilized seawater collected from the South Bay portion of Kaneohe Bay, HI (21° 26.181'N, 157° 46.642'W). Surface seawater was collected in a clean polycarbonate bucket and transferred to acid washed and seawater rinsed 20L carboys. Seawater was sterilized by in-line pre-filtration through a 142 mm 0.8 µm Supor-800 membrane (Pall Corporation), followed by a 0.1 µm Supor AcroPak 1000 PES capsule filter (Pall Corporation) before tangential flow filtration (TFF) through a 50 cm² 30 kDa regenerated cellulose Pellicon 2 mini filter (Millipore) according to existing protocols (Becker et al. 2007). The sterile permeate was collected in acid washed polycarbonate bottles fitted with ported caps inside a laminar flow hood (Figure 1). Sterilized seawater was stored in the dark at 4 °C. A subsample of each collection bottle (20 ml) was placed at 26 °C for several weeks and periodically tested for sterility via microscopy. Collection bottles with subsamples displaying any signs of microbial growth were not used for control medium preparation. Sterile seawater was amended with small additions of inorganic nitrogen (1 µM NH₄⁺; 1 µM NO₃⁻) and phosphorus (0.1 µM PO₄³⁻) to produce the control medium, referred to hereafter as NP medium.

High-molecular weight DOM amendments

High-molecular weight (> 1kDa) DOM (HMW-DOM) was isolated and concentrated from the Natural Energy Laboratory of Hawaii Authority (NELHA) in

Kona, HI in 2006, an oxygen deficient zone off the coast of Chile ($20^{\circ} 04.999'S$, $70^{\circ} 48.001'W$) and two locations in the South Pacific subtropical gyre ($23^{\circ} 27.510'S$, $88^{\circ} 46.089'W$; $26^{\circ} 14.844'S$, $103^{\circ} 57.676'W$) during the C-MORE BiG RAPA cruise in 2010 (RAPA01, RAPA04, and RAPA07 respectively). Seawater drawn from the 15 m intake pipe at the NELHA was filtered to remove bacteria and small particles using an acid cleaned (10% HCl) Suporflow dual-stage (0.8 mm and 0.2 mm) Gelman polyether sulfone cartridge filter (Chisolm Corp.) fitted to an Advanta stainless-steel housing. HMW-DOM samples were collected using a cross-flow ultrafiltration system consisting of a stainless-steel centripetal pump and membrane housings and a fluorinated high-density polyethylene reservoir. The system was plumbed with Teflon tubing and fitted with polyvinyl difluoride valves. Membranes were cleaned using isopropanol, detergent (0.01% Micro 90; Cole-Parmer), HCl (0.01 M), and NaOH (0.01 M); stored in sodium azide (0.55 mM); and rinsed with seawater immediately before use. 24,000 L of seawater were concentrated to ca. 20 L, frozen in fluorinated high-density polyethylene containers, and returned to the Woods Hole Oceanographic Institution for further processing. Samples were desalted by diafiltration with water, reduced to 2 L, lyophilized to a fluffy white powder, and stored in the dark at $4^{\circ}C$.

South Pacific surface seawater (15m) was collected using acid-cleaned Teflon tubing connected to a compressed air-driven diaphragm pump (Wilden) and concentrated ca. 100-fold using a single thin-film ultrafiltration membrane element (Separation Engineering) in a custom-built polycarbonate membrane housing after pre-filtration through a $0.2 \mu m$ Polycap TC filter (Whatman). Both ultrafiltration membranes

(Separation Engineering) retained organic matter with a nominal molecular weight of >1 kDa (>99% rejection of vitamin B12). South Pacific samples underwent a final 30 kDa filtration step to immediately remove viral particles before freezing in the dark at -20 °C. Samples were taken for DOC analysis at various stages throughout the concentration process to track the partitioning of organic carbon (Figure 2). 150 mg of lyophilized HMW-DOM (38% carbon as determined by DOC measurements) from NELHA was resuspended in 15 ml of control medium (vortex mixed) and filter sterilized through a 0.2 µm Steriflip filter (Millipore). Various amounts (190 µl, 380 µl, 760 µl) were then added to 30 ml of NP medium to produce treatment medium (NELHA-S, -M, -L, respectively). 20 ml of HMW-DOM from each South Pacific location was thawed and added to 180 ml of control medium, then filter sterilized through a pre-rinsed 0.1 µm Anotop syringe filter (Whatman) to produce treatment media (RAPA01, RAPA04 & RAPA07). Samples of all treatment media were taken for dissolved organic carbon analysis.

Phytoplankton-derived DOM amendments

DOM produced and released by five axenic marine phytoplankton cultures was isolated by solid-phase extraction (SPE) and characterized for use as media amendments. DOM obtained from phytoplankton cultures in this manner is hereafter referred to as DOM_P. Three strains of cyanobacteria (*Prochlorococcus* str. MIT9313, *Synechococcus* str. WH8102, *Crocospaera* str. WH8501) and two eukaryotic phytoplankton strains (the diatoms *Thalassiosira pseudonana* strain CCMP1335 and *Phaeodactylum tricorutum* strain CCMP632) were chosen to examine a diverse range of DOM sources. Large

volume (20L) cultures of each strain were prepared to obtain a sufficient quantity of DOM_P for use in growth assays. The base for all phytoplankton growth media was oceanic surface water collected in the Sargasso Sea and filtered (0.2 μm; Whatman Polycap 36 TC capsule filter) before transportation back to the laboratory. This water was then sub-sampled for each growth experiment and ultra-filtered by 500 Da using a pre-rinsed large volume ultrafiltration unit to reduce background levels of DOC. Ultra-filtered (< 500 Da) seawater was then sterilized by 0.1 μm filtration (Whatman Polycap 36 TC capsule filter) and/or autoclaving just prior to media preparation. A negative control (20 L of medium only with no cells added) was also prepared and processed along with each culture in order to distinguish between DOM_P and DOM present in the background seawater media. *Prochlorococcus* str. MIT9313 was grown in Pro99 medium prepared according to existing protocols (Moore et al. 2007). The culture was maintained at 22 °C in 24 h light (ca. 20 μE/m²/s) and monitored for growth using fluorescence. *Synechococcus* str. WH8102 and *Crocospaera* str. WH8501 were grown in SN medium prepared according to existing protocols (Waterbury et al. 1986). *Synechococcus* and *Crocospaera* cultures were maintained at 24 °C and 30 °C respectively and monitored for growth via direct cell counts. *Thalassiosira pseudonana* str. CCMP1335 and *Phaeodactylum tricornutum* str. CCMP632 were grown in f/2 medium prepared according to existing protocols (Guillard 2002), maintained at 15 °C and monitored for growth via direct cell counts.

All phytoplankton cultures were monitored for both particulate and dissolved organic carbon production (POC and DOC respectively). 50 ml of culture was vacuum

filtered through a combusted 47 mm 0.7 μm glass fiber filter (Whatman GF/F) using combusted glassware for POC analysis and the filtrates were collected in combusted glass vials for DOC analysis. Upon reaching stationary phase (7 - 13 days), cells were removed by centrifugation, followed by gentle filtration through a 0.1 μm Polycap 36 TC capsule filter (Whatman). Filtrates were stored briefly in the dark at 4 °C until solid-phase extraction of metabolites. Filtrates were removed from the fridge and acidified to pH 2-3 by adding trace metal grade hydrochloric acid (HCl) before loading onto a custom packed column loaded with soxhlet extracted octadecyl (C_{18}) functionalized silica gel (Sigma-Aldrich) at a rate of 2 mL/min. Columns were washed with 500 mL of acidified ultrapure water (pH 2-3) at a flow rate of 1 mL/min to remove salts before eluting with 10 column volumes of acidified HPLC-grade methanol (pH 2-3) at a rate of 1 mL/min. Salt removal was confirmed using a silver nitrate solution and the eluted material was collected in a combusted flask and concentrated down to 1.5 mL using a rotary evaporator. Subsamples (10 μL) were placed onto combusted 25 mm 0.7 μm glass fiber filters (Whatman GF/F) for POC analysis to quantify the amount of organic material recovered via solid-phase extraction. A 100 μL subsample was also taken for chemical analysis using high performance liquid chromatography coupled to a single quadrupole mass spectrometer (HPLC-MS). The remaining sample was dried using filtered high purity nitrogen gas and rinsed several times with ultrapure water to remove residual methanol before a final drying step. Dried DOM_p was stored in a combusted amber vial in the dark prior to resuspension in 10 ml of control medium. Resuspended DOM_p was then sterile filtered through a pre-rinsed 0.1 μm Anotop filter (Whatman) before various

amounts (1.7 ml, 8.3 ml) were added to 20 ml of control media to produce treatment media (Pro-, Syn-, Cro-, Pse-, Tri- DOM_p; -S, -L, respectively).

DOM-derived from a small volume subsample (1L) of *Crocospaera* str. WH8501 was also prepared as an amendment prior to SPE. Cell biomass was removed via centrifugation and filtration as previously described and DOM was stored in the dark at -20°C. *Crocospaera* DOM was thawed and filter sterilized through a 0.2 µm Steriflip filter (Millipore) before various amounts (7 ml, 9.5 ml) were added to 30 ml of NP medium to produce treatment media (CroDOM -S, -L, respectively). Samples of all treatment media were taken for dissolved organic carbon analysis and control media corresponding to each DOM_p sample was also prepared via SPE of cell-free phytoplankton media (Pro99-, SN-, f/2- DOM_p; -S, -L).

Fractionation of Prochlorococcus-derived DOM_p

Three additional 20 L cultures of *Prochlorococcus* str. MIT9313 were grown in Pro99 medium prepared according to existing protocols (Moore et al. 2007) for use in a final round of DOM screens. The replicate cultures were maintained at 22 °C in 24 h light (ca. 20 µE/m²/s) and monitored for growth using bulk fluorescence and flow cytometry counts alongside a 20 L cell-free Pro99 medium control. Samples for direct cell counts by flow cytometry were fixed with 0.125% final concentration of grade I glutaraldehyde (Sigma) and stored at -80°C prior to analysis. Flow cytometry samples were obtained four times during the growth of the 20L batch cultures. Fixed samples were diluted in 0.2 µm filtered seawater and *Prochlorococcus* were enumerated using an

Influx Cell Sorter (BD Biosciences) following (Olson et al. 1985, Cavender-Bares et al. 1999). ProDOM_P was obtained via SPE as previously described with the following modification: SPE was performed using Discovery DSC-18 SPE tubes (Sigma-Aldrich) with a 10g bed weight and 60 ml volume capacity (rinsed with HPLC grade methanol and ultrapure water prior to use) instead of custom packed columns. Each replicate culture was subsampled for POC analysis and for characterization via HPLC-MS before the 3 replicates were pooled into one ProDOM_P sample. A portion of this pooled sample (5%) was removed and HPLC was used to separate the ProDOM_P into 5 distinct fractions based on polarity. Chromatographic separation and detection was achieved using an Agilent 1200 series liquid chromatograph comprised of a G1379B degasser, G1312A binary pump, G1367C automatic liquid sampler and F1315C diode array detector coupled to an Agilent G1364B fraction collector. The mobile phases were 0.1% formic acid in water and 0.1% formic acid in methanol. 25 μ L the pooled ProDOM_P was injected onto an Agilent ZORBAX SB-C18 column (3.5 μ m 4.6x150 mm) at a flow of 1 mL min⁻¹ (starting with 100% A, ramping to 80% B at 28 min, ramping to 100% B at 39.2 min and holding until 61.6 min, ramping to 0% B at 72.8 min and holding until 84 min). Full scan absorbance data were acquired from 210 to 800 nm with a 2.0 nm step and 4 nm slit width. Fractions were collected from 0-15 min (0-38.2% B), 15-25 min (38.2-63.9% B), 25-36 min (63.9-87.5% B), 36-50 min (87.5-100% B), and 50-84 min (100-0% B). The unfractionated pooled sample and each fraction were concentrated down to 1.5 mL using a rotary evaporator, subsampled for POC analysis (10 μ L), dried using filtered high purity nitrogen gas and rinsed several times with ultrapure water to remove residual

methanol before a final drying step. Dried ProDOM_p fractions and the pooled sample were stored in combusted amber vials in the dark prior to resuspension in 10 ml of control medium. Resuspended DOM_p was then sterile filtered through a pre-rinsed 0.1 µm Anotop filter (Whatman) before adding it to 125 ml of control medium to produce treatment media (ProDOM_p-pooled; ProDOM_p-F1 through F5). Samples of all treatment media were taken for dissolved organic carbon analysis and control medium prepared from the cell-free Pro99 media control was also prepared in the same fashion (Pro99DOM_p-2).

Quantification of organic carbon

Combusted glassware (450 °C for 8 h) was used for all sampling. Duplicate or triplicate sub-samples of 30 mL for dissolved organic carbon (DOC) analysis were collected in glass vials and acidified with 150 µl of a 25% phosphoric acid solution before sealing with acid-washed Teflon lined septa and storage at 4 °C until processing. TOC analysis was performed using the high temperature combustion method on a Shimadzu TOC-V_{CSH} with platinized aluminum catalyst coupled to a TNM-1 total nitrogen detector. Sample concentrations were determined alongside potassium hydrogen phthalate and potassium nitrate standards and consensus reference materials provided by the DOC-CRM program (www.rsmas.miami.edu/groups/biogeochem/CRM.html). Duplicate samples for particulate organic carbon (POC) analysis were collected onto combusted 25 mm 0.7 µm glass fiber filters (Whatman GF/F). Filters were placed inside combusted glass petri dishes, wrapped in foil, and immediately frozen. Blank filters were

also prepared at every sampling point. Filters were then thawed and put in a drying oven (60 °C) overnight to ensure they were thoroughly dry before encapsulation into 9x10 mm tin capsules and shipped to the University of California Davis Stable Isotope Facility for analysis.

Chromatographic separation and detection of DOM_P metabolites

Chromatographic separation and detection was achieved using an Agilent 1200 series liquid chromatograph comprised of a G1379B degasser, G1312A binary pump, G1367C automatic liquid sampler and F1315C diode array detector coupled to an Agilent 6130 mass spectrometer. The mobile phases were 0.1% formic acid in water and 0.1% formic acid in methanol. 25 µL of each DOM_P sample and its respective media control was injected onto an Agilent ZORBAX SB-C18 column (3.5 µm 4.6x150 mm) at a flow of 1 mL min⁻¹ (starting with 100% A, ramping to 80% B at 25 min, ramping to 100% B at 35 min and holding until 55 min, ramping to 0% B at 65 min and holding until 75 min). Full scan absorbance data were acquired from 210 to 800 nm with a 2.0 nm step and 4 nm slit width. Mass spectrometry was performed using an atmospheric electrospray ionization source. Source conditions were as follows: drying gas at 11.5 L min⁻¹, nebulizer at 60 psig, drying gas temperature at 300 °C, capillary voltages at + or - 4000 V. Acquisition ranges were from 100-2000 Da in the positive mode and used a fragmentor set at 4.0, threshold at 150 and a step size of 0.1. Positive mode mass spectral data was processed by converting the native Agilent ".ms" format files to the industry standard common data format files using GCMS File Translator Pro (ChemSW). Data

was then processed using MZmine 2 (Pluskal et al. 2010). Ions with a minimum signal intensity that was 5-fold greater than the maximum noise level were detected using a centroid mass detector. Chromatograms were then built from the raw data using the same minimum signal intensity, a retention time tolerance of 5 s, and a m/z tolerance of 0.3. Features (defined here as an individual m/z arriving at a particular retention time) that were also present in the corresponding cell-free media control were subtracted from each DOM_P sample.

Bacterioplankton growth assay

Fifteen axenic bacterioplankton isolates were chosen to screen for growth effects due to various DOM sources (Table 1). All bacterioplankton strains were originally isolated from Kaneohe Bay between 2005 and 2010 using extinction culturing techniques (Connon & Giovannoni 2002, Becker et al. 2007) and included representatives from several phylogenetic lineages known to be abundant in Kaneohe Bay (Yeo et al. 2013). Tissue culture-grade polystyrene or polycarbonate plasticware (Nalge Nunc Int.) soaked overnight in 10% HCl and rinsed thoroughly with ultrapure water before autoclaving was used for all culturing and media storage. Measures taken to minimize the risk of contamination during media preparation, inoculation and screening included: frequent changing of gloves and sterile pipette tips, minimized exposure of human skin during manipulations, and minimized time in which cultures and media were open and exposed to the air. All culture manipulations were performed under a laminar flow hood with prior sterilization by ethanol and ultraviolet light. NELHA-S, -M, -L and CroDOM -S, -

L treatment media were prepared for 30 ml cultures, while all other treatment media were prepared for 4 ml cultures in 5ml 48-well plates (Axygen Scientific Inc.) covered with a permeable adhesive film (VWR International). 30 ml seed cultures were prepared and subcultured multiple times in NP medium to provide inoculate after identity confirmation via sequencing of the small subunit (SSU) rRNA gene. All cultures were maintained at 26-27 °C on a light-dark (12:12) cycle with periodic gentle mixing to mimic origin conditions. All media types were preconditioned in this environment prior to inoculation. Cultures were periodically checked visually for obvious contamination via epifluorescence microscopy and combinations exhibiting enhanced growth were checked for purity and proper identification at the end of the experiment (i.e. late stationary phase) via sequencing of the small subunit (SSU) rRNA gene. Genomic DNA was extracted and the SSU rRNA gene was amplified for sequence identification as previously described (Becker et al. 2007).

Cell enumeration was determined by either direct epifluorescence microscopy in a 48-well format or using a Guava EasyCyte Plus cytometry system (Millipore Corp.) in a 96-well format. To screen by epifluorescence microscopy, 125 µl of multiple cultures were arrayed on polycarbonate membranes (GE Osmonics Inc.) via vacuum filtration using a custom designed 48-well dot blot manifold containing small-diameter (ca. 1 mm) wells (HyTek Plastics). The number of cells in each well was determined by direct cell counts after fixation with 1% buffered formaldehyde and staining with the DNA-binding dye 4',6-Diamidino-2-phenylindole (DAPI). Fluorescent microspheres were also added to each well (ca. 20 beads per field of view at 1000x magnification) to aid in well

location and size estimation. Before and after each use, the top portion of the 48-well manifold was thoroughly cleaned with ultrapure water and sonication to prevent sample carryover. Membranes were mounted on 75 x 50 mm microscope slides using 48 x 60 mm coverslips and high viscosity immersion oil. In some cases, slides were stored temporarily at 4 °C before enumeration. Screening using the Guava EasyCyte Plus cytometry system was performed by adding 1 µl of a freshly prepared SYBR Safe DNA gel stain (Life Technologies) working stock (1:20 dilution in ultrapure water) to 99 µl of each culture arrayed into a 96-well polycarbonate plate (BD Falcon), and incubating for 1 h in the dark. Plates were counted using the Guava ExpressPro software with custom settings for each strain (determined empirically). General settings that worked for most strains included 147 V forward scatter (x128), 504 V side scatter, and 649 V for the green wavelength with a threshold of 16. Cell counts were performed only after accurate counting of the bead check solution. The cytometer was cleaned with ultrapure water prior to each use and with both ultrapure water and instrument cleaning fluid at the end of each day. Dilution of growing cultures using sterile NP medium was performed as necessary to stay below the cytometer's upper limit of accurate counting (ca. 5.5×10^5 cells/ml). Inoculations were performed to achieve a starting cell concentration equal to the cytometer's lower limit of accurate counting (8×10^3 cells/ml).

Results and Discussion

Growth responses to DOM amendments from a wide variety of sources were monitored for more than a dozen bacterioplankton strains isolated from the Kaneohe Bay,

Hawaii (Table 1). The culture collection employed in this study provided a unique opportunity to test DOM amendments on a diverse array of bacterioplankton isolates that had not previously been domesticated for manipulation in the laboratory. These isolates were obtained using a high-throughput dilution to extinction cultivation approach (Becker et al. 2007) and comprised axenic strains from multiple lineages within the Gamma-, Beta- and Alphaproteobacteria, including several genetically diverse strains of the ubiquitous SAR11 clade (Figure 3). High-resolution cell counts were taken to examine the effects of DOM enrichments on the growth rate and yield of these naturally abundant marine bacteria. The testing of various sources, concentrations and controls in triplicate resulted in over 1,000 individual batch cultures and more than 15,000 individually enumerated samples.

We found a broad range of responses, from inhibitory effects to significant enhancement of both growth rate and yield. These responses depended on both the bacterioplankton strain and on the source and quantity of the DOM amendment, indicating that both the identity of the consumer and the source of the material play a role in determining DOM lability and that heterotrophic (and mixotrophic) bacterioplankton diversity is related to the composition of marine DOM. Here we are not considering lability as an intrinsic property that holds under all scenarios, but rather as a dynamic parameter that depends on many factors, including the identity of a potential consumer. Our findings support a large body of research, dating back several decades, indicating that different bacterioplankton taxa preferentially utilize DOM from different sources (e.g. (Bell et al. 1974, Cottrell & Kirchman 2000, Carlson et al. 2004, Lau et al. 2007,

Moreira et al. 2011, Tada et al. 2011, Trabelsi & Rassoulzadegan 2011, Nelson & Carlson 2012, Sarmiento & Gasol 2012, Gómez-Consarnau et al. 2012). Testing a diverse suite of DOM sources prepared in various ways allowed for investigating response dynamics of individual bacterioplankton strains to DOM sources of distinct size, chemistry and origin. Amendments with DOM derived from axenic cultures of marine phytoplankton prepared by SPE allowed for a rare look at the consumption of DOM from a single-organism source without any prior biological degradation. Chemical interrogation of these DOM_P sources by mass spectrometry revealed significant compositional variation among different phytoplankton sources (Tables 2 & 3, also Chapter 2) and differential responses to different DOM_P sources by the same bacterioplankton strain indicate that these compositional differences translate into lability differences. Taken as a whole, our results demonstrate that DOM lability is a complex and dynamic parameter that depends heavily on its biological source and sink and highlight the utility of amending individual bacterioplankton strains with various DOM sources to capture strain-specific responses under controlled conditions.

Natural Energy Laboratory of Hawaii Authority HMW-DOM

HMW-DOM prepared from NELHA in Kona, Hawaii was used to amend control medium at three DOC concentrations: 247 μM , 929 μM , and 2 mM above control medium DOC (NELHA-S, -M, -L, respectively). These concentrations correspond to roughly 3X, 10X, and 20X control medium DOC levels. For reference, previous experiments that have been conducted off the Oregon Coast typically involved DOC

additions on the order of ca. 2 - 3 mM (35X) above baseline levels (Rappé et al. 2002, Cho & Giovannoni 2004). Three concentrations of NELHA media were prepared to examine dose effects on the growth of several marine bacterioplankton strains (Table 4). Direct cell counts via epifluorescence microscopy taken every 48 hr for a period of 4 weeks revealed strain-specific responses ranging from slightly suppressed growth yield (Figure 4a) to enhanced growth rate and yield (Figure 4b). The most striking effect of the NELHA additions was unexpectedly found for marine *Synechococcus* str. HIMB12, which exhibited a dose specific increase in both growth rate and maximum cell density (Figure 4c). Increased growth rate and biomass production of HIMB12 in response to additions of HMW-DOM support previous observations of mixotrophy in marine *Synechococcus* spp. (Berman & Chava 1999, Sowell et al. 2009, Baran et al. 2011) and further investigation into the uptake of HMW-DOM by this and other cyanobacteria could alter current thinking on microbial DOM cycling and how bacterial production is currently measured. A less extreme dose specific positive growth response to NELHA HMW-DOM was also found for HIMB30, an isolate from the bacterial order *Oceanospirillales*, and dose specific suppressed growth yield was found in 3 of the 4 strains tested from the SAR11 clade.

South Pacific HMW-DOM

HMW-DOM prepared from three locations in the South Pacific was used to amend control medium at a target DOC concentration of 1.5X above control medium. HMW-DOM from 890 L of 15m water from an oxygen deficient zone off the coast of

Chile (RAPA01) was concentrated to 8.9 L and a portion was added to control medium, resulting in a 37 μM increase in DOC (1.3X). HMW-DOM from 940 L of 15m water from a transition zone between the coast of Chile and the South Pacific subtropical gyre (RAPA04) was concentrated to 9.4 L and a portion was added to control medium, resulting in a 57 μM increase in DOC (1.5X). HMW-DOM from 1,055 L of 15m water from an area near the center of the South Pacific subtropical gyre (RAPA07) was concentrated to 10.5 L and a portion was added to control medium, resulting in a 64 μM increase in DOC (1.5X). Ten strains of heterotrophic bacterioplankton were inoculated into each HMW-DOM treatment media in triplicate alongside triplicate control medium. High-resolution cell enumeration performed over a 10 – 17 day time period via flow cytometry revealed strain-specific responses ranging from suppressed (Figure 5a) to enhanced growth (Figure 5b). All five strains of the SAR11 clade, representing three different subgroups exhibited either suppressed growth or no effect from all South Pacific HMW-DOM amendments. HIMB11, the single *Rhodobacteriales* strain tested, also exhibited suppressed growth in the presence of all three amendments, indicating an overall negative effect of South Pacific HMW-DOM addition on all Alphaproteobacteria strains tested (Figure 5c). One strain of gammaproteobacteria from the OM60 clade, HIMB55, displayed an increase in growth rate and maximum cell density when grown with the RAPA01 and RAPA04 amendments, and an increase in growth rate when grown with the RAPA07 amendment. Strain HIMB30 also exhibited an increase in growth rate and maximum cell density due to the RAPA04 amendment, indicative of an overall positive effect of the South Pacific HMW-DOM addition on the Gammaproteobacteria

tested. Of the three South Pacific HMW-DOM additions, RAPA04 induced the most growth enhancement, while RAPA07 was the least bioactive (Table 5). Genomic DNA extraction and subsequent sequencing of the SSU 16S rRNA gene confirmed both the purity and strain identity of all isolates that exhibited a positive growth response due to the addition of South Pacific HMW-DOM.

Crocospaera str. WH8501 DOM and DOM_p

A large volume (20 L) axenic culture of *Crocospaera* str. WH8501 was prepared in SN medium with a natural seawater base to obtain *Crocospaera*-derived DOM for use in growth screens. WH8501 reached a maximum cell density of 1.26×10^6 cells/ml over a growth period of 8 days and purity was confirmed via direct microscopic observations. A maximum quantity of 0.5 mM POC was produced after 6 days of growth (388.9 fmol/cell) and a maximum quantity of 83.8 μ M DOC was produced and released after 8 days of growth (66.5 fmol/cell). 15% of the total organic carbon produced by WH8501 was released as DOC into the surrounding medium and ca. 8% of this DOC was obtained via SPE after biomass removal (Table 6; Figure 6).

Prior to SPE, 1L of WH8501 spent medium (i.e. after biomass removal) was removed and filter sterilized for preparation of CroDOM treatment medium at two concentration levels to examine dose effect responses of various bacterioplankton to additions of total *Crocospaera*-derived DOM. CroDOM-S and CroDOM-L media elevated control medium DOC by 7.9 μ M and 29.6 μ M, respectively. These concentrations correspond to roughly 1.1X, and 1.3X control medium DOC levels.

Direct cell counts via epifluorescence microscopy taken every 48 hr for a period of 8 days (plus a final count after 14 days) revealed strain-specific responses including slight suppression of growth rate (Figure 7a) to enhanced growth (Figure 7b). Lack of replication and sufficient temporal resolution of cell counting rendered screens with other bacterioplankton strains inconclusive; however, despite these being relatively small additions of DOC, growth rate enhancement was observed in 2 out of the 3 isolates amended with CroDOM relative to controls, suggesting the presence of reactive organic compounds in CroDOM (Table 7).

The remaining 19 L of WH8501 spent medium (i.e. after biomass removal) was processed by SPE to obtain hydrophobic *Crocospaera*-derived DOM for preparation of CroDOM_P medium and for chemical interrogation by mass spectrometry. HPLC-MS analysis of DOM_P from WH8501 revealed the presence of only 13 features not found in the background SN medium control sample (Table 2), suggesting that WH8501 did not release a significant amount of DOC capable of SPE in this manner. The 13 features detected were small (308 – 363 m/z) and relatively nonpolar with retention times corresponding to 90 – 100% methanol. SPE material from both the WH8501 culture and SN medium control were used to prepare treatment and control medium at two concentration levels to examine dose effect responses of various bacterioplankton to additions of *Crocospaera*-derived hydrophobic DOM. The SNDOM_P controls were an important addition as they allow for distinguishing between bacterioplankton responses due to DOM present in the natural seawater-based SN medium from responses due to DOM produced and released by the WH8501 culture. CroDOM_P-S and CroDOM_P-L

media elevated control medium DOC by 1.9 μM and 6.7 μM , respectively. SNDOM_{P-S} and SNDOM_{P-L} media elevated control medium DOC by 0.2 μM and 6.7 μM , respectively. For both WH9501 and SN treatments, these concentrations correspond to roughly 1.01X, and 1.06X control medium DOC levels and the similarity in these DOC increases supports the conclusion from the mass spectral data that the majority of the DOC obtained by SPE from the WH8501 culture was also present in the SN medium. High-resolution cell counts of triplicate cultures for each media type taken over a 10 – 17 day time period via flow cytometry revealed strain-specific responses ranging from suppressed (Figure 8a) to enhanced growth (Figure 8b). Overall, CroDOM_P suppressed the growth of more isolates (6) than it enhanced (2). Suppressed growth was also found in five of the ten strains amended with SNDOM_P, indicating toxic effects from material present in the background SN medium. All Alphaproteobacteria strains tested exhibited either suppressed growth or no effect from the CroDOM_P additions. Only two strains demonstrated enhanced growth effects from CroDOM_P – HIMB55 displayed enhanced growth rates, while HIMB1363, an unclassified *Flavobacterium* from the Bacteroidetes phylum, displayed enhanced maximum cell yield. Genomic DNA extraction and sequencing of the SSU 16S rRNA gene confirmed both the purity and strain identity of all isolates that exhibited a positive growth response due to the addition of CroDOM_P.

Synechococcus str. WH8102 DOM_P

A large volume (20 L) axenic culture of *Synechococcus str. WH8102* was prepared in SN medium with a natural seawater base to obtain *Synechococcus*-derived

DOM for use in growth screens. WH8102 reached a maximum cell density of 3.49×10^7 cells/ml over a growth period of 13 days and purity was confirmed via direct microscopic observations. A maximum quantity of 0.4 mM POC was produced after 10 days of growth (11.5 fmol/cell) and a maximum quantity of 59.1 μ M DOC was produced and released after 13 days of growth (1.7 fmol/cell). 13% of the total organic carbon produced by WH8102 was released as DOC into the surrounding medium and ca. 3% of this DOC was obtained via SPE after biomass removal (Table 6).

WH8102 spent medium (i.e. after biomass removal) was processed by SPE to obtain hydrophobic *Synechococcus*-derived DOM for preparation of SynDOM_P medium and for chemical interrogation by mass spectrometry. HPLC-MS analysis of DOM_P from WH8102 revealed the presence of 302 features not found in the background SN medium control sample (Table 2), indicating that WH8102 released a significant amount of DOC capable of acquisition by SPE. The 302 features detected ranged from 114 – 854 m/z over a broad polarity spectrum (retention times corresponding to 20 – 100% methanol). SPE material from both the WH8102 culture and SN medium control were used to prepare treatment and control media at two concentration levels to examine dose effect responses of various bacterioplankton to additions of *Synechococcus*-derived hydrophobic DOM. SynDOM_P-S and SynDOM_P-L media elevated control medium DOC by 1.8 μ M and 7.7 μ M, respectively, corresponding to roughly 1.01X and 1.06X control media DOC levels. SNDOM_P-S and SNDOM_P-L media elevated control medium DOC by 2.1 μ M and 6.0 μ M, respectively. High-resolution cell counts of triplicate cultures for each media type taken over a 10 – 17 day time period via flow cytometry revealed no

conclusive strain-specific responses. Overall, SynDOM_p caused a slight suppression in the growth of more isolates (4) than it enhanced (3), but none of these results were clearly reproducible or dose specific. A unique diauxic growth pattern with significant secondary growth enhancement was detected for a *Flavobacteriales* strain amended with SynDOM_p, however subsequent genomic DNA extraction and sequencing of the SSU 16S rRNA gene revealed the presence of an Alphaproteobacteria contaminant in these cultures.

Prochlorococcus str. MIT9313 DOM_p

A large volume (20 L) axenic culture of *Prochlorococcus* str. MIT9313 was prepared in Pro99 medium with a natural seawater base to obtain *Prochlorococcus*-derived DOM for use in growth screens. MIT9313 reached a maximum cell density of 6.29×10^7 cells/ml over a growth period of 7 days and purity was confirmed via purity broth tests for the presence of other bacteria. A maximum quantity of 0.8 mM POC was produced (12.7 fmol/cell) and a maximum quantity of 75.4 μ M DOC was produced and released (1.2 fmol/cell) after 13 days of growth. 9% of the total organic carbon produced by MIT9313 was released as DOC into the surrounding medium and ca. 3% of this DOC was obtained via SPE after biomass removal (Table 6).

MIT9313 spent medium (i.e. after biomass removal) was processed by SPE to obtain hydrophobic *Prochlorococcus*-derived DOM for preparation of ProDOM_p medium and for chemical interrogation by mass spectrometry. HPLC-MS analysis of DOM_p from MIT9313 revealed the presence of 82 features not found in the background Pro99

medium control sample (Table 2). The 82 features detected ranged from 139 – 536 m/z over a broad polarity spectrum (retention times corresponding to 4 – 100% methanol). SPE material from both the MIT9313 culture and Pro99 medium control were used to prepare treatment and control media at two concentration levels to examine dose effect responses of various bacterioplankton to additions of *Prochlorococcus*-derived hydrophobic DOM. ProDOM_P-S and ProDOM_P-L media elevated control medium DOC by 1.2 μM and 3.7 μM, respectively, corresponding to roughly 1.01X and 1.03X control media DOC levels. Pro99DOM_P-S and Pro99DOM_P-L media elevated control medium DOC by 0.5 μM and 3.1 μM, respectively. High-resolution cell counts of triplicate cultures for each media type taken over a 10 – 17 day time period via flow cytometry revealed more positive growth effects due to ProDOM_P than for any other DOM source tested. This is especially interesting due to the extremely small increase in total DOC caused by the ProDOM_P additions and the relatively small number of features found in ProDOM_P that were not present in Pro99DOM_P. Enhanced growth rate and increased maximum cell density was found in 5 of the 11 strains amended with ProDOM_P. While all three of the strains tested from the IA subgroup of the SAR11 clade demonstrated either no effect or slight growth suppression due to ProDOM_P amendment, both strains tested from the IIIA subgroup of the SAR11 clade displayed enhanced growth rate and maximum cell yield in response to ProDOM_P additions (Figure 9a & 9b).

Rhodobacterales strains and an *Oceanospirillales* strain also exhibited enhanced growth rate and increased maximum cell yield when amended with ProDOM_P. Both of these strains also displayed suppressed growth relative to control medium due to Pro99DOM_P

amendments, indicating that material produced and released by MIT9313 not only enhanced the growth of these strains, but was also able to overcome toxicity effects from material present in the background Pro99 medium (Figures 10a & 10b). Isolates from both of these groups have also been shown to enhance the growth of MI9313 in co-culture, suggesting that many of our findings may be indicative of mutualistic relationships (Sher et al. 2011). Genomic DNA extraction and subsequent sequencing of the SSU 16S rRNA gene confirmed both the purity and strain identity of all isolates that exhibited a positive growth response due to the addition of ProDOM_p.

Our findings are not in agreement with recent work that showed *Prochlorococcus*-derived DOM amendments stimulated the growth of members of the Bacteroidetes and Alteromonas groups, but not member of the SAR11 or Roseobacter clades in field populations (Sarmiento & Gasol 2012). There are many experimental differences that could account for these discrepancies. For instance, DOM in that study was obtained from a different strain of *Prochlorococcus* (high-light adapted MED4 from the Mediterranean Sea prepared in f/2 medium) and was not isolated by SPE. Also, Sarmiento and Gasol used microautoradiography to track the uptake of radiolabeled DOM in a mixed coastal community in which grazing processes and a myriad of complex microbial interactions existed. The SAR11 population in that community was likely quite distinct from the SAR11 strains employed in this study and were certainly under different stressors than the pure cultures studied here. Finally, no control amendments were made with DOM derived from the f/2 medium used to grow MED4, so it is impossible to distinguish community responses due to *Prochlorococcus*-derived DOM

from responses induced by nutrients in the f/2 medium or background seawater-base in their study. The experimental differences and contrasting conclusions of these two recent studies exemplifies the need for additional research on bacterioplankton utilization of phytoplankton-derived DOM and suggests that specific heterotrophic taxa may vary their response to a given DOM source under different conditions. Finally, these disparate results caution against the use of findings from a single DOM-microbe study to make generalizations about relationships between individual phytoplankton and bacterioplankton groups.

Thalassiosira pseudonana str. CCMP1335 DOM_P

A large volume (20 L) axenic culture of *Thalassiosira pseudonana* str. CCMP1335 was prepared in f/2 medium with a natural seawater base to obtain *T. pseudonana*-derived DOM for use in growth screens. CCMP1335 reached a maximum cell density of 3.44×10^6 cells/ml over a growth period of 11 days and purity was confirmed via direct microscopic observations. A maximum quantity of 0.6 mM POC was produced (174.4 fmol/cell) and a maximum quantity of 48.4 μ M DOC was produced and released after 11 days of growth (14.1 fmol/cell). 7% of the total organic carbon produced by CCMP1335 was released as DOC into the surrounding medium and ca. 4% of this DOC was obtained via SPE after biomass removal (Table 6).

CCMP1335 spent medium (i.e. after biomass removal) was processed by SPE to obtain hydrophobic *T. pseudonana*-derived DOM for preparation of PseDOM_P medium and for chemical interrogation by mass spectrometry. HPLC-MS analysis of DOM_P from

CCMP1335 revealed the presence of 52 features not found in the background f/2 medium control sample (Table 2). The 52 features detected ranged from 243 – 399 m/z with retention times corresponding to 82 – 100% methanol, indicating predominantly nonpolar material. SPE material from both the CCMP1335 culture and f/2 medium control were used to prepare treatment and control media at two concentration levels to examine dose effect responses of various bacterioplankton to additions of *T. pseudonana*-derived hydrophobic DOM. PseDOM_p-S elevated control medium DOC by 0.9 μM, corresponding to roughly 1.01X control medium DOC levels. f/2DOM_p-S elevated control medium DOC by 0.4 μM. DOC data for PseDOM_p-L and f/2DOM_p-L were not acquired due to an instrument error. High-resolution cell counts of triplicate cultures for each media type taken over a 10 – 17 day time period via flow cytometry revealed either suppressed growth or no effect for eight of the ten heterotrophic bacterioplankton strains tested. PseDOM_p and f/2DOM_p amendments were responsible for more suppressed growth than any other amendment tested, indicating the presence of toxic material in either the background f/2 phytoplankton medium or produced by *T. pseudonana* or both. Enhanced growth rate and yield was found in a strain from the OM60 clade when amended with PseDOM_p, however these positive effects were also found when amended with f/2DOM_p, indicating a response to material in the background f/2 phytoplankton medium rather than material produced and released by *T. pseudonana* (Figure 11). A unique diauxic growth pattern with significant secondary growth enhancement similar to that found with SynDOM_p amendments was detected for a *Flavobacteriales* amended with PseDOM_p, however subsequent genomic DNA extraction and sequencing of the

SSU 16S rRNA gene revealed the presence of an Alphaproteobacteria contaminant in these cultures.

Phaeodactylum tricornutum str. CCMP632 DOM_P

A large volume (20 L) axenic culture of *Phaeodactylum tricornutum* str. CCMP632 was prepared in f/2 medium with a natural seawater base to obtain *P. tricornutum*-derived DOM for use in growth screens. CCMP632 reached a maximum cell density of 5.13×10^6 cells/ml over a growth period of 7 days and purity was confirmed via direct microscopic observations. A maximum quantity of 1.4 mM POC was produced (272.9 fmol/cell) and a maximum quantity of 92.9 μ M DOC was produced and released after 7 days of growth (18.1 fmol/cell). 6% of the total organic carbon produced by CCMP632 was released as DOC into the surrounding medium and ca. 4% of this DOC was obtained via SPE after biomass removal (Table 6).

CCMP632 spent medium (i.e. after biomass removal) was processed by SPE to obtain hydrophobic *P. tricornutum*-derived DOM for preparation of TriDOM_P medium and for chemical interrogation by mass spectrometry. HPLC-MS analysis of DOM_P from CCMP632 revealed the presence of 273 features not found in the background f/2 medium control sample (Table 2), indicating that CCMP632 released a significant amount of DOC capable of acquisition by SPE. The 273 features detected ranged from 146 – 1,220 m/z over a broad polarity range (retention times corresponding to 12 – 100% methanol). SPE material from both the CCMP632 culture and f/2 medium control were used to prepare treatment and control media at two concentration levels to examine dose effect

responses of various bacterioplankton to additions of *P. tricornutum*-derived hydrophobic DOM. TriDOM_P-S and TriDOM_P-L elevated control medium DOC by 2.0 μM and 6.4 μM respectively, corresponding to roughly 1.02X and 1.05X control medium DOC levels. f/2DOM_P-S and f/2DOM_P-L elevated control medium DOC by 2.0 μM and 5.6 μM, respectively. High-resolution cell counts of triplicate cultures for each media type taken over a 10 – 17 day time period via flow cytometry revealed either suppressed growth or no effect for seven of the ten heterotrophic bacterioplankton strains tested. Enhanced growth rate and yield was found for one *Rhodobacteriales* strain when amended with TriDOM_P, however these positive effects were also found when amended with f/2DOM_P, indicating a response to material in the background f/2 phytoplankton medium rather than material produced and released by *P. tricornutum* (Figure 12). A unique diauxic growth pattern with extreme secondary growth enhancement similar to that found with SynDOM_P and PseDOM_P amendments was detected for a *Flavobacteriales* amended with TriDOM_P, however subsequent genomic DNA extraction and sequencing of the SSU 16S rRNA gene revealed the presence of an Alphaproteobacteria contaminant in these cultures that at least partially contributed to the detected increase in cell numbers. Previous connections between diatoms and members of the *Flavobacteriaceae* family have been identified (Pinhassi et al. 2004, Tada et al. 2012), suggesting that growth enhancement of this *Flavobacterium* due to DOM derived from both *T. pseudonana* and *P. tricornutum* may be worthy of additional investigation despite complications due to contamination. Neither of the diatom-derived DOM amendments induced enhanced growth for any of the strains tested from the SAR11

clade, which is consistent with recent observations of a mesocosm diatom bloom (Tada et al. 2012), but not with recent experiments using microautoradiography to track DOM uptake from various phytoplankton sources (Sarmiento & Gasol 2012). Discrepancies in these studies highlight the need for additional research involving source-specific DOM uptake by different bacterioplankton taxa under a variety of conditions.

Prochlorococcus str. MIT9313 DOM_P fractions

Three additional large volume (20 L) axenic cultures of *Prochlorococcus str. MIT9313* (-A, -B, and -C, respectively) were prepared in Pro99 medium with a natural seawater base to obtain *Prochlorococcus*-derived DOM for fractionation by HPLC and use in growth screens. Fractionation was performed to test subsets of ProDOM_P in screens with various bacterioplankton strains that previously displayed enhanced growth due to ProDOM_P additions in an effort to refine the list of target features responsible for the observed growth effects. ProDOM_P from strain MIT9313 was chosen for fractionation testing due to its consumption by more bacterioplankton strains than any other DOM source tested. Co-culture experiments with MIT9313 and a multitude of bacterioplankton strains provide additional evidence for abundant interactions between this strain and heterotrophic bacteria, especially when compared to *Prochlorococcus str. MED4* (Sher et al. 2011). Strain MIT9313 was also recently shown to possess genetic machinery for the production of a variety of secondary metabolites (Li et al. 2010). MIT9313A, B, and C reached a maximum cell density of 1.94×10^8 cells/ml, 2.11×10^8 cells/ml, and 1.95×10^8 cells/ml respectively over a growth period of 13.5 days, and a

maximum quantity of 2.3 – 2.5 mM POC was produced (11.4 – 12.8 fmol/cell) (Table 6). The purity of all three cultures was confirmed via purity broth tests for the presence of other bacteria. DOC quantification of these cultures was not possible due to the addition of HEPES buffer to mitigate changes in pH level during growth.

MIT9313 spent medium (i.e. after biomass removal) from all 3 cultures was processed by SPE to obtain hydrophobic *Prochlorococcus*-derived DOM for HPLC fractionation and chemical interrogation by mass spectrometry. HPLC-MS analysis of DOM_p from MIT9313A, B, and C revealed the presence of 767, 813, and 750 features not found in the background Pro99 medium control sample (Table 3). Features ranged from 136 – 1423 m/z over a broad polarity spectrum (retention times corresponding to 5 – 100% methanol). The fact that an order of magnitude greater number of features was detected in these MIT9313 cultures compared to the previous MIT9313 culture is likely due to the use of commercially available columns for SPE with tighter resin packing than the custom packed columns used previously. These commercial tubes appear to extract phytoplankton-derived DOC, particularly larger (> 1,000 m/z) material, with greater efficiency than hand packed columns (Tables 2 and 3; also see chapter 2). SPE material from all three replicate MIT9313 cultures was pooled into 1 sample prior to fractionation by HPLC. Five percent of the pooled sample was divided into five distinct fractions based on polarity, with fraction 1 being the most polar and fraction 5 the most nonpolar. HPLC-MS analysis of each fraction revealed a variable number of features found in each fraction that were not present in the background Pro99 medium control sample. Fraction 4 contained the most features (224), while fraction 5 contained the least (4). Fraction 3

contained the largest features (182 – 1150 m/z) and fraction 1 contained the smallest (136 – 431 m/z). The retention time range for features found in each fraction corresponded closely to the retention time cutoffs used to create these fractions (Table 3).

SPE material from the pooled MIT9313 replicate cultures, the 5 fractions, and the Pro99 medium control were used to prepare treatment and control media to examine growth responses of various bacterioplankton that exhibited enhanced growth due to previous additions of *Prochlorococcus*-derived hydrophobic DOM. ProDOM_p-pooled and Pro99DOM_p-2 media elevated control medium DOC by 37.4 μM and 21.3 μM, respectively, corresponding to roughly 1.5X and 1.3X control medium DOC levels. Although growth media for all 5 ProDOM_p fractions (ProDOM_p-F1 through F5) were prepared in equal proportions, the fractions elevated control medium DOC by widely varying amounts (3 – 171.9 μM, corresponding to 1.04X – 3.1X control medium DOC levels) (Figure 13). ProDOM_p-F3 contained 31% of the total features found in all fractions and contained the most DOC (171.9 μM above control medium). ProDOM_p-F5 contained only 1% of the total features found in all fractions and contained the least amount of DOC (3 μM above control medium). Surprisingly, ProDOM_p-F4 contained 48% of the total features found in all fractions, yet only increased DOC by 17.4 μM above control medium, while ProDOM_p-F1 contained only 6% of the total features found in all fractions and increased DOC by 60.3 μM above control medium. Therefore, the number of features detected by HPLC-MS did not correlate with DOC quantity.

High-resolution cell counts of triplicate cultures for each media type taken over an 8 – 17 day time period via flow cytometry revealed general trends regarding the lability

of ProDOM_p and the ProDOM_p fractions. The ProDOM_p-pooled treatment enhanced the growth of all heterotrophic bacterioplankton strains tested that had previously displayed enhanced growth when supplemented with ProDOM_p, indicating robust responses due to excellent agreement between replicate bacteria-DOM combinations from multiple rounds of screening. These strains included members of the SAR11 clade (subgroup IIIA), *Rhodobacterales*, and *Oceanospirillales* order. Strain HIMB1291 from the 1A subgroup of the SAR11 clade displayed inconclusive results during previous ProDOM_p additions and was therefore tested again with the ProDOM_p-pooled amendment, but displayed no change in growth relative to control medium. HIMB95 from the SAR11 clade (subgroup IIIA) was tested with the ProDOM_p-pooled amendment due to enhanced growth with previous ProDOM_p additions to HIMB114, a closely related strain. The addition of ProDOM_p-pooled material enhanced the growth of HIMB95 as well, indicating a potential connection between this particular subgroup of the SAR11 clade (IIIA) and the consumption of DOM derived from *Prochlorococcus* str. MIT9313.

While no correlation was found between ProDOM_p fraction growth enhancement and the number of features or amount of DOC in each fraction, a connection between fraction polarity and growth enhancement was apparent. The two least polar fractions (ProDOM_p-F4 and ProDOM_p-F4) suppressed growth in four of the five strains amended with these fractions, indicating the presence of toxic material in these nonpolar fractions (Table 8). Three of the strains displaying suppressed growth due to these fractions exhibited enhanced growth when amended with the ProDOM_p-pooled material and other fractions (Figure 14). This finding reveals the presence of material

that is both labile and toxic to a particular bacterioplankton strain within ProDOM_P and highlights the complexity of these bacteria-DOM interactions. The most striking positive responses to any amendment were found with additions of the most polar fraction (ProDOM_P-F1), which enhanced growth rate and maximum cell yield in four of the five strains tested and never suppressed the growth of any strain (Table 8). ProDOM_P-F1 amendments resulted in extreme growth enhancement for strains from the SAR116 and Roseobacter clades (Figure 15a & 15b). Fractions of intermediate polarity (ProDOM_P-F2 and ProDOM_P-F3) also enhanced the growth of the gammaproteobacterium HIMB30.

Screens with individual sub-fractions of DOM_P demonstrated the ability of this screening technique to more precisely identify the compounds utilized by a particular bacterial strain, and reveal the presence of both labile and toxic material in the same DOM_P sample. While the overall effect of a given DOM_P amendment on a particular bacterioplankton strain may be a suppression of growth, there is still the potential for the presence of labile compounds within that amendment. The finding that the less polar ProDOM_P fractions frequently suppressed the growth of an isolate, while the more polar ProDOM_P fractions frequently enhanced the growth of that same isolate highlights a previously undocumented connection between the polarity and reactivity of phytoplankton-derived DOM and is worthy of additional investigation with DOM derived from phytoplankton other than *Prochlorococcus* str. MIT9313. HPLC fractionation and subsequent screening of phytoplankton media controls (i.e. Pro99DOM_P) would also be a useful tool to distinguish between growth suppression due to less polar phytoplankton-derived material and toxic material that may arise from the leaching of octadecyl carbon

resins during SPE of DOM_P. Overall, our findings reveal polar DOM_P to be an ideal candidate for future studies investigating DOM consumption and uncover several novel connections between DOM from a known biological source and its consumption by a particular strain of heterotrophic (or mixotrophic) bacterioplankton.

Conclusions

The screening assay employed here was able to identify specific bacterioplankton types responsible for the consumption of DOM from a multitude of sources, revealing novel connections between the origin and chemical composition of marine DOM and its consumption. Our results provide strong evidence for strain-specific uptake of DOM compound suites derived from both natural seawater and a diverse suite of axenic phytoplankton cultures. Variable growth responses by individual bacterioplankton strains to different DOM sources indicate that compositional differences in DOM derived from different sources (such as those identified in chapter 2) affect the uptake and utilization of these substrates. Variable growth responses to the same DOM source by different bacterioplankton strains demonstrate that DOM lability also varies according to the consumer. Collectively, these observations highlight the importance of identifying robust connections between DOM from a known source and particular bacterioplankton groups in order to elucidate why DOM lability varies due to biological source and sink. Metaproteomic and metatranscriptomic analysis of oligotrophic surface waters has shown that variability in transporters and periplasmic substrate binding proteins among different bacterioplankton taxa may at least partially explain differential DOM uptake, particularly

in oligotrophic environments (Sowell et al. 2009, Ottesen et al. 2013). Model systems for the degradation of marine DOM would be an invaluable resource toward a greater understanding of secondary production and carbon cycling in the sea. The large matrix of replicate substrates and isolates presented here allowed for the identification of several potential model systems worthy of further investigation, such as the utilization of *Prochlorococcus*-derived DOM by organisms from the SAR11 clade (subgroup IIIA) and *Oceanospirillales* order.

A longstanding challenge in marine DOM research has been the chemical identification of reactive DOM components. Despite advances in instrumentation, the complexity of marine DOM makes a systematic compositional assessment of a particular DOM sample both impractical and arduous. Recently, it has been suggested that the marine DOM pool is too complex to accurately test the hypothesis that DOM compositional diversity can select for and maintain bacterioplankton diversity (Sarmiento & Gasol 2012). One way around this complexity is to develop means of identifying reactive components within a given DOM pool, reducing the task of characterizing an immensely complex mixture down to a more manageable and biologically relevant exercise. HPLC fractionation and subsequent screening of DOM_P for heterotrophic utilization provides a unique method of identifying reactive DOM components within a complex chemical mixture. By first distinguishing DOM features produced and released by a specific phytoplankton strain from those present in the background medium and then identifying subsets of these features that enhance bacterioplankton growth, one can reduce thousands of compounds down to < 100 of biological interest. These manageable

feature lists can then be further reduced through additional fractionation and screening or directly targeted for chemical characterization. Positive growth responses induced by a specific sub-fraction of phytoplankton-derived DOM, as opposed to the entire suite, greatly reduces the list of potential compounds causing the response and allows for a more informed investigation of the most reactive compounds in any particular DOM-microbe combination. The screening assay we described provided a rapid means of testing a large number of bacteria-substrate combinations in order to identify combinations of physiological interest for further studies regarding the chemical nature of this material and the metabolic pathways involved in its consumption.

In chapter 2, a direct connection between DOM composition and its biological source was established, suggesting that marine DOM could act as a link between autotrophic and heterotrophic microbial community structure. This link is only plausible if the speciation of DOM supplied by various marine phytoplankton translate into variations in lability. Here we present findings that demonstrate marine DOM lability is a factor of both biological source and sink, further implicating marine DOM as a driver of marine microbial diversity. Implementation of a large biological screening assay uncovered novel relationships between autotrophic and heterotrophic bacteria and identified potential model systems for future studies on DOM cycling. Cultivation-based approaches are not without caveats, and genomic variation between cultured isolates and their relatives in the wild could translate into differences in the production and consumption of DOM in the laboratory and how these processes occur in nature (Luo et al. 2011). Further efforts are now needed to determine how and why DOM lability varies

among different sources and sinks in the laboratory and link these findings to in-situ observations in order to gain a more complete understanding of microbial physiology and the passage of fixed carbon from autotrophic microbes to heterotrophic microbes in the sea. Metatranscriptomic analyses of DOM utilization by complex bacterioplankton communities in microcosm experiments are beginning to reveal taxa-specific patterns in metabolic processes used by various heterotrophic taxa to consume marine DOM (McCarren et al. 2010, Rinta-Kanto et al. 2011, Teeling et al. 2012) and similar studies are needed for individual bacterioplankton strains. This information will be necessary to predict how microbes will continue to partition and cycle carbon and other micro- and macronutrients under changing ocean conditions.

Acknowledgments

I would like to thank M. Rappé, J. Grote, O. Sosa, E. Omori, and K. Manoi for help with bacterioplankton culturing and media preparation. Thanks to D. Moran, K. Munson, and M. Miller for assistance with cell counts. I am also grateful to P. Berube and S. Chisholm for providing the *Prochlorococcus* culture, and to J. Waterbury and S. Dyhrman for providing the *Crocospaera* and *T. Pseudonana* inoculates. Finally, I would like to acknowledge M. Nieto-Cid for assistance with DOC analysis and C. Johnson for help with early mass spectrometry analysis development. Funding for this work was provided by a grant to D.J.R from the Gordon and Betty Moore Foundation (#1711), National Science Foundation Science and Technology Center Award

#EF0424599 and #DBI0424599 to E.F.D and D.J.R., and a grant from the Coastal Ocean Institute at the Woods Hole Oceanographic Institution awarded to J.W.B.

Table 1. Bacterioplankton strains screened in this study and their corresponding strain number from the Hawaii Institute of Marine Biology (HIMB) culture collection. All strains were previously isolated from coastal seawater taken from Kaneohe Bay, Oahu. The dominant phytoplankton type in Kaneohe Bay (Yeo et al. 2013) was also chosen for screening due to its mixotrophic tendencies (Baran et al. 2011).

<i>HIMB Strain #</i>	<i>Phylogenetic Affiliation</i>
12	Cyanobacteria, <i>Synechococcus spp.</i>
1363	Bacteroidetes, unclassified <i>Flavobacteriales</i>
4	<i>Alphaproteobacteria</i> , SAR11 subgroup IA
83	<i>Alphaproteobacteria</i> , SAR11 subgroup IA
1291	<i>Alphaproteobacteria</i> , SAR11 subgroup IA
114	<i>Alphaproteobacteria</i> , SAR11 subgroup IIIA
95	<i>Alphaproteobacteria</i> , SAR11 subgroup IIIA
136	<i>Alphaproteobacteria</i> , SAR11 subgroup IIIA
59	<i>Alphaproteobacteria</i> , SAR11 subgroup V
100	<i>Alphaproteobacteria</i> , SAR116
1	<i>Alphaproteobacteria</i> , <i>Rhodobacteriales</i>
11	<i>Alphaproteobacteria</i> , <i>Rhodobacteriales</i>
137	<i>Betaproteobacteria</i> , OM43
1416	<i>Gammaproteobacteria</i> , <i>Pseudoalteromonas</i>
30	<i>Gammaproteobacteria</i> , OM252
55	<i>Gammaproteobacteria</i> , OM60

Table 2. Chemical composition of DOM_P as determined by high performance liquid chromatography couple to a single quadrupole mass spectrometry with an electrospray ionization source (HPLC-MS). DOM_P features were defined as a unique *m/z* and retention time pair that was not present in any control sample. Ranges for size and retention time of features detected is shown for each phytoplankton culture examined in this study. Features were all detected in positive ion mode.

<i>Phytoplankton strain</i>	<i>Total features</i>	<i>m/z range</i>	<i>Retention time range (min)</i>
<i>Crocospaera watsonii</i> (WH8501)	13	308-363	30.8-36.5
<i>Synechococcus sp.</i> (WH8102)	302	114-854	6.4-40.3
<i>Prochlorococcus marinus</i> (MIT9313)	82	139-536	1.3-36.5
<i>Thalassiosira pseudonana</i> (CCMP1335)	52	243-399	26.2-41.7
<i>Phaeodactylum tricornutum</i> (CCMP632)	273	146-1220	3.6-38.4
<i>Prochlorococcus marinus</i> (MIT9313-A)	777	136-1423	1.6-59.4
<i>Prochlorococcus marinus</i> (MIT9313-B)	813	136-1423	1.6-59.6
<i>Prochlorococcus marinus</i> (MIT9313-C)	750	136-1423	1.6-59.4

Table 3. Chemical composition for five distinct fractions of DOM_p derived from *Prochlorococcus* str. MIT9313 that were divided based on polarity using high performance liquid chromatography (HPLC). Features were identified by HPLC couple to a single quadrupole mass spectrometry with an electrospray ionization source (HPLC-MS). Features were identified in an identical fashion to unfractionated DOM_p samples. The retention time range corresponds to the actual time ranges detected during HPLC-MS analysis, while the fraction range corresponds to the time span of collection during the separation of each fraction prior to HPLC-MS analysis.

<i>Sample</i>	<i>Total features</i>	<i>m/z range</i>	<i>Retention time range (min)</i>	<i>Fraction range (min)</i>
<i>Prochlorococcus marinus</i> (MIT9313-Fraction1)	28	136-431	1.6-15	0-15
<i>Prochlorococcus marinus</i> (MIT9313-Fraction2)	66	166-922	15.1-25.1	15-25
<i>Prochlorococcus marinus</i> (MIT9313-Fraction3)	142	182-1150	25.3-35.9	25-36
<i>Prochlorococcus marinus</i> (MIT9313-Fraction4)	224	191-944	36.0-48.6	36-50
<i>Prochlorococcus marinus</i> (MIT9313-Fraction5)	4	280-565	50.2-55	50-84

Table 4. Growth responses to additions of HMW-DOM from the 15m intake pipe at the Natural Energy Laboratory of Hawaii Authority (NELHA). Three concentrations were tested to examine dose-specific responses, corresponding to 247 μM (S), 929 μM (M), and 2 mM (L) elevations in DOC above control levels. Effects on the growth rate (rate) and maximum cell density reached (yield) were evaluated for each isolate. Arrows indicate increases and decreases, while “=” indicates no change was observed relative to controls.

	SAR11-IA (HIMB83)		SAR11-IIIA (HIMB114)		SAR11-IIIA (HIMB95)		SAR11-IIIA (HIMB136)		<i>Rhodobacteriales</i> (HIMB1)		OM252 (HIMB30)		OM60 (HIMB55)		OM43 (HIMB137)		<i>Synechococcus</i> (HIMB12)	
	rate	yield	rate	yield	rate	yield	rate	yield	rate	yield	rate	yield	rate	yield	rate	yield	rate	yield
NELHA-S	=	=	=	=	=	=	=	↓	↑	↑	=	=	=	=	↑	↑	↑	↑
NELHA-M	=	=	=	↓	=	↓	=	↓	↑	↑	=	=	=	↓	↑	↑	↑	↑
NELHA-L	=	=	=	↓	=	↓	=	↓	↑	↑	↑	↑	=	↓	↑	↑	↑	↑

Table 5. Growth responses to additions of HMW-DOM from several locations in the South Pacific (RAPA-01, -04, -07) and DOM_P obtained from five different strains of marine phytoplankton along with controls. Small (S) and large (L) concentrations of DOM_P from phytoplankton cultures and media controls were added to examine dose-specific responses (see text for DOC concentrations). Effects on the growth rate (rate) and maximum cell density reached (yield) were evaluated for each isolate. Arrows indicate increases and decreases, while “=” indicates no change was observed relative to controls.

	SAR11-IA (HIMB4)		SAR11-IA (HIMB83)		SAR11-IA (HIMB1291)		SAR11-IIIA (HIMB114)		SAR11-V (HIMB59)		<i>Rhodobacteriales</i> (HIMB11)		OM252 (HIMB30)		OM60 (HIMB55)		<i>Pseudoalteromonas</i> (HIMB1416)		<i>Bacteroidetes</i> (HIMB1363)	
	rate	yield	rate	yield	rate	yield	rate	yield	rate	yield	rate	yield	rate	yield	rate	yield	rate	yield	rate	yield
RAPA01	=	↓	=	↓	=	=	=	=	↓	↓	↓	↓	=	=	↑	↑	=	=	=	=
RAPA04	=	↓	=	=	=	=	=	=	=	=	↓	↓	↑	↑	↑	↑	=	=	=	=
RAPA07	=	↓	=	=	=	↓	↓	↓	↓	↓	↓	↓	=	=	↑	=	=	=	=	=
ProDOM _P -S	=	=	=	=	=	=	↑	↑	=	=	↑	↑	=	=	=	=	=	=	=	=
ProDOM _P -L	=	↓	=	↓	=	=	↑	↑	↑	↑	↑	↑	↑	↑	=	=	=	=	=	=
Pro99DOM _P -S	=	↓	=	↓	=	=	=	=	=	↓	↓	↓	↓	↓	=	=	=	=	=	=
Pro99DOM _P -L	=	↓	=	↓	=	=	=	↑	=	=	↓	↓	↓	↓	=	=	=	=	=	=
SynDOM _P -S	=	=	=	=	=	=	=	=	=	=	↓	↓	↓	↓	=	=	=	=	=	↑
SynDOM _P -L	=	↓	=	↓	=	↑	=	=	=	=	=	↓	=	=	↑	=	=	=	=	↑
SNDOM _P -S	↓	↓	=	↓	=	=	↑	↑	=	↓	↓	↓	↓	↓	=	=	=	=	=	=
SNDOM _P -L	=	↓	=	↓	=	↑	↑	↑	=	↓	↓	↓	=	↓	↑	↑	=	=	=	↑
CroDOM _P -S	=	↓	=	=	=	=	=	=	=	=	=	↓	↓	↓	=	=	=	=	=	=
CroDOM _P -L	=	↓	=	↓	=	↓	=	=	↓	↓	=	↓	↓	=	↓	↑	=	=	=	↑
PseDOM _P -S	=	↓	=	↓	↓	↓	↓	↓	=	=	↓	↓	↓	↓	↑	↑	=	=	↑	↑
PseDOM _P -L	=	↓	↓	↓	↓	↓	↓	↓	=	=	↓	↓	↓	↓	↑	↑	=	=	↑	↑
f/2DOM _P -S	=	↓	=	↓	↓	↓	↓	↓	=	=	↓	↓	=	↑	↓	↓	↑	↑	=	↑
f/2DOM _P -L	=	↓	↓	↓	↓	↓	↓	↓	=	=	↓	↓	↑	↑	↓	↓	↑	↑	=	↑
TriDOM _P -S	=	↓	=	=	=	=	=	=	↓	↓	=	↑	↓	↓	=	=	=	=	=	↑
TriDOM _P -L	=	↓	=	↓	=	=	=	=	=	↓	↑	↑	↑	↑	↓	↓	=	=	↑	↑

Table 6. Growth and organic carbon production for marine phytoplankton cultures from which DOM_p was isolated. Five different strains were chosen for DOM_p isolation and *Prochlorococcus* str. MIT9313 was re-grown multiple times (MIT9313-A, -B, -C) to obtain additional DOM_p for further screening. DOC data for these cultures was not available due to the additional of HEPES buffer during growth to mitigate pH changes. %NPP released as DOC = (DOC/(DOC+POC))*100. %DOC obtained by SPE = (DOC recovered/DOC produced)*100. NPP: net primary production. SPE: solid-phase extraction.

Phytoplankton strain	Growth period (days)	Maximum density (cells/ml)	DOC produced (μM)	POC produced (mM)	DOC produced per cell (fmol)	POC produced per cell (fmol)	% NPP released as DOC	% DOC obtained by SPE
<i>Crocospaera watsonii</i> (WH8501)	8	1.26E+06	83.8	0.5	66.5	388.9	15	8
<i>Synechococcus</i> sp. (WH8102)	13	3.49E+07	59.1	0.4	1.7	11.5	13	3
<i>Prochlorococcus marinus</i> (MIT9313)	7	6.29E+07	75.4	0.8	1.2	12.7	9	3
<i>Thalassiosira pseudonana</i> (CCMP1335)	11	3.44E+06	48.4	0.6	14.1	174.4	7	4
<i>Phaeodactylum tricornutum</i> (CCMP632)	7	5.13E+06	92.9	1.4	18.1	272.9	6	4
<i>Prochlorococcus marinus</i> (MIT9313-A)	13.5	1.94E+08	NA	2.3	NA	11.9	NA	NA
<i>Prochlorococcus marinus</i> (MIT9313-B)	13.5	2.11E+08	NA	2.4	NA	11.4	NA	NA
<i>Prochlorococcus marinus</i> (MIT9313-C)	13.5	1.95E+08	NA	2.5	NA	12.8	NA	NA

Table 7. Growth responses to additions of total DOC produced by *Crocospaera watsonii* (str. WH8501). Two concentrations were tested to examine dose-specific responses, corresponding to 7.9 μM (S) and 29.6 μM (L) elevations in DOC above control levels. Effects on the growth rate (rate) and maximum cell density reached (yield) were evaluated for each isolate. Arrows indicate increases and decreases, while “=” indicates no change was observed relative to controls.

	SAR11-IA (HIMB83)		<i>Rhodobacterales</i> (HIMB1)		OM43 (HIMB137)	
	rate	yield	rate	yield	rate	yield
CroDOM-S	↑	=	↑	↑	=	=
CroDOM-L	↑	=	↑	↑	↓	=

Table 8. Growth responses to additions of five DOM_p polarity fractions derived from *Prochlorococcus* str. MIT9313 alongside unfractionated DOM_p and a media control. Fractions are listed in order of decreasing polarity (i.e. fraction 1 (F1) is the most polar fraction, while fraction 5 (F5) is the most nonpolar fraction). Effects on the growth rate (rate) and maximum cell density reached (yield) were evaluated for each isolate. Arrows indicate increases and decreases, while “=” indicates no change was observed relative to controls, and N/A indicated data not available.

	SAR11-IA (HIMB1291)		SAR11-IIIA (HIMB114)		SAR11-IIIA (HIMB95)		SAR116 (HIMB100)		<i>Rhodobacteriales</i> (HIMB11)		OM252 (HIMB30)	
	rate	yield	rate	yield	rate	yield	rate	yield	rate	yield	rate	yield
ProDOM _p -pooled	=	=	↑	↑	↑	↑	↑	=	↑	=	↑	↑
Pro99DOM _p -2	=	=	↓	↓	=	=	=	=	=	=	=	=
ProDOM _p -F1	=	=	=	↑	N/A	N/A	↑	↑	↑	↑	=	↑
ProDOM _p -F2	=	=	↓	↓	N/A	N/A	↓	↓	=	=	↑	↑
ProDOM _p -F3	↓	↓	↓	↓	N/A	N/A	↓	↓	↓	↓	↓	↑
ProDOM _p -F4	↓	↓	↓	↓	N/A	N/A	↓	↓	↑	=	↓	↓
ProDOM _p -F5	↓	↓	↓	↓	N/A	N/A	↓	↓	↑	=	↓	↓

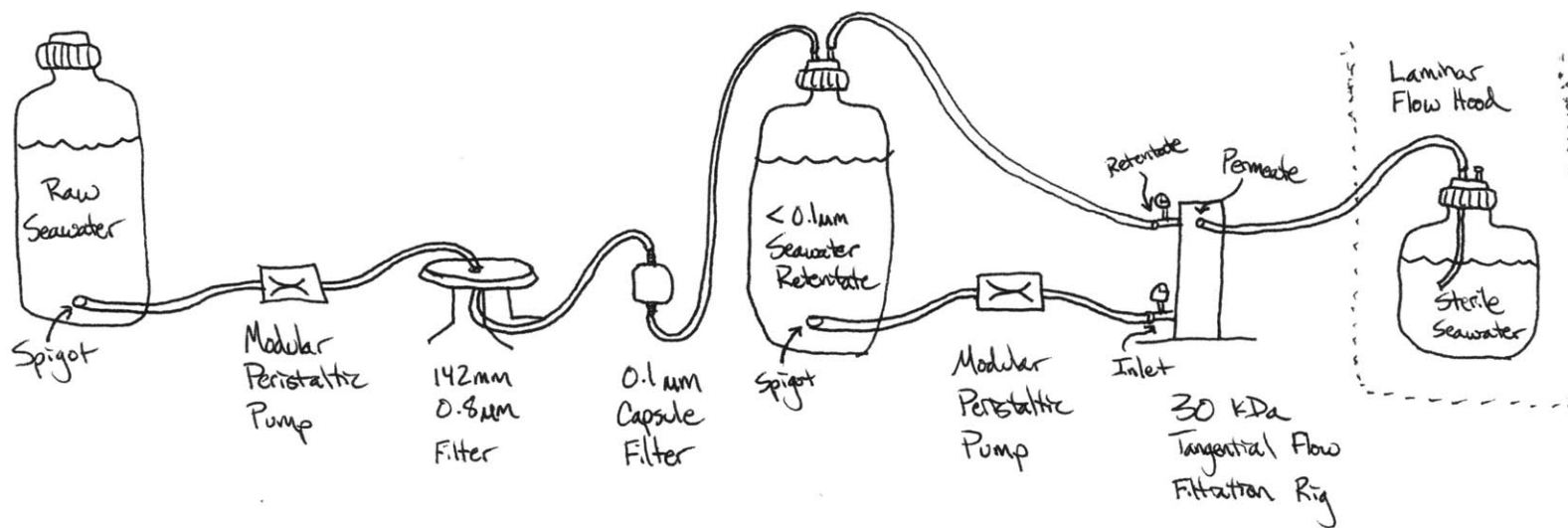


Figure 1. Schematic drawing depicting the sterilization of Kaneohe Bay seawater for use as base media for all cultivation experiments. Workflow moves from left to right and includes 0.8 μm , followed by 0.1 μm prefiltration of raw seawater prior to sterilization using a 30 kDa tangential flow ultrafiltration rig. Specifics regarding labware, including manufacturer information, are provided in the text.

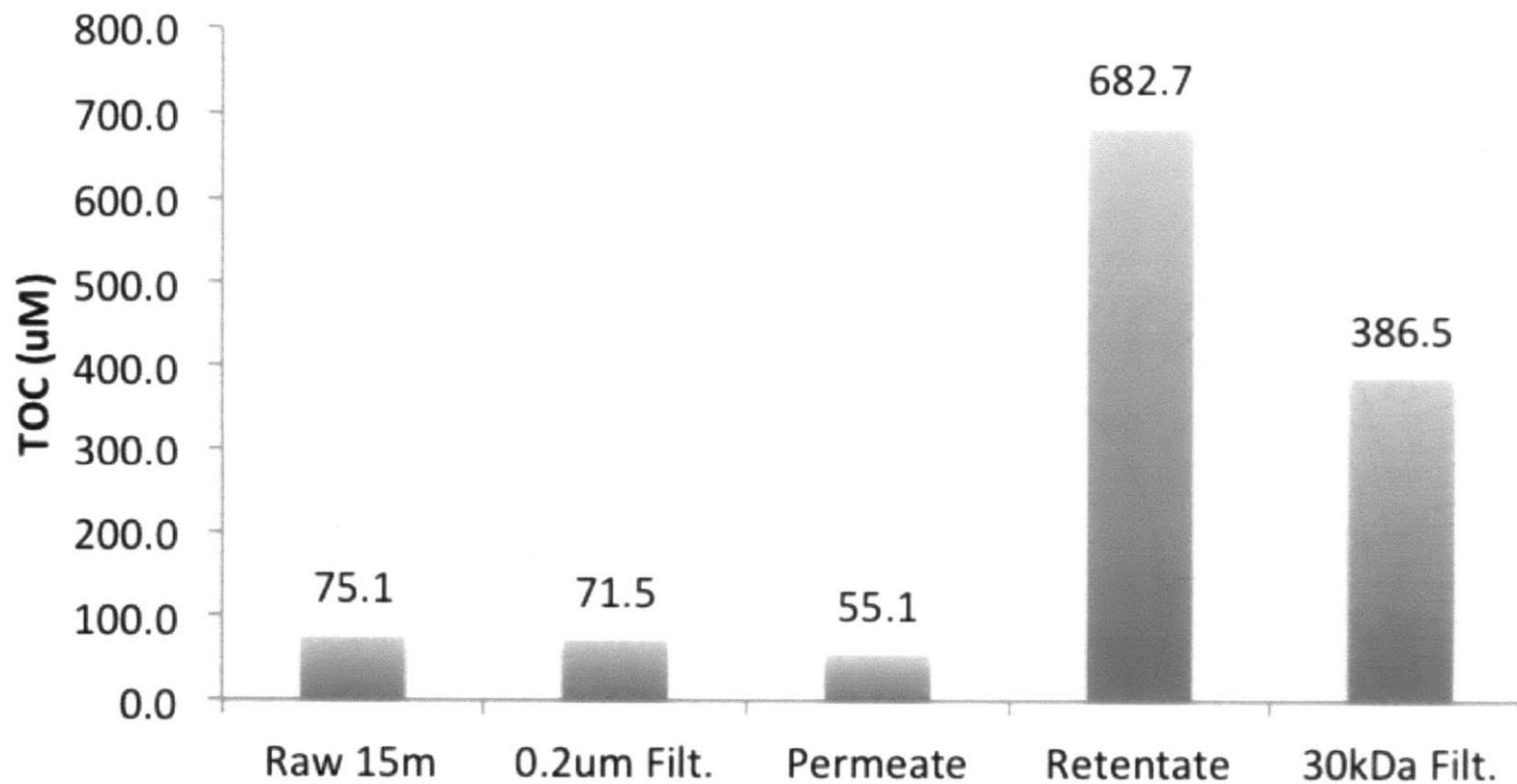


Figure 2. Quantification of dissolved organic carbon during preparation of high-molecular weight dissolved organic matter (HMW-DOM) from natural seawater. Surface seawater was collected in clean barrels (Raw 15m) and pre-filtered (0.2 µm), which reduced raw seawater DOC by ca. 3.6 µM. The filtrate was then processed using tangential flow filtration to concentrate DOC ≥ 1 kDa. On average, permeate DOC was 16.4 µM lower than the input and the DOC content of the final concentrate (Retentate) was increased almost 10-fold above input levels. A final filtration step (30 kDa) to reduce viral particles decreased the concentrate DOC by ca. 300 µM, resulting in a final concentration of 386.5 µM DOC (30kDa).

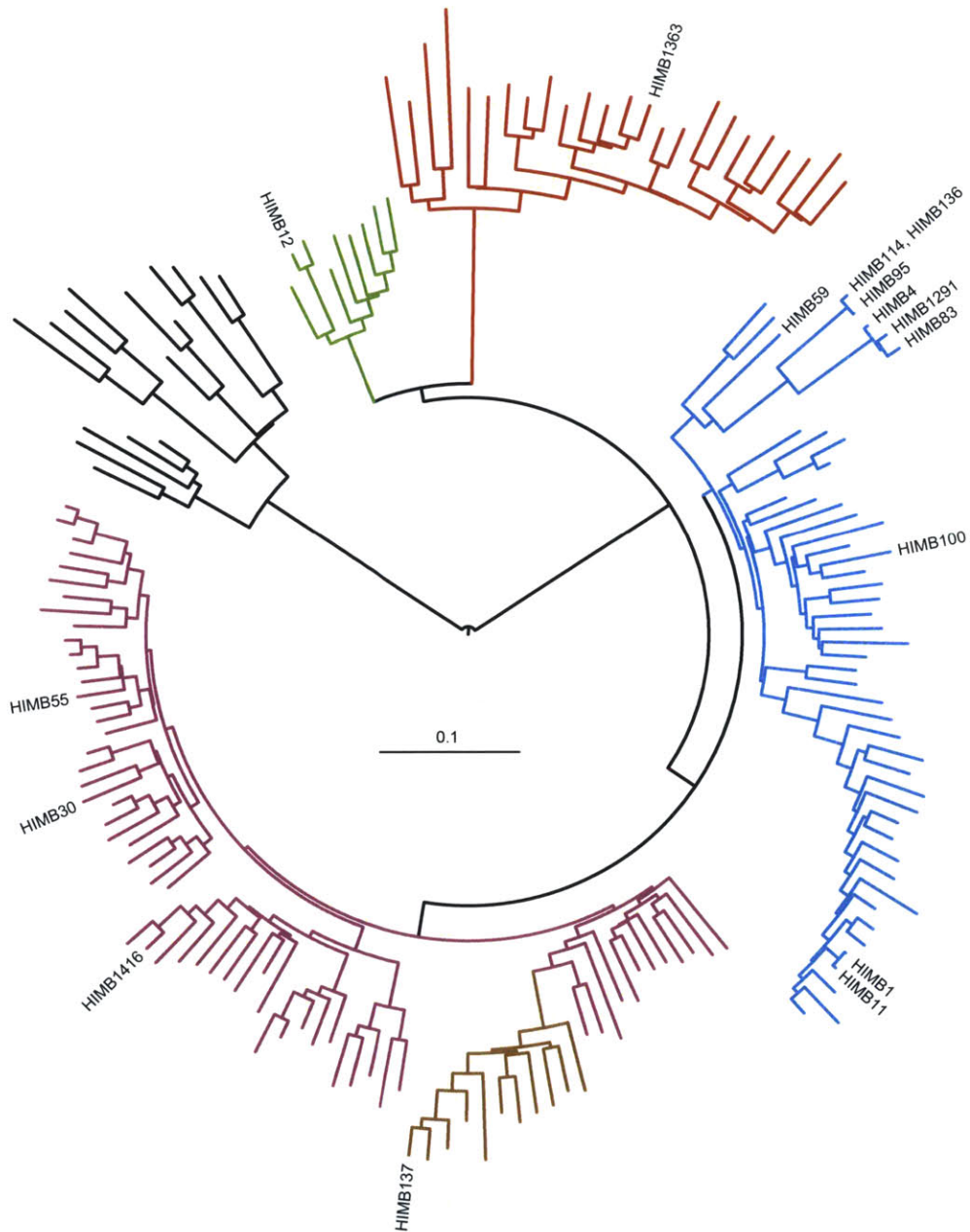


Figure 3. Phylogenetic breadth of the bacterial strains tested in this study. SSU rRNA gene sequences were aligned with version 111 of the ‘All-Species Living Tree’ project SSU rRNA gene database (Yarza et al. 2008) using the ARB software package (Ludwig 2004). The phylogeny was constructed from nearly full-length gene sequences using the RAxML maximum likelihood method (Stamatakis 2006) within ARB. The scale bar corresponds to 0.1 substitutions per nucleotide position. Green: *Cyanobacteria*; red: *Bacteroidetes*; blue: *Alphaproteobacteria*; violet: *Gammaproteobacteria*; brown: *Betaproteobacteria*; black: *Archaea* (outgroup).

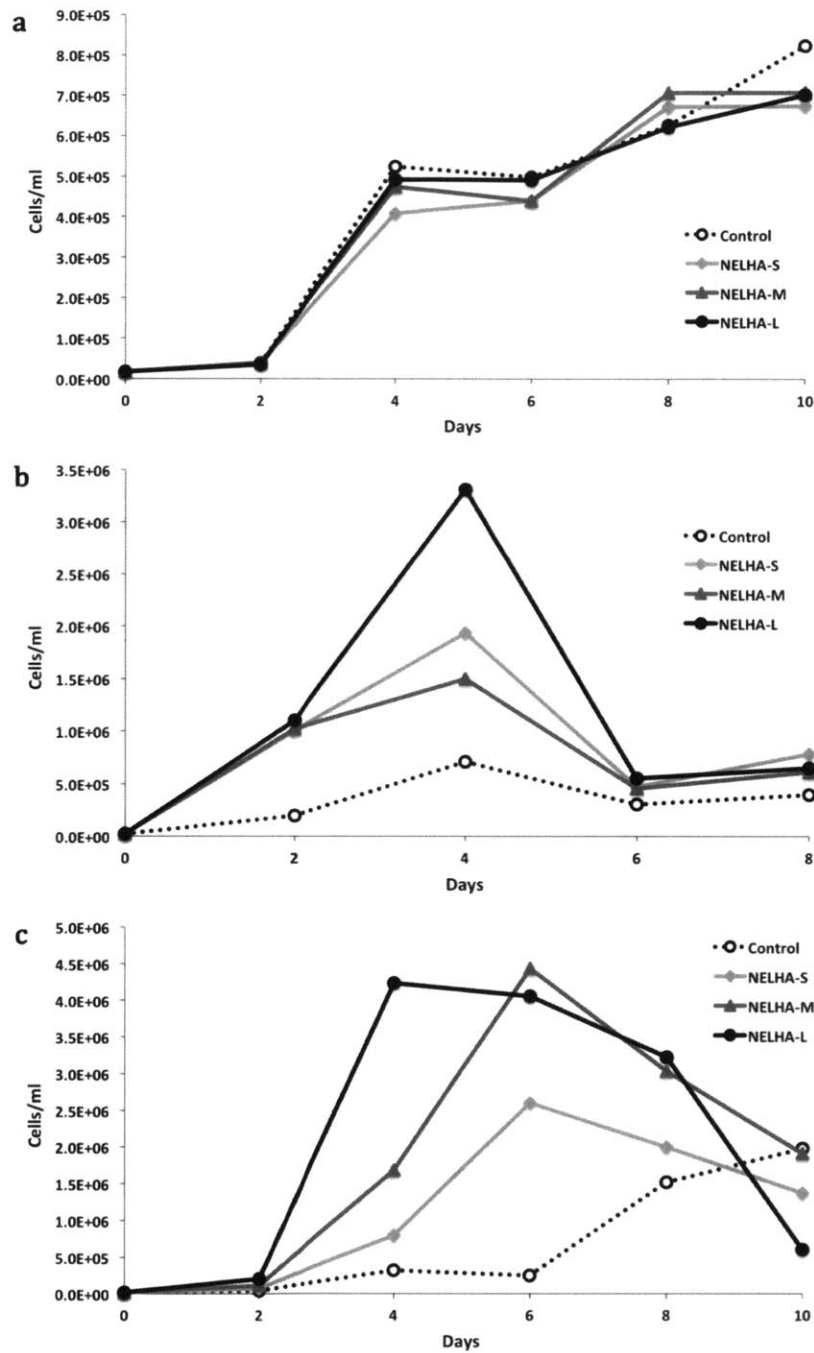


Figure 4. Growth response of HIMB136 (SAR11 subgroup IIIA; a), HIMB1 (*Rhodobacteriales*; b), and HIMB12 (*Synechococcus*; c) to amendments with high-molecular weight dissolved organic matter (HMW-DOM) from the 15m intake pipe at the Natural Energy Laboratory of Hawaii Authority (NELHA). DOM was added in small (S), medium (M), and large (L) quantities to examine dose specific effects on growth rate and maximum cell density. Certain strains tested exhibited either no response or even suppressed growth due to DOM additions (a), while others exhibited increases in both growth rate and maximum cell density (b,c).

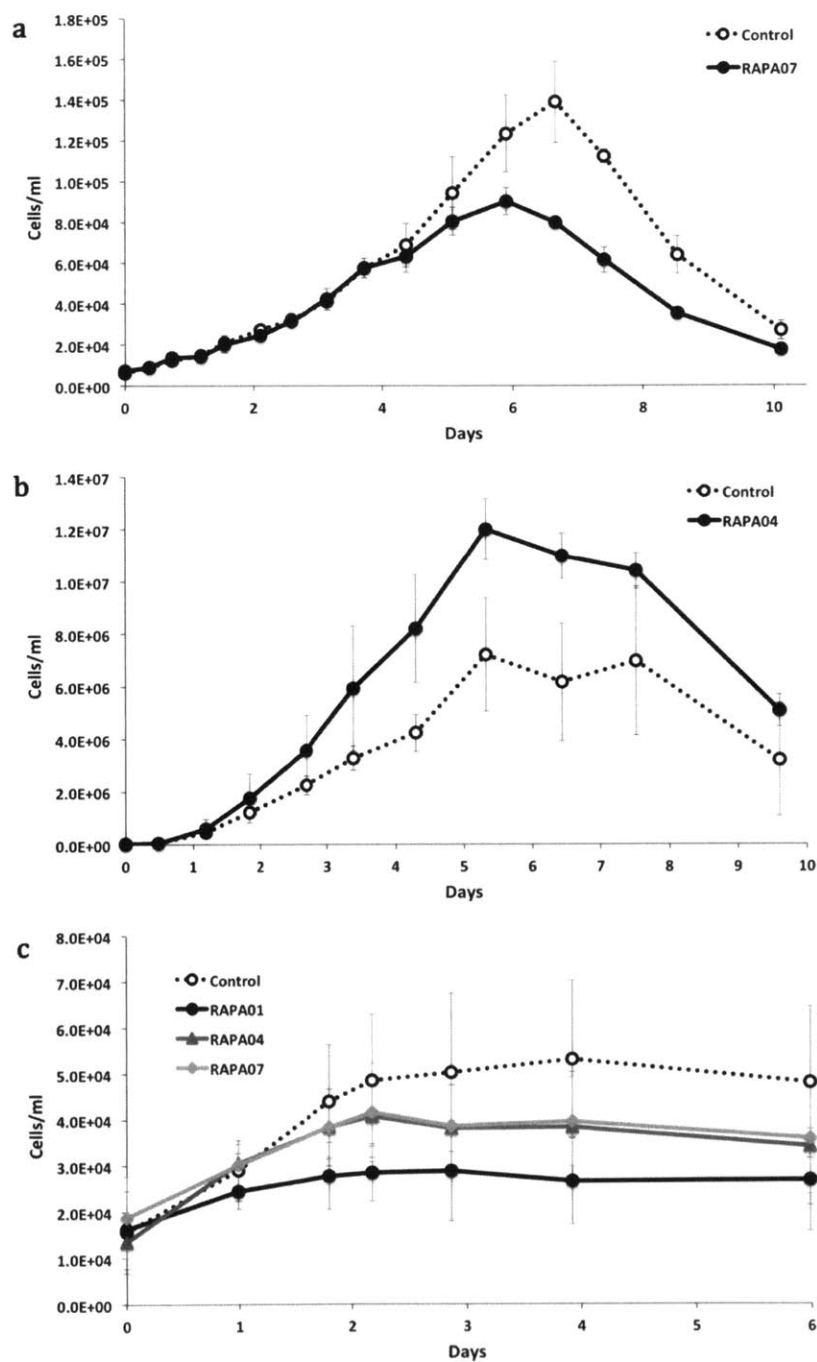


Figure 5. Growth response of HIMB4 (SAR11 subgroup IA; a), HIMB55 (OM60; b), and HIMB11 (*Rhodobacteriales*; c) to amendments with high-molecular weight dissolved organic matter (HMW-DOM) from S. Pacific surface seawater. DOM was prepared at three locations, a coastal oxygen minimum zone (RAPA01), inside the South Pacific subtropical gyre (RAPA07) and a transition zone in between (RAPA04). Certain strains tested exhibited suppressed growth due to DOM additions (a,c), while others exhibited increases in both growth rate and maximum cell density (b). Error bars indicate plus or minus one standard deviation of triplicate cultures.

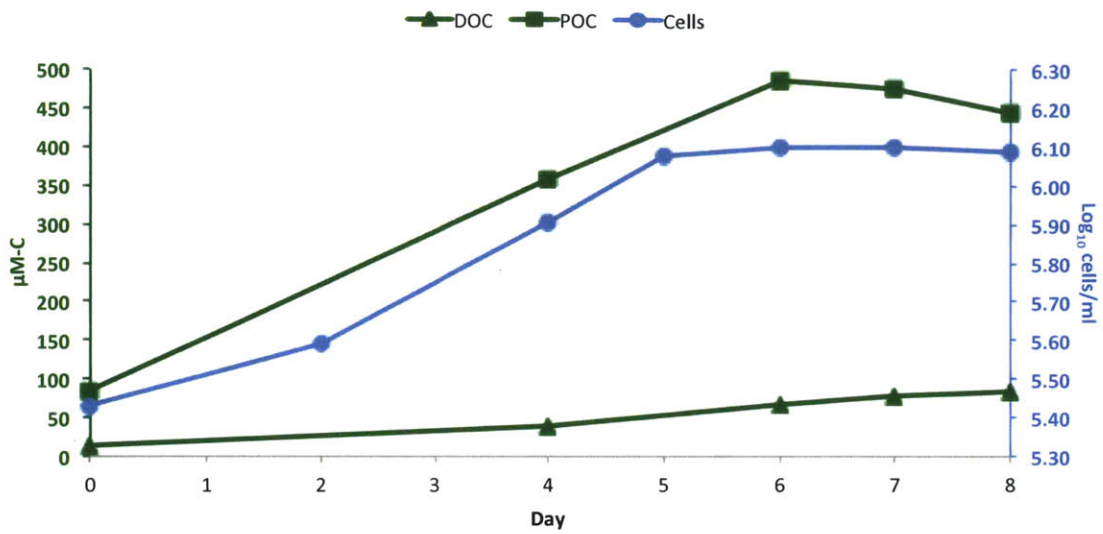


Figure 6. Growth and organic carbon production of *Crocosphaera watsonii* str. WH8501. Cell counts are presented in blue (right axis), while particulate (POC; squares) and dissolved (DOC; triangles) organic carbon production are presented at micromolar concentrations in green (left axis). Organic carbon values are blank corrected to account for organic carbon present in the background medium.

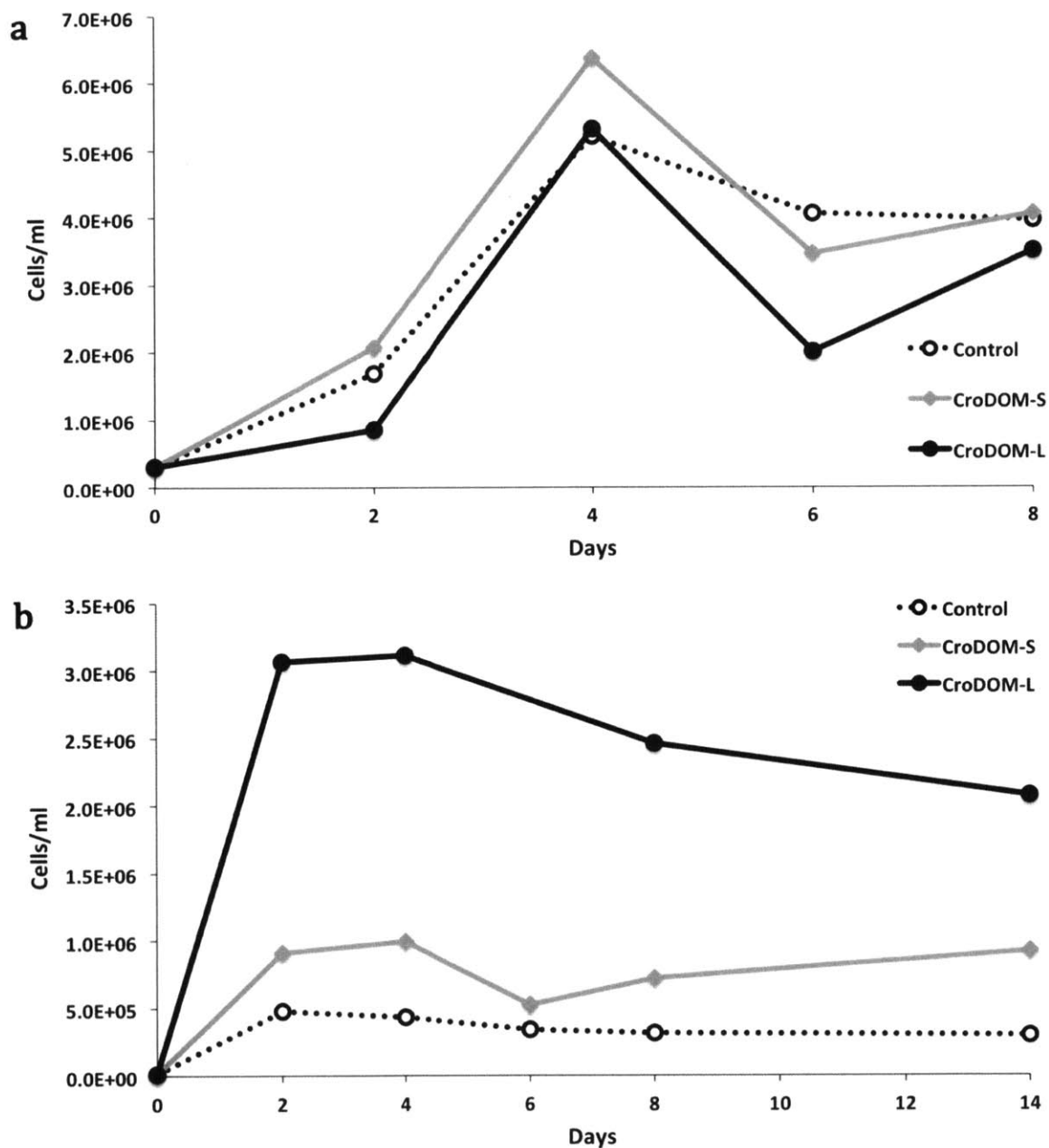


Figure 7. Growth response of HIMB137 (OM43; a) and HIMB1 (*Rhodobacterales*; b) to amendments with total DOM derived from *Crocosphaera watsonii* str. WH8501. DOM was prepared from WH8501 spent medium after removal of cell material and other particulate organic carbon. DOM was added in small (S) and large (L) quantities to examine dose specific effects on growth rate and maximum cell density. Certain strains tested exhibited either no response or even suppressed growth due to DOM additions (a), while others exhibited significant increases in both growth rate and maximum cell density (b).

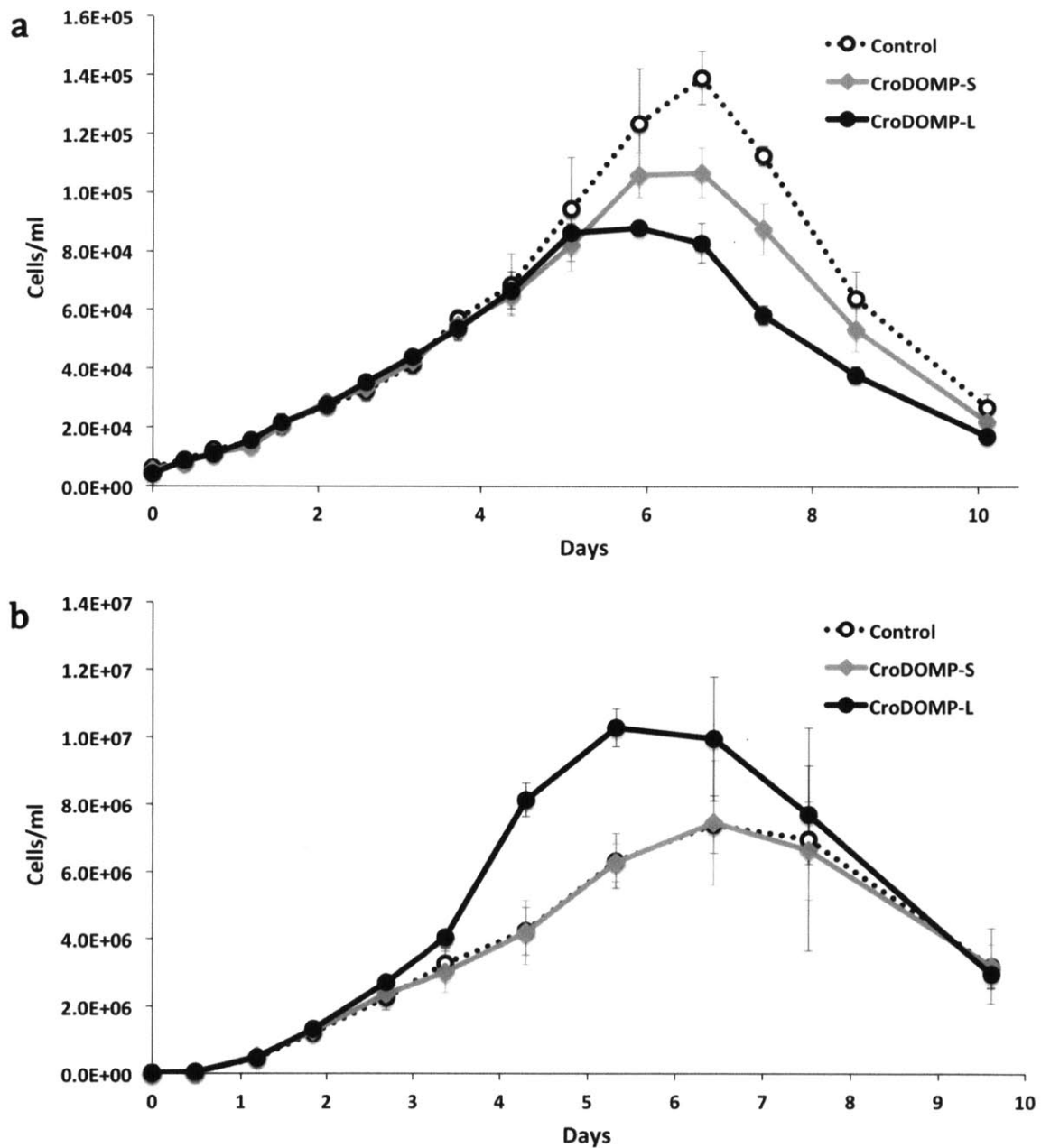


Figure 8. Growth response of HIMB4 (SAR11 subgroup IA; a) and HIMB55 (OM60; b) to amendments with DOM_p derived from *Crocospaera watsonii* str. WH8501. DOM was added in small (S) and large (L) quantities to examine dose specific effects on growth rate and maximum cell density. Certain strains tested exhibited either no response or even suppressed growth due to DOM additions (a), while others exhibited increases in both growth rate and maximum cell density (b). Error bars indicate plus or minus one standard deviation of triplicate cultures.

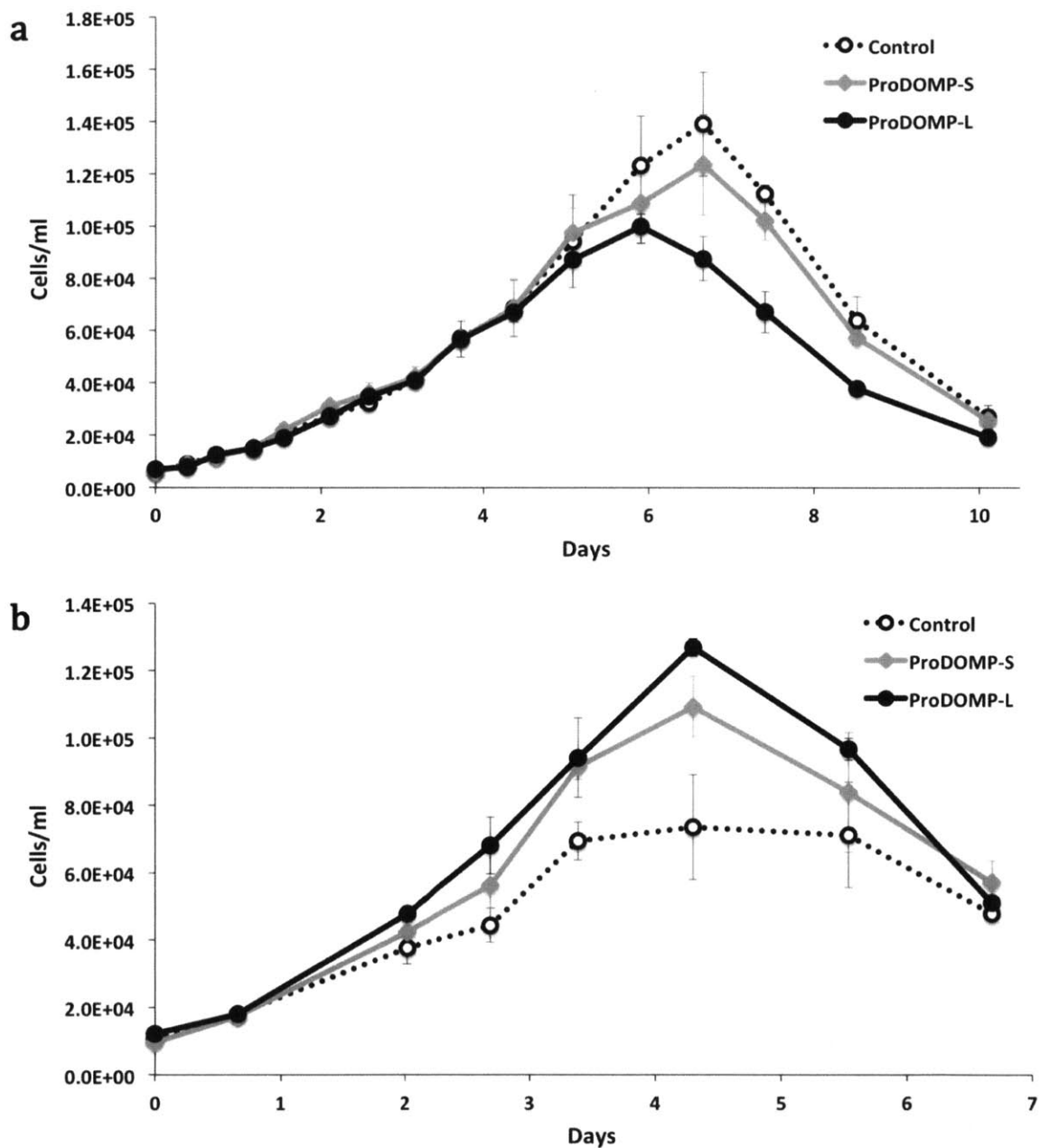


Figure 9. Growth response of HIMB4 (SAR11 subgroup IA; a) and HIMB114 (SAR11 subgroup IIIA; b) to amendments with DOM_p derived from *Prochlorococcus* str. MIT9313. DOM was added in small (S) and large (L) quantities to examine dose specific effects on growth rate and maximum cell density. Certain strains tested exhibited either no response or even suppressed growth due to DOM additions (a), while others exhibited increases in both growth rate and maximum cell density (b). Error bars indicate plus or minus one standard deviation of triplicate cultures.

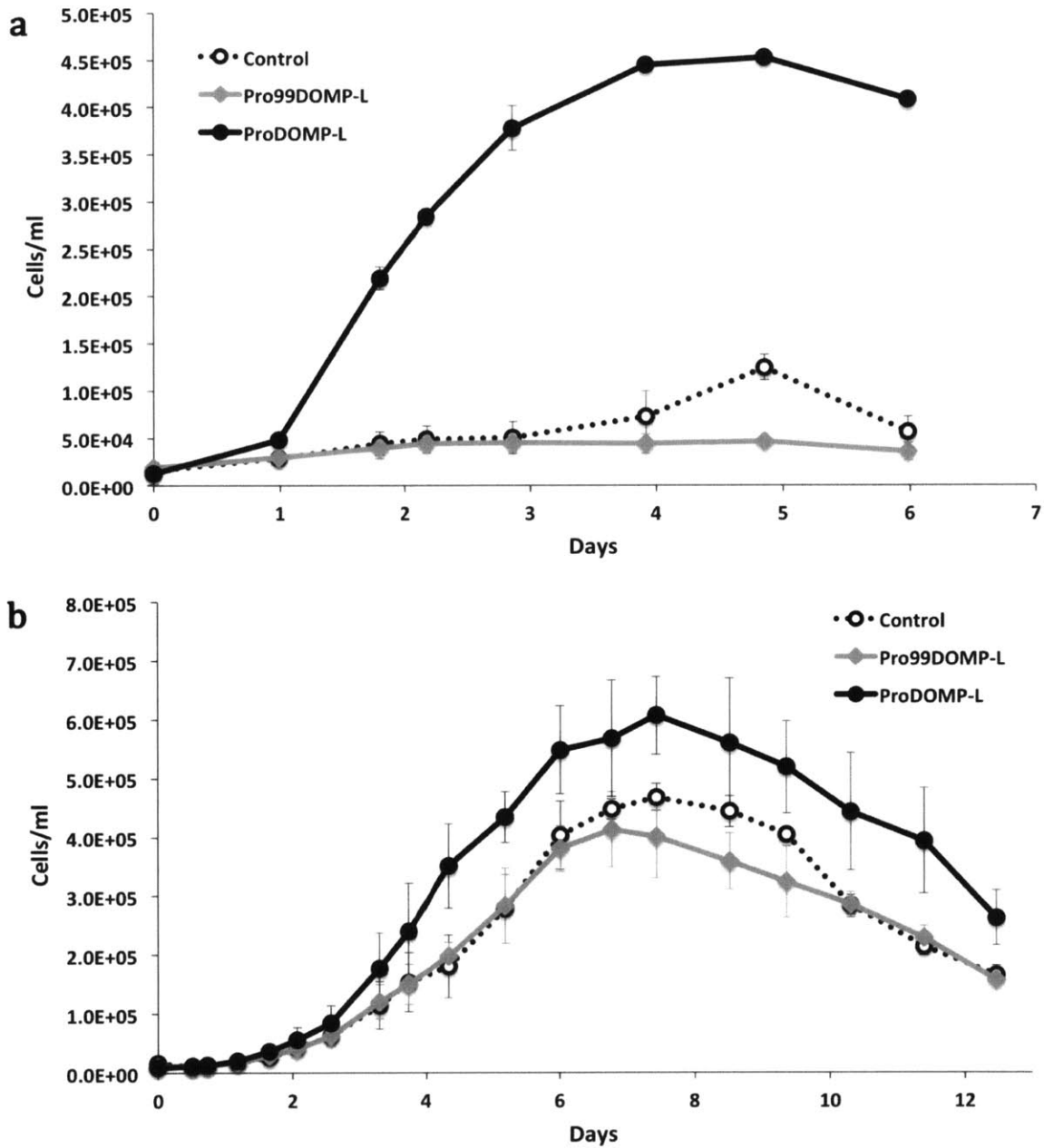


Figure 10. Growth response of HIMB11 (*Rhodobacteriales*; a) and HIMB30 (OM252; b) to amendments with DOM_P derived from *Prochlorococcus* str. MIT9313 and Pro99 media controls. The addition of DOM_P prepared from phytoplankton media was an important control to correct for growth effects caused by material present in the phytoplankton growth medium. Both strains shown here exhibited increases in both growth rate and maximum cell density due to DOM_P derived from MIT9313 but not from DOM_P derived from the Pro99 medium, indicating these responses are due to material produced by the phytoplankton culture. Error bars indicate plus or minus one standard deviation of triplicate cultures.

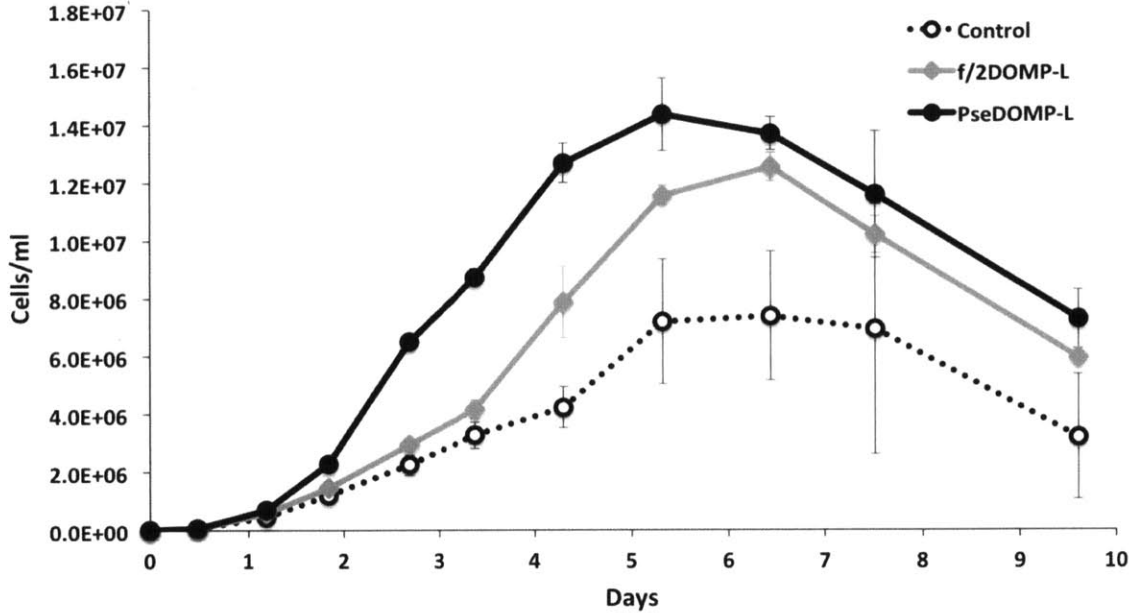


Figure 11. Growth response of HIMB55 (OM60) to amendments with DOM_p derived from *Thalassiosira pseudonana* str. CCMP1335 and the f/2 medium control. The addition of DOM_p prepared from phytoplankton media was an important control to correct for growth effects caused by material present in the phytoplankton growth medium. HIMB55 exhibited increases in both growth rate and maximum cell density due to DOM_p derived from CCMP1335, but also displayed increased growth due to the f/2 medium control, indicating that a portion of the observed response was likely due to material present in the medium used to grow CCMP1335. Error bars indicate plus or minus one standard deviation of triplicate cultures.

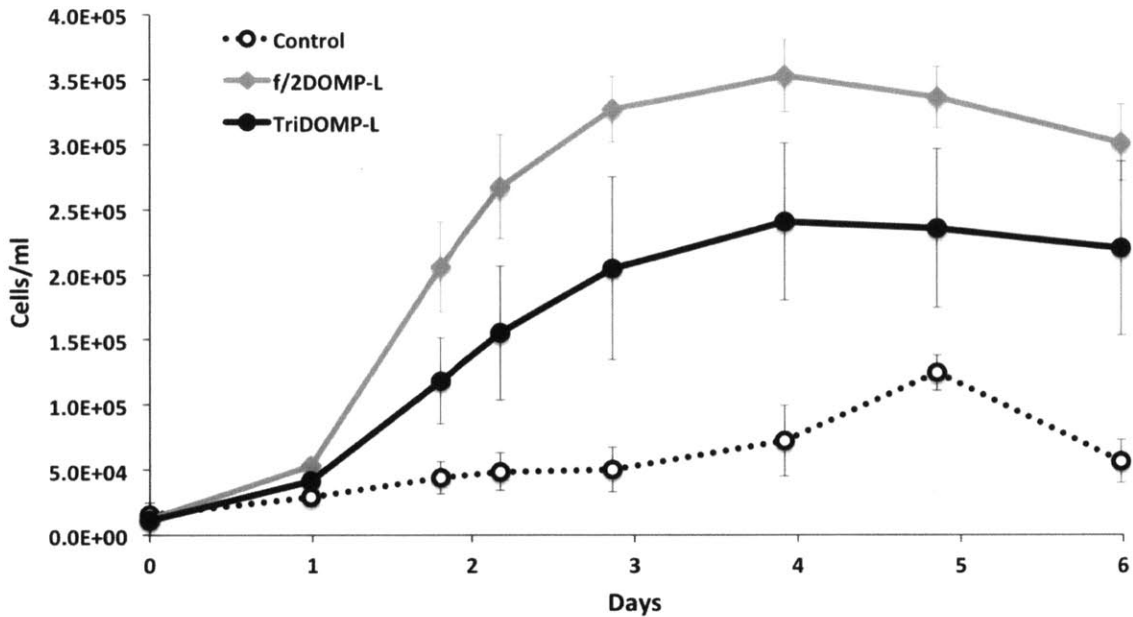


Figure 12. Growth response of HIMB11 (*Rhodobacteriales*) to amendments with DOM_p derived from *Phaeodactylum tricornutum* str. CCMP632 and the f/2 medium control. The addition of DOM_p prepared from phytoplankton media was an important control to correct for growth effects caused by material present in the phytoplankton growth medium. HIMB11 exhibited increases in both growth rate and maximum cell density due to DOM_p derived from CCMP632, but also displayed similar increases in growth due to the f/2 medium control, indicating that the observed response was due to material present in the medium used to grow CCMP632 and not necessarily from material produced by CCMP632. Error bars indicate plus or minus one standard deviation of triplicate cultures.

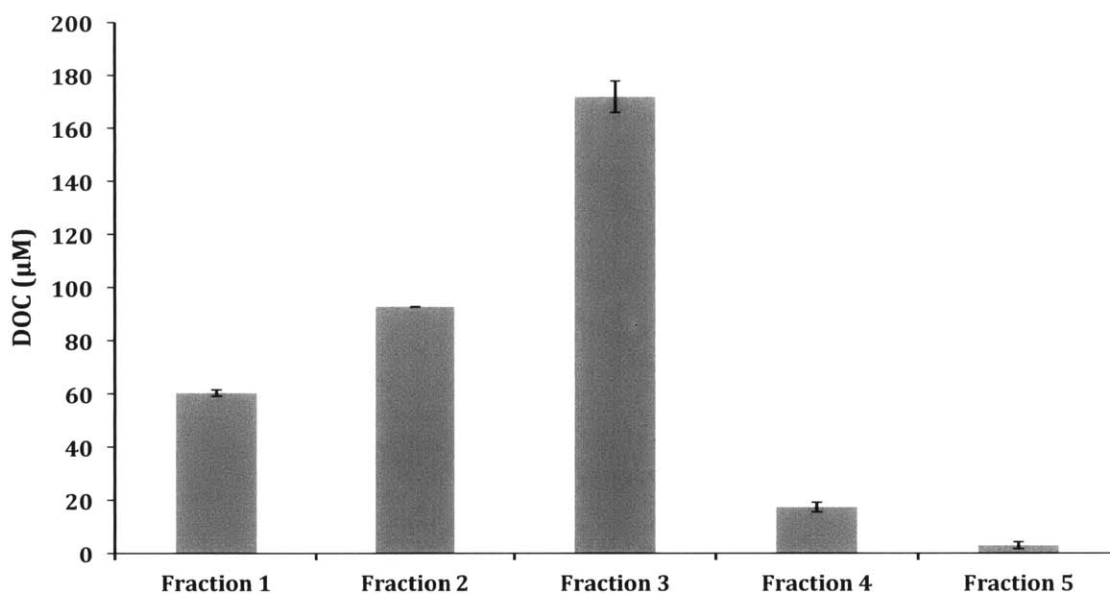


Figure 13. Micromolar concentrations of dissolved organic carbon (DOC) in treatment media prepared from five polarity fractions of DOM_P derived from *Prochlorococcus* str. MIT9313. DOC values are corrected to account for DOC in the control medium and error bars indicate plus or minus one standard deviation of duplicate samples.

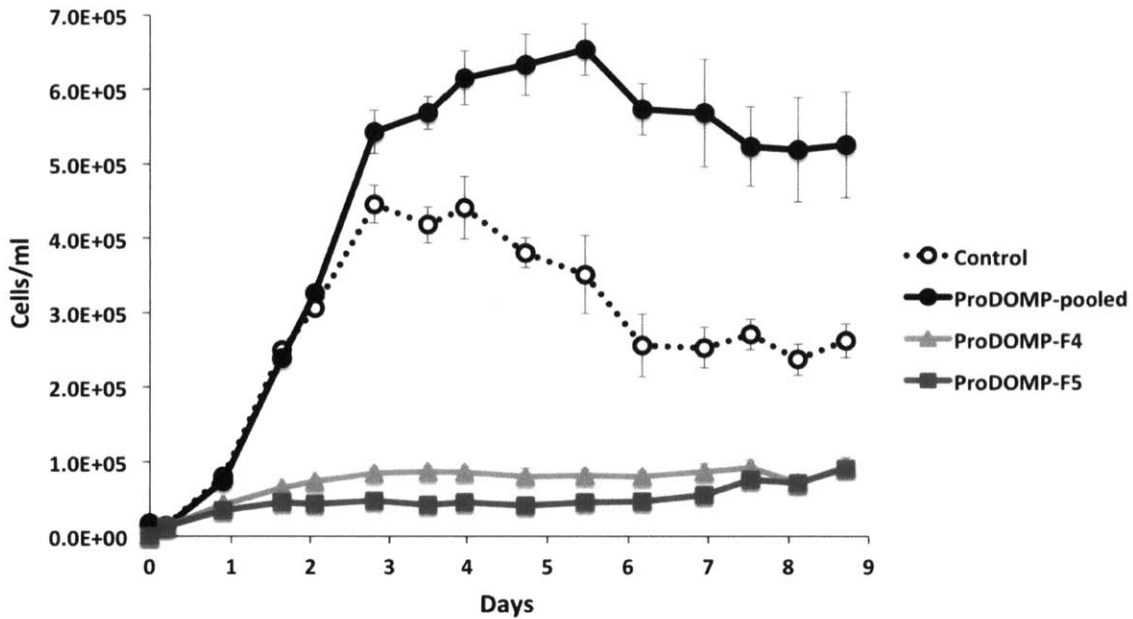


Figure 14. Growth response of HIMB30 (OM252) to amendments with DOM_P fractions derived from *Prochlorococcus* str. MIT9313. HIMB30 exhibited increases in both growth rate and maximum cell density due to unfractionated DOM_P derived from MIT9313, but displayed almost no growth when given more nonpolar fractions (F4, F5) of MIT9313 DOM_P, indicating toxic effects from these fractions and suggesting opposing effects in the unfractionated material. Error bars indicate plus or minus one standard deviation of triplicate cultures.

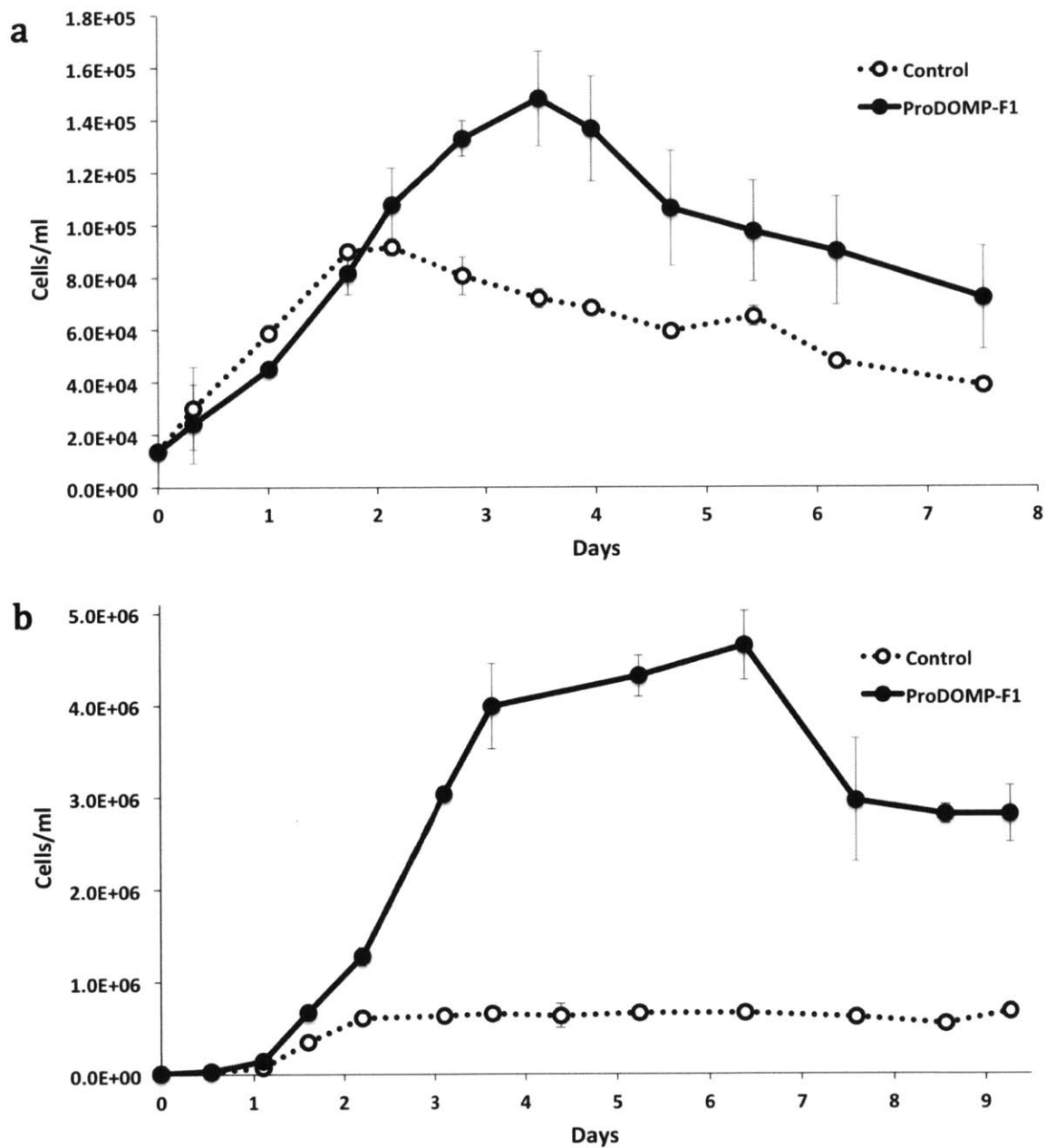


Figure 15. Growth response of HIMB100 (SAR116; a) and HIMB11 (*Rhodobacterales*; b) to amendments with the most polar DOM_p fraction (F1) derived from *Prochlorococcus* str. MIT9313. HIMB100 reached greater cell densities due this fraction edition, while HIMB11 displayed significant increases in both growth rate and maximum cell density. Error bars indicate plus or minus one standard deviation of triplicate cultures.

References

- Azam F, Malfatti F (2007) Microbial structuring of marine ecosystems. *Nature Publishing Group* 5:782–791
- Baran R, Bowen BP, Northen TR (2011) Untargeted metabolic footprinting reveals a surprising breadth of metabolite uptake and release by *Synechococcus* sp. PCC 7002. *Mol BioSyst* 7:3200
- Becker JW, Brandon ML, Rappé MS (2007) Cultivating microorganisms from dilute aquatic environments: melding traditional methodology with new cultivation techniques and molecular methods. In: Hurst CJ (ed) *Manual of Environmental Microbiology*, 3rd edn. ASM Press, p 399–406
- Bell WH, Lang JM, Mitchell R (1974) Selective stimulation of marine bacteria by algal extracellular products. *Limnol Oceanogr*:833–839
- Berman T, Chava S (1999) Algal growth on organic compounds as nitrogen sources. *Journal of Plankton Research* 21:1423–1437
- Carlson CA, Giovannoni SJ, Hansell DA, Goldberg SJ, Parsons R, Vergin K (2004) Interactions among dissolved organic carbon, microbial processes, and community structure in the mesopelagic zone of the northwestern Sargasso Sea. *Limnol Oceanogr*:1073–1083
- Cavender-Bares KK, Mann EL, Chisholm SW, Ondrusek ME, Bidigare RR (1999) Differential response of equatorial Pacific phytoplankton to iron fertilization. *Limnol Oceanogr*:237–246
- Cho J-C, Giovannoni SJ (2004) Cultivation and growth characteristics of a diverse group of oligotrophic marine Gammaproteobacteria. *Applied and Environmental Microbiology* 70:432–440
- Connon SA, Giovannoni SJ (2002) High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. *Applied and Environmental Microbiology* 68:3878–3885
- Cottrell MT, Kirchman DL (2000) Natural assemblages of marine proteobacteria and members of the Cytophaga-Flavobacter cluster consuming low- and high-molecular-weight dissolved organic matter. *Applied and Environmental Microbiology* 66:1692–1697
- Gómez-Consarnau L, Lindh MV, Gasol JM, Pinhassi J (2012) Structuring of bacterioplankton communities by specific dissolved organic carbon compounds. *Environmental Microbiology* 14:2361–2378

- Guillard RRL (2002) Culture of phytoplankton for feeding marine invertebrates. In: Smith WL, H CM (eds) Biogeochemistry of Marine Dissolved Organic Matter. Academic Press, p 26–60
- Hedges JI (2002) Why Dissolved Organics Matter. In: Hansell DA, Carlson CA (eds) Biogeochemistry of Marine Dissolved Organic Matter. Academic Press, p 1–33
- Lau WWY, Keil RG, Armbrust EV (2007) Succession and Diel Transcriptional Response of the Glycolate-Utilizing Component of the Bacterial Community during a Spring Phytoplankton Bloom. *Applied and Environmental Microbiology* 73:2440–2450
- Li B, Sher D, Kelly L, Shi Y, Huang K, Knerr PJ, Joewono I, Rusch D, Chisholm SW, Van Der Donk WA (2010) Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. *Proceedings of the National Academy of Sciences* 107:10430
- Ludwig W (2004) ARB: a software environment for sequence data. *Nucleic Acids Research* 32:1363–1371
- Luo H, Löytynoja A, Moran MA (2011) Genome content of uncultivated marine Roseobacters in the surface ocean. *Environmental Microbiology* 14:41–51
- McCarren J, Becker JW, Repeta DJ, Shi Y, Young CR, Malmstrom RR, Chisholm SW, DeLong EF (2010) Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. *Proc Natl Acad Sci USA* 107:16420–16427
- Moore LR, Coe A, Zinser ER, Saito MA, Sullivan MB, Lindell D, Frois-Moniz K, Waterbury J, Chisholm SW (2007) Culturing the marine cyanobacterium *Prochlorococcus*. *Limnol Oceanogr-Meth* 5:353–362
- Moreira IC, Bianchini I Jr, Vieira AAH (2011) Decomposition of dissolved organic matter released by an isolate of *Microcystis aeruginosa* and morphological profile of the associated bacterial community. *Braz J Biol* 71:57–63
- Nelson CE, Carlson CA (2012) Tracking differential incorporation of dissolved organic carbon types among diverse lineages of Sargasso Sea bacterioplankton. *Environmental Microbiology* 14:1500–1516
- Olson RJ, Vaultot D, Chisholm SW (1985) Marine phytoplankton distributions measured using shipboard flow cytometry. *Deep Sea Research Part A* 32:1273–1280
- Ottesen EA, Young CR, Eppley JM, Ryan JP, Chavez FP, Scholin CA, DeLong EF (2013) Pattern and synchrony of gene expression among sympatric marine microbial populations. *PNAS* 110(6): E488-E497

- Pinhassi J, Sala MM, Havskum H, Peters F, Guadayol OS, Malits A, Marrase C (2004) Changes in Bacterioplankton Composition under Different Phytoplankton Regimens. *Applied and Environmental Microbiology* 70:6753–6766
- Pluskal T, Castillo S, Villar-Briones A, Orešič M (2010) MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 11:395
- Pomeroy LR, Williams PJJ, Azam F, Hobbie JE (2007) The Microbial Loop. *Oceanography* 20:28–33
- Rappé MS, Giovannoni SJ (2003) The Uncultured Microbial Majority. *Annu Rev Microbiol* 57:369–394
- Rappé MS, Connon SA, Vergin KL, Giovannoni SJ (2002) Cultivation of the ubiquitous SAR 11 marine bacterioplankton clade. *Nature* 418:630–633
- Rink B, Seeberger S, Martens T, Duerselen CD, Simon M, Brinkhoff T (2007) Effects of phytoplankton bloom in a coastal ecosystem on the composition of bacterial communities. *Aquat Microb Ecol* 48:47
- Rinta-Kanto JM, Sun S, Sharma S, Kiene RP, Moran MA (2011) Bacterial community transcription patterns during a marine phytoplankton bloom. *Environmental Microbiology* 14:228–239
- Romera-Castillo C, Sarmiento H, Alvarez-Salgado XA, Gasol JM, Marrase C (2011) Net Production and Consumption of Fluorescent Colored Dissolved Organic Matter by Natural Bacterial Assemblages Growing on Marine Phytoplankton Exudates. *Applied and Environmental Microbiology* 77:7490–7498
- Sarmiento H, Gasol JM (2012) Use of phytoplankton-derived dissolved organic carbon by different types of bacterioplankton. *Environmental Microbiology* 14:2348–2360
- Sher D, Thompson JW, Kashtan N, Croal L, Chisholm SW (2011) Response of *Prochlorococcus* ecotypes to co-culture with diverse marine bacteria. *ISME*:1–8
- Sowell SM, Wilhelm LJ, Norbeck AD, Lipton MS, Nicora CD, Barofsky DF, Carlson CA, Smith RD, Giovannoni SJ (2009) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *The ISME Journal* 3:93–105
- Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690

- Tada Y, Taniguchi A, Nagao I, Miki T, Uematsu M, Tsuda A, Hamasaki K (2011) Differing Growth Responses of Major Phylogenetic Groups of Marine Bacteria to Natural Phytoplankton Blooms in the Western North Pacific Ocean. *Applied and Environmental Microbiology* 77:4055–4065
- Tada Y, Taniguchi A, Sato-Takabe Y, Hamasaki K (2012) Growth and succession patterns of major phylogenetic groups of marine bacteria during a mesocosm diatom bloom. *J Oceanogr* 68:509–519
- Teeling H, Fuchs BM, Becher D, Klockow C, Gardebrecht A, Bennke CM, Kassabgy M, Huang S, Mann AJ, Waldmann J, Weber M, Klindworth A, Otto A, Lange J, Bernhardt J, Reinsch C, Hecker M, Peplies J, Bockelmann FD, Callies U, Gerdts G, Wichels A, Wiltshire KH, Glockner FO, Schweder T, Amann R (2012) Substrate-Controlled Succession of Marine Bacterioplankton Populations Induced by a Phytoplankton Bloom. *Science* 336:608–611
- Trabelsi A, Rassoulzadegan F (2011) Effect of bacterial community dynamics on DOC seasonal changes in the north-western Mediterranean Sea. *Journal of Plankton Research* 33:1249–1262
- Watanabe K, Komatsu N, Kitamura T, Ishii Y, Park H-D, Miyata R, Noda N, Sekiguchi Y, Satou T, Watanabe M, Yamamura S, Imai A, Hayashi S (2012) Ecological niche separation in the Polynucleobacter subclusters linked to quality of dissolved organic matter: a demonstration using a high sensitivity cultivation-based approach. *Environmental Microbiology* 14:2511–2525
- Waterbury JB, Watson SW, Valois FW, Franks DG (1986) Biological and Ecological Characterization of the Marine Unicellular Cyanobacterium *Synechococcus*. In: Platt T, Li W (eds) *Photosynthetic Picoplankton*. Can. J. Fish. Aquat. Sci. Bull., Hoboken, NJ, USA, p 71–120
- Yarza P, Richter M, Peplies J, Euzeby J, Amann R, Schleifer K-H, Ludwig W, Glöckner FO, Rosselló-Mora R (2008) The All-Species Living Tree project: A 16S rRNA-based phylogenetic tree of all sequenced type strains. *Systematic and Applied Microbiology* 31:241–250
- Yeo SK, Huggett MJ, Eiler A, Rappé MS (2013) Coastal Bacterioplankton Community Dynamics in Response to a Natural Disturbance (DL Kirchman, Ed.). *PLoS ONE* 8:e56207

CHAPTER FOUR:

Transcriptional profiling of a marine *Oceanospirillales* reveals increased expression of genes related to anabolism and proteorhodopsin-based phototrophy, and decreased expression of genes related to motility and chemotaxis in the presence of *Prochlorococcus*-derived dissolved organic matter

Abstract

A large quantity of Earth's organic carbon exists as nonliving material dissolved in seawater (DOM), a vast reservoir almost exclusively accessible to marine microorganisms. Microbial uptake of marine DOM and its allocation into catabolic and anabolic processes largely dictates whether nutrients within the DOM pool become available to higher trophic levels or remain inaccessible via their conversion to inorganic forms or recalcitrant DOM. Despite the importance of these processes to the global cycling of carbon and other nutrients relevant to both biology and climate, details regarding metabolic pathways involved in the consumption of DOM are scarce. The current paucity of information regarding the physiology of DOM consumption limits our understanding of microbial ecology in the sea and precludes the inclusion of heterotrophic microbial lineages in most marine biogeochemical models. Here, we employ a combined transcriptomic and untargeted metabolomic approach to investigate the effect of DOM produced by *Prochlorococcus* str. MIT9313 on a marine *Oceanospirillales*, str. HIMB30. The transcriptional response of HIMB30 to the addition of *Prochlorococcus*-derived DOM revealed the upregulation of several genes indicative of anabolic metabolism and cellular growth, included elevated transcription of tripartite ATP-independent periplasmic (TRAP) transporters and sugar ABC transporters for DOM acquisition, gluconeogenesis, the glyoxylate cycle, and the non-oxidative portion of the pentose phosphate pathway. Interestingly, the *Prochlorococcus* DOM addition also elicited decreased expression of genes related to flagellar assembly and chemotaxis, and an increase in proteorhodopsin and associated genes for retinal biosynthesis.

Metabolomic analysis of the same *Prochlorococcus*-derived DOM suggested the presence of carboxylic acids, glucosides, and terpenoids that in part may explain the observed patterns of gene expression. Our findings highlight the potential for linking detailed chemical analysis of complex DOM from a known biological source with information regarding the microbial physiology underlying consumption of this DOM via gene expression.

Introduction

Marine phytoplankton produce approximately 50 petagrams (10^{15} g) of organic carbon each year (Chavez et al. 2011), about half of which ends up dissolved in seawater as a non-living pool of organic material referred to as dissolved organic matter, or DOM (Hansell & Carlson 1998). Heterotrophic bacterioplankton rely on marine DOM for survival, and the activity of these microbes determines the fate of organic nutrients in the DOM pool, including whether organic carbon from phytoplankton is ultimately stored in seawater (as either biomass or refractory DOM) or respired back the atmosphere as carbon dioxide (CO_2) (Azam et al. 1983, Pomeroy et al. 2007, Azam & Malfatti 2007). Understanding the physiological processes underlying uptake and utilization of DOM by heterotrophic microbes is challenging due to the functional and genomic diversity of marine microbes and the chemical complexity of marine DOM. Investigating these processes is important however; as the partitioning of marine DOM into catabolic vs. anabolic pathways by heterotrophic bacterioplankton affects global biogeochemical

cycles due to the large reservoir size of marine DOM and the numerical abundance of these organisms in seawater.

In marine systems, studies that link detailed information regarding cellular physiology with DOM utilization have been sparse, and have largely focused on the consumption of defined compounds such as glucose or free amino acids rather than more complex suites of environmentally relevant organic matter (e.g. Keil & Kirchman 1999, Cottrell & Kirchman 2000, Malmstrom et al. 2005). While informative, these previous experiments, as well as more recent work involving more complex DOM sources derived from marine phytoplankton cultures (Romera-Castillo et al. 2011, Nelson & Carlson 2012, Sarmiento & Gasol 2012), have focused on how DOM manipulations affect the community composition of microbial assemblages, with little attention on the underlying metabolic response. These studies support a growing body of evidence that different heterotrophic bacterioplankton taxa vary in their response to different DOM sources. However, the physiology and chemical transformations underlying these observations remain largely unknown.

Measuring the number of RNA transcripts of a gene by direct sequencing (RNA-seq) has become the method of choice for assessing gene expression in microorganisms (Wang et al. 2009, Costa et al. 2010, Croucher & Thomson 2010). In microcosm experiments, this approach has been used to examine the transcriptional response of mixed microbial communities to several different sources of DOM, including defined polyamines (Mou et al. 2011), organic sulfur (Vila-Costa et al. 2010), size-fractionated seawater DOM (McCarren et al. 2010), and phytoplankton-derived DOM (Poretsky et al.

2010, Rinta-Kanto et al. 2011). These experiments have yielded novel insight into metabolic processes relevant to marine DOM consumption, including the importance of transporters for specific DOM components (e.g. carbohydrates, carboxylic acids, polyamines), syntrophic interactions, and temporal succession dynamics. Collectively, these studies demonstrate the utility of a RNA-seq approach to interrogate the complex physiology of microbial DOM consumption. While this work has expanded our knowledge of how marine DOM is consumed in the natural environment, two factors limit the ability of the aforementioned studies to unequivocally identify connections between DOM and individual heterotrophs. First, the extreme complexity of natural microbial assemblages, including a plethora of cooperative and antagonistic interactions, limits the possibility of generating robust models of DOM degradation for any individual taxa. Second, the lack of compositional information regarding the DOM sources limits interpretation of gene expression signals, as there are often multiple potential explanations for the transcriptional response of a particular gene or pathway. A lack of chemical composition information hinders interpretation of signals derived from different DOM sources.

Recently, a global transcriptome sequencing approach has been employed in manipulation experiments involving pure cultures of marine bacteria (Giannoukos et al. 2012) and used to study both diel cycling (Waldbauer et al. 2012) and photoheterotrophy (Kimura et al. 2011) in light manipulation experiments. Cultivation-based application of RNA-seq offer the advantage of a much simpler and well defined system to identify physiological responses at the gene expression level, especially when working with

isolates for which an annotated genome sequence is available. The use of an untargeted metabolomic approach to study marine bacteria in culture has also recently been shown to be an effective way to obtain chemical composition information on both the intracellular and extracellular metabolome associated with a particular organism (Kujawinski et al. 2009, Baran et al. 2011). Untargeted metabolomics has the added benefit of requiring no *a priori* knowledge of DOM composition, unlike more traditional chemical assays for bulk DOM constituents (e.g. neutral sugars, amino sugars), and allows for putative identification of compounds at the molecular level. The combination of cultivation-based RNA-seq and untargeted metabolomics offers an ideal platform to investigate the consumption of complex DOM suites by representative marine bacteria as gene expression signals and DOM composition information can each be used to aid in the interpretation of the other.

A combined transcriptomic and metabolomic approach was used here to study the consumption of DOM isolated from *Prochlorococcus* str. MIT9313 (DOM_P) by a marine *Oceanospirillales* (str. HIMB30). HIMB30, likely representing a novel family and genus within the *Oceanospirillales* order, is a motile marine gammaproteobacterium isolated from the coast of Oahu with close relatives found in many genomic surveys of coastal environments (Huggett & Rappé 2012). The recent sequencing of the HIMB30 genome revealed genes involved in the Entner-Doudoroff pathway, proteorhodopsin-based phototrophy, sulfur oxidation, and CO₂ fixation, highlighting its diverse metabolic capabilities. *Prochlorococcus* was chosen as the DOM source organism as it is the dominant phytoplankton clade in oligotrophic surface waters and likely the most

abundant photosynthetic organism on Earth (Partensky et al. 1999, Mühling 2011). DOM derived from strain MIT9313 was previously found to enhance the growth of several marine bacterial isolates, including HIMB30 (see chapter 3).

In this study, gene expression in HIMB30 grown in the presence or absence of *Prochlorococcus*-derived DOM revealed several physiological responses to the DOM addition responsible for the observed increase in growth, including uptake of carboxylic acids, glucosides, and glycerolipids and utilization of these compounds for anabolic processes including biosynthesis of pentoses, cell membrane, and cell wall material. DOM addition also resulted in decreased expression of genes involved in flagellar assembly and chemotaxis and increased expression of genes related to proteorhodopsin-based phototrophy. Untargeted metabolomic analysis of *Prochlorococcus*-derived DOM resulted in putative compound identifications that supported transcription signals. Our findings provide novel information on microbial physiology related to the consumption of marine DOM and challenge several current paradigms regarding the relationship between organic nutrient concentrations, motility and photoheterotrophy in other microbial systems.

Materials and Methods

Isolation of Prochlorococcus DOM_p

Triplicate 20 L batch cultures of *Prochlorococcus* str. MIT9313 were grown to early stationary phase in order to obtain a large quantity of ProDOM_p for use in DOM characterization and heterotrophic growth assays. Details of *Prochlorococcus* growth

conditions and preparation of the pooled ProDOM_P are described in Chapter 3 methods under “Fractionation of *Prochlorococcus*-derived DOM_P” as this same DOM_P was also used in pooled and fractionated small volume growth assays. Previous work demonstrated a high degree of compositional variation in DOM_P from replicate cultures, likely due to minor variations in growth conditions (see chapter 2 for details), so DOM_P from triplicate cultures was pooled in an effort to obtain a more representative suite of DOM_P derived from *Prochlorococcus* str. MIT9313. Particulate organic carbon (POC) production was measured for *Prochlorococcus* str. MIT9313 prior to isolation of DOM_P according to protocols outlined in chapter 3 methods under ‘DOM_P amendments’. Measurements of dissolved organic carbon (DOC) production were not possible due to additions of HEPES buffer during MIT9313 culturing to mitigate changes in pH.

Chromatographic separation and detection of ProDOM_P-pooled metabolites

Replicate injections of 25 µl subsamples of the ProDOM_P-pooled material used to prepare DOM medium and Pro99DOM_P medium control were separated using high performance liquid chromatography (HPLC; Surveyor MS Pump, Thermo Scientific Corp.) and chemically analyzed using a hybrid 7 Tesla linear ion trap Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer (LTQ FT Ultra, Thermo Electron Corp.). Mass spectrometry was performed using a micro-electrospray ionization (ESI) source and HPLC effluent was infused into the ESI source at a rate of 0.8 ml/min. ESI spray voltage was 4.2 kV (positive mode) and the capillary temperature was 330 °C. Data was acquired using conventional full scan mode (m/z 100-1200) and collected as a

RAW file using Xcalibur 2.0 software. Features of interest previously identified using HPLC coupled to low resolution mass spectrometry as described in chapter 2 were used as targets for fragmentation by collision-induced dissociation (CID) during ultrahigh resolution (± 2 ppm) mass spectrometry runs (Table 1). Targets consisted of the top 120 most intense features in the *Prochlorococcus* culture samples relative to the Pro99 medium control that were not found to be produced by other phytoplankton, including multiple *Synechococcus* strains and the diatom strains *T. pseudonana*, *T. rotula*, and *P. tricornutum* (see chapter 2 for details). HPLC conditions (column, solvents, flow rate, etc.) were replicated from low-resolution runs (see chapter 2 methods, ‘extraction and analysis of DOM_p’ for details).

XCMS online (Tautenhahn et al. 2012a) was used to analyze ultrahigh resolution mass data and MZmine 2 molecular profiling software (Pluskal et al. 2010) was used to analyze tandem mass spectrometry (MS/MS) data. For XCMS online analysis, Thermo Fisher RAW files were converted to centroid mode mzXML files using ProteoWizard (Chambers et al. 2012) before uploading both control and treatment datasets. Features were detected using a 2 ppm maximum m/z deviation in consecutive scans and a signal to noise ratio of 5:1. Retention times were corrected based on correlations of the raw data using a step size of 1 m/z . Features with p -values <0.01 (calculated from two sample Welch’s t test) and fold change >4.0 were included in analyses. Putative compound identifications for several target features were made using the METLIN Metabolite Database (Smith et al. 2005) and further supported with MS/MS fragmentation information when possible. A recent test of the physiological relevance of this database

revealed that ca. 50% of the total features detected in cultures of *Escherichia coli* had matches in METLIN, which was higher than other available databases such as LipidMaps, the Human Metabolome Database (HMDB), and MassBank (Tautenhahn et al. 2012b). Here, high resolution mass information from *Prochlorococcus* str. MIT9313 was used to locate matching metabolites in the METLIN database (± 2 ppm mass difference) using multiple adducts ($[M+H]^+$, $[M+Na]^+$) for data acquired in positive ion mode and a previously described annotation algorithm (Kuhl et al. 2012). It is important to note that these putative identifications are limited to metabolites that exist in the METLIN database and influenced by the parameters one sets when performing matches (e.g. the annotation algorithm employed, adducts included, Δ ppm window, etc.).

Heterotrophic growth media preparation

Control medium was sterilized seawater from Kaneohe Bay, Hawaii, collected and prepared using existing protocols (Becker et al. 2007) and as previously described in chapter 3 methods under ‘Control medium preparation’. Un-inoculated medium was also incubated and checked periodically for sterility as described in chapter 3. Resuspended pooled ProDOM_p was sterile filtered through a pre-rinsed 0.1 μ m Anotop filter (Whatman) before adding it to 10 L of control medium in order to produce the treatment medium (hereafter referred to as DOM medium). DOC was quantified in both control and DOM media prior to addition of heterotrophic cells according to methods described in chapter 3 methods under ‘Quantification of organic carbon’.

HIMB30 cultivation and sampling

Cells of strain HIMB30 were pre-conditioned in control medium for several generations at 26-27 °C on a light-dark (12:12) cycle (ca. 60-80 $\mu\text{E}/\text{m}^2/\text{s}$) with periodic gentle mixing. Tissue culture-grade polystyrene or polycarbonate plasticware (Nalge Nunc Int.) was prepared as previously described (see chapter 3 methods, 'DOM lability assay' for details) and used for all culture growth and media storage. Exponentially growing cells were inoculated into 10 L of both control and DOM media that was preconditioned overnight at 27 °C. Inoculations targeted an initial concentration of 8×10^3 cells/ml in order to meet the lower limit of accurate quantification. Small samples (<1 ml) were immediately removed from both cultures in order to measure initial cell concentrations before carboys were sealed with pre-sterilized ported caps. Ports included an inlet for sparging with filter-sterilized (0.2 and 0.1 μm HEPA filters) air to promote mixing and oxygenation, a capped sampling port to withdraw subsamples, and a filtered (0.1 μm HEPA filter) exhaust vent. Carboys were incubated at 26 - 27 °C on a light-dark (12:12) cycle and constant sparging with filter-sterilized air. Light levels ranged from 60-80 $\mu\text{E}/\text{m}^2/\text{s}$ during light cycles and both sampling and vent ports were covered with bags for an added layer of protection. Cellular enumeration occurred approximately every 6 hr during exponential and stationary growth using a Guava EasyCyte Plus cytometry system (Millipore Corp.) according to methods described previously (see chapter 3 methods, 'Bacterioplankton growth assay' for details).

Subsamples for transcriptomic characterization were taken at specific points during growth and filtered for subsequent RNA extraction. Culture carboys were moved

to a laminar flow hood, subsamples filtered onto 0.2 μm Sterivex filters (Millipore), and returned to the incubator in ca. 8 min. Two liters of each culture was filtered at time points during exponential and late exponential growth phase, while 1 L of each was filtered at points during early stationary and stationary phase. In addition, the remaining culture that was used for initial inoculations (824 ml) was also filtered for subsequent RNA extraction (referred to hereafter as the T0 sample) to examine the physiological state of the inoculum. Once filtration was complete, 1.8 ml of RNAlater (Applied Biosystems) was immediately added to each filter before capping, flash freezing in liquid nitrogen, and storage at -80°C . Samples were periodically taken for epifluorescence microscopy (see chapter 3 for protocol) in order to monitor the cellular morphology of strain HIMB30 (Figure 1) in an effort to rapidly identify potential contamination. Due to a pleomorphic cell morphology, microscopy checks for purity and identification were not definitive. Thus, small (20 ml) samples were also taken periodically for small subunit (SSU) rRNA gene sequencing using previously described methods (Becker et al. 2007).

RNA extraction, processing and pyrosequencing

Sterivex filters containing cellular biomass were processed to extract total RNA using the *mirVana* miRNA kit (Ambion), modified as previously described (Shi et al. 2009, McCarren et al. 2010). In brief, RNAlater was removed from filters after thawing on ice, replaced with 1.5 ml of lysis buffer and 150 μl miRNA homogenate additive (Ambion), and vortexed multiple times to lyse cells prior to acid-phenol:chloroform extraction following the manufacturer's protocol. Total extracted RNA was treated with

DNase using the TURBO DNA-free kit (Ambion) to remove carryover genomic DNA prior to sample purification using the RNeasy MinElute kit (Qiagen).

HIMB30 ribosomal RNA (rRNA) transcripts were reduced in the bulk RNA extracts using rRNA probes in a subtraction hybridization protocol as previously described (Stewart et al. 2010). In brief, amplicons of the 16S and 23S rRNA genes of HIMB30 were created using universal primers for the 16S rRNA gene and custom designed primers for the 23S rRNA gene, with both reverse primers containing a T7-appended promoter sequence (DeLong et al. 1999). Amplicons were used for in vitro transcription of biotin-labeled anti-sense rRNA (aRNA) probes, which were used to bind HIMB30 rRNA in each sample before hybridization to streptavidin-coated magnetic beads (New England Biolabs), allowing for magnetic separation of rRNA from non-rRNA in each sample. The presence of 16S and 23S rRNA peaks were examined to assess subtraction efficiency on a 2100 Bioanalyzer (Agilent). Primer sequences were as follows (T7 promoter sequence underlined):

Eub16S_27F:*AGAGTTTGATCCTGGCTCAG*

Eub16S_1492R_T7:*GCCAGTGAATTGTAATACGACTCACTATAGGGACGGCT
ACCTTGTTACGACTT*

HIMB3023S_28F:*CCTTGGCAATCAGAGGCGATGAA*

HIMB3023S_2807R_T7:*GCCAGTGAATTGTAATACGACTCACTATAGGGCTC
ACAACGCTTACACACCCAAC*

Bulk RNA remaining after rRNA subtraction was polyadenylated, reverse transcribed to cDNA, converted to double-stranded DNA, and purified before in vitro transcription to

produce multiple copies of aRNA using the MessageAmp II-Bacteria kit (Ambion) as previously described (Frias-Lopez et al. 2008). Double-stranded cDNA was then synthesized from purified aRNA using the SuperScript Double-Stranded cDNA Synthesis Kit (Life Technologies) according to the manufacturer's instructions. Double-stranded cDNA was purified using the QIAquick PCR Purification Kit (Qiagen) before digesting with *BpmI* to remove poly(A) tails. Digested cDNA was purified and subsequently sequenced on a Roche Genome Sequencer FLX machine using the AMPure kit (Agencourt) according to the manufacturer's protocol. Single-stranded DNA libraries generated from purified cDNA (454 Life Sciences, Roche) were sequenced using Titanium series chemistry (half-plate per sample except the T0 sample, which was a full plate).

Bioinformatic analysis

Transcriptomes were sequenced from both sterile seawater (control) and DOM-treated HIMB30 batch cultures at 37, 60, 81, and 108 hours after inoculation (T1, T2, T3, and T4, respectively). Pyrosequencing yielded an average of 482,000 reads per sample after removal of technical replicates. A rRNA blast database comprised of the 5S, 16S, and 23S rRNA genes for HIMB30 was generated based on annotated genome information from NCBI, and rRNA reads in the dataset were subsequently identified by blasting the generated sequence data against this database. In silico removal of the remaining rRNA reads in each sample left an average of 459,000 non-rRNA reads per half-plate. A database comprising all HIMB30 annotated proteins was generated using the KEGG

Automatic Annotation Server (KAAS). The HIMB30 genome from NCBI was blasted against related gammaproteobacteria in the manually curated KEGG GENES database to produce KEGG Orthology (KO) assignments. Non-rRNA sequences were then blasted against this database (bit score cutoff of 50) to identify significant transcript matches and evaluate coverage. On average, 334,000 reads per sample (73%) had protein hits, matching an average of 2,180 proteins per sample, or 98% of the 2,227 NCBI predicted HIMB30 proteins. If a read had multiple hits, the more common hit was assigned, and assignments remained consistent among all samples. Protein hits were then organized by KO into KEGG hierarchy levels.

Statistically significant changes in gene expression in the presence of *Prochlorococcus*-derived DOM were identified using baySeq, a program in R that employs empirical Bayesian methods to estimate posterior likelihoods of differential expression within a sample set (Hardcastle 2010). Normalized read counts were compared between control and DOM cultures at all four time points to estimate the probability of both differential expression (DE; posterior likelihood ≥ 0.9) and non-differential expression (NDE; posterior likelihood < 0.9) for each gene. Reads within each sample were also normalized to the total reads for that sample to provide relative expression information and to calculate fold changes in relative expression between the control and treatment. Fold changes were used to indicate whether DE genes were enriched or underrepresented in the DOM treatment at each time point. DE genes were organized by their inclusion in KEGG level 3 pathways to uncover patterns of physiological variation between the treatment and control over time. DE genes were also

organized by genome position to examine syntenic clusters of increased or decreased expression in the DOM treatment relative to the control at each time point. Scatter plots comparing relative gene expression levels between the control and treatment were generated for each time point to visualize fold changes in both DE and NDE genes. Finally, expression data was overlaid on the annotated HIMB30 genome from NCBI to visualize genome coverage and to identify regions of unexpected expression, including areas annotated as non-coding regions. Read statistics and DE counts are summarized for each sample in Tables 2 and 3.

Results and Discussion

Strain HIMB30, a unique lineage of Gammaproteobacteria within the order *Oceanospirillales* (Huggett & Rappé 2012), repeatedly exhibited enhanced growth rates and maximum cell densities when cultured in the presence of DOM_P derived from *Prochlorococcus* str. MIT9313 (see chapter 3, figure 9b). The response was robust, occurring in biological triplicate cultures during multiple rounds of testing, including additions of independently prepared DOM_P from strain MIT9313 batch cultures grown at separate times. Dose specific effects were evident and no growth enhancement was found for additions of DOM from Pro99 medium controls (see chapter 3, figure 10b). These findings indicate that HIMB30 is capable of utilizing metabolites produced and released by *Prochlorococcus* str. MIT9313, an ability that was not ubiquitous among all of the heterotrophic marine bacterial strains tested in this study (see chapter 3), providing a controlled system in which to examine the metabolic utilization of marine DOM from a

known biological source in detail. The coupling of this metabolic investigation with information regarding the chemical composition of the DOM source offers an unparalleled look at physiological processes pertinent to the microbial loop. Coupled knowledge of the source, sink, and composition of marine DOM can begin to produce a holistic view of DOM cycling by marine microbes.

Prochlorococcus str. MIT9313 DOM

Triplicate 20 L cultures of *Prochlorococcus str. MIT9313* reached an average density of 2×10^8 cells/ml and produced an average of 2.4 mM POC (12 fmol/cell) over a growth period of 13.5 days (see chapter 3, Table 3 for details). Solid-phase extraction of DOM_P from each culture yielded many features observed using low resolution mass spectrometry coupled to HPLC. An average of 777 features ranging from 136 to 1,423 *m/z* were found to be present in culture samples, but absent from all control and blank samples, including the Pro99 medium control (see chapter 3, Table 5 for details). DOM_P from triplicate cultures was pooled and analyzed alongside Pro99DOM_P media controls using ultrahigh resolution mass spectrometry. A total of 673 features were found to be unique to the pooled culture sample (*p*-value <0.01), 661 of which had intensities >4.0 fold the mean intensity of the control features, while 229 features were found unique to the media controls, 141 with a fold change >4.0 (Figure 2). Unique features ranged from 147.101 to 1,088.583 *m/z* over a broad polarity spectrum (retention times corresponding to 4.5 – 100% methanol). A total of 238 unique features had putative matches (+/- 2 ppm) in the METLIN database, including 21 features from the list of 120 that were

targeted for fragmentation based on low resolution mass spectrometry data (Table 1). Putative DOM_P compound identifications are discussed below alongside related information regarding the effects of DOM_P addition on gene expression in str. HIMB30.

HIMB30 growth on Prochlorococcus-derived DOM

Initial experiments with replicate small volume (4 ml) HIMB30 batch cultures grown in the presence of 4 μ M DOC derived from *Prochlorococcus* str. MIT9313 DOM_P and 3 μ M DOC derived from Pro99 medium DOM_P revealed a 30% increase in maximum cell density in the presence of the *Prochlorococcus* culture-derived DOM_P (see chapter 3, Figure 10b). DOM medium produced from the pooled MIT9313 DOM_P sample for larger volume (10 L) HIMB30 batch cultures consisted of a larger increase in DOC (37 μ M), or a 1.5 fold increase over the control medium concentration of 82 μ M DOC. For reference, traditional media used for cultivation of marine bacteria such as R-2A and Marine Broth 2216 (Difco Laboratories) contain millimolar concentrations of DOC, typically ca. 4 orders of magnitude greater than the *Prochlorococcus* culture-derived DOM_P medium employed in this study. Large volume batch culturing of str. HIMB30 in *Prochlorococcus* culture-derived DOM_P medium induced an extended exponential phase, resulting in a 46% increase in maximum cell density over control medium (Figure 3). Periodic checks via epifluorescence microscopy and direct sequencing of the SSU rRNA gene revealed no contamination of str. HIMB30 batch cultures throughout the experiment regardless of media type.

Transcriptional response of str. HIMB30 to Prochlorococcus-derived DOM

An analysis of the HIMB30 transcriptome at four time points during growth in the presence and absence of *Prochlorococcus*-derived DOM_p revealed a number of significant differences in gene expression (Table 3). For all time point samples but one (Control T1), there was a very small percentage of rRNA remaining after the subtraction hybridization protocol (< 1% on average). The control T1 sample contained 32.1% rRNA, which was likely due to a reduction in the amount of probes added to this sample during the subtraction hybridization step. The control T1 sample contained the least amount of total RNA (76 ng vs. ca. 250 ng for all other samples), therefore probe concentrations were reduced to maintain a consistent RNA to probe ratio and this appears to have affected the efficiency of the subtraction. All samples contained nearly 400,000 non-rRNA reads, which were used to evaluate changes in gene expression between the control and treatment batch cultures at each time point.

Expression levels indicate that, even during exponential growth, HIMB30 cells grown in the presence of *Prochlorococcus*-derived DOM_p medium are able to devote more resources to pathways involved in anabolic processes such as gluconeogenesis, glyoxylate cycle, and the pentose phosphate pathway compared to cells in the control medium, suggesting an increased bacterial growth efficiency due to DOM treatment (Figure 4). In late exponential phase, when HIMB30 cell numbers in the DOM treatment began to exceed those in the control, many genes related to cellular growth and replication (e.g. DNA replication, translation, ribosomal proteins, cell division, cell wall and membrane biosynthesis, etc.) were overexpressed in the DOM treatment and the

control was more focused on combating oxidative stress. The presence of *Prochlorococcus*-derived DOM_P also affected the expression of genes related to motility and chemotaxis, as well as proteorhodopsin-based phototrophy (Figure 4). Genes and associated KEGG level 3 pathways involved in these potential physiologies are discussed below along with putative identifications for relevant compounds detected in the DOM treatment by ultrahigh-resolution mass spectrometry.

Carboxylic acid metabolism

Many of the genes exhibiting significant differential expression between the control and DOM treatment were either directly or indirectly related to the metabolism of carboxylic acids, particularly fatty acids, and the use of this material in anabolic processes (Figures 5-8). Although distant relatives, the closest cultivated relatives to HIMB30, *Litoricola lipolytica* and *Litoricola marina*, have been shown to have esterase lipase enzyme activity for hydrolyzing C8 lipids (Kim et al. 2007, Choi et al. 2010), suggesting that the degradation of lipid material may be an important process for many organisms within this order. HIMB30 genes encoding for four important enzymes involved in β -oxidation of fatty acids were overexpressed early in the DOM treatment (acyl-CoA synthetase, polyhydroxylalkanoate synthase, acetoacetyl-CoA reductase, β -ketothiolase), with significant overexpression exhibited at T1 for polyhydroxylalkanoate synthase and β -ketothiolase. Acyl-CoA synthetase converts fatty acids to fatty acyl-CoA in the first step of β -oxidation, while the other 3 enzymes are involved in the final steps, which liberate acetyl-CoA (Table 4). Metabolism of linoleic acid in particular was

evident by the significant overexpression (T1, T5) of a gene involved in the oxidation of hydroxyoctadecadienoic acid (L-sorbose dehydrogenase).

The increased expression of transport-related genes indicated the potential for increased acquisition of carboxylic acids by HIMB30 cells in the DOM treatment. A gene coding for a putative tricarboxylic transport membrane protein (TctC) was significantly overexpressed in the treatment at T1 and T5. Twelve genes encoding tripartite ATP-independent periplasmic (TRAP) transporters, a diverse group united by their transport of substrates containing carboxylate groups (Forward et al. 1997), were also significantly overexpressed in the treatment for at least one time point. These genes primarily encode for solute binding proteins from the DctP and TAXI families and two fusion proteins likely involved in fusing the small and large TRAP integral membrane proteins of the DctPQM TRAP system, a phenomenon known to occur for some TRAP transporters (Mulligan et al. 2010). Also, it is thought that TRAP transporters are often dependent on Na^+ as a coupling ion (Mulligan et al. 2010) and a Na^+/H^+ antiporter was found to be significantly overexpressed in the treatment at T3. Genes encoding TRAP transporters and Na^+ symporters have previously been shown to be significantly overexpressed in marine communities amended with phytoplankton-derived DOM (Poretsky et al. 2010) and during marine phytoplankton blooms (Rinta-Kanto et al. 2011). TRAP transporters have also been found in abundance in proteomic analyses of the ubiquitous marine bacteria clade SAR11 both in culture (Sowell et al. 2008) and *in situ* (Sowell et al. 2009) and a survey of prokaryotic genomes revealed that TRAP-related

gene densities were greatest in marine microbes (Mulligan et al. 2010), highlighting the likely importance of this transporter family in the marine environment.

In contrast to the β -oxidation signals found in the DOM treatment, genes involved in the biosynthesis of fatty acids, particularly for initial synthesis steps, were often more highly expressed in the control (Table 4). Genes FabBDFG, which were all significantly more highly expressed in the control for at least one time point, encode enzymes that first convert malonyl-CoA to malonyl-acyl carrier protein (ACP) en route to producing various 3-hydroxy-acyl-ACPs. Six genes encoding TRAP-type transport proteins were found to be significantly overexpressed in the control, two of which are likely related to malonate-specific TRAP transport (Mulligan et al. 2010), as these genes have synteny in the HIMB30 genome with malonyl-CoA decarboxylase and malate dehydrogenase. These nearby genes suggest that malonate may have been transported into cells for use in fatty acid biosynthesis. Interestingly, enzymes carrying out the final two steps in the biosynthesis of saturated fatty acids (FabAI) were both overexpressed in the treatment, and could indicate that carboxylic acid precursors of saturated fatty acids were present in the DOM medium, required just the final dehydration and reduction steps to produce saturated acyl-ACPs.

Expression data for several central metabolic pathways support the idea that HIMB30 cells in the control had to generate acetyl-CoA for biosynthesis of fatty acids (and likely the citric acid cycle as well), while cells in the DOM treatment acquired acetyl-CoA through other means, likely oxidation of exogenous carboxylic acids. Several genes involved in the production of acetyl-CoA from the metabolism of glucose,

pyruvate and acetate were overexpressed in the control at multiple time points. Nine genes involved in the conversion of glucose to pyruvate via glycolysis were significantly overexpressed in the control for at least one time point, including triosephosphate dehydrogenase, phosphoglycerate kinase, and pyruvate kinase. Four genes involved in the conversion of glucose to pyruvate specifically via the Entner-Doudoroff pathway (glucose-6-phosphate dehydrogenase, 6-phosphogluconolactonase, phosphogluconate dehydratase, KDPG aldolase) were also significantly overexpressed in the control for at least one time point. Genes encoding enzymes involved in the conversion of pyruvate to acetyl-CoA, including pyruvate dehydrogenase, were all significantly overexpressed in the control at multiple time points (and always at T1).

In contrast, the gene encoding the enzyme fructose-1,6-bisphosphatase, which conducts a key irreversible step in gluconeogenesis, was significantly overexpressed in the DOM treatment at the first 3 time points, indicating the ability to funnel pyruvate toward anabolic rather than catabolic processes. This additional source of glucose was likely used for anabolic purposes as seven genes involved in the generation of ribose-5P and erythrose-4P from glucose via the non-oxidative phase of the pentose phosphate pathway (Fbp, TktAB, TalB, and RpiA) were significantly overexpressed in the DOM treatment. This finding may be of particular importance given the fact that the HIMB30 genome contains lesions in the oxidative portion of the pentose phosphate pathway (Huggett & Rappé 2012). The extra pentoses generated appear to have been used to synthesize aromatic amino acids, particularly tryptophan, as six genes involved in the conversion of erythrose-4P to tryptophan were significantly overexpressed expressed in

the treatment for at least one time point, including genes encoding for the alpha and beta chains of tryptophan synthase (TrpABF).

Additional signals for the use of acetyl-CoA for anabolism by HIMB30 cells in the DOM treatment include significant overexpression of the gene encoding isocitrate lyase, a key enzyme in the glyoxylate cycle, which diverts isocitrate away from the catabolic reactions of the citric acid cycle toward the synthesis of various amino acids and intermediates for several biosynthetic pathways, including gluconeogenesis and carbohydrate synthesis. Six genes involved in the biosynthesis of valine, leucine, and isoleucine from pyruvate and acetyl-CoA were also significantly overexpressed expressed early in the treatment, including *LeuBCD* and *IlvBCE* along with significant overexpression of aminoacyl tRNA synthetases for these three amino acids. In general, genes involved in ABC transport of amino acids exhibited lower expression in the treatment, suggesting a reduced need for exogenous amino acids. Finally, several genes involved in the citric acid cycle exhibited significant overexpression in the control at T1, indicating that additional acetyl-CoA derived from fatty acid oxidation in the DOM treatment was not initially used for catabolic generation of energy by this major pathway. HIMB30 cells in the DOM treatment may have also supplemented their reducing power needs through other means, such as the metabolism of linoleic acid, which generates NADH. At later time points, expression of genes for the conversion of pyruvate to acetyl-CoA and genes involved in the citric acid cycle were generally overexpressed in the DOM treatment, likely due to slowed growth in the control sample later in the experiment (Figure 3).

Collectively, these expression signals indicate the uptake and metabolism of carboxylic acids by HIMB30 cells in the DOM treatment to generate acetyl-CoA for use in anabolic reactions. Putative compound identifications based on ultrahigh-resolution mass spectrometry data indicated the existence of carboxylic acids in the DOM medium, lending further support to the previous physiological interpretations. Low molecular weight carboxylic acids have been shown to comprise up to 20% of DOM produced by strains of *Prochlorococcus* in culture (Bertilsson et al. 2005) and organic acids are likely to be obtained by SPE, so the presence of these compounds in the DOM medium is not necessarily surprising; however here we identify carboxylic acids that are much larger than those previously examined in the study by Bertilsson et al. More than 30 putative carboxylic acids (both saturated and unsaturated fatty acids; ranging from 169-381 m/z) were identified in DOM_P derived from *Prochlorococcus*, including isomers of oleic acid, lauric acid, DHA, and palmitic acid (Table 5). Many of these features were not included in the low resolution fragmentation target list (Table 1) as they were not unique to str. MIT9313, but were still automatically chosen for fragmentation based on their intensities. Putative identifications were further supported using MS/MS fragmentation data when possible, such as for one heptadecanoic acid (Figure 9). While many of these identifications matched C16, C17, C18, and C20 fatty acids that are commonly found in cyanobacteria (Singh et al. 2002), some will acquire additional research to confirm their production by str. MIT9313, including several smaller features that could be improperly identified fragments of larger fatty acids (Table 5).

The apparent abundance of carboxylic acids in the DOM medium along with the gene expression signals mentioned above suggest an important role for these compounds in the cycling of marine DOM. These compounds appear to be easily acquired by heterotrophic marine bacteria via ATP-independent transporters for use in anabolic processes, thereby increasing secondary production and likely bacterial growth efficiency as well. Positive correlations between marine primary production and bacterial growth efficiency have been identified in the natural environment (del Giorgio et al. 1997) and our findings suggest that labile carboxylic acids in the DOM pool could be partly responsible for this observation. Factors affecting phytoplankton contributions of carboxylic acids to the DOM pool may directly affect the growth efficiency of the surrounding heterotrophic community and the overall metabolic balance (i.e. net autotrophy vs. net heterotrophy) of the marine system.

Glucoside and glycolipid metabolism

A suite of gene expression signals related to metabolism of various glucosides were found to be differentially expressed between the control and DOM treatment (Figures 5-8). Several genes involved in the production of phosphatidic acids (PAs), precursor for the biosynthesis of many other glycerolipids, were significantly overexpressed in the DOM treatment at T2, including GlpAK and DgkA (Table 6). DgkA was also significantly overexpressed in the treatment at T1 and encodes a kinase that converts diacylglycerol (DG) to PA. Several genes related to amino sugar metabolism involved in peptidoglycan synthesis were also found to be differentially

expressed between the control and treatment. Two genes involved in the synthesis of the UDP derivatives of peptidoglycan were significantly overexpressed in the control at T1 (GlmU and MiaA), however 5 genes involved in downstream peptidoglycan biosynthesis (MurCEF, DdlAB, FtsI) were significantly overexpressed in the DOM treatment at either T1 or T2. Also, the gene encoding the integral membrane protein responsible for translocating lipid precursors across the cytoplasmic membrane during peptidoglycan synthesis (MviN) was significantly overexpressed in the treatment at T4 and also highly overexpressed (1.88 fold; posterior likelihood 0.78) at T1. This discrepancy in expression signals suggests that HIMB30 cells may have obtained compounds similar to the UDP derivatives of peptidoglycan from the DOM treatment for use in cell wall biosynthesis. Interestingly, one gene (MraY), which liberates uridine monophosphate (UMP) during peptidoglycan synthesis was significantly overexpressed in the control at T3 and T4. It appears that control cells required additional UMP for pyrimidine synthesis as a gene involved in the conversion of UMP to uracil (PyrR) was also significantly overexpressed in the control at T3.

Differential expression was also found for genes involved in inositol phosphate metabolism (Table 6). Five genes involved in the conversion of inositol phosphates to glyceraldehyde-3P (SuhB, IolBCD, TpiA) were found to be significantly overexpressed in the DOM treatment for at least one time point, most often at T4. The SuhB gene, which was significantly overexpressed in the treatment at T4, encodes for a phosphatase that cleaves phosphate groups from inositol phosphates to produce inositol, liberating inorganic phosphorus for use in other processes. Interestingly, the malonate

semialdehyde produced during this process was not used by HIMB30 cells in the DOM treatment to produce acetyl-CoA as much as the control cells are as the gene that encodes this step (*IolA*) was significantly overexpressed in the control at multiple time points (T1 and T5). This observation lends further support to the hypothesis that the DOM treatment and control cells generated acetyl-CoA via disparate processes.

Aside from one sugar ABC transporter system likely involved in the transport of the simple sugar dihydroxyacetone (DHA) based on gene synteny with DHA kinase, all sugar ABC transporters detected were significantly overexpressed for at least one time point in the DOM treatment (Table 6). One such transporter was significantly overexpressed in the treatment at T2, T3 and T4 and linked to a putative murein hydrolase gene (*LrgA*), hypothesized to be involved in the synthesis of peptidoglycan (Brunskill & Bayles 1996). The only sugar ABC transporter that was significantly overexpressed in the treatment at all time points was linked to a gene coding for an epimerase involved in the production of N-acyl-D-glucosamine, a major component of peptidoglycan. Another sugar ABC transporter significantly overexpressed in the treatment at the final 3 time points was linked to a gene encoding a glycerol kinase involved in glycerolipid metabolism. At the final time point, statistically significant overexpression of a sugar ABC transporter was found linked to a gene encoding a transport ATP-binding protein for glycerol-3-phosphate (*UgpC*), a component of glycerophospholipids. A gene encoding an extracellular solute-binding protein with synteny to a glycerol-3-phosphate dehydrogenase gene (*GlpA*) was also found to be significantly overexpressed in the treatment at all time points. Collectively, these

transport signals indicate that HIMB30 cells in the DOM treatment acquired sugars associated with cell membrane and cell wall components and the expression signals discussed above suggest that some of these compounds were utilized to synthesize peptidoglycan, while others were broken down to glyceraldehyde-3P for use in gluconeogenesis and pentose phosphate pathways.

Putative compound identifications in the DOM_P sample revealed a number of glycerolipids and glucosides that may support the expression signals discussed above (Table 7). Many of these features were not included in the low resolution fragmentation target list (Table 1) as they were not unique to str. MIT9313, but were still automatically chosen for fragmentation based on their intensities. Among the most intense features in the culture DOM_P sample that were absent in the Pro99 medium control was a compound putatively identified as a PA (1-(9Z-tetradecenoyl)-glycero-3-phosphate) and a compound putatively identified as a phosphatidylcholine (1-hexyl-2-acetyl-sn-glycero-3-phosphocholine). Four additional features were putatively identified as either 1,2-Dioctanoyl-sn-glycerol or 1-dodecanoyl-sn-glycerol, which are both known precursors for the synthesis of PA. Several features with putative identifications related to cell wall material were also identified. These included 2-aminoheptanedioic acid, which closely resembles meso-diaminoheptanedioate, a compound involved in lysine biosynthesis and a precursor for peptidoglycan, and a N-acetyl amino acid amide (NAc-LLY-amide), which closely resembles an unphosphorylated version of the UDP derivatives of peptidoglycan. Seven putative identifications were also made for larger glucosides (> 500 *m/z*), all of which contain one or multiple inositol groups. Two of these glucosides were

glucuronides and one was a glucopyranosyl, which all contain aminoheptanedioic acid in addition to an inositol group. Putative identifications were further supported using fragmentation data when possible, such as for one phosphatidylinositol (Figure 10).

Although astaxanthin is not known to be a major carotenoid produced by *Prochlorococcus* (Cailliau et al. 1996), the putative identification here of two astaxanthin diglucosides in DOM produced by *Prochlorococcus* str. MIT9313 may indicate the ability of this strain to synthesize astaxanthins. While the genome of MIT9313 does contain a gene encoding a beta-carotene hydroxylase (CrtR) to synthesize zeaxanthin, it appears to be lacking the beta-carotene ketolase (CrtW) necessary to produce astaxanthin. However, a gene encoding a putative beta-carotene ketolase (PMT1558) has been identified in MIT9313 by the comparative genomics project SEED (<http://www.theseed.org>), suggesting that biosynthesis of astaxanthins by MIT9313 is plausible.

Taken together, this chemical and physiological information indicates that *Prochlorococcus*-derived DOM constituted a source of labile glycerol, inositol, phosphate, and amino sugars to HIMB30 cells, which were then taken up and used in the biosynthesis of cell membrane and cell wall material as well as the production of compounds like glyceraldehyde-3P for use in glycolysis and gluconeogenesis. Our findings suggest that constituents of cell membranes and cell walls may be important components of labile DOM as they appear to be easily transported into cells and inserted into various pathways for the biosynthesis of these essential cell structures. As MIT9313 cells have been shown to contain more cell wall material than related cells with smaller

genomes such as *Prochlorococcus* str. MED4 (Ting et al. 2007), the abundance of these constituents in the DOM pool may vary according to the genome size of the dominant phytoplankton in a particular region with concurrent effects on the surrounding heterotrophic community.

Motility and chemotaxis

Genes encoding bacterial motility proteins related to flagellar assembly and chemotaxis were among the most differentially expressed genes detected in our study (Figures 5-8). At T1, 11 such genes were significantly overexpressed in the DOM treatment while none were found to be significantly overexpressed in the control (Table 8). After T1, however the opposite trend was strikingly clear. At T2, 19 motility-related genes, including assembly genes for the flagellar basal body, M ring, P ring, L ring, hook, filament, filament cap, motor, and motor switch were all found to be significantly overexpressed in the control, with this number rising to 24 total genes at T3 and T4 (Table 8). After T1, only 1 gene (FlhG) related to motility was significantly overexpressed in the DOM treatment (at T3) and it encodes a protein with a region classified as an antiactivator of flagellar biosynthesis. While this finding appears to be at odds with previous research indicating that increases in organic nutrients generally correspond to increases in the percentage of motile cells in marine systems (Mitchell et al. 1995, Grossart et al. 2001), it can likely be explained by regulation via the OmpR family two-component system. Phosphorylation of OmpR, which occurs both under high osmolarity conditions and due to high intracellular levels of acetyl phosphate (AcP), has

been shown to repress transcription of flagellar biosynthesis genes in *E. coli* (Shin & Park 1995, Klein et al. 2007). Wild-type *E. coli* cells have been found to be less motile in early exponential phase when intracellular concentrations of AcP are high, compared to stationary phase when AcP levels are lower (Prüss & Wolfe 1994). Although intracellular AcP levels for HIMB30 were not measured in our study, control cells increased transcription of genes related to flagellar motility in later stages of growth relative to treatment cells, suggesting that intracellular AcP levels may have remained high in the treatment for an extended period of time. Two genes in the OmpR family (PhoR and PhoB) were found to be significantly overexpressed in the treatment at either T3 or T4, indicating the involvement of OmpR regulation at later stages of growth in the DOM medium.

While we cannot conclusively determine the cause of OmpR phosphorylation in HIMB30 treatments cells, increased osmotic stress, increased intracellular AcP levels derived from organophosphates in the DOM addition, and catabolic repression due to the presence of labile DOM substrates are all reasonable possibilities. Significant overexpression of ProX, a gene involved in the transport of the common cyanobacterial osmoticum glycine betaine (Reed et al. 1984, WARR et al. 1984), in the treatment at T3 and T4 and the presence of putative organophosphates and glucosides in the DOM treatment suggest that all of the above possibilities may be contributing factors.

Whatever the cause, this finding has significant implications for understanding the relationship between DOM and bacterial motility. While increases in DOM may help to alleviate the high energy costs associated with motility (Grossart et al. 2001), there

appear to be counteracting affects that may ultimately lead to reduced motility for certain marine bacteria under increased organic nutrient conditions. Motility affects due to the addition of phytoplankton derived DOM as opposed to more traditional defined sources (e.g. peptone, glucose, tryptic soy broth) may be more representative of the marine environment, therefore DOM-related factors that repress transcription of flagellar biosynthesis genes in marine bacteria warrant additional investigation. Decreased motility in the presence of labile DOM substrates may be an adaptation allowing marine bacteria to stay within transient DOM patches (Stocker 2012), while cells in more dilute microenvironments must increase swimming motility in an effort to locate a DOM patch.

Proteorhodopsin-based phototrophy

HIMB30 contains genes related to proteorhodopsin-based phototrophy, many of which were found to be statistically differentially expressed between the control and DOM treatment conditions (Figures 5-8). The proteorhodopsin photoprotein (PR) is a retinal bound light-driven proton pump that has now been found in a multitude of marine bacteria and archaea (Giovannoni et al. 2005, McCarren & DeLong 2007, DeLong & Béjà 2010) and several studies have recently demonstrated a light-dependent growth advantage for PR-containing marine organisms (Gómez-Consarnau et al. 2007, 2010, Kimura et al. 2011), likely due to the synthesis of ATP and enhancement of sodium ion gradients related to transport functions (Martinez et al. 2007, Kimura et al. 2011).

The HIMB30 gene encoding PR exhibited some of the highest expression levels of any gene detected in this study and was found to be significantly more expressed in the

DOM treatment at all time points (Table 9). Five genes involved in retinal biosynthesis were found to also be statistically overexpressed in the treatment at either T2, T3, or T4, including a retinal biosynthesis gene (CrtE) found to be significantly overexpressed by a marine flavobacterium when grown in the light as described by Kimura et al. (2011). An additional farnesyltransferase (CyoE) likely involved in retinal biosynthesis was also significantly overexpressed in the treatment at T2. The HIMB30 PR gene was also found to be linked to a gene significantly overexpressed in the treatment at T2, T3, and T4 and annotated as a polyprenyl synthetase, which are enzymes involved in the biosynthesis of terpenoids such as farnesyl diphosphate. The overexpression of this gene in the treatment and its synteny with PR suggest polyprenyl synthetase is also an important enzyme for the biosynthesis of retinal by HIMB30.

Several additional genes involved in light absorption were also found to be more expressed in the treatment at T1, T2 and T4, including genes encoding a deoxyribodipyrimidine photolyase and a multi-sensor hybrid histidine kinase. These two genes were also found to be similarly overexpressed under light conditions in the Kimura study. Two genes encoding quinone oxidoreductases similar to those described by Kimura et al. were also detected as significantly overexpressed in the DOM treatment at T3 or T4 along with two genes encoding pyrroloquinoline quinone biosynthesis proteins (PqqBC) significantly overexpressed at T1 and T4. However, most of these DNA photolyase and quinone oxidoreductase genes, as well as a few additional genes, were all found to be significantly overexpressed in the control at T3, which could relate to

growth-phase variability as cells in the control were entering stationary phase by T3, while treatment cells were still replicating exponentially.

Three genes involved in solute transport, one annotated as a putative sodium symporter and another as a Na^+/H^+ antiporter, were significantly overexpressed in the DOM treatment along with these PR-related signals at T2, T3, and T4, supporting the idea proposed by Kimura et al. that Na^+ ion gradients are an important component of PR-based phototrophy, though one of these genes was also significantly overexpressed in the control at T1 (Table 9). We identified the same six Na^+ -translocating NADH-quinone oxidoreductase genes (NqrABCDEF) as Kimura et al., and four of the five genes they found to be significantly overexpressed in the light were also overexpressed here in the DOM treatment (NqrABCF). Unlike the study by Kimura et al., we also detected significant overexpression of genes related to ATP synthase. Five genes encoding various subunits of ATP synthase (AtpADEFH) were significantly overexpressed in the control at T1. Four of these genes (AtpADFH) were then found to be significantly overexpressed in the DOM treatment by T3, along with three additional ATP synthase related genes (AtpGI) and one additional gene overexpressed at T4 (AtpC), indicating increased ATP production via ATP synthase in the DOM treatment relative to the control at later growth stages. Interestingly, the overexpression of these ATP synthase genes in the treatment at T3 corresponds to an overexpression of genes related to DNA photolyase activity in the control as previously mentioned, perhaps indicating a temporal offset in light absorption PR-driven ATP synthesis between the control and treatment. It is also important to note that genes encoding for the b, c1, and iron-sulfur subunits of the

cytochrome bc₁ complex (complex III in the electron transport chain) followed a similar pattern of significant differential expression (i.e. overexpressed at T1 in control and at T3 in the treatment), indicating that the potential increase in ATP synthase activity in the treatment at T3 and T4 was likely due to a combination of photo- and organoheterotrophic activity and it is not possible to discern which of the ATP synthase signals are truly PR-related. However, the constant significant overexpression of the PR gene in the treatment at all time points and accompanied overexpression of retinal biosynthesis genes, suggests a potential increase in PR-based phototrophy due to DOM addition.

Putative identifications were made for several compounds containing isoprene units in the DOM used to prepare treatment medium that were absent in Pro99 medium, including several compounds resembling carotenoid derivatives as well as retinal (Table 10). Many of these features were not included in the low resolution fragmentation target list (Table 1) as they were not unique to str. MIT9313, but were still automatically chosen for fragmentation based on their intensities. Seven features detected in DOM_P derived from *Prochlorococcus* MIT9313 were putatively identified as vitamin D₃ derivatives, isoprenoids synthesized from triterpenoids. A terpenoid-quinone compound was also putatively identified, which is synthesized from four isoprene units. Farnesyl phosphate and two compounds resembling unphosphorylated versions of geranyl diphosphate were also identified. Both farnesyl phosphate and geranyl diphosphate are important terpenoids involved in the biosynthesis of carotenoids, including β -carotene. Finally, we identified three retinoic acids (RAs): retinoyl glucuronide (RAG), All-trans-5,6-epoxy-

RA, and All-*trans*-4-oxo-RA, the last of which is identical to retinal except for having a carboxylic acid group in place of the aldehyde group. Putative identifications were further supported using MS/MS fragmentation data when possible, such as for one of the RAs (Figure 11).

Several RAs, including the All-*trans*-4-oxo-RA detected in this study, were recently shown to be produced and released in significant quantities (ng/L) by freshwater cyanobacteria and a marine *Synechococcus* strain in culture, however the function of these compounds in cyanobacteria is currently unknown (Wu et al. 2012). A strain of the cyanobacterium *Synechocystis* has been shown to synthesize retinal from β -apo-carotenals instead of β -carotene using a previously unknown carotenoid oxygenase (Ruch et al. 2004), the structure of which was later determined (Kloer 2005). While genes involved in traditional retinal biosynthesis pathways have not been identified thus far for *Prochlorococcus*, there is some evidence that other green-light photoreceptors, such as β phycoerythrin, are produced by certain *Prochlorococcus* strains (Steglich et al. 2005). A survey of the current annotated *Prochlorococcus* genomes revealed that many *Prochlorococcus* strains, including strain MIT9313, contain a gene annotated as retinal pigment epithelial membrane protein (PMT1879), the expression of which has been shown to be light-induced in *Prochlorococcus* str. MED4 (Steglich et al. 2006). This protein contains a lignostilbene- α,β -dioxygenase (LSD) region, which is typically thought to be involved in the degradation of lignin to vanillin. Pathways for lignin degradation are not common in cyanobacteria however, and the novel carotenoid oxygenase involved in retinal production discovered in *Synechocystis* was originally annotated as a LSD

based on gene homology (Ruch et al. 2004), suggesting that PMT1879 in *Prochlorococcus* may also be involved in the cleavage of carotenoids to produce retinal in a manner similar to the enzyme described for *Synechocystis*. PMT1879 could help explain the presence of the putative RAs detected here in DOM produced by *Prochlorococcus* str. MIT9313.

Our findings that HIMB30 expression of PR and related genes involved in the synthesis of retinal and other light associated proteins was significantly higher in the DOM treatment may initially seem contrary to previous reports of PR providing a greater advantage to cells under low organic nutrient conditions for certain microbial systems (Gómez-Consarnau et al. 2007, Kimura et al. 2011). However, the conclusions that PR benefits diminish at higher DOC concentrations needs to first be placed in environmental context. The DOC concentration reported by Gómez-Consarnau et al. (2007) at which light-stimulated PR benefits were no longer apparent for a marine flavobacterium (242 mM C) is >3 orders of magnitude higher than typical values for even the most productive coastal regions (Hansell 1993) and likely toxic to many oligotrophic marine bacteria. Increases in bacterial cell yields in light vs. dark cultures have been reported to decrease from ca. 4-fold to ca. 2.5-fold when DOC is increased between 0.35 – 0.96 mM C (Gómez-Consarnau et al. 2007) and growth rate increases in light vs. dark cultures have been reported to decrease from ca. 1.6-fold to 1.2-fold when DOC was increased by 0.25 mM C (Kimura et al. 2011) in cultures of PR-containing marine flavobacteria. The magnitude of these DOC increases are much greater than reported fluctuations in the marine environment, especially in the oligotrophic ocean (Hansell & Carlson 1998, Avril

2002, Church et al. 2002), and an order of magnitude larger than the DOC difference between the control and treatment examined in this study (32 $\mu\text{M C}$).

Our results indicate that environmentally relevant increases in DOM from marine phytoplankton may in fact enhance PR expression due to the uptake and utilization of compounds related to carotenoid and retinal biosynthesis. DOC increases in the aforementioned studies were done using peptone and yeast extract and would therefore not contain the terpenoids or polyenes putatively identified here in *Prochlorococcus*-derived DOM (Table 10). As several of the putatively identified compounds are not known to be produced by *Prochlorococcus*, it will be important to conduct further experiments in order to confirm the production of these compounds by MIT9313. These findings also highlight the need for more studies that test for PR growth advantages under environmentally relevant conditions, which may help elucidate discrepancies in laboratory experiments that test for light-stimulated growth related to PR in other microbial systems (Giovannoni et al. 2005, Stingl et al. 2007, Lami et al. 2009, Gómez-Consarnau et al. 2010). Furthermore, the idea that *Prochlorococcus* and other cyanobacteria may produce retinoids that could enter the DOM pool and be utilized by photoheterotrophic bacteria represents a new finding that should be further investigated.

HIMB30 physiology and growth phase

In addition to signals related to the DOM treatment, genes with consistently higher or lower expression during different phases of growth in both the DOM treatment and control media were also observed. Relative expression levels were averaged for the

first two control time points to obtain mean relative expression of genes in early growth phase, while relative expression levels were averaged for the last two control time points to obtain mean relative expression of genes in late growth phase.

Genes with higher expression during early growth phase included several involved in pyruvate metabolism, including genes that encode malate dehydrogenase (MaeB) and lactate dehydrogenase (LldD), which respectively convert malate and lactate to pyruvate, which can then be used to synthesis glucose or fatty acids or generate energy via the TCA cycle. Genes encoding the E1 and E2 components of the pyruvate dehydrogenase enzyme (AceEF), which converts pyruvate to acetyl-CoA were also overexpressed early, suggesting a focus on energy generation via the TCA cycle during exponential growth. Many genes encoding ATP-binding cassette (ABC) transporters were also up-regulated during early growth including 3 involved in iron transport (AfuABC) along with an iron siderophore transporter (K02016). ABC transport of amino acids, particularly branched-chain amino acids (LivMFGHK) and a putative glutamine transporter (K10040) were also overexpressed early, along with transporters for sulfonates/nitrate (SsuCB) and glycine betaine (ProV). Collectively, these signals indicate the central importance of nutrient acquisition, in particular nitrogen and iron for HIMB30 cells during exponential growth.

Genes potentially involved in carbon fixation via the Calvin Benson Bassham cycle were also among the most overexpressed genes during early growth, including genes coding for the small and large chain ribulose-bisphosphate carboxylase (RuBisCo) enzyme (RbcSL), which produces of 2 molecules of glycerate 3-phosphate from carbon

dioxide (CO₂) and ribose 5-phosphate. This increase in RuBisCo gene expression may have been related to the oxygenase activity of RuBisCo, and does not necessarily indicate carbon fixation by HIMB30. Several genes involved in the regeneration of ribose 5-phosphate were also overexpressed early including phosphoribulokinase (PrkB), fructose-1,6-bisphosphatase (Fbp), fructose-bisphosphate aldolase (FbaA) and transketolases (TktAB). The net generation of glycerate 3-phosphate can be used in many central metabolic pathways including gluconeogenesis or biosynthesis of amino acids such as phenylalanine, tyrosine and tryptophan. Notably, 2 genes (AroA, AroI), which encode for enzymes that convert erythrose 4-phosphate (another potential product of carbon fixation) to 3-dehydroquinate for use in biosynthesis of these 3 amino acids were also up-regulated during early growth. Finally, several genes involved in early steps for the biosynthesis of pyridoxine (vitamin B₆; an important cofactor in the biosynthesis of many compounds including amino acids, glucose, and sphingolipids) from erythrose 4-phosphate were overexpressed during early growth (PdxAB, ThrC), indicating an important role for this vitamin in facilitating exponential growth of HIMB30.

Genes with higher expression during later stages of growth included transcripts involved in DNA mismatch repair (HolCB, XseB, MutS, and single-strand DNA-binding protein ssb), homologous recombination (RecOF), and RNA degradation (DnaK, GroEL, VacB), demonstrating that HIMB30 cells are more focused on nucleic acid maintenance when in stationary phase. Genes involved in nucleotide excision repair were also overexpressed (UvrBC) at later time points, suggesting the presence of damaged DNA. There is indication that this damage could be due to oxidative stress as genes encoding a

manganese containing superoxide dismutase (SOD2) and a glutathione peroxidase (K00432), both of which are known to reduce lipid hydroperoxides (Zhang et al. 2009), are up-regulated later in growth. Another apparent focus during later growth stages is on cell defense as exhibit by increased expression of genes for transport of antibiotics (YadG) and the biosynthesis of ansamycins (TktAB) and vancomycin group antibiotics (RfbB, RffG).

Genes related to proteorhodopsin based phototrophy, including the gene encoding the HIMB30 proteorhodopsin photoprotein as well as genes involved in carotenoid (CrtBP), terpenoid backbone (IDI, DXS, LytB), and terpenoid-quinone (HppD, UbiAEH) biosynthesis also exhibited increased expression at later time points. Collectively, these signals suggest that HIMB30 cells may rely more heavily on photoheterotrophic generation of energy during later stages of growth as organic substrates and nutrients become depleted. Genes involved in the synthesis of other porphyrins (HemCFHX) were also overexpressed at later time points. Enzymes encoded by these genes convert amino acids like glycine and glutamate into protoheme that may be involved in both cytochrome activity related to proteorhodopsin and/or peroxidase activity to combat oxidative stress as discussed above.

Finally, genes involved in both cell wall and outer membrane biosynthesis were found to be overexpressed at later stages of growth. A multitude of genes that code for amino acid ligases, GTases and other enzymes involved in the synthesis of precursors for peptidoglycan synthesis (DapABE, MurCEFG) as well as peptidoglycan glycosyltransferase were all overexpressed in stationary phase growth. Several genes

involved in biosynthesis of lipopolysaccharide, particularly lipid IV_a (LpxKL, KdsBC, KdtA) were also overexpressed along with an outer membrane channel protein (TolC) and a lipopolysaccharide export system permease protein (LptG). HIMB30 cells in stationary phase growth appear to invest more resources into maintaining cellular structural integrity and may also be related to defense. It is likely these two signals are coupled as outer membrane proteins have been shown to regulate peptidoglycan synthesis and while the average glycan chain length of *E. coli* peptidoglycan is known to decrease in stationary-phase cells, the number of crosslinkages and attachments to lipoprotein increase (Typas et al. 2011).

Highly expressed tmRNA

One small (333 bp) read was found to be expressed at a level that was almost an order of magnitude greater than any other gene detected, with consistently high expression at all time points. This read was part of a non-coding region of contig 28 according to the annotated HIMB30 genome on NCBI. Further investigation revealed this short transcript to be a putative transfer-messenger RNA (tmRNA) based on homology to other tmRNAs and gene synteny within the HIMB30 genome. The main function of tmRNA is to repair ribosomes that have stalled during translation in a process known as *trans*-translation (Keiler 2008). Small RNA (sRNA) reads, including tmRNAs, have been shown to constitute a large proportion of metatranscriptomic datasets from marine microbial communities where they exhibit extremely high levels of expression when compared to protein-coding genes due to their important regulatory functions (Shi

et al. 2009). In the study by Shi et al., the most common group of putative sRNAs identified was connected primarily to gammaproteobacteria and contained a conserved sequence motif indicative of regulation at the translation level, which could be indicative of tmRNA activity. tmRNAs have also been found to be the dominant sRNA family expressed in pure culture of a marine flavobacterium (Kimura et al. 2011). Our findings support the idea that sRNAs are likely also important regulatory elements during growth in vitro and highlight the potential of using RNA-seq and cultured isolates to identify and functionally validate sRNAs of interest for individual microbial taxa.

Conclusions

Using RNA-seq to evaluate physiology at the level of gene expression and ultrahigh-resolution mass spectrometry to make putative compound identifications, we identified several likely factors responsible for the enhanced growth response of HIMB30 to the presence of MIT9313 DOM_p. Transcriptomic profiling of HIMB30 in multiple phases of growth revealed a suite of metabolic responses related to growth phase and to the addition of DOM, highlighting metabolic pathways likely involved in the uptake and utilization of labile DOM substrates. These included the transport and utilization of carboxylic acids and glucosides that were primarily inserted into anabolic pathways including gluconeogenesis, pentose phosphate, and the glyoxylate cycle and ultimately used to enhance biosynthesis of cell membrane and cell wall material.

Ultrahigh-resolution mass spectrometry data coupled to targeted MS/MS fragmentation information provided putative identifications for several compounds

present in *Prochlorococcus*-derived DOM, yet absent from Pro99 seawater-based medium, that were used to aid in interpretation of transcriptomic data, including isomers of many carboxylic acids (e.g. oleic acid, lauric acid, palmitic acid), glucosides, glycerolipids, and amino sugars. As these putative identifications were limited by the contents of the METLIN database and occasionally matched compounds that are not currently known to be produced by *Prochlorococcus*, it is essential that future work be devoted to confirming the production of these metabolites by strain MIT9313. Overall however, our findings reveal specific metabolic information regarding how labile constituents in the DOM pool increase the growth efficiency of heterotrophic bacteria.

In addition to these substrate responses, gene expression data also indicated that the *Prochlorococcus*-derived DOM addition might have resulted in decreased motility and increased proteorhodopsin-based heterotrophy. Further investigation into these signals and support from putative chemical composition data revealed likely OmpR regulated suppression of flagellar assembly genes and increased biosynthesis of carotenoids and retinal, potentially due to the uptake of terpenoids and retinoids present in the DOM treatment. Both of these findings challenge current paradigms established for other marine microbial systems regarding the effects of organic nutrient concentrations on heterotrophic motility and phototrophic activity and highlight the need for additional studies on interactions between heterotrophic microbes and various DOM sources. It is also important to note, that trace metals and other material present in *Prochlorococcus* str. MIT9313 DOM for which putative identification were not possible may have relieved specific inorganic nutrient limitations in HIMB30 cells, which could

have contributed to the increased growth signals and would be independent of the organic carbon uptake signals discussed here.

By combining next generation RNA-seq data with ultrahigh-resolution mass spectral data, we were able to go beyond the simple observation that DOM produced by *Prochlorococcus* str. MIT9313 enhanced the growth of HIMB30 and investigate *why* this observation may have occurred through identification of the metabolic processes and chemical compounds involved. Our results highlight the potential for linking detailed chemical analysis of DOM from a defined biological source with physiological information regarding its metabolism by individual taxa derived from gene expression analyses. This holistic approach can provide information regarding the production, composition, and degradation of labile DOM and generate new hypotheses regarding DOM-mediated interactions between autotrophic and heterotrophic microbes in the sea.

Acknowledgements

Thanks M. Rappé, O. Sosa, E. Omori, and J. Grote for help with culturing and P. Berube and S. Chisholm for providing the *Prochlorococcus* culture. Also thanks to J. Eppley, A. Sharma, E. Ottesen, S. Gifford, J. Bryant, and T. Palden for assistance with RNA sequencing and analysis and M. Soule for mass spectrometry support. Funding for this work was provided by the Gordon and Betty Moore Foundation (#1711 to D.J.R.) and National Science Foundation Awards #DBI0424599 and #EF0424599 to D.J.R and E.F.D.

Table 1. Features detected in DOM produced and released by *Prochlorococcus* str. MIT9313 using low-resolution mass spectrometry (MS) that were then targeted for MS/MS fragmentation during ultrahigh-resolution mass spectrometry performed on the same sample. Details regarding low-resolution MS can be found in chapter 2. Targets were chosen based on intensity fold change between the MIT9313 culture sample and control medium (Pro99) sample and also by their uniqueness to str. MIT9313 when compared to other phytoplankton strains. Features with a large fold change indicate material that was produced by MIT9313 and not present in the background natural seawater-based Pro99 medium.

<i>m/z</i>	Retention Time (min)	MIT9313 Intensity	Pro99 Intensity	Fold Increase
267.2	43.2	2.76E+06	1.10E+04	249.9
711.3	42.9	8.00E+05	5.10E+03	156.6
268.2	43.2	5.50E+05	3.72E+03	147.9
265.2	41.9	1.55E+06	1.54E+04	100.5
712.3	42.9	3.69E+05	4.40E+03	84.0
266.2	41.9	2.91E+05	4.07E+03	71.5
268.2	39.5	7.38E+05	1.47E+04	50.1
235.2	43.2	2.30E+05	4.60E+03	50.1
420.2	33.3	6.34E+05	1.30E+04	48.7
267.2	39.5	3.90E+06	8.08E+04	48.3
721.4	48.6	1.47E+05	3.20E+03	46.1
265.2	42.3	3.76E+05	9.28E+03	40.6
270.2	41.1	8.46E+05	2.37E+04	35.7
284.2	35.7	2.39E+05	6.89E+03	34.6
653.2	40.4	1.97E+05	5.69E+03	34.6
269.2	41.1	4.56E+06	1.33E+05	34.3
273.2	36.4	2.11E+05	6.57E+03	32.1
434.3	45.2	1.61E+05	5.06E+03	31.8
667.2	41.9	1.78E+05	5.63E+03	31.6
769.3	42.6	1.14E+05	3.70E+03	30.9
774.2	35.9	1.22E+05	4.11E+03	29.7
333.2	34.1	3.57E+05	1.25E+04	28.7
229.2	35.6	5.29E+05	1.86E+04	28.5
713.3	43.0	1.07E+05	3.97E+03	26.9
447.3	36.7	2.33E+05	1.02E+04	22.8
268.2	35.6	1.59E+05	7.76E+03	20.5
230.2	35.5	8.25E+04	4.26E+03	19.4
321.2	43.2	2.38E+05	1.24E+04	19.1
393.2	30.9	1.89E+05	9.93E+03	19.0
349.2	27.3	1.51E+05	8.26E+03	18.3
241.2	38.7	2.32E+05	1.34E+04	17.4
403.2	41.4	1.51E+05	8.71E+03	17.3
287.2	40.4	1.96E+05	1.13E+04	17.3
776.2	36.0	6.85E+04	4.15E+03	16.5
347.2	27.0	1.20E+05	7.30E+03	16.5
313.2	31.8	2.11E+05	1.33E+04	15.9
237.2	41.1	1.46E+05	9.48E+03	15.4
421.2	33.3	1.63E+05	1.10E+04	14.8
217.1	39.6	6.06E+04	4.15E+03	14.6
461.2	37.0	1.15E+05	8.05E+03	14.3
331.2	34.0	2.03E+05	1.44E+04	14.1

Table 1 continued.

332.2	38.8	1.09E+05	7.95E+03	13.7
331.2	28.8	1.14E+05	8.41E+03	13.6
331.2	28.6	1.08E+05	8.00E+03	13.5
361.2	32.1	1.26E+05	1.02E+04	12.4
264.2	25.9	7.95E+04	6.42E+03	12.4
668.2	41.9	7.80E+04	6.31E+03	12.4
249.2	35.7	1.08E+05	8.83E+03	12.3
333.2	33.7	1.69E+05	1.38E+04	12.2
321.2	25.9	9.01E+04	7.45E+03	12.1
350.2	34.3	8.07E+04	6.75E+03	11.9
375.1	24.2	9.39E+04	7.91E+03	11.9
201.1	31.1	8.77E+04	7.42E+03	11.8
331.2	27.9	9.21E+04	7.91E+03	11.6
832.2	55.3	3.57E+05	3.08E+04	11.6
607.2	45.6	5.64E+04	4.88E+03	11.6
274.2	38.1	7.80E+04	6.88E+03	11.3
315.2	34.1	2.60E+05	2.31E+04	11.2
833.1	55.3	3.08E+05	2.78E+04	11.1
303.2	35.7	1.43E+05	1.30E+04	11.0
834.2	55.3	2.70E+05	2.47E+04	10.9
257.2	38.1	3.05E+05	2.79E+04	10.9
259.2	39.5	7.87E+04	7.25E+03	10.8
306.2	29.8	9.35E+04	8.92E+03	10.5
716.3	25.1	6.67E+04	6.39E+03	10.4
237.1	28.3	1.03E+05	9.91E+03	10.4
267.2	36.4	8.67E+05	8.39E+04	10.3
315.2	34.7	1.62E+05	1.58E+04	10.3
331.2	39.6	1.18E+05	1.20E+04	9.8
209.1	25.6	9.41E+04	1.02E+04	9.3
320.2	35.7	6.36E+04	6.89E+03	9.2
303.2	40.1	8.03E+04	9.44E+03	8.5
235.1	39.5	3.02E+05	3.71E+04	8.1
286.2	39.3	1.25E+05	1.55E+04	8.1
181.1	27.3	2.11E+05	2.77E+04	7.6
229.2	34.4	1.91E+05	2.55E+04	7.5
236.1	26.1	1.24E+05	1.84E+04	6.7
391.2	30.7	8.16E+04	1.22E+04	6.7
379.2	25.3	5.22E+04	7.82E+03	6.7
241.2	38.3	2.30E+05	3.83E+04	6.0
334.2	34.2	9.33E+04	1.59E+04	5.9
347.2	34.4	1.21E+05	2.11E+04	5.7
216.1	26.9	5.92E+04	1.08E+04	5.5
610.1	49.4	7.09E+05	1.31E+05	5.4

Table 1 continued.

760.2	52.8	3.37E+05	6.48E+04	5.2
197.1	20.4	3.00E+05	5.80E+04	5.2
230.1	29.6	5.22E+05	1.02E+05	5.1
347.2	37.6	9.67E+04	1.90E+04	5.1
255.2	36.5	1.34E+05	2.65E+04	5.1
759.2	52.7	3.86E+05	8.11E+04	4.8
758.1	52.7	5.08E+05	1.08E+05	4.7
213.2	36.1	8.37E+04	1.87E+04	4.5
315.2	33.6	1.17E+05	2.71E+04	4.3
265.2	38.5	1.14E+05	3.08E+04	3.7
611.1	49.4	2.72E+05	7.46E+04	3.7
214.1	20.4	6.87E+04	1.94E+04	3.5
329.2	34.5	7.96E+04	2.28E+04	3.5
225.2	38.0	6.85E+04	1.96E+04	3.5
612.1	49.4	2.08E+05	5.98E+04	3.5
273.2	39.3	5.83E+05	1.68E+05	3.5
314.2	42.0	8.49E+04	2.55E+04	3.3
241.2	39.4	1.85E+05	5.57E+04	3.3
241.2	39.4	1.81E+05	5.57E+04	3.3
685.1	50.8	3.84E+05	1.22E+05	3.1
684.1	50.8	5.68E+05	1.82E+05	3.1
274.2	39.3	1.01E+05	3.23E+04	3.1
171.2	22.7	2.04E+05	6.61E+04	3.1
255.2	38.6	1.14E+05	3.71E+04	3.1
686.1	50.8	3.05E+05	9.95E+04	3.1
290.2	39.3	1.16E+05	3.85E+04	3.0
413.2	45.8	2.90E+05	9.77E+04	3.0
301.2	42.2	2.24E+05	7.55E+04	3.0
301.2	37.3	9.02E+04	3.15E+04	2.9
297.2	46.4	1.19E+05	4.28E+04	2.8
313.2	32.7	2.05E+05	7.48E+04	2.7
243.2	45.5	1.37E+05	5.10E+04	2.7
283.2	39.4	1.04E+05	3.94E+04	2.7
580.3	37.2	1.89E+05	7.34E+04	2.6
550.5	45.7	1.22E+05	4.79E+04	2.6
371.2	18.6	1.54E+05	6.07E+04	2.5

Table 2. RNA read information for each transcriptome sample presented in this study. Growth phase determination based off figure 3. Total reads are reported after initial quality control of raw reads. rRNA content was reduced in each sample prior to sequencing using a subtraction hybridization protocol described in the methods and %rRNA indicates the proportion of rRNA remaining after subtraction determined in silico. The final column indicates the number of different coding DNA sequences (CDSs) detected in this study out of the possible 2,227 annotated CDSs for the HIMB30 genome on NCBI.

Sample	Growth Phase	Total RNA Reads	Non-rRNA Reads	% rRNA	Protein hits (out of 2227)
Control T1	Exp.	551,538	374,544	32.1	2,213
DOM T1	Exp.	544,961	541,167	0.7	2,219
Control T2	Late exp.	476,460	472,944	0.74	2,215
DOM T2	Late exp.	569,242	566,016	0.57	2,217
Control T3	Early stat.	460,309	456,021	0.93	2,213
DOM T3	Early stat.	486,220	480,226	1.23	2,212
Control T4	Stat.	442,616	435,519	1.6	2,215
DOM T4	Stat.	401,682	398,169	0.87	2,209

Table 3. The number of reads found to be significantly overexpressed in either the DOM treatment (*Up*) or the control (*Down*) at each of the four time points. Significant expression was determined using empirical Bayesian methods to estimate posterior likelihoods of differential expression (Hardcastle 2010). A posterior likelihood ≥ 0.9 was defined as significant.

Time Point	Up	Down	Total
T1	187	255	442
T2	149	140	289
T3	237	337	574
T4	261	185	446

Table 4. Read numbers and ratios of expression in the DOM treatment compared to the control medium (*DOM/Ctrl*) for coding DNA sequences (CDSs) associated with carboxylic acid metabolism and potentially related anabolic pathways. *DOM/Ctrl* ratios are derived from read numbers for each sample that are normalized to the total number of CDSs in each sample. Posterior likelihoods ≥ 0.9 indicate significant differential expression and are highlighted in bold.

Name	Locus tag	Read number								<i>DOM/Ctrl</i> ratio (posterior likelihood)			
		T1		T2		T3		T4		T1	T2	T3	T4
		Ctrl	DOM	Ctrl	DOM	Ctrl	DOM	Ctrl	DOM				
<i>β-oxidation of fatty acids</i>													
Long-chain acyl-CoA synthetase (<i>FadD</i>)	ZP_09994715	45	102	62	88	77	51	84	64	1.51 (0.4347)	1.24 (0.0717)	0.69 (0.3642)	1.03 (0.0509)
Polyhydroxyalkanoate synthase (<i>PhaC</i>)	ZP_09993065	396	1066	522	649	296	504	169	77	1.80 (1)	1.08 (0.0145)	1.77 (1)	0.61 (0.9730)
Acetoacetyl-CoA reductase (<i>PhbB</i>)	ZP_09993064	290	504	133	126	84	100	142	94	1.16 (0.1290)	0.82 (0.0890)	1.24 (0.1435)	0.89 (0.0632)
β -ketothiolase (<i>AtoB</i>)	ZP_09995046	440	963	596	619	361	280	156	95	1.46 (1)	0.90 (0.0245)	0.81 (0.5100)	0.82 (0.1285)
L-sorbose dehydrogenase	ZP_09993994	1543	2913	650	865	501	501	1274	1373	1.26 (1)	1.16 (0.1492)	1.04 (0.0147)	1.45 (1)
<i>Fatty acid biosynthesis</i>													
3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase (<i>FabA</i>)	ZP_09992936	106	186	113	146	112	143	145	120	1.17 (0.0844)	1.12 (0.0441)	1.33 (0.3991)	1.11 (0.0594)
3-oxoacyl-[acyl-carrier-protein] synthase I (<i>FabB</i>)	ZP_09992937	325	428	310	346	194	155	486	164	0.88 (0.0804)	0.97 (0.0161)	0.83 (0.1685)	0.45 (1)
[Acyl-carrier-protein] S-malonyltransferase (<i>FabD</i>)	ZP_09994653	135	134	238	203	135	377	119	132	0.66 (0.9182)	0.74 (0.7419)	2.90 (1)	1.49 (0.8683)
3-oxoacyl-[acyl-carrier-protein] synthase II (<i>FabF</i>)	ZP_09994650	159	149	71	68	67	128	100	51	0.63 (0.9941)	0.83 (0.0569)	1.98 (0.9997)	0.69 (0.3863)
3-oxoacyl-[acyl-carrier protein] reductase (<i>FabG</i>)	ZP_09994652	88	106	139	156	100	204	99	84	0.80 (0.1197)	0.98 (0.0278)	2.12 (1)	1.14 (0.0694)
Enoyl-[acyl-carrier protein] reductase I (<i>FabI</i>)	ZP_09994681	319	719	369	514	359	387	467	613	1.50 (1)	1.21 (0.3638)	1.12 (0.0649)	1.77 (1)
<i>Carboxylic acid transporters</i>													
Putative tricarboxylic transport membrane protein (<i>TctC</i>)	ZP_09994879	278	718	125	151	88	111	134	201	1.72 (1)	1.05 (0.0305)	1.31 (0.2615)	2.02 (1)
TRAP dicarboxylate transporter subunit (<i>DctM</i>)	ZP_09992986	32	20	8	20	10	9	13	18	0.42 (0.8657)	2.18 (0.1785)	0.94 (0.0610)	1.86 (0.1977)
TRAP dicarboxylate transporter subunit (<i>DctP</i>)	ZP_09992988	357	843	177	182	132	151	353	187	1.58 (1)	0.90 (0.0420)	1.19 (0.1269)	0.71 (0.9772)
TRAP dicarboxylate transporter subunit (<i>DctP</i>)	ZP_09993170	4762	9480	2060	2907	2295	3205	1869	1857	1.33 (1)	1.23 (1)	1.45 (1)	1.34 (1)
TRAP dicarboxylate transporter subunit (<i>DctQ</i>)	ZP_09993171	423	734	214	273	197	210	82	97	1.16 (0.1692)	1.11 (0.0384)	1.11 (0.0645)	1.59 (0.8602)
TRAP transporter solute receptor, TAXI family protein	ZP_09993399	429	730	99	175	52	118	90	73	1.14 (0.0896)	1.54 (0.9262)	2.36 (1)	1.09 (0.0570)
TRAP transporter 4TM/12TM fusion protein	ZP_09993400	61	69	23	32	13	16	16	18	0.75 (0.1431)	1.21 (0.0430)	1.28 (0.0758)	1.51 (0.1044)
TRAP dicarboxylate transporter subunit (<i>DctP</i>)	ZP_09993600	194	228	167	200	210	320	206	109	0.78 (0.4019)	1.04 (0.0265)	1.58 (1)	0.71 (0.7358)
TRAP dicarboxylate transporter subunit (<i>DctM</i>)	ZP_09993602	43	50	79	44	42	30	24	40	0.78 (0.0922)	0.48 (0.9875)	0.74 (0.1299)	2.24 (0.9026)
TRAP C4-dicarboxylate transport system permease (<i>DctM</i>)	ZP_09993832	49	44	28	35	74	34	57	45	0.60 (0.5012)	1.09 (0.0359)	0.48 (0.9839)	1.06 (0.0555)
TRAP dicarboxylate transporter subunit (<i>DctQ</i>)	ZP_09993833	30	26	24	15	18	8	15	12	0.58 (0.2848)	0.54 (0.1629)	0.46 (0.2805)	1.08 (0.0531)
TRAP transporter subunit (<i>DctP</i>)	ZP_09993834	338	209	189	156	284	450	341	219	0.41 (1)	0.72 (0.7373)	1.65 (1)	0.86 (0.0916)
TRAP dicarboxylate transporter subunit (<i>DctP</i>)	ZP_09993962	44	88	28	43	56	32	24	45	1.33 (0.1424)	1.34 (0.0681)	0.59 (0.5419)	2.52 (0.9858)
TRAP dicarboxylate transporter subunit (<i>DctM</i>)	ZP_09993964	11	32	21	43	28	29	18	29	1.94 (0.2569)	1.78 (0.2968)	1.08 (0.0659)	2.17 (0.6402)
TRAP dicarboxylate transporter subunit (<i>DctP</i>)	ZP_09994229	246	433	189	184	160	103	152	208	1.17 (0.1515)	0.85 (0.0813)	0.67 (0.9034)	1.84 (1)
TRAP dicarboxylate transporter subunit (<i>DctQ</i>)	ZP_09994230	26	17	15	13	18	9	10	7	0.44 (0.6517)	0.75 (0.0416)	0.52 (0.2042)	0.94 (0.0500)
TRAP dicarboxylate transporter subunit (<i>DctM</i>)	ZP_09994231	41	48	41	59	44	31	42	52	0.78 (0.0867)	1.25 (0.0615)	0.73 (0.1425)	1.67 (0.5420)
TRAP-type C4-dicarboxylate transport protein, permease component	ZP_09994769	41	64	57	85	63	40	45	49	1.04 (0.0469)	1.30 (0.0999)	0.66 (0.3667)	1.47 (0.2376)
TRAP C4-dicarboxylate transport system permease (<i>DctM</i>)	ZP_09994844	230	131	54	50	79	57	18	25	0.38 (1)	0.81 (0.0599)	0.75 (0.2043)	1.87 (0.3130)
TRAP transporter, 4TM/12TM fusion protein	ZP_09994864	13	37	3	22	7	13	16	13	1.90 (0.2953)	6.38 (0.9671)	1.93 (0.1499)	1.09 (0.0539)
TRAP transporter solute receptor	ZP_09994865	54	92	22	18	11	27	39	26	1.14 (0.0582)	0.71 (0.0568)	2.55 (0.7394)	0.90 (0.0590)
TRAP dicarboxylate transporter subunit (<i>DctM</i>)	ZP_09995039	9	24	17	9	17	6	4	15	1.78 (0.1329)	0.46 (0.1701)	0.37 (0.4467)	5.05 (0.9037)
Dihydroxypropanesulfonate TRAP transporter (<i>DctQ</i>)	ZP_09995040	3	2	7	11	10	4	2	6	0.44 (0.0594)	1.37 (0.0376)	0.42 (0.1720)	4.04 (0.2051)
Dihydroxypropanesulfonate TRAP transporter	ZP_09995041	71	122	98	139	124	87	101	203	1.15 (0.0635)	1.23 (0.1001)	0.73 (0.4333)	2.71 (1)
Na ⁺ /H ⁺ antiporter	ZP_09993678	64	67	66	74	53	94	72	57	0.70 (0.2741)	0.98 (0.0333)	1.84 (0.9785)	1.07 (0.0548)

Table 4 continued.

Name	Locus tag	Read number								DOM/Ctrl ratio (posterior likelihood)			
		T1		T2		T3		T4		T1	T2	T3	T4
		Ctrl	DOM	Ctrl	DOM	Ctrl	DOM	Ctrl	DOM				
<i>Anabolic pathways</i>													
Putative fructose-1,6-bisphosphatase (<i>Fbp</i>)	ZP_09993133	75	510	27	117	14	38	4	9	4.45 (1)	3.77 (1)	2.82 (1)	3.03 (0.2450)
Transketolase (<i>TktAB</i>)	ZP_09994301	88	203	82	162	138	102	89	71	1.54 (0.9420)	1.72 (0.9926)	0.77 (0.3012)	1.07 (0.0544)
Transketolase (<i>TktAB</i>)	ZP_09993131	282	702	157	306	62	79	35	49	1.66 (1)	1.70 (1)	1.32 (0.2000)	1.88 (0.7852)
Transaldolase B (<i>TalB</i>)	ZP_09993495	75	122	90	156	83	101	53	78	1.09 (0.0490)	1.51 (0.8143)	1.26 (0.1732)	1.98 (0.9902)
Ribose-5-phosphate isomerase A (<i>RpiA</i>)	ZP_09994308	49	40	56	107	43	149	49	42	0.54 (0.7417)	1.66 (0.8278)	3.60 (1)	1.15 (0.0665)
Tryptophan synthase subunit alpha (<i>TrpA</i>)	ZP_09993429	138	270	174	197	210	117	168	179	1.31 (0.4883)	0.99 (0.0245)	0.58 (0.9998)	1.43 (0.9075)
Tryptophan synthase subunit beta (<i>TrpB</i>)	ZP_09993430	271	520	235	271	252	264	252	279	1.28 (0.8385)	1.00 (0.0195)	1.09 (0.0498)	1.49 (0.9989)
N-(5'-phosphoribosyl)anthranilate isomerase (<i>TrpF</i>)	ZP_09993431	48	137	57	67	35	32	72	60	1.90 (0.9931)	1.02 (0.0337)	0.95 (0.0649)	1.12 (0.0629)
Cyclohexadienyl dehydrogenase/3-P-shikimate,1-carboxyvinyltransferase (<i>AroA</i>)	ZP_09993093	156	175	131	305	132	175	180	132	0.75 (0.5239)	2.03 (1)	1.38 (0.7051)	0.99 (0.0384)
Aromatic amino acid aminotransferase (<i>AspC</i>)	ZP_09994396	50	70	53	51	49	66	23	42	0.93 (0.0485)	0.84 (0.0498)	1.40 (0.2445)	2.46 (0.9699)
Phospho-2-dehydro-3-deoxyheptonate aldolase (<i>AroFGH</i>)	ZP_09994966	55	61	89	120	46	206	48	54	0.74 (0.1484)	1.17 (0.0578)	4.65 (1)	1.51 (0.3320)
isocitrate lyase (<i>AceA</i>)	ZP_09992865	375	1489	1112	1541	1904	1173	2522	2475	2.65 (1)	1.21 (0.9700)	0.64 (1)	1.32 (1)
3-isopropylmalate dehydrogenase (<i>LeuB</i>)	ZP_09993290	73	182	172	224	166	152	150	174	1.66 (0.9824)	1.13 (0.0493)	0.95 (0.0501)	1.56 (0.9911)
Isopropylmalate isomerase small subunit (<i>LeuD</i>)	ZP_09993291	90	203	113	259	207	168	175	175	1.51 (0.9038)	2.00 (1)	0.84 (0.1429)	1.35 (0.6248)
Isopropylmalate isomerase large subunit (<i>LeuC</i>)	ZP_09993292	190	325	205	247	237	299	176	329	1.14 (0.0768)	1.05 (0.0242)	1.31 (0.8061)	2.52 (1)
Branched-chain amino acid aminotransferase (<i>IlvE</i>)	ZP_09994157	346	787	380	405	284	189	283	300	1.52 (1)	0.93 (0.0205)	0.69 (0.9895)	1.43 (0.9951)
Ketol-acid reductoisomerase (<i>IlvC</i>)	ZP_09993185	240	343	355	568	334	363	279	259	0.95 (0.0287)	1.39 (0.9995)	1.13 (0.0771)	1.25 (0.3943)
Acetolactate synthase 3 catalytic subunit (<i>IlvBGL</i>)	ZP_09993187	314	641	440	602	535	511	335	533	1.36 (0.9981)	1.19 (0.2755)	0.99 (0.0113)	2.14 (1)
Isoleucyl-tRNA synthetase (<i>IlvS</i>)	ZP_09993202	371	962	485	619	509	376	416	478	1.73 (1)	1.11 (0.0297)	0.77 (0.9706)	1.55 (1)
Valyl-tRNA synthetase (<i>ValS</i>)	ZP_09993724	370	701	560	639	603	394	415	543	1.26 (0.9213)	0.99 (0.0056)	0.68 (1)	1.76 (1)
Leucyl-tRNA synthetase (<i>LeuS</i>)	ZP_09993862	266	517	262	371	229	246	165	197	1.30 (0.8983)	1.23 (0.3235)	1.12 (0.0680)	1.61 (0.9988)

Table 5. Putative compound identification for features detected in *Prochlorococcus* str. MIT9313 DOM_P related to carboxylic acid metabolism. Putative identifications are based on ultrahigh-resolution mass information (confidence level ± 2 ppm) with additional verification based on MS/MS fragmentation information when possible. The difference in the detected mass and the putative identification in the METLIN database (Δ ppm) along with the median m/z and retention time for multiple sample injections. The mean intensity level of each MIT9313 DOM_P feature was divided by the mean intensity of any matching feature in the Pro99 medium control as is reported here as the fold change. Significant intensity differences were determined using Welch's t test (unequal variances) and only features for which p -values were ≤ 0.01 and fold changes were ≥ 4 were included in the analysis and shown here.

METLIN ID#	Putative name	Formula	Δ ppm	m/z	Retention time (min)	Fold change	p -Value
35505	11,12,13-trihydroxy-9-octadecenoic acid	C ₁₈ H ₃₄ O ₅	0	353.22991	31.98	27423	0.0069
82357	Phosphatidic acid (14:1(9Z)/0:0)	C ₁₇ H ₃₃ O ₇ P	0	381.20369	38.84	1185	0.0041
45886	9-nitrooctadec-9E-enoic acid	C ₁₈ H ₃₃ NO ₄	0	328.24826	23.70	792	0.0050
74735	3-oxo-tridecanoic acid	C ₁₃ H ₂₄ O ₃	0	229.17992	25.27	703	0.0097
34780	2Z,4E-decadienoic acid	C ₁₀ H ₁₆ O ₂	0	169.12230	20.45	388	0.0093
62998	17-keto-4(Z),7(Z),10(Z),13(Z),15(E),19(Z)-docosahexaenoic acid	C ₂₂ H ₃₀ O ₃	0	365.20875	38.95	254	0.0061
35879	12,13-dihydroxy-11-methoxy-9-octadecenoic acid	C ₁₉ H ₃₆ O ₅	0	367.24554	34.35	213	0.0033
34630	7-methyl-6E-hexadecenoic acid	C ₁₇ H ₃₂ O ₂	0	269.24751	39.56	205	0.0049
35797	16-oxo-heptadecanoic acid	C ₁₇ H ₃₂ O ₃	0	307.22435	37.83	164	0.0061
86504	(Z)-15-Oxo-11-eicosenoic acid	C ₂₀ H ₃₆ O ₃	0	347.25570	43.15	133	0.0001
35561	12-hydroxy-5,8,10-heptadecatrienoic acid	C ₁₇ H ₂₈ O ₃	0	303.19305	36.86	121	0.0076
35430	2-hydroxy-heptadecanoic acid	C ₁₇ H ₃₄ O ₃	0	287.25802	40.62	115	0.0038
35537	12-hydroxy-3Z,6Z-dodecadienoic acid	C ₁₂ H ₂₀ O ₃	0	213.14852	26.22	57	0.0098
35782	9,12-dioxo-dodecanoic acid	C ₁₂ H ₂₀ O ₄	0	229.14345	31.53	42	0.0061
87733	(R)-3-Hydroxy-5-phenylpentanoic acid	C ₁₁ H ₁₄ O ₃	0	195.10158	25.80	35	0.0097
35631	8-hydroxy-13,17-octadecadiene-9,11-diynoic acid	C ₁₈ H ₂₄ O ₃	0	289.17983	30.00	34	0.0013
34495	2-hydroxyhexadecanoic acid	C ₁₆ H ₃₂ O ₃	0	295.22437	40.47	29	0.0015
35576	13-hydroxy-docosanoic acid	C ₂₂ H ₄₄ O ₃	0	379.31822	47.68	28	0.0029
35023	3,6,9,12,15-octadecapentaenoic acid	C ₁₈ H ₂₆ O ₂	0	275.20063	33.61	26	0.0031
35423	2-hydroxy-pentadecanoic acid	C ₁₅ H ₃₀ O ₃	0	259.22679	36.59	24	0.0018
5748	Octadecanedioic acid	C ₁₈ H ₃₄ O ₄	0	337.23489	37.13	22	0.0087
5578	Decenedioic acid	C ₁₀ H ₁₆ O ₄	0	201.11213	18.22	16	0.0035
35515	12(S)-hydroxy-16-heptadecynoic acid	C ₁₇ H ₃₀ O ₃	0	305.20875	39.22	16	0.0024
35592	6-hydroxy-nonanoic acid	C ₉ H ₁₈ O ₃	0	175.13285	24.68	14	0.0019
35530	2-hydroxy-10-undecenoic acid	C ₁₁ H ₂₀ O ₃	0	201.14853	29.77	14	0.0021
35550	3,12-dihydroxy-pentadecanoic acid	C ₁₅ H ₃₀ O ₄	0	297.20366	35.23	11	0.0038
35818	10-oxo-nonadecanoic acid	C ₁₉ H ₃₆ O ₃	0	313.27375	40.60	10	0.0049
35523	10-hydroxy-2E-decenoic acid	C ₁₀ H ₁₈ O ₃	0	187.13287	24.92	10	0.0048
5642	Hexadecanedioic acid	C ₁₆ H ₃₀ O ₄	0	309.20367	33.95	9	0.0022
35560	12-hydroxy-8E,10E-heptadecadienoic acid	C ₁₇ H ₃₀ O ₃	0	305.20872	36.81	6	0.0062
35520	6-hydroxy-decanoic acid	C ₁₀ H ₂₀ O ₃	0	189.14852	27.60	5	0.0089
35645	8-hydroxy-13Z-octadecene-9,11-diynoic acid	C ₁₈ H ₂₆ O ₃	0	291.19544	33.84	5	0.0064
35415	2-hydroxy hendecanoic acid	C ₁₁ H ₂₂ O ₃	0	203.16421	29.94	5	0.0064

Table 6. Read numbers and ratios of expression in the DOM treatment compared to the control medium (*DOM/Ctrl*) for coding DNA sequences (CDSs) associated with glucoside and glycerolipid metabolism, including biosynthesis of peptidoglycan. *DOM/Ctrl* ratios are derived from read numbers for each sample that are normalized to the total number of CDSs in each sample. Posterior likelihoods ≥ 0.9 indicate significant differential expression and are highlighted in bold.

Name	Locus tag	Read number								DOM/Ctrl ratio (posterior likelihood)			
		T1		T2		T3		T4		T1	T2	T3	T4
		Ctrl	DOM	Ctrl	DOM	Ctrl	DOM	Ctrl	DOM				
<i>Glycerolipid biosynthesis</i>													
Glycerol-3-phosphate dehydrogenase (<i>GlpA</i>)	ZP_09993697	1282	2202	898	1533	1149	637	26	40	1.15 (0.3583)	1.49 (1)	0.58 (1)	2.07 (0.8069)
Glycerol kinase (<i>GlpK</i>)	ZP_09993705	1948	1375	324	701	424	480	168	144	0.47 (1)	1.88 (1)	1.18 (0.2293)	1.15 (0.0793)
Diacylglycerol kinase (<i>DgkA</i>)	ZP_09993907	21	75	51	105	77	69	86	98	2.38 (0.9836)	1.79 (0.9415)	0.93 (0.0647)	1.53 (0.7654)
<i>Peptidoglycan biosynthesis</i>													
UDP-N-acetylglucosamine pyrophosphorylase (<i>GlmU</i>)	ZP_09994217	54	39	28	44	33	30	35	30	0.48 (0.9569)	1.37 (0.0768)	0.94 (0.0652)	1.15 (0.0634)
tRNA delta(2)-isopentenylpyrophosphate transferase (<i>MiaA</i>)	ZP_09994361	37	59	27	66	49	65	30	53	1.06 (0.0483)	2.13 (0.9308)	1.38 (0.2174)	2.38 (0.9907)
UDP-N-acetylmuramate-L-alanine ligase (<i>MurC</i>)	ZP_09993243	51	114	41	112	61	63	94	28	1.49 (0.4663)	2.38 (0.9999)	1.07 (0.0653)	0.40 (0.9994)
UDP-N-acetylmuramyl-tripeptide synthetase (<i>MurE</i>)	ZP_09993249	76	221	101	114	114	105	160	93	1.94 (1)	0.98 (0.0307)	0.96 (0.0563)	0.78 (0.2151)
UDPmuramoylalanylNAc-D-glutamyl-2,6-diaminopimelate-D-alanyl ligase (<i>MurF</i>)	ZP_09993248	102	253	99	115	131	99	180	101	1.66 (0.9986)	1.01 (0.0306)	0.79 (0.2267)	0.76 (0.3681)
D-alanine-D-alanine ligase (<i>DdlAB</i>)	ZP_09993242	37	93	26	68	49	51	71	46	1.68 (0.6738)	2.28 (0.9754)	1.08 (0.0668)	0.87 (0.0671)
Peptidoglycan glycosyltransferase (<i>FtsI</i>)	ZP_09993250	121	410	192	268	227	174	276	193	2.26 (1)	1.22 (0.1557)	0.80 (0.3493)	0.94 (0.0344)
Integral membrane protein (<i>MviN</i>)	ZP_09993204	27	76	84	100	63	45	30	51	1.88 (0.7827)	1.04 (0.0325)	0.74 (0.1771)	2.29 (0.9797)
Phospho-N-acetylmuramoyl-pentapeptide-transferase (<i>MraY</i>)	ZP_09993247	111	190	130	143	100	57	136	48	1.14 (0.0677)	0.96 (0.0299)	0.59 (0.9159)	0.48 (0.9996)
<i>Inositol phosphate metabolism</i>													
Myo-inositol-1-monophosphotase (<i>SuhB</i>)	ZP_09994972	23	50	48	65	44	45	37	64	1.45 (0.1318)	1.18 (0.0484)	1.06 (0.0652)	2.33 (0.9972)
Methylmalonate-semialdehyde dehydrogenase (<i>IoIA</i>)	ZP_09994561	180	91	61	115	70	50	197	83	0.34 (1)	1.64 (0.8436)	0.74 (0.1950)	0.57 (0.9990)
5-keto-2-deoxygluconokinase (<i>IoBI</i>)	ZP_09994462	6	10	6	11	12	15	11	29	1.11 (0.0451)	1.60 (0.0463)	1.30 (0.0762)	3.55 (0.9879)
PfkB protein [gamma proteobacterium (<i>IoC</i>)	ZP_09994950	8	38	17	27	16	10	18	31	3.17 (0.9186)	1.38 (0.0570)	0.65 (0.1061)	2.32 (0.7923)
Acetolactate synthase [gamma proteobacterium (<i>IoD</i>)	ZP_09994951	18	59	17	32	18	20	21	37	2.19 (0.8498)	1.64 (0.1246)	1.15 (0.0689)	2.37 (0.9148)
Triosephosphate isomerase (<i>TpiA</i>)	ZP_09993909	102	136	146	219	180	447	177	182	0.89 (0.0580)	1.31 (0.3746)	2.58 (1)	1.38 (0.7963)
<i>Sugar ABC transporters</i>													
Sugar ABC transporter permease (<i>LrgA</i> synteny)	ZP_09993144	33	47	37	90	23	54	30	27	0.95 (0.0478)	2.12 (0.9906)	2.44 (0.9865)	1.21 (0.0693)
Sugar ABC transporter ATP-binding protein (<i>GlpA</i> synteny)	ZP_09993700	1274	1067	240	353	261	334	21	22	0.56 (1)	1.28 (0.5893)	1.33 (0.9188)	1.41 (0.0968)
Sugar ABC transporter permease	ZP_09994754	4	19	8	17	18	15	5	25	3.17 (0.4094)	1.85 (0.0886)	0.87 (0.0680)	6.73 (0.9995)
Putative sugar ABC transporter periplasmic sugar-binding protein	ZP_09994755	58	486	52	182	122	181	144	351	5.59 (1)	3.05 (1)	1.54 (0.9809)	3.28 (1)
Sugar ABC transporter permease (<i>UgpC</i> synteny)	ZP_09994858	12	11	10	16	16	16	3	21	0.61 (0.0875)	1.39 (0.0445)	1.04 (0.0629)	9.42 (0.9994)
Sugar-binding protein (<i>UgpC</i> synteny)	ZP_09994857	65	143	53	50	47	38	44	70	1.47 (0.5773)	0.82 (0.0543)	0.84 (0.0849)	2.14 (0.9944)
Sugar ABC transporter, ATP-binding protein (DHA synteny)	ZP_09993021	62	35	35	31	48	14	13	8	0.38 (0.9997)	0.77 (0.0575)	0.30 (0.9985)	0.83 (0.0549)
Sugar ABC transporter, permease (DHA synteny)	ZP_09993024	68	58	37	33	50	31	7	8	0.57 (0.8715)	0.8 (0.0577)	0.64 (0.3108)	1.54 (0.0684)
Sugar ABC transporter, substrate-binding protein (DHA synteny)	ZP_09993025	735	952	570	352	469	236	164	90	0.86 (0.2407)	0.54 (1)	0.52 (1)	0.74 (0.4100)

Table 7. Putative compound identification for features detected in *Prochlorococcus* str. MIT9313 DOM_p related to glucoside and glycerolipid metabolism. Putative identifications are based on ultrahigh-resolution mass information (confidence level ± 2 ppm) with additional verification based on MS/MS fragmentation information when possible. The difference in the detected mass and the putative identification in the METLIN database is indicated (Δppm) along with the median m/z and retention time for multiple sample injections. The mean intensity level of each MIT9313 DOM_p feature was divided by the mean intensity of any matching feature in the Pro99 medium control as is reported here as the fold change. Significant intensity differences were determined using Welch's t test (unequal variances) and only features for which p -values were ≤ 0.01 and fold changes were ≥ 4 were included in the analysis and shown here.

METLIN ID#	Putative name	Formula	Δppm	m/z	Retention time (min)	Fold change	p-Value
57941	Glycocholic acid 3-glucuronide	C ₃₂ H ₅₃ NO ₁₂	2	644.36266	25.78	4455	0.0044
82357	Phosphatidic acid (14:1(9Z)/0:0)	C ₁₇ H ₃₃ O ₇ P	0	381.20369	38.84	1185	0.0041
57947	D-glucopyranosiduronic acid	C ₃₂ H ₅₁ NO ₁₂	2	642.34969	22.23	688	0.0092
90017	Medicagenic acid 3-O- β -D-glucoside	C ₃₆ H ₅₆ O ₁₁	1	687.37187	21.68	681	0.0025
68969	N-Acetyl-leu-leu-tyr-amide	C ₂₃ H ₃₆ N ₄ O ₅	0	449.27593	21.12	525	0.0007
41358	Astaxanthin diglucoside/ Astaxanthin- β -D-diglucoside	C ₅₂ H ₇₂ O ₁₄	1	921.49816	24.46	225	0.0042
414	1,2-Dioctanoyl-sn-glycerol	C ₁₉ H ₃₆ O ₅	0	367.24554	34.35	213	0.0033
40332	Phosphatidylcholine (8:0/0:0)	C ₁₆ H ₃₄ NO ₇ P	2	406.19711	16.56	146	0.0036
817	Indinavir-N-glucuronide	C ₄₂ H ₅₅ N ₅ O ₁₀	1	812.38296	19.85	99	0.0010
92393	3-Hydroxy- β -ionol 3-[glucosyl-(1- \rightarrow 6)-glucoside]	C ₂₅ H ₄₂ O ₁₂	2	535.27600	21.05	26	0.0069
80996	Phosphatidylinositol (0-16:0/19:1(9Z))	C ₄₄ H ₈₅ O ₁₂ P	2	836.43531	46.82	20	0.0072
4254	1-dodecanoyl-sn-glycerol	C ₁₅ H ₃₀ O ₄	0	297.20366	35.23	11	0.0038
89805	2-Aminoheptanedioic acid	C ₇ H ₁₃ NO ₄	0	176.09169	2.30	6	0.0098

Table 8. Read numbers and ratios of expression in the DOM treatment compared to the control medium (*DOM/Ctrl*) for coding DNA sequences (CDSs) associated with motility and chemotaxis as well as CDSs involved in OmpR family regulation. *DOM/Ctrl* ratios are derived from read numbers for each sample that are normalized to the total number of CDSs in each sample. Posterior likelihoods ≥ 0.9 indicate significant differential expression and are highlighted in bold.

Name	Locus tag	Read number								<i>DOM/Ctrl</i> ratio (posterior likelihood)			
		T1		T2		T3		T4		T1	T2	T3	T4
		Ctrl	DOM	Ctrl	DOM	Ctrl	DOM	Ctrl	DOM				
<i>Flagellar assembly</i>													
Cobyrinic acid ac-diamide synthase (<i>FlhG</i>)	ZP_09992977	44	50	71	61	66	105	73	68	0.76 (0.1054)	0.75 (0.1202)	1.65 (0.9205)	1.25 (0.1154)
Flagellar hook protein (<i>FlgE</i>)	ZP_09993565	1807	2920	1825	2070	1451	823	2487	806	1.08 (0.0008)	0.99 (0.0000)	0.59 (1)	0.44 (1)
Flagellar basal body rod protein (<i>FlgF</i>)	ZP_09993566	947	1951	1489	1046	918	470	1708	757	1.37 (1)	0.61 (1)	0.53 (1)	0.60 (1)
Flagellar basal-body rod protein (<i>FlgG</i>)	ZP_09993567	1096	2066	1263	867	785	490	1758	583	1.26 (1)	0.60 (1)	0.65 (1)	0.45 (1)
Flagellar basal body L-ring protein	ZP_09993568	930	1769	1234	953	730	342	1386	470	1.27 (1)	0.67 (1)	0.49 (1)	0.46 (1)
Flagellar P-ring protein	ZP_09993569	791	1480	1251	848	720	236	1257	390	1.25 (0.9989)	0.59 (1)	0.34 (1)	0.42 (1)
Flagellar hook-associated protein 1 (<i>Hap1</i>)	ZP_09993571	1852	2753	1880	1243	1235	512	2552	709	0.99 (0.0000)	0.58 (1)	0.43 (1)	0.37 (1)
Flagellar hook-associated protein (<i>FlgI</i>)	ZP_09993572	736	1007	776	522	516	181	885	280	0.91 (0.0212)	0.59 (1)	0.36 (1)	0.43 (1)
Flagellin domain-containing protein (<i>FliC</i>)	ZP_09993573	3184	5144	3975	3070	3050	1828	5191	2443	1.08 (0.0000)	0.67 (1)	0.62 (1)	0.63 (1)
Flagellar hook-associated 2 domain-containing protein	ZP_09993574	724	1161	868	637	543	307	997	488	1.07 (0.0092)	0.64 (1)	0.59 (1)	0.66 (1)
Flagellar protein (<i>FlaG</i>)	ZP_09993575	52	87	54	39	35	32	68	23	1.12 (0.0541)	0.63 (0.2917)	0.95 (0.0649)	0.46 (0.9546)
Flagellin (<i>FliC</i>)	ZP_09993578	240	561	215	255	232	160	346	298	1.56 (1)	1.03 (0.0220)	0.72 (0.8971)	1.16 (0.0970)
Flagellar protein (<i>FliS</i>)	ZP_09994417	66	136	68	41	51	29	144	53	1.38 (0.3118)	0.52 (0.8936)	0.59 (0.4900)	0.50 (0.9994)
Flagellar regulatory protein (<i>FleQ</i>)	ZP_09994420	1658	3278	2498	2464	2206	1398	2672	1703	1.32 (1)	0.86 (0.8538)	0.66 (1)	0.86 (0.4360)
Flagellar hook-basal body complex protein (<i>FliE</i>)	ZP_09994421	370	665	515	438	433	234	426	179	1.20 (0.4206)	0.74 (0.9975)	0.56 (1)	0.57 (1)
Flagellar M-ring protein (<i>FliF</i>)	ZP_09994422	1406	1867	1864	1340	1175	672	1695	595	0.89 (0.0851)	0.63 (1)	0.59 (1)	0.47 (1)
Flagellar assembly protein FliH/type III secretion system (<i>HrpE</i>)	ZP_09994424	1237	1648	746	556	600	379	1175	461	0.89 (0.0643)	0.65 (1)	0.66 (1)	0.53 (1)
Flagellar protein export ATPase (<i>FliI</i>)	ZP_09994425	816	1277	559	623	616	317	965	427	1.04 (0.0037)	0.97 (0.0066)	0.53 (1)	0.60 (1)
Flagellar hook-length control protein	ZP_09994427	460	962	481	395	408	239	1060	295	1.40 (1)	0.71 (0.9995)	0.61 (1)	0.37 (1)
Flagellar basal body-associated protein (<i>FliL</i>)	ZP_09994428	643	1112	451	394	476	400	1369	477	1.15 (0.2285)	0.76 (0.9666)	0.87 (0.1068)	0.47 (1)
Flagellar protein (<i>FliO</i>)	ZP_09994431	307	508	280	272	256	141	534	129	1.10 (0.0450)	0.85 (0.1125)	0.57 (1)	0.33 (1)
Flagellar biosynthesis protein (<i>FliP</i>)	ZP_09994432	385	725	780	686	618	190	414	227	1.26 (0.9084)	0.77 (0.9993)	0.32 (1)	0.74 (0.9559)
Flagellar biosynthesis protein (<i>FliQ</i>)	ZP_09994433	179	509	419	310	238	83	288	162	1.90 (1)	0.64 (1)	0.36 (1)	0.76 (0.6527)
Flagellar biosynthetic protein (<i>FliR</i>)	ZP_09994434	149	289	283	172	128	41	152	58	1.29 (0.4832)	0.53 (1)	0.33 (1)	0.51 (0.9992)
Flagellar biosynthesis protein (<i>FliB</i>)	ZP_09994435	236	371	318	256	172	96	303	116	1.05 (0.0284)	0.70 (0.9932)	0.58 (0.9986)	0.52 (1)
RNA polymerase sigma factor for flagellar operon (<i>FliA</i>)	ZP_09994441	321	559	408	409	269	146	372	140	1.16 (0.1472)	0.87 (0.0701)	0.56 (1)	0.51 (1)
Cobyrinic acid a,c-diamide synthase (<i>FlhG</i>)	ZP_09994442	646	1100	913	879	721	382	750	367	1.14 (0.1103)	0.84 (0.6551)	0.55 (1)	0.66 (1)
Flagellar biosynthesis regulator (<i>FliH</i>)	ZP_09994443	841	1919	1285	1117	1075	448	978	421	1.52 (1)	0.76 (1)	0.43 (1)	0.58 (1)
Flagellar biosynthesis protein (<i>FliA</i>)	ZP_09994444	989	2153	2038	1807	1434	512	1347	531	1.45 (1)	0.77 (1)	0.37 (1)	0.53 (1)
Flagellar basal body P-ring biosynthesis protein-like protein	ZP_09994453	19	23	17	26	14	13	26	18	0.81 (0.0586)	1.33 (0.0503)	0.96 (0.0622)	0.93 (0.0551)
Flagellar basal-body rod protein (<i>FlgB</i>)	ZP_09994454	215	494	417	416	338	219	564	377	1.53 (1)	0.87 (0.0803)	0.67 (0.9992)	0.90 (0.0353)
Flagellar basal-body rod protein (<i>FlgC</i>)	ZP_09994455	353	556	435	311	264	186	485	230	1.05 (0.0200)	0.62 (1)	0.73 (0.8905)	0.64 (1)
Flagellar basal body rod modification protein	ZP_09994456	464	885	622	537	389	208	661	254	1.27 (0.9865)	0.75 (0.9987)	0.56 (1)	0.52 (1)
Flagellar hook protein (<i>FlgE</i>)	ZP_09994457	828	1273	934	957	559	371	1047	339	1.03 (0.0027)	0.89 (0.0310)	0.69 (1)	0.44 (1)

Table 8 continued.

Name	Locus tag	Read number								DOM/Ctrl ratio (posterior likelihood)			
		T1		T2		T3		T4		T1	T2	T3	T4
		Ctrl	DOM	Ctrl	DOM	Ctrl	DOM	Ctrl	DOM				
<i>Chemotaxis</i>													
Flagellar motor protein (<i>PomA</i>)	ZP_09994414	172	274	211	238	168	143	302	172	1.06 (0.0370)	0.98 (0.0218)	0.88 (0.0799)	0.77 (0.6058)
Flagellar protein (<i>MotB</i>)	ZP_09994415	224	352	243	238	234	140	267	210	1.05 (0.0294)	0.85 (0.0868)	0.62 (0.9991)	1.06 (0.0327)
Flagellar motor switch protein (<i>FliG</i>)	ZP_09994423	1281	2050	1093	927	616	429	1252	624	1.07 (0.0018)	0.74 (1)	0.72 (0.9999)	0.67 (1)
Flagellar motor switch protein (<i>FliM</i>)	ZP_09994429	829	1529	852	666	681	395	1339	474	1.23 (0.9958)	0.68 (1)	0.60 (1)	0.48 (1)
Flagellar motor switch protein (<i>FliN</i>)	ZP_09994430	324	446	229	205	175	145	490	121	0.92 (0.0352)	0.78 (0.3928)	0.86 (0.1041)	0.33 (1)
<i>OmpR family two-component system</i>													
Signal transduction histidine kinase (<i>PhoR</i>)	ZP_09994253	73	135	90	152	114	104	128	190	1.23 (0.1128)	1.47 (0.6942)	0.95 (0.0577)	2.00 (1)
Two-component transcriptional regulator (<i>PhoB</i>)	ZP_09994256	22	37	14	28	14	69	13	14	1.12 (0.0512)	1.74 (0.1343)	5.12 (1)	1.45 (0.0807)

Table 9. Read numbers and ratios of expression in the DOM treatment compared to the control medium (*DOM/Ctrl*) for coding DNA sequences (CDSs) associated with proteorhodopsin-based phototrophy, electron transport chain and ATP synthase. *DOM/Ctrl* ratios are derived from read numbers for each sample that are normalized to the total number of CDSs in each sample. Posterior likelihoods ≥ 0.9 indicate significant differential expression and are highlighted in bold.

Name	Locus tag	Read number								<i>DOM/Ctrl</i> ratio (posterior likelihood)			
		T1		T2		T3		T4		T1	T2	T3	T4
		Ctrl	DOM	Ctrl	DOM	Ctrl	DOM	Ctrl	DOM				
<i>Opsin</i>													
Proteorhodopsin	ZP_09993982	1460	3192	2196	3636	9398	13666	2703	3747	1.46 (1)	1.44 (1)	1.51 (1)	1.87 (1)
<i>Retinal biosynthesis</i>													
Geranylgeranyl pyrophosphate synthase (<i>CrtE</i>)	ZP_09993210	28	35	40	45	16	49	33	20	0.83 (0.0598)	0.98 (0.0342)	3.18 (0.9987)	0.82 (0.0689)
Protoheme IX farnesyltransferase (<i>CyoE</i>)	ZP_09993674	161	222	165	271	145	140	267	174	0.92 (0.0447)	1.43 (0.9489)	1.00 (0.0482)	0.88 (0.0716)
Brp/Blh family beta-carotene 15,15'-monooxygenase	ZP_09993977	12	4	9	34	30	12	10	3	0.22 (0.7217)	3.29 (0.9377)	0.42 (0.7192)	0.40 (0.1235)
Lycopene beta and epsilon cyclase (<i>CrtY</i>)	ZP_09993978	25	35	35	71	163	144	84	67	0.93 (0.0485)	1.77 (0.6549)	0.92 (0.0606)	1.07 (0.0547)
Phytoene synthase (<i>CrtB</i>)	ZP_09993979	42	40	57	134	297	277	99	94	0.64 (0.2838)	2.05 (0.9995)	0.97 (0.0307)	1.28 (0.1715)
Phytoene desaturase (<i>CrtP</i>)	ZP_09993980	52	59	83	148	450	339	131	97	0.76 (0.1213)	1.55 (0.8627)	0.78 (0.8697)	1.00 (0.0445)
Polyprenyl synthetase	ZP_09993981	74	115	117	288	687	881	191	200	1.04 (0.0437)	2.14 (1)	1.33 (1)	1.41 (0.9093)
<i>Light absorption</i>													
Deoxyribodipyrimidine photolyase	ZP_09993439	74	196	146	305	412	234	207	386	1.77 (0.998)	1.82 (1)	0.59 (1)	2.51 (1)
Deoxyribodipyrimidine photolyase	ZP_09993411	12	13	22	25	107	34	60	27	0.72 (0.0626)	0.99 (0.0335)	0.33 (1)	0.61 (0.3990)
Deoxyribodipyrimidine photo-lyase type I	ZP_09993412	27	23	51	49	219	63	96	61	0.57 (0.2640)	0.84 (0.0495)	0.30 (1)	0.86 (0.0768)
Deoxyribodipyrimidine photolyase-related protein	ZP_09993414	31	33	84	77	415	124	150	129	0.71 (0.1108)	0.80 (0.0846)	0.31 (1)	1.16 (0.0805)
PAS/PAC/GAF sensor-containing diguanylate cyclase/P-diesterase	ZP_09992866	4	12	9	17	6	13	9	6	2.00 (0.0905)	1.64 (0.0656)	2.25 (0.2098)	0.90 (0.0498)
Multi-sensor hybrid histidine kinase	ZP_09993376	34	51	32	90	47	29	55	50	1.00 (0.0466)	2.45 (0.9993)	0.64 (0.2947)	1.22 (0.0867)
<i>Electron transport chain</i>													
Pyrrroquinoline quinone biosynthesis protein (<i>PqqB</i>)	ZP_09993998	657	1442	710	947	809	614	757	1025	1.46 (1)	1.16 (0.1827)	0.79 (0.9907)	1.82 (1)
Pyrrroquinoline quinone biosynthesis protein (<i>PqqC</i>)	ZP_09993999	84	200	41	57	43	36	21	42	1.59 (0.9706)	1.21 (0.0518)	0.87 (0.0757)	2.69 (0.9890)
Quinone oxidoreductase	ZP_09992848	161	283	123	130	119	49	132	71	1.17 (0.1074)	0.92 (0.0357)	0.43 (1)	0.72 (0.3688)
<i>YhdH/YhfP</i> family quinone oxidoreductase	ZP_09992933	173	239	186	208	157	204	180	250	0.92 (0.0432)	0.97 (0.0243)	1.35 (0.7128)	1.87 (1)
Na ⁺ -translocating NADH-quinone reductase subunit F (<i>NqrF</i>)	ZP_09994527	414	529	376	517	256	360	451	304	0.85 (0.2042)	1.20 (0.2578)	1.46 (0.9993)	0.91 (0.0357)
Na ⁺ -translocating NADH-quinone reductase subunit E (<i>NqrE</i>)	ZP_09994528	184	193	165	231	136	144	147	94	0.70 (0.9265)	1.22 (0.1386)	1.10 (0.0655)	0.86 (0.0808)
Na ⁺ -translocating NADH-quinone reductase subunit D (<i>NqrD</i>)	ZP_09994529	138	186	231	287	121	158	162	91	0.90 (0.0542)	1.08 (0.0282)	1.36 (0.5598)	0.76 (0.3168)
Na ⁺ -translocating NADH-quinone reductase subunit C (<i>NqrC</i>)	ZP_09994530	195	264	171	156	87	151	196	116	0.90 (0.0509)	0.79 (0.1881)	1.80 (0.9991)	0.80 (0.2154)
Na ⁺ -translocating NADH-quinone reductase subunit B (<i>NqrB</i>)	ZP_09994531	439	615	443	652	377	506	412	344	0.93 (0.0207)	1.28 (0.9619)	1.39 (0.9996)	1.12 (0.0498)
Na ⁺ -translocating NADH-quinone reductase subunit A (<i>NqrA</i>)	ZP_09994532	450	634	494	572	275	640	447	337	0.94 (0.0183)	1.01 (0.0072)	2.42 (1)	1.01 (0.0143)
NAD(P)H quinone oxidoreductase	ZP_09994721	66	109	86	119	98	50	86	98	1.10 (0.0521)	1.20 (0.0708)	0.53 (0.9861)	1.53 (0.7654)

Table 9 continued.

Name	Locus tag	Read number								DOM/Ctrl ratio (posterior likelihood)			
		T1		T2		T3		T4		T1	T2	T3	T4
		Ctrl	DOM	Ctrl	DOM	Ctrl	DOM	Ctrl	DOM				
<i>Solute transporters</i>													
Putative sodium symporter protein	ZP_09993374	529	570	114	197	116	100	237	147	0.72 (1)	1.5 (0.9315)	0.83 (0.1351)	0.83 (0.1351)
Solute symporter protein	ZP_09993375	70	102	24	58	26	28	41	69	0.97 (0.0439)	2.1 (0.8523)	1.12 (0.0688)	2.27 (0.9977)
Na ⁺ /H ⁺ antiporter	ZP_09993678	64	67	66	74	53	94	72	57	0.7 (0.2741)	0.98 (0.0333)	1.84 (0.9785)	1.07 (0.0548)
<i>ATP synthase</i>													
ATP synthase subunit I	ZP_09994208	87	135	111	124	87	136	103	59	1.04 (0.0422)	0.97 (0.0303)	1.62 (0.9710)	0.77 (0.1600)
ATP synthase subunit A (<i>AtpB</i>)	ZP_09994209	238	272	240	231	179	224	147	108	0.76 (0.7328)	0.84 (0.1168)	1.30 (0.5626)	0.99 (0.0426)
ATP synthase subunit C (<i>AtpE</i>)	ZP_09994210	130	113	132	157	101	138	89	59	0.58 (0.9968)	1.04 (0.0287)	1.42 (0.6712)	0.89 (0.0628)
ATP synthase subunit B (<i>AtpF</i>)	ZP_09994211	312	331	226	187	120	250	148	137	0.71 (0.9965)	0.72 (0.8538)	2.16 (1)	1.25 (0.1852)
ATP synthase F1 subunit delta (<i>AtpH</i>)	ZP_09994212	285	293	229	231	141	300	205	186	0.69 (0.9984)	0.88 (0.0535)	2.21 (1)	1.22 (0.1861)
ATP synthase subunit alpha (<i>AtpA</i>)	ZP_09994213	697	653	530	597	459	788	434	406	0.63 (1)	0.98 (0.0067)	1.78 (1)	1.26 (0.7481)
ATP synthase F1 subunit gamma (<i>AtpG</i>)	ZP_09994214	240	283	192	220	173	237	178	180	0.79 (0.5183)	1.00 (0.0228)	1.42 (0.9568)	1.36 (0.7072)
ATP synthase subunit beta (<i>AtpD</i>)	ZP_09994215	511	508	388	424	251	496	291	258	0.66 (1)	0.95 (0.0146)	2.05 (1)	1.19 (0.1632)
ATP synthase subunit epsilon (<i>AtpC</i>)	ZP_09994216	342	419	385	447	258	254	214	223	0.82 (0.4469)	1.01 (0.0111)	1.02 (0.0332)	1.40 (0.9369)

Table 10. Putative compound identification for features detected in *Prochlorococcus* str. MIT9313 DOM_p related to proteorhodopsin-based phototrophy. Putative identifications are based on ultrahigh-resolution mass information (confidence level ± 2 ppm) with additional verification based on MS/MS fragmentation information when possible. The difference in the detected mass and the putative identification in the METLIN database is indicated (Δppm) along with the median m/z and retention time for multiple sample injections. The mean intensity level of each MIT9313 DOM_p feature was divided by the mean intensity of any matching feature in the Pro99 medium control as is reported here as the fold change. Significant intensity differences were determined using Welch's t test (unequal variances) and only features for which p -values were ≤ 0.01 and fold changes were ≥ 4 were included in the analysis and shown here.

METLIN ID#	Putative name	Formula	Δppm	m/z	Retention time (min)	Fold change	p-Value
64007	3'-Hydroxy-geranylhydroquinone	C ₁₆ H ₂₂ O ₃	0	263.16414	30.75	1188	0.0052
2280	1-O-all-trans-retinoyl- β -glucuronic acid	C ₂₆ H ₃₆ O ₈	2	477.24723	19.16	881	0.0100
41945	1 α ,21-dihydroxy-20-oxo-22,23,24,25,26,27-hexanorvitamin D3	C ₂₁ H ₃₀ O ₄	0	369.20368	33.79	442	0.0025
87169	(E,E)-Boviquinone 3	C ₂₁ H ₂₈ O ₄	0	345.20610	30.87	122	0.0017
2279	5,6-Epoxyretinoic acid	C ₂₀ H ₂₈ O ₃	0	339.19307	38.39	87	0.0014
41944	1 α -hydroxy-20-oxo-22,23,24,25,26,27-hexanorvitamin D3	C ₂₁ H ₃₀ O ₃	0	353.20878	41.00	42	0.0080
95320	O-Geranylvanillin	C ₁₈ H ₂₄ O ₃	0	289.17983	30.00	34	0.0013
73539	Trans,trans-farnesyl phosphate	C ₁₅ H ₂₇ O ₄ P	1	325.15412	34.65	20	0.0093
88087	Methyl (9Z)-10'-oxo-6,10'-diapo-6-carotenoate	C ₂₀ H ₂₄ O ₃	0	313.17990	35.21	20	0.0080
6966	All-trans-retinoic acid	C ₂₀ H ₃₀ O ₂	0	325.21393	38.94	9	0.0059
41978	(24R,25S)-25,26-epoxy-1 α ,24-dihydroxy-27-norvitamin D3	C ₂₆ H ₄₀ O ₄	0	439.28183	44.12	5	0.0048

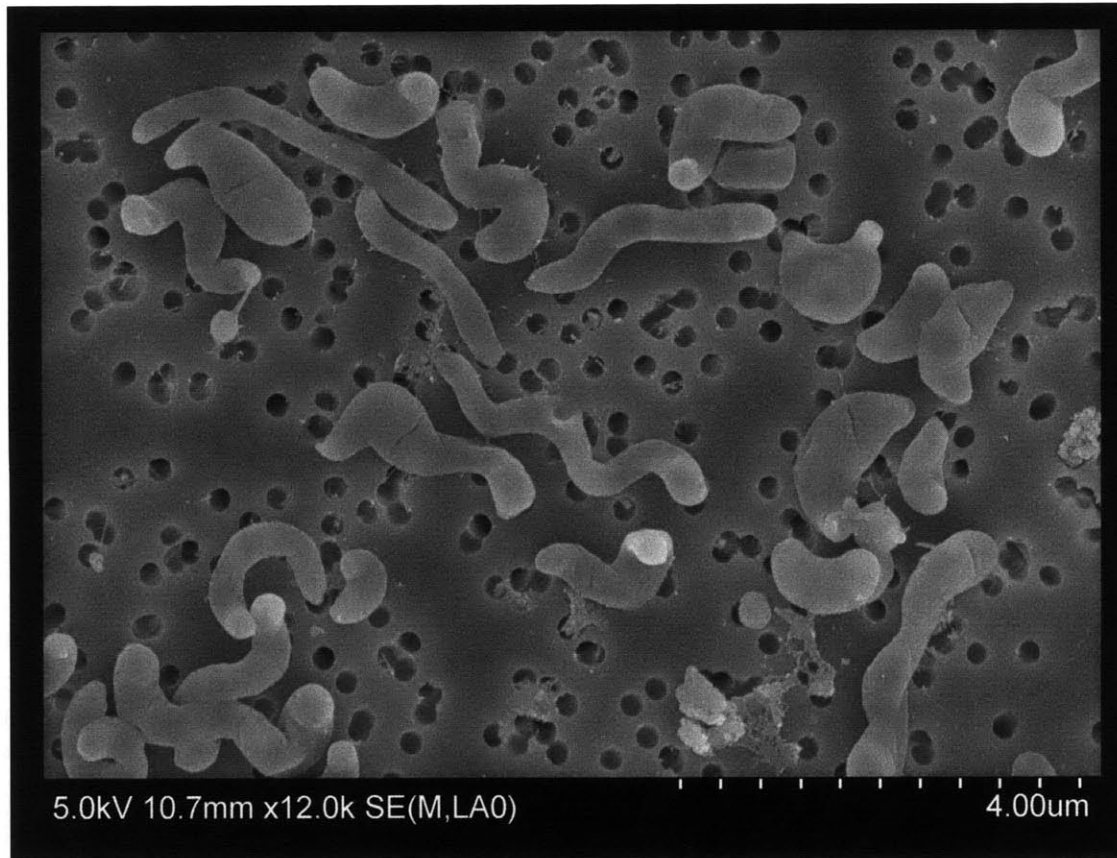


Figure 1. Scanning electron micrograph of HIMB30 cells displaying spirillum cell morphology. Scale bar is 4.0 μm (each hash, 0.4 μm). Image courtesy of M. Huggett and M. Rappé.

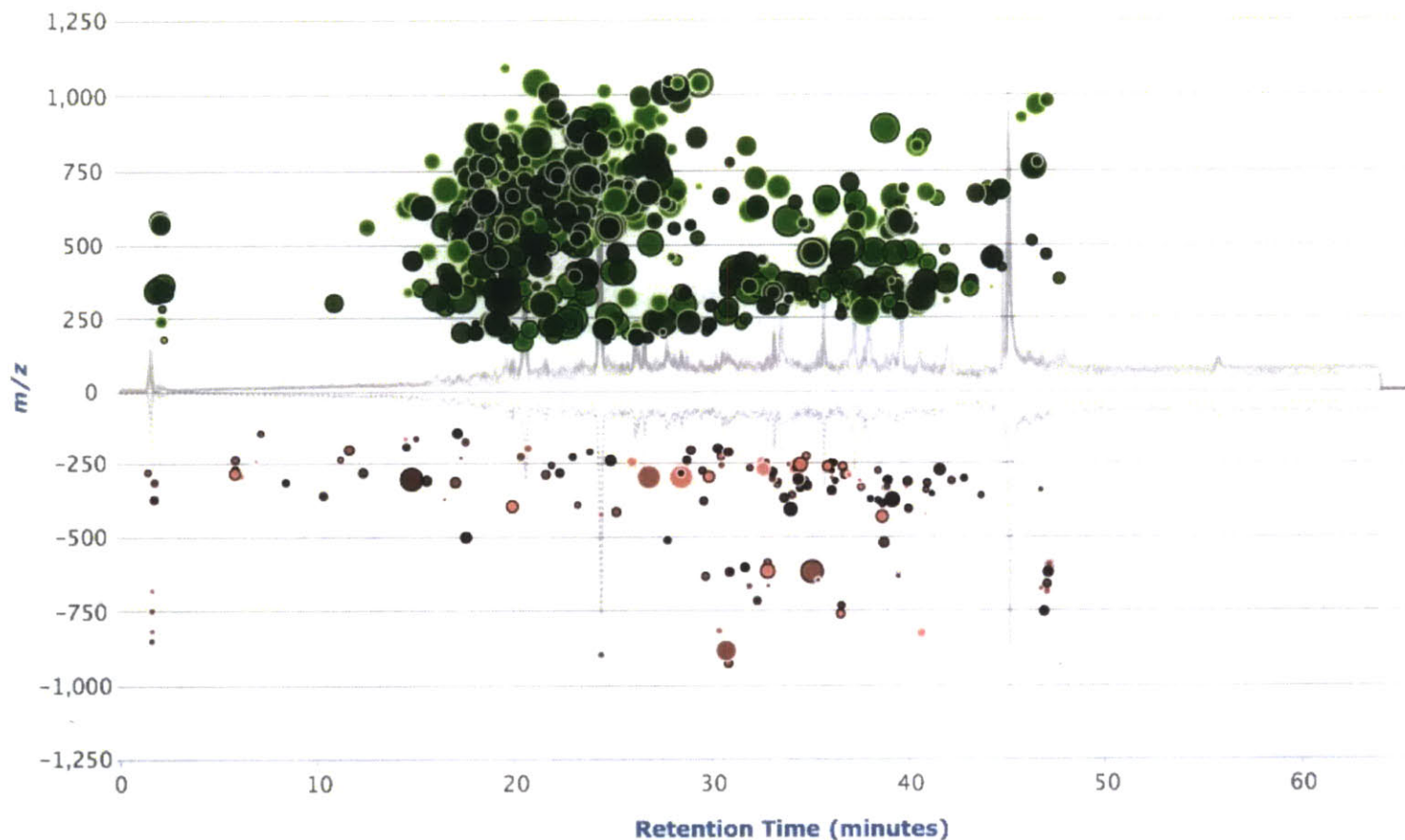


Figure 2. Cloud plot of MIT9313 DOM_p and Pro99 medium, 802 features depicted with p -value ≤ 0.01 and fold change ≥ 4 . Total ion currents for samples are shown in the background. Each circle indicates a feature with a unique m/z and retention time combination (displayed on the Y- and X-axis respectively). Features with increased intensity in the MIT9313 culture sample are shown on the top in green, while features with greater intensity in the Pro99 medium control are shown on the bottom in red. The area of each feature circle is proportional to the log fold change (i.e. larger area corresponds to greater fold change) and the color of each feature is proportional to the statistical significance of this change (Welch's t test, unequal variances), where brighter features are more significant (i.e. lower p -value) than darker features (i.e. higher p -value). Features outlined in black indicate putative identifications exist in the METLIN database.

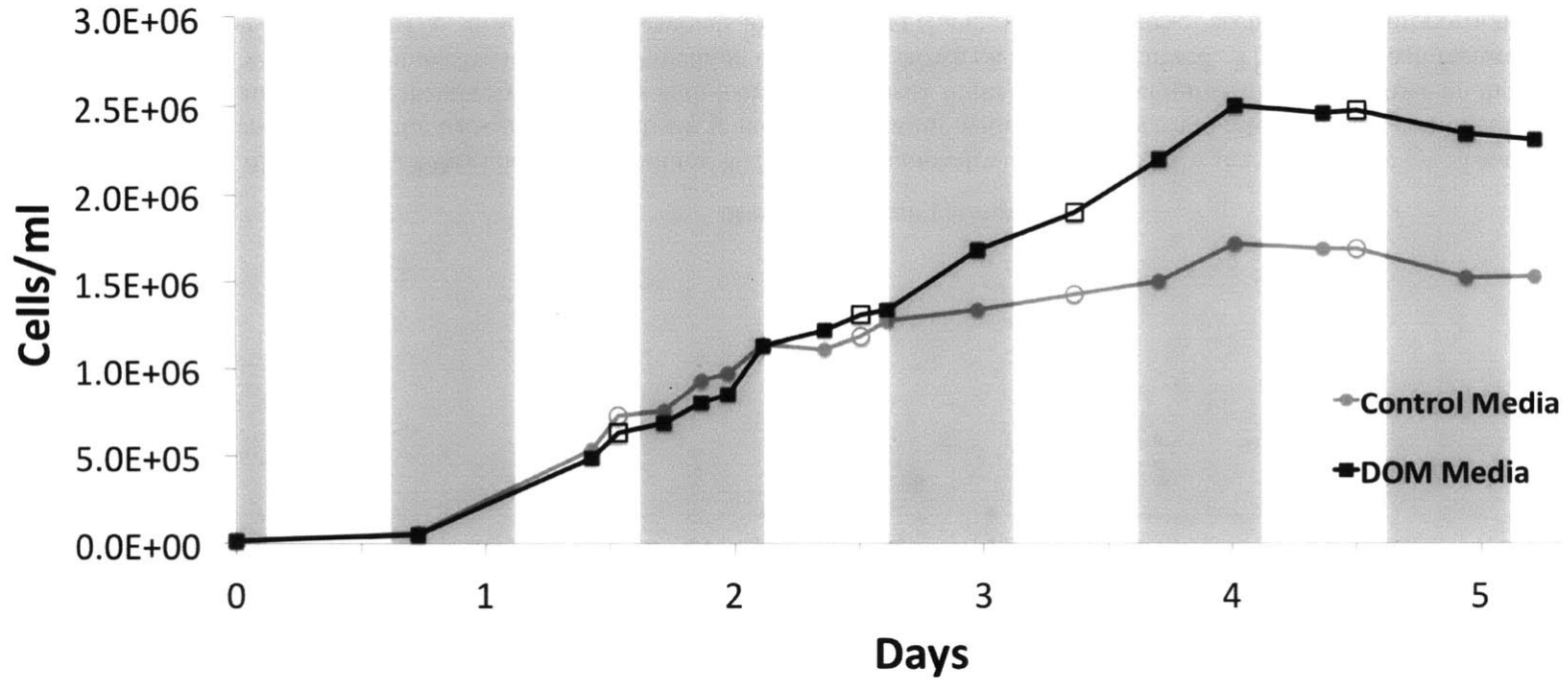


Figure 3. Cell enumerations determined via flow cytometry for HIMB30 grown in the presence (black squares) and absence (gray circles) of DOM_P produced by *Prochlorococcus* str. MIT9313. Open symbols indicate points at which samples were taken for RNA sequencing from each culture (i.e. T1, T2, T3 ad T4) and gray shading indicates periods of dark incubation.

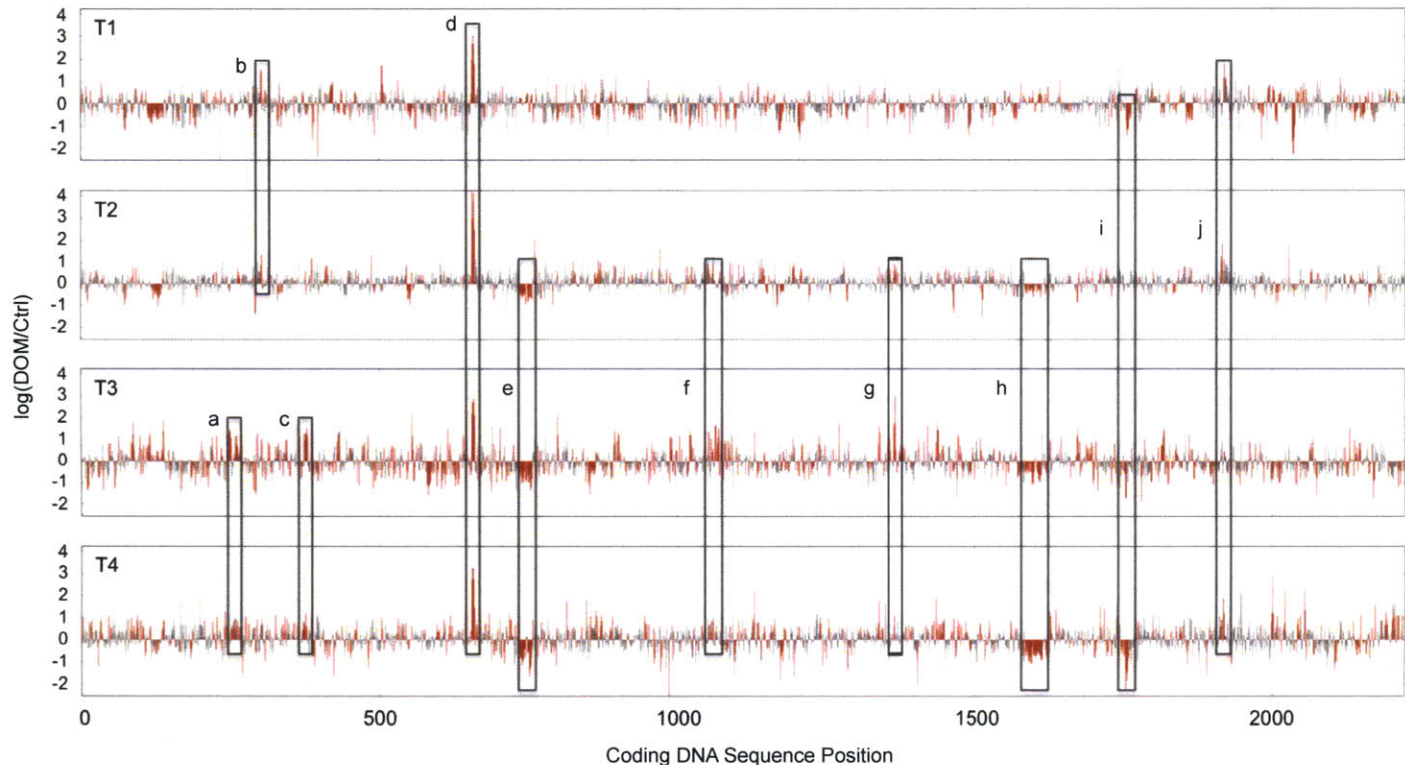


Figure 4. Linearized representation of the HIMB30 genome, where each of the 2,223 coding genes detected in this study are lined up according to genome position (non-coding regions are excluded). Each vertical line represents a coding gene and positive lines indicate overexpression of that gene in the DOM treatment, while negative lines indicate overexpression in the control. Genes that were significantly overexpressed (posterior likelihood ≥ 0.9) are colored in red and plots are shown for all four time points. Expression was calculated as the log of the ratio of read numbers for each sample that were first normalized to the total number of reads in each sample ($\log(\text{DOM}/\text{Ctrl})$). Regions of interest are highlighted in boxes and include genes involved in amino acid biosynthesis (a), the non-oxidative portion of the pentose phosphate pathway (b), tripartite ATP-independent periplasmic (TRAP) carboxylic acid transporters (c), ABC transporters for sugars and organic sulfur (d,j), proteorhodopsin and retinal biosynthesis (f), and ATP synthase (g) that were significantly overexpressed in the treatment and genes involved in motility and chemotaxis (e,h), and degradation pathways that liberate acetyl-CoA (i) that were significantly overexpressed in the control. The majority of clusters showing significant overexpression that are not highlighted belong to genes encoding for various ribosomal proteins.

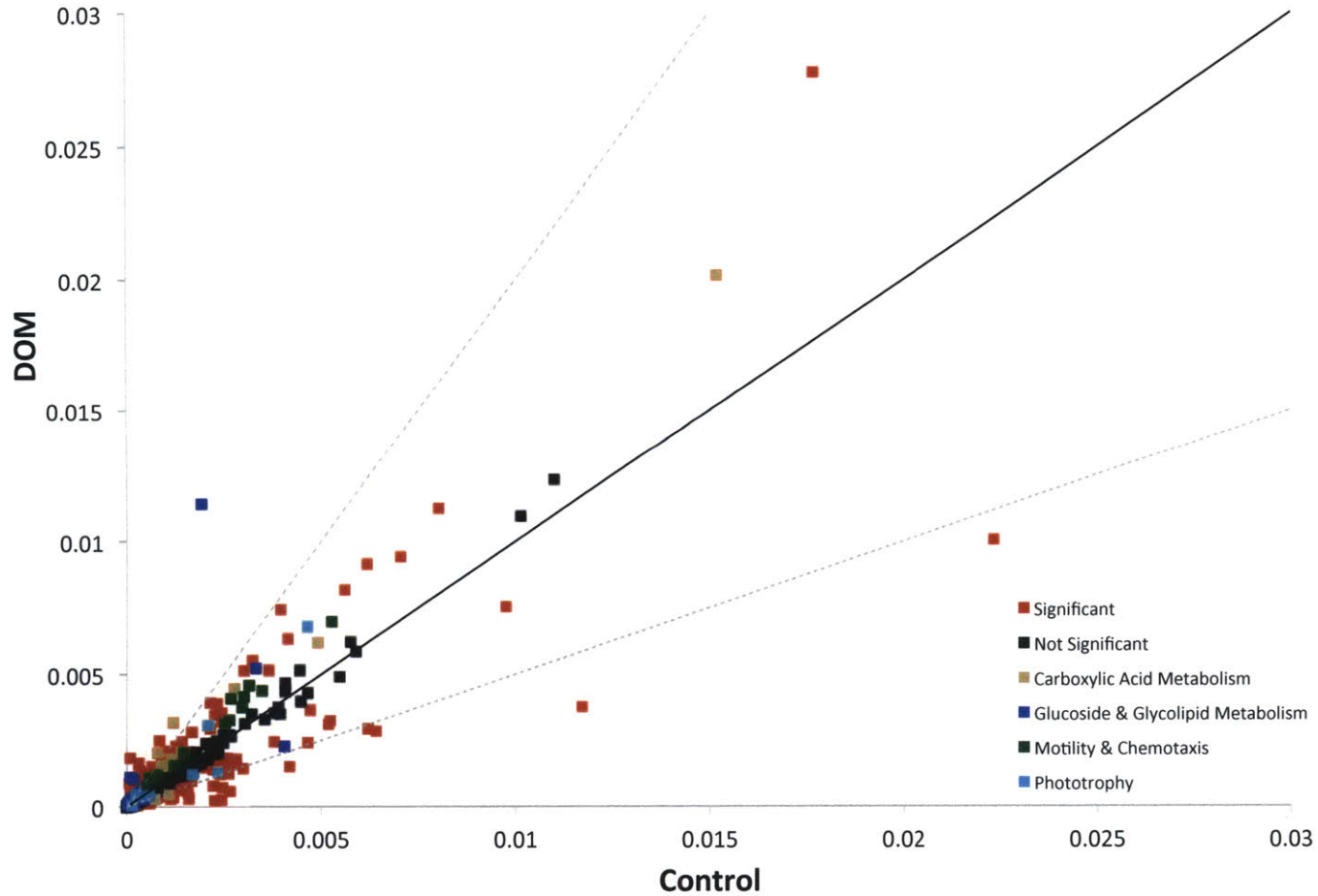


Figure 5. Scatter plot depicting differences in gene expression levels between the DOM treatment and control at T1. Each coding DNA sequence (CDS) is indicated by a square, where the position is determined by normalized expression of that CDS in each sample alongside 1:1 (solid) and 2-fold (broken) lines. CDSs that were not significantly overexpressed are shown in black. Significantly overexpressed genes of interest are color coded by group according to the figure legend.

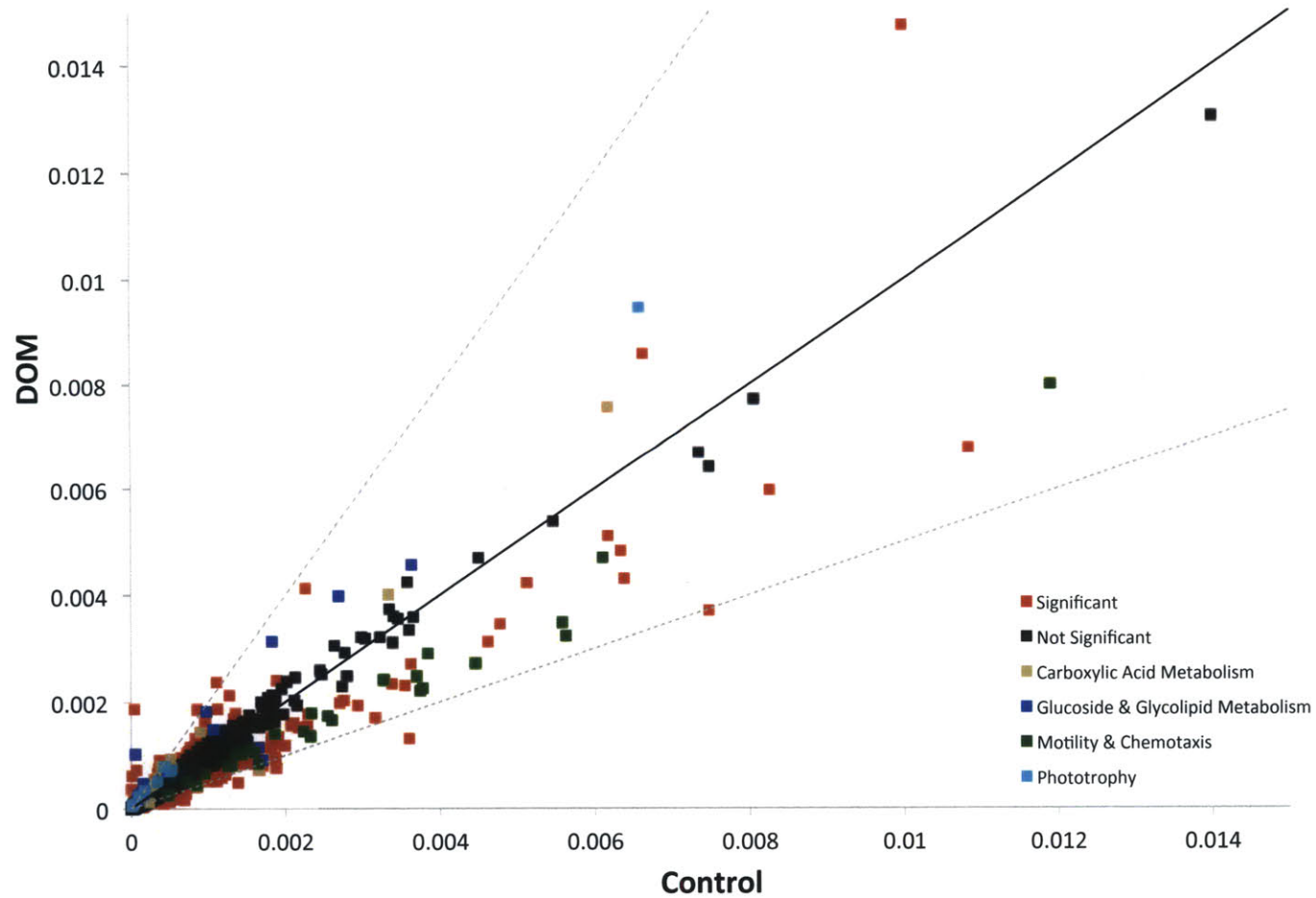


Figure 6. Scatter plot depicting differences in gene expression levels between the DOM treatment and control at T2. Each coding DNA sequence (CDS) is indicated by a square, where the position is determined by normalized expression of that CDS in each sample alongside 1:1 (solid) and 2-fold (broken) lines. CDSs that were not significantly overexpressed are shown in black. Significantly overexpressed genes of interest are color coded by group according to the figure legend.

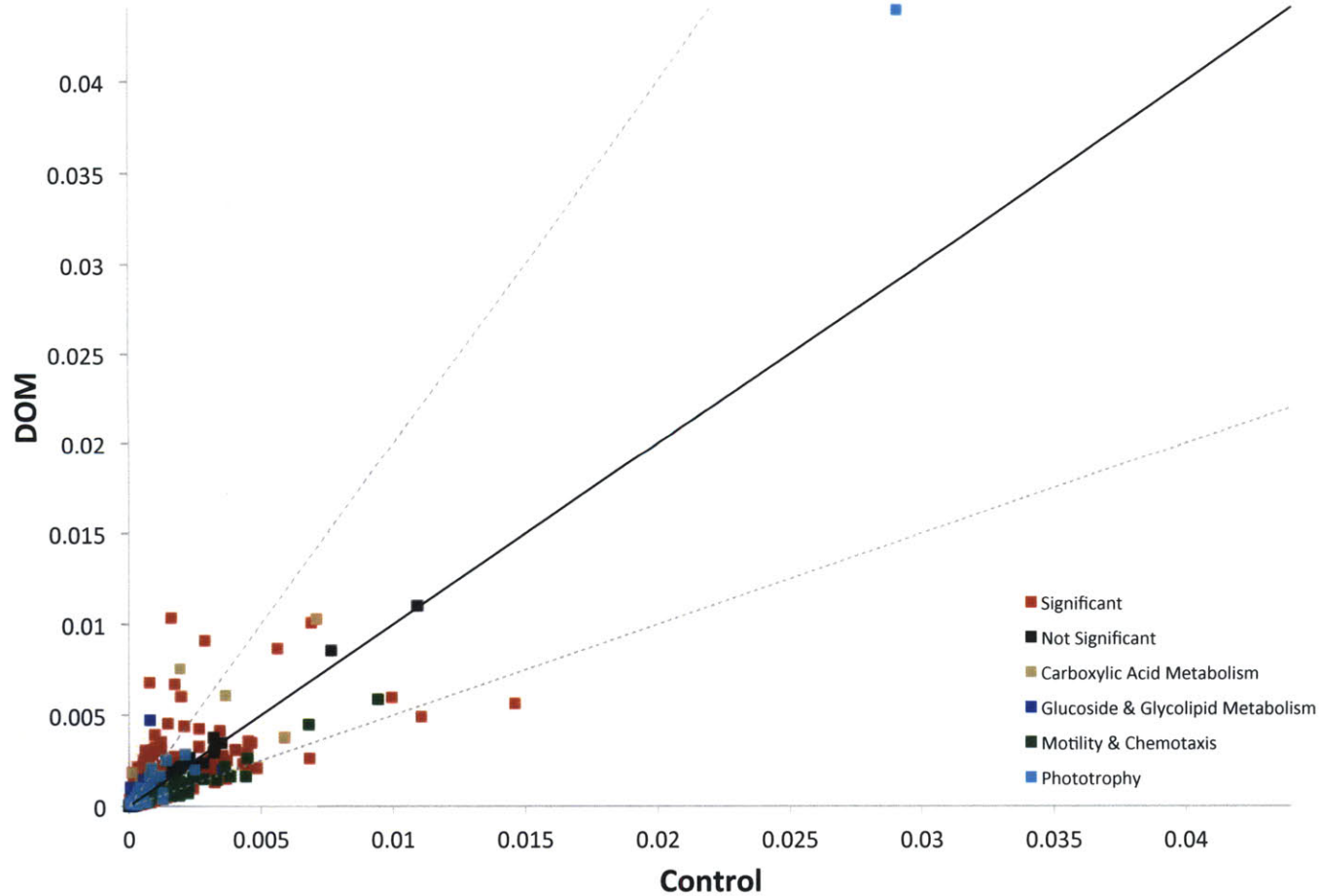


Figure 7. Scatter plot depicting differences in gene expression levels between the DOM treatment and control at T3. Each coding DNA sequence (CDS) is indicated by a square, where the position is determined by normalized expression of that CDS in each sample alongside 1:1 (solid) and 2-fold (broken) lines. CDSs that were not significantly overexpressed are shown in black. Significantly overexpressed genes of interest are color coded by group according to the figure legend. The cyan CDS displaying high expression in both samples and significant overexpression in the treatment encodes proteorhodopsin.

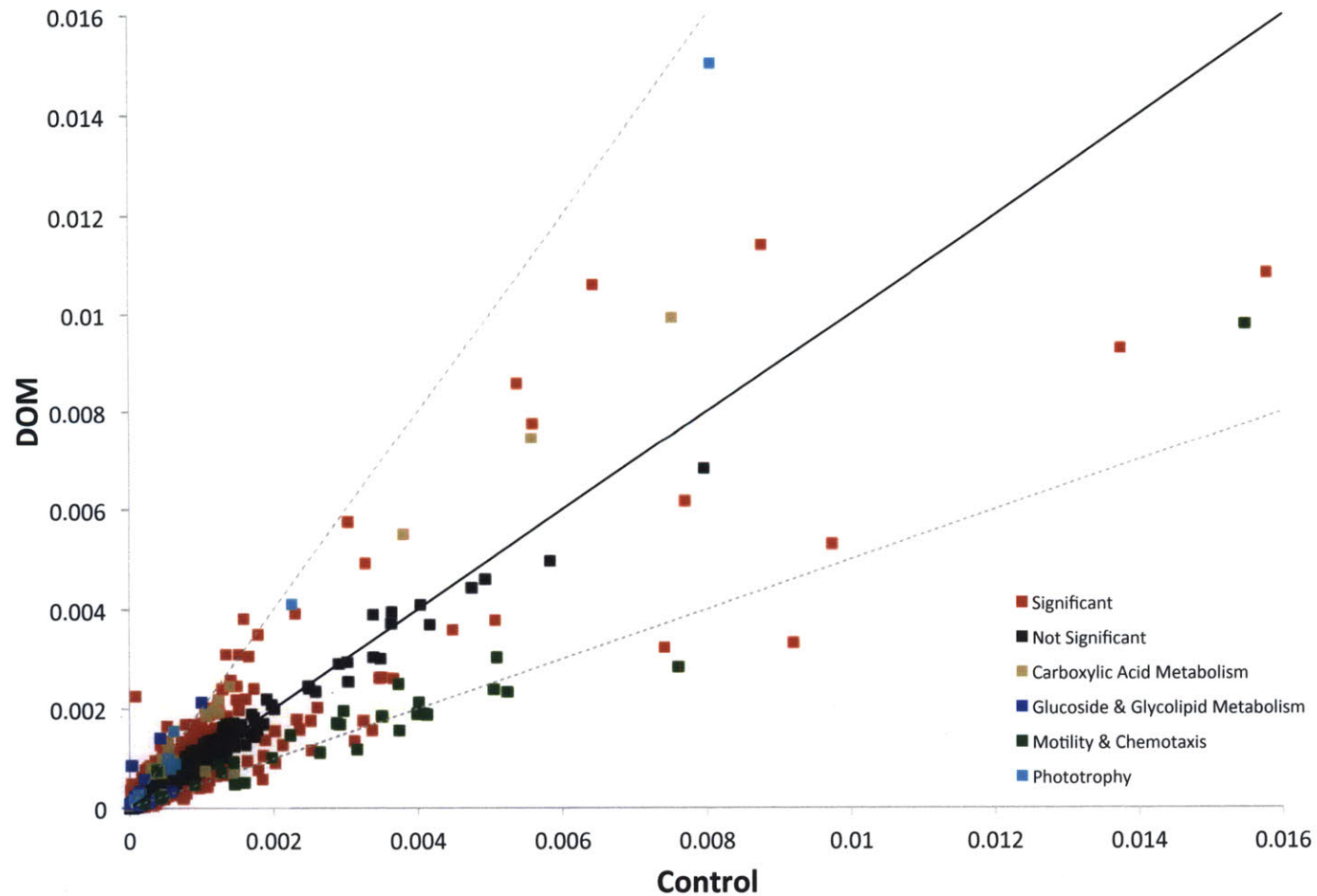


Figure 8. Scatter plot depicting differences in gene expression levels between the DOM treatment and control at T4. Each coding DNA sequence (CDS) is indicated by a square, where the position is determined by normalized expression of that CDS in each sample alongside 1:1 (solid) and 2-fold (broken) lines. CDSs that were not significantly overexpressed are shown in black. Significantly overexpressed genes of interest are color coded by group according to the figure legend.

TIC for m/z 287.25802

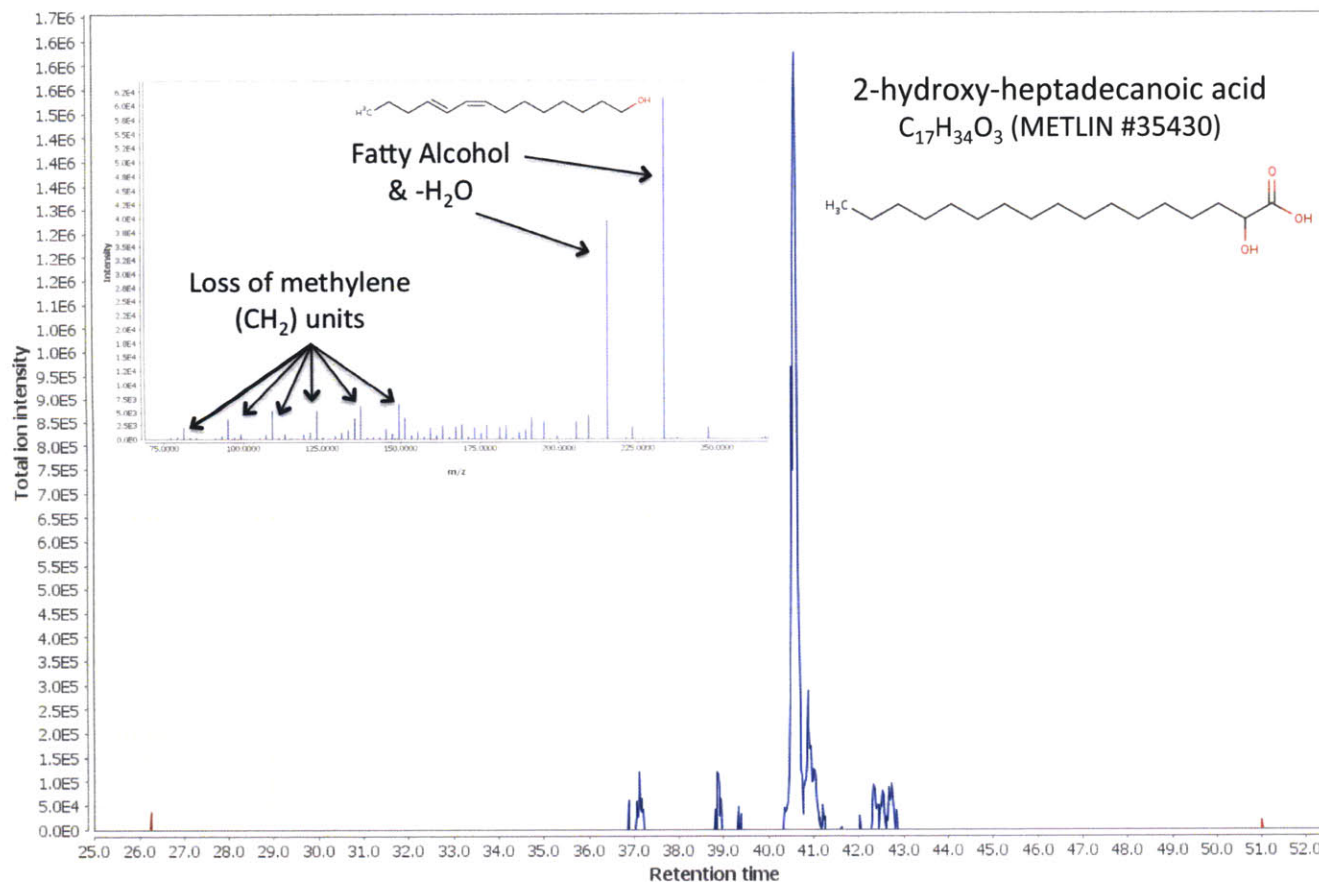


Figure 9. Mass chromatogram for a feature detected in DOM_p derived from MIT9313 (m/z 287.25802), putatively identified in the METLIN database as a fatty acid (1-hydroxy-heptadecanoic acid). Both the MIT9313 sample (blue) and the Pro99 control (red) ion currents are shown, although the control signal is mostly below detection. MS/MS fragmentation of this feature (inset) supports the putative identification, revealing fragments of a fatty alcohol (M-propynal), dehydrated fatty alcohol (M-propynal- H_2O), and the successive loss of methylene (CH_2) units.

TIC for m/z 836.43531

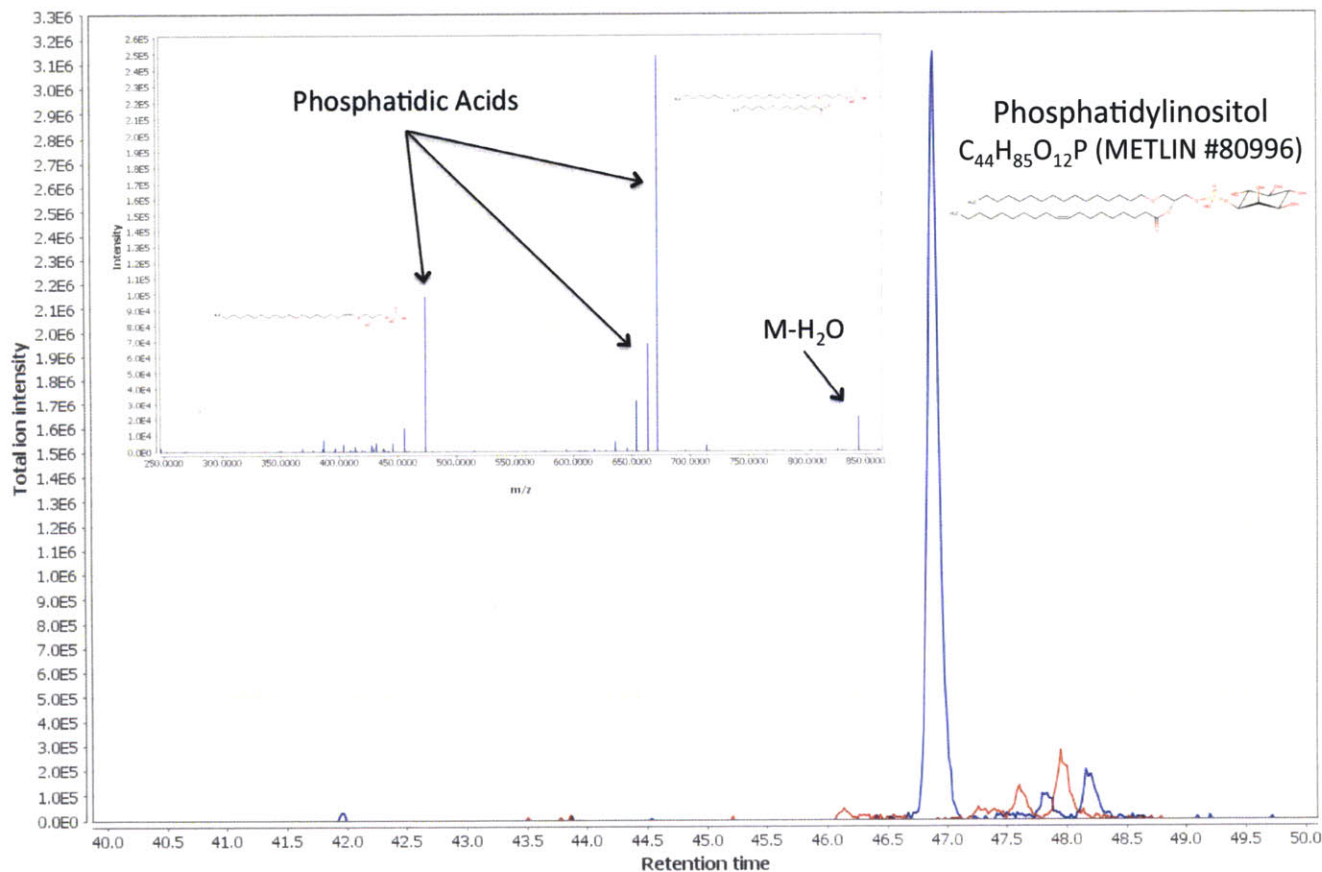


Figure 10. Mass chromatogram for a feature detected in DOM_P derived from MIT9313 (m/z 836.43531), putatively identified in the METLIN database as a glycerophospholipid (phosphatidylinositol). Both the MIT9313 sample (blue) and the Pro99 control (red) ion currents are shown. MS/MS fragmentation of this feature (inset) supports the putative identification, revealing several putative phosphatidic acid fragments.

TIC for m/z 325.21393

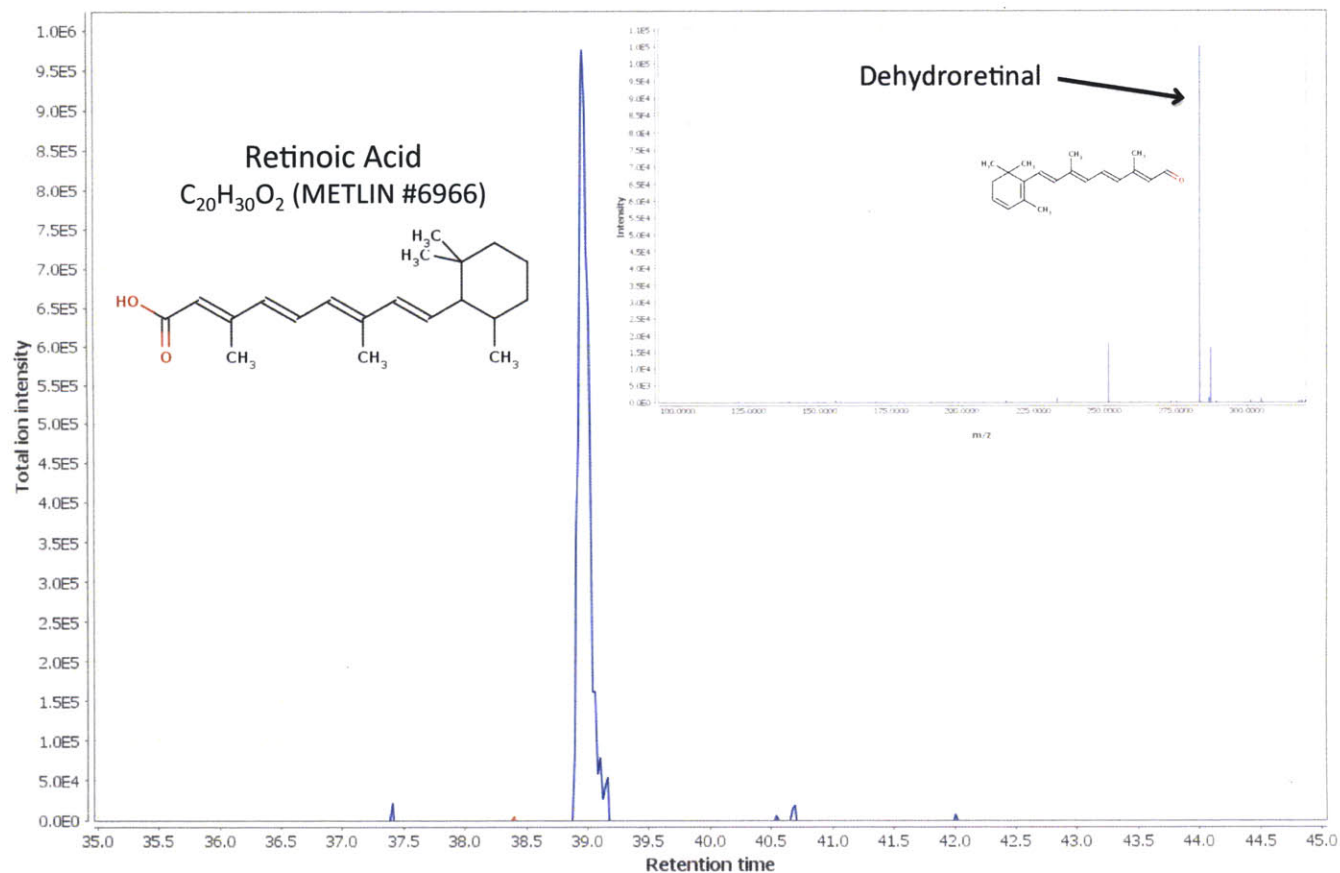


Figure 11. Mass spectral total ion current for a feature detected in DOM_p derived from MIT9313 (m/z 325.21393), putatively identified in the METLIN database as a retinoic acid. Both the MIT9313 sample (blue) and the Pro99 control (red) ion currents are shown, although the control signal is mostly below detection. MS/MS fragmentation of this feature (inset) supports the putative identification, revealing a putative dehydroretinal fragment.

References

- Avril B (2002) DOC dynamics in the northwestern Mediterranean Sea (DYFAMED site). *Deep Sea Research Part II: Topical Studies in Oceanography* 49:2163–2182
- Azam F, Malfatti F (2007) Microbial structuring of marine ecosystems. *Nature Publishing Group* 5:782–791
- Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA, Thingstad F (1983) The Ecological Role of Water-Column Microbes in the Sea. *Mar Ecol Prog Ser* 10:257–263
- Baran R, Bowen BP, Northen TR (2011) Untargeted metabolic footprinting reveals a surprising breadth of metabolite uptake and release by *Synechococcus* sp. PCC 7002. *Mol BioSyst* 7:3200–3206
- Becker JW, Brandon ML, Rappé MS (2007) Cultivating microorganisms from dilute aquatic environments: melding traditional methodology with new cultivation techniques and molecular methods. In: Hurst CJ (ed) *Manual of Environmental Microbiology*, 3rd edn. ASM Press, p 399–406
- Bertilsson S, Berglund O, Pullin MJ, Chisholm SW (2005) Release of dissolved organic matter by *Prochlorococcus*. *Vie et Milieu* 55:225–232
- Brunskill EW, Bayles KW (1996) Identification of *LytSR*-regulated genes from *Staphylococcus aureus*. *Journal of Bacteriology* 178:5810–5812
- Cailliau C, Claustre H, Vidussi F, Marie D, Vaulot D (1996) Carbon biomass, and gross growth rates as estimated from ^{14}C pigment labeling, during photoacclimation in *Prochlorococcus* CCMP 1378. *Mar. Ecol. Prog. Ser.* 145:209–221
- Chambers MC, Maclean B, Burke R, Amode D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J, Hoff K, Kessner D, Tasman N, Shulman N, Frewen B, Baker TA, Brusniak M-Y, Paulse C, Creasy D, Flashner L, Kani K, Moulding C, Seymour SL, Nuwaysir LM, Lefebvre B, Kuhlmann F, Roark J, Rainer P, Detlev S, Hemenway T, Huhmer A, Langridge J, Connolly B, Chadick T, Holly K, Eckels J, Deutsch EW, Moritz RL, Katz JE, Agus DB, MacCoss M, Tabb DL, Mallick P (2012) A cross-platform toolkit for mass spectrometry and proteomics. *Nature Publishing Group* 30:918–920
- Chavez FP, Messié M, Pennington JT (2011) Marine Primary Production in Relation to Climate Variability and Change. *Annu Rev Marine Sci* 3:227–260
- Choi A, Lee K, Oh HM, Feng J, Cho JC (2010) *Litoricola marina* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* 60(6): 1303–1306

- Church MJ, Ducklow HW, Karl DM (2002) Multiyear increases in dissolved organic matter inventories at Station ALOHA in the North Pacific Subtropical Gyre. *Limnol Oceanogr* 47:1–10
- Costa V, Angelini C, De Feis I, Ciccodicola A (2010) Uncovering the Complexity of Transcriptomes with RNA-Seq. *Journal of Biomedicine and Biotech.* 2010:1–19
- Cottrell MT, Kirchman DL (2000) Natural assemblages of marine proteobacteria and members of the Cytophaga-Flavobacter cluster consuming low-and high-molecular-weight dissolved organic matter. *Applied and Environmental Microbiology* 66:1692–1697
- Croucher NJ, Thomson NR (2010) Studying bacterial transcriptomes using RNA-seq. *Current Opinion in Microbiology* 13:619–624
- del Giorgio PA, Cole JJ, Cimbleris A (1997) Respiration rates in bacteria exceed phytoplankton production in unproductive aquatic systems. *Nature* 385:148–151
- DeLong EF, Bèjà O (2010) The Light-Driven Proton Pump Proteorhodopsin Enhances Bacterial Survival during Tough Times. *PLoS Biol* 8:e1000359
- DeLong EF, Taylor LT, Marsh TL, Preston CM (1999) Visualization and enumeration of marine planktonic archaea and bacteria by using polyribonucleotide probes and fluorescent in situ hybridization. *Applied and Environmental Microbiology* 65:5554–5563
- Forward JA, Behrendt MC, Wyborn NR, Cross R, Kelly DJ (1997) TRAP transporters: a new family of periplasmic solute transport systems encoded by the *dctPQM* genes of *Rhodobacter capsulatus* and by homologs in diverse gram-negative bacteria. *Journal of Bacteriology* 179:5482–5493
- Frias-Lopez J, Shi Y, Tyson GW, Coleman ML, Schuster SC, Chisholm SW, DeLong EF (2008) Microbial community gene expression in ocean surface waters. *Proc Natl Acad Sci USA* 105:3805–3810
- Giannoukos G, Ciulla DM, Huang K, Haas BJ, Izard J, Levin JZ, Livny J, Earl AM, Gevers D, Ward DV, Nusbaum C, Birren BW, Gnirke A (2012) Efficient and robust RNA-seq process for cultured bacteria and complex community transcriptomes. *Genome Biology* 13:r23
- Giovannoni SJ, Bibbs L, Cho J-C, Stapels MD, Desiderio R, Vergin KL, Rappé MS, Laney S, Wilhelm LJ, Tripp HJ, Mathur EJ, Barofsky DF (2005) Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature* 438:82–85

- Gómez-Consarnau L, Akram N, Lindell K, Pedersen A, Neutze R, Milton DL, González JM, Pinhassi J (2010) Proteorhodopsin Phototrophy Promotes Survival of Marine Bacteria during Starvation (JA Eisen, Ed.). PLoS Biol 8:e1000358
- Gómez-Consarnau L, González JM, Coll-Lladó M, Gourdon P, Pascher T, Neutze R, Pedrós-Alió C, Pinhassi J (2007) Light stimulates growth of proteorhodopsin-containing marine Flavobacteria. Nature 445:210–213
- Grossart H-P, Riemann L, Azam F (2001) Bacterial motility in the sea and its ecological implications. Aquat Microb Ecol 25:247–258
- Hansell DA (1993) Results and observations from the measurement of DOC and DON in seawater using a high-temperature catalytic oxidation technique. Marine Chemistry 41:195–202
- Hansell DA, Carlson CA (1998) Net community production of dissolved organic carbon. Global Biogeochem Cycles 12:443–453
- Hardcastle TJ (2010) baySeq: Empirical Bayesian analysis of patterns of differential expression in count data. BMC Bioinformatics 11:442
- Huggett MJ, Rappé MS (2012) Genome Sequence of Strain HIMB30, a Novel Member of the Marine Gammaproteobacteria. Journal of Bacteriology 194:732–733
- Keil RG, Kirchman DL (1999) Utilization of dissolved protein and amino acids in the northern Sargasso Sea. Aquat Microb Ecol 18:293–300
- Keiler KC (2008) Biology of trans-Translation. Annu Rev Microbiol 62:133–151
- Kim H, Choo YJ, Cho JC (2007) *Litoricolaceae* fam. nov., to include *Litoricola lipolytica* gen. nov., sp. nov., a marine bacterium belonging to the order *Oceanospirillales*. International Journal of Systematic and Evolutionary Microbiology 57(8): 1793-1798
- Kimura H, Young CR, Martinez A, DeLong EF (2011) Light-induced transcriptional responses associated with proteorhodopsin-enhanced growth in a marine flavobacterium. The ISME Journal 5:1641–1651
- Klein AH, Shulla A, Reimann SA, Keating DH, Wolfe AJ (2007) The Intracellular Concentration of Acetyl Phosphate in Escherichia coli Is Sufficient for Direct Phosphorylation of Two-Component Response Regulators. Journal of Bacteriology 189:5574–5581
- Kloer DP (2005) The Structure of a Retinal-Forming Carotenoid Oxygenase. Science 308:267–269

- Kuhl C, Tautenhahn R, Bottcher C, Larson TR, Neumann S (2012) CAMERA: An integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Anal. Chem.* 84(1): 283-289
- Kujawinski EB, Longnecker K, Blough NV, Del Vecchio R, Finlay L, Kitner JB, Giovannoni SJ (2009) Identification of possible source markers in marine dissolved organic matter using ultrahigh resolution mass spectrometry. *Geochimica et Cosmochimica Acta* 73:4384–4399
- Lami R, Cottrell MT, Campbell BJ, Kirchman DL (2009) Light-dependent growth and proteorhodopsin expression by Flavobacteria and SAR11 in experiments with Delaware coastal waters. *Environmental Microbiology* 11:3201–3209
- Malmstrom RR, Cottrell MT, Elifantz H, Kirchman DL (2005) Biomass production and assimilation of dissolved organic matter by SAR11 bacteria in the Northwest Atlantic Ocean. *Applied and Environmental Microbiology* 71:2979–2986
- Martinez AA, Bradley ASA, Waldbauer JRJ, Summons RER, DeLong EFE (2007) Proteorhodopsin photosystem gene expression enables photophosphorylation in a heterologous host. *Proc Natl Acad Sci USA* 104:5590–5595
- McCarren J, DeLong EF (2007) Proteorhodopsin photosystem gene clusters exhibit co-evolutionary trends and shared ancestry among diverse marine microbial phyla. *Environmental Microbiology* 9:846–858
- McCarren J, Becker JW, Repeta DJ, Shi Y, Young CR, Malmstrom RR, Chisholm SW, DeLong EF (2010) Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. *Proc Natl Acad Sci USA* 107:16420–16427
- Mitchell JG, Pearson L, Bonazinga A, Dillon S, Khouri H, Paxinos R (1995) Long lag times and high velocities in the motility of natural assemblages of marine bacteria. *Applied and Environmental Microbiology* 61:877–882
- Mou X, Vila-Costa M, Sun S, Zhao W, Sharma S, Moran MA (2011) Metatranscriptomic signature of exogenous polyamine utilization by coastal bacterioplankton. *Environmental Microbiology Reports* 3:798–806
- Mulligan C, Fischer M, Thomas GH (2010) Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. *FEMS Microbiology Reviews* 35:68–86
- Mühling M (2011) On the culture-independent assessment of the diversity and distribution of *Prochlorococcus*. *Environmental Microbiology* 14:567–579

- Nelson CE, Carlson CA (2012) Tracking differential incorporation of dissolved organic carbon types among diverse lineages of Sargasso Sea bacterioplankton. *Environmental Microbiology* 14:1500–1516
- Partensky F, Hess WR, Vault D (1999) *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiology and Molecular Biology Reviews* 63:106–127
- Pluskal T, Castillo S, Villar-Briones A, Orešič M (2010) MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 11:395–406
- Pomeroy LR, Williams PJL, Azam F, Hobbie JE (2007) The Microbial Loop. *Oceanography* 20:28–33
- Poretsky RS, Sun S, Mou X, Moran MA (2010) Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon. *Environmental Microbiology* 12:616–627
- Prüss BM, Wolfe AJ (1994) Regulation of acetyl phosphate synthesis and degradation, and the control of flagellar expression in *Escherichia coli*. *Molecular Microbiology* 12:973–984
- Reed RH, Chudek JA, Foster R, Stewart WDP (1984) Osmotic adjustment in cyanobacteria from hypersaline environments. *Archives of Microbiology* 138:333–337
- Rinta-Kanto JM, Sun S, Sharma S, Kiene RP, Moran MA (2011) Bacterial community transcription patterns during a marine phytoplankton bloom. *Environmental Microbiology* 14:228–239
- Romera-Castillo C, Sarmiento H, Alvarez-Salgado XA, Gasol JM, Marrase C (2011) Net Production and Consumption of Fluorescent Colored Dissolved Organic Matter by Natural Bacterial Assemblages Growing on Marine Phytoplankton Exudates. *Applied and Environmental Microbiology* 77:7490–7498
- Ruch S, Beyer P, Ernst H, Al-Babili S (2004) Retinal biosynthesis in Eubacteria: in vitro characterization of a novel carotenoid oxygenase from *Synechocystis* sp. PCC 6803. *Molecular Microbiology* 55:1015–1024
- Sarmiento H, Gasol JM (2012) Use of phytoplankton-derived dissolved organic carbon by different types of bacterioplankton. *Environmental Microbiology* 14:2348–2360
- Shi Y, Tyson GW, DeLong EF (2009) Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column. *Nature* 459:266–269

- Shin S, Park C (1995) Modulation of flagellar expression in *Escherichia coli* by acetyl phosphate and the osmoregulator OmpR. *Journal of Bacteriology* 177:4696–4702
- Singh SC, Sinha RP, Hader DP (2002) Role of lipids and fatty acids in stress tolerance in cyanobacteria. *Acta Protozool.* 41:297-308
- Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, Custodio DE, Abagyan R, Siuzdak G (2005) METLIN: a metabolite mass spectral database. *Ther Drug Monit* 27:747–751
- Sowell SM, Norbeck AD, Lipton MS, Nicora CD, Callister SJ, Smith RD, Barofsky DF, Giovannoni SJ (2008) Proteomic Analysis of Stationary Phase in the Marine Bacterium “*Candidatus Pelagibacter ubique*.” *Applied and Environmental Microbiology* 74:4091–4100
- Sowell SM, Wilhelm LJ, Norbeck AD, Lipton MS, Nicora CD, Barofsky DF, Carlson CA, Smith RD, Giovanonni SJ (2009) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *The ISME Journal* 3:93–105
- Steglich C, Frankenberg-Dinkel N, Penno S, Hess WR (2005) A green light-absorbing phycoerythrin is present in the high-light-adapted marine cyanobacterium *Prochlorococcus* sp. MED4. *Environmental Microbiology* 7:1611–1618
- Steglich C, Futschik M, Rector T, Steen R, Chisholm SW (2006) Genome-Wide Analysis of Light Sensing in *Prochlorococcus*. *Journal of Bacteriology* 188:7796–7806
- Stewart FJ, Ottesen EA, DeLong EF (2010) Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. *The ISME Journal* 4:896–907
- Stingl U, Desiderio RA, Cho JC, Vergin KL, Giovannoni SJ (2007) The SAR92 Clade: an Abundant Coastal Clade of Culturable Marine Bacteria Possessing Proteorhodopsin. *Applied and Environmental Microbiology* 73:2290–2296
- Stocker R (2012) Marine Microbes See a Sea of Gradients. *Science* 338:628–633
- Tautenhahn R, Patti GJ, Rinehart D, Siuzdak G (2012a) XCMS Online: A Web-Based Platform to Process Untargeted Metabolomic Data. *Anal Chem* 84:5035–5039
- Tautenhahn R, Cho K, Uritboonthai W, Zhu Z, Patti GJ, Siuzdak G (2012b) An accelerated workflow for untargeted metabolomics using the METLIN database. *Nature Biotechnology* 30: 826-828

- Ting CS, Hsieh C, Sundararaman S, Mannella C, Marko M (2007) Cryo-Electron Tomography Reveals the Comparative Three-Dimensional Architecture of *Prochlorococcus*, a Globally Important Marine Cyanobacterium. *Journal of Bacteriology* 189:4485–4493
- Typas A, Banzhaf M, Gross CA, Vollmer W (2011) From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nature Publishing Group* 10:123–136
- Vila-Costa M, Rinta-Kanto JM, Sun S, Sharma S, Poretsky R, Moran MA (2010) Transcriptomic analysis of a marine bacterial community enriched with dimethylsulfoniopropionate. *The ISME Journal* 4:1410–1420
- Waldbauer JR, Rodrigue S, Coleman ML, Chisholm SW (2012) Transcriptome and Proteome Dynamics of a Light-Dark Synchronized Bacterial Cell Cycle (S Lin, Ed.). *PLoS ONE* 7:e43432
- Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10:57–63
- Warr SR, Reed RH, Stewart WD (1984) Osmotic adjustment of cyanobacteria: the effects of NaCl, KCl, sucrose and glycine betaine on glutamine synthetase activity in a marine and a halotolerant strain. *Journal of general microbiology* 130:2169–2175
- Wu X, Jiang J, Wan Y, Giesy JP, Hu J (2012) Cyanobacteria blooms produce teratogenic retinoic acids. *Proc Natl Acad Sci USA* 109:9477–9482
- Zhang Y, Ikeno Y, Qi W, Chaudhuri A, Li Y, Bokov A, Thorpe SR, Baynes JW, Epstein C, Richardson A, Van Remmen H (2009) Mice Deficient in Both Mn Superoxide Dismutase and Glutathione Peroxidase-1 Have Increased Oxidative Damage and a Greater Incidence of Pathology but No Reduction in Longevity. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 64A:1212–1220

CHAPTER FIVE

Summary and Future Directions

Since the recognition of microorganisms as important members of marine food webs (Pomeroy 1974, Azam et al. 1983), publications on microbial oceanography frequently begin with introductory comments regarding the importance of marine microbes in the global cycling of energy, carbon, nutrients and other biologically relevant material. With particular regard to carbon cycling, marine microbes are often classified as either mitigators (Williams et al. 2013) or exacerbators (Duarte et al. 2013) of anthropogenic climate change and there has been recent attention placed on the need to include microbial activities in climate models (DeLong & Harwood 2012).

Despite this attention, the draft of newest (2013) National Climate Assessment, which includes a chapter on ocean health, makes no mention of marine microbes, except in regards to human and coral pathogens. I do not believe this to be an oversight by the authors (indeed, one of the lead authors is a member of a National Science Foundation center devoted specifically to the study of microbial oceanography), but rather a consequence of our limited understanding of microbial organic matter cycling in the sea. The biogeochemical cycles that occur in our oceans and impact our climate are all influenced by microbial activity, yet we currently lack enough information about how marine microbes produce and consume organic matter to include them in our assessments.

In this thesis, I utilized a cultivation-based approach combining biological and chemical techniques to develop simplified systems aimed at addressing several longstanding questions regarding marine microbes and dissolved organic matter (DOM) cycling, which included:

How are microbial diversity and marine DOM composition related?

How are DOM composition and microbial diversity related to DOM lability?

How (physiologically) is DOM metabolized by marine microbes?

Insight into these questions will be essential if we are to include marine microbes in climate models and at the same time, will provide basic knowledge regarding how microbes interact with each other and their surrounding environment for a richer understanding of the marine ecosystem. Information regarding the last question has particular value for a deeper understanding of microbial physiology and is a promising area of research for generating new hypotheses regarding how microbes survive in various marine habitats and under changing environmental conditions.

In Chapter 2, I assayed the composition of DOM produced by a phylogenetically diverse array of marine phytoplankton cultures to investigate if and how marine DOM composition depends on the identity of its producer. Results showed that while each phytoplankton strain tested produced a highly unique suite of DOM, there was a tendency for more closely related phytoplankton to produce more highly similar suites of DOM. This relationship appears to extend beyond the simple qualitative presence/absence of particular DOM components to a quantitative comparison as well, indicating that the amount of different DOM components produced is also related to producer identity.

These findings reveal that the chemical composition of marine DOM, and therefore the substrates available to surrounding heterotrophic bacterioplankton, depends largely on the organisms that produce it, a concept that has often been assumed but not conclusively demonstrated previously. These findings also imply that factors affecting marine phytoplankton community composition will likely vary the composition (and potentially the lability) of the surrounding DOM pool.

A logical next step was to utilize the same phytoplankton-derived DOM that was chemically interrogated in Chapter 2 as a medium amendment in cultures of various heterotrophic bacterioplankton. In Chapter 3, a large screening matrix of more than a dozen bacterioplankton strains supplied with varying doses of DOM from multiple phytoplankton and natural seawater sources was set up in triplicate alongside controls and monitored at a high temporal resolution via flow cytometry for changes in growth rate. This matrix was repeated in various forms on multiple occasions to identify robust responses for specific DOM-microbe combinations. DOM additions were carried out using dissolved organic carbon increases in the low micromolar range in order to mimic fluctuations that are known to occur in nature, which is not always the case in DOM amendment experiments, particularly when “model” DOM compounds are being used. Growth responses ranged from suppressed to enhanced growth and lability was a factor of both the DOM source and the identity of the consumer. Many of the bacterioplankton strains tested exhibited enhanced growth when supplied with DOM produced by *Prochlorococcus* str. MIT9313, including several strains from the SAR11 clade, demonstrating a potential trophic connection for these two numerically dominant groups.

The affect of DOM polarity on its reactivity was also explored using polarity fractions of *Prochlorococcus*-derived DOM and it was found that more polar fractions are generally more bioactive.

An additional benefit of the screening matrix described above was the identification of DOM-microbe target combinations for further study. In Chapter 4, one such target combination was used to study the underlying physiology of microbial DOM metabolism via a novel combined transcriptomic and untargeted metabolomic approach. The growth of strain HIMB30, an isolate representing a novel family in the *Oceanospirillales* order, was observed to be reproducibly enhanced by additions of DOM derived from *Prochlorococcus* str. MIT9313. Direct sequencing of HIMB30 protein coding genes at multiple phases of growth in the presence and absence of the DOM revealed a suite of physiological responses involved in the uptake and metabolism of individual components. Transport and utilization of carboxylic acids for anabolic processes including gluconeogenesis, the glyoxylate cycle, and the non-oxidative portion of the pentose phosphate pathway was apparent in the DOM treatment. Transport and utilization of glucosides, glycolipids, and amino sugars for biosynthesis of cell wall and cell membrane material was also identified in the treatment.

Completely unexpected signals due to the DOM addition were also found including reduced motility and chemotaxis and increased proteorhodopsin-based phototrophy. An untargeted metabolomic assessment of DOM from strain MIT9313 using ultrahigh-resolution mass spectrometry revealed putative identifications for many compounds that supported the HIMB30 transcriptome signals including an abundance of

fatty acids, glucosides, terpenoids, and several retinoic acids. Collectively, these findings highlight metabolic pathways involved in the consumption of environmentally relevant DOM that result in increased growth efficiency and describe a model platform to chemically identify reactive compounds in DOM from a known origin, thereby tracking labile DOM all the way from biological source to sink.

Efforts were also made to investigate the turnover of environmentally relevant DOM sources by mixed microbial assemblages in microcosm experiments as discussed in the appendices. Once again, the use of RNA sequencing (in this case, a metatranscriptomic approach) was shown to be a valuable tool for uncovering metabolic pathways pertinent to microbial degradation of marine DOM. Oligotrophic microbial communities from the North Pacific subtropical gyre were found to employ syntrophic degradation of concentrated high-molecular weight DOM and the use of dissolved organic nitrogen compounds present in both high-molecular weight and *Prochlorococcus*-derived DOM was found to be an important alleviator of nitrogen stress in these communities. The microbial assemblages employed in the two microcosm experiments did not respond to similar DOM additions in the same manner, which highlights the functional complexity of these natural communities and likely also the importance of founding community composition in DOM turnover.

Implications

The findings discussed throughout this thesis indicate that the composition of marine DOM is closely related to the identity of its producer and that phylogenetically

similar marine phytoplankton produce DOM of similar composition. In addition, the reactivity of marine DOM was shown to depend both on the identity of its producer and the identity of its consumer, indicating that DOM turnover is related to its composition and to bacterioplankton physiology. Finally, a novel approach employing transcriptomics and untargeted metabolomics was shown to be a successful method for uncovering metabolic pathways involved in the consumption of DOM and revealed the importance of carboxylic acids and glucosides in secondary production as well as novel information regarding the effects of organic nutrients on motility and phototrophy. Several of the findings described in this thesis are inconsistent with previous work, including the consumption of DOM sources by particular taxa as discussed in Chapter 3 and paradigms that exist for other microbial systems regarding organic carbon and proteorhodopsin-based phototrophy discussed in Chapter 4. Much of these discrepancies may be explained by the specific efforts made throughout this thesis to conduct manipulation experiments with environmentally relevant DOM sources at reasonable quantities. Conclusions drawn from manipulations in which specific model compounds are supplied in quantities that are likely orders of magnitude greater than natural concentrations ever reach may elicit substantial microbial responses, however the relevance of these responses to the marine environment needs to be closely evaluated.

Collectively, this thesis supports a scenario more akin to the image on the right of figure 2 in Chapter 1, where individual marine phytoplankton types produce and release unique suites of material into the marine DOM pool, portions of which are then preferentially consumed by different heterotrophic bacterioplankton taxa. This scenario

implicates DOM as a driver of marine microbial diversity by forming a direct link between autotrophic and heterotrophic organisms. Thus, compositional changes in DOM may be responsible for many of the spatial and temporal variations observed in marine microbial communities (Giovannoni & Vergin 2012) and this thesis suggests that greater efforts should be made by microbial oceanographers to incorporate chemical analyses of DOM into both time-series and short term manipulation studies.

Future directions

As is often the case with basic research, for every question answered by this thesis, several new questions arose. The strong connection between phytoplankton phylogeny and DOM composition identified in Chapter 2 is clouded a bit by the high degree of variation in DOM composition among biological replicate cultures. Whether these differences arise from uncontrollable variations in methodology or minor variations in growth conditions is unknown and should be further evaluated. Previous work has demonstrated that the composition of DOM produced by some marine phytoplankton in culture depends heavily on growth phase (Barofsky et al. 2009) and other factors (e.g. temperature, light, pH, nutrient levels) likely also play a role. Exploring how changes in these growth conditions affect the composition and reactivity of phytoplankton-derived DOM would be a useful undertaking, particularly for manipulations that mimic future climate scenarios. The application of additional chemical analyses such as ultrahigh-resolution mass spectrometry, tandem mass spectrometry, and nuclear magnetic resonance spectroscopy to make putative identifications of features found to be unique to

certain phytoplankton groups has the potential to identify novel biomarkers and uncover valuable information regarding phytoplankton physiology and the composition of marine DOM.

The growth screening assays conducted in Chapter 3 could easily be expanded to incorporate DOM from and cultivable autotroph and evaluate consumption by any cultivable heterotroph or mixotroph. As evolving isolation methods for marine microbes continue to bring more representative organisms into the laboratory (Stingl et al. 2007), the possibilities for screening expand greatly. Growth assays could be expanded to include syntrophic interactions through the use of co-cultures or by growing bacterioplankton strains in spent DOM media obtained from a different strain in an effort to examine successive degradation patterns. Chemical analyses such as those mentioned above could also be performed during screening assays to assess which DOM components are consumed by specific organisms. These experiments could be aided by the screening of individual DOM polarity fractions as described in Chapter 3, as this approach can greatly reduce the number of DOM features potentially responsible for a particular growth response.

Finally, the large datasets generated in Chapter 4 from both RNA sequencing and ultrahigh-resolution mass spectrometry provide a significant amount of information that could be analyzed without any additional experimental effort. While most of the major differential expression signals have been discussed here along with supporting chemical identifications, more minor signals may exist in both datasets that reveal new information regarding the consumption of *Prochlorococcus*-derived DOM by HIMB30. As the

approach outlined in Chapter 4 proved to be successful at uncovering novel information regarding both DOM composition and microbial physiology pertinent to DOM turnover, an obvious next step would be to employ this same approach with additional DOM-microbe combinations of interest identified in Chapter 3. This would likely result in novel information regarding the composition of DOM produced by marine phytoplankton other than *Prochlorococcus* str. MIT9313 and microbial physiology information related to the turnover of marine DOM by strains other than HIMB30. As additional connections between strains involved in the production and consumption of marine DOM are revealed, it will become increasingly important to look for these relationships among related microorganisms that co-occur in the marine environment in order to put these cultivation-based findings in a meaningful context.

References

- Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA, Thingstad F (1983) The Ecological Role of Water-Column Microbes in the Sea. *Mar Ecol Prog Ser* 10:257–263
- Barofsky A, Vidoudez C, Pohnert G (2009) Metabolic profiling reveals growth stage variability in diatom exudates. *Limnol Oceanogr-Meth* 7:382–390
- DeLong EF, Harwood CS (2012) Incorporating microbial processes into climate models (S Chisholm, DM Karl, MA Moran, TM Schmidt, JM Tiedje, KK Treseder, and AZ Worden, Eds.). American Society for Microbiology
- Duarte CM, Regaudie-de-Gioux A, Arrieta JM, Delgado-Huertas A, Agustí S (2013) The Oligotrophic Ocean Is Heterotrophic. *Annu Rev Marine Sci* 5:551–569
- Giovannoni SJ, Vergin KL (2012) Seasonality in Ocean Microbial Communities. *Science* 335:671–676
- Pomeroy LR (1974) The ocean's food web, a changing paradigm. *Bioscience*:499–504
- Stingl U, Tripp HJ, Giovannoni SJ (2007) Improvements of high-throughput culturing yielded novel SAR11 strains and other abundant marine bacteria from the Oregon coast and the Bermuda Atlantic Time Series study site. *The ISME Journal*
- Williams PJLB, Quay PD, Westberry TK, Behrenfeld MJ (2013) The Oligotrophic Ocean Is Autotrophic. *Annu Rev Marine Sci* 5:535–549

APPENDICES

Dissolved organic matter consumption by mixed microbial assemblages

The primary focus of this thesis was to investigate the production and turnover of marine dissolved organic matter (DOM) using in-vitro cultivation based approaches. Although extra challenging due to the complexity of natural systems and microbial interactions, examining DOM turnover in-situ is also an important undertaking, as one can interrogate cells in their natural setting. These in-situ studies can be a valuable tool for placing cultivation-based experiments in an environmental context. Two such studies in which microcosms of natural microbial assemblages were incubated with DOM amendments alongside controls are included here as appendices. Appendix A describes an addition of concentrated high-molecular weight DOM, while appendix B describes a similar experiment as well as an addition of DOM derived from *Prochlorococcus* str. MIT9313. Both studies were conducted at sea with microbial assemblages from the North Pacific Subtropical Gyre and functionally assayed using a metatranscriptomic approach.

APPENDIX A

**Microbial community transcriptomes reveal microbes and metabolic pathways
associated with dissolved organic matter turnover in the sea**

Jay McCarren^{a,b}

Jamie W. Becker^{a,c}

Daniel J. Repeta^d

Yanmei Shi^a

Curtis R. Young^a

Rex R. Malmstrom^{a,c}

Sallie W. Chisholm^a

Edward F. DeLong^{a,f,1}

Departments of ^aCivil and Environmental Engineering and ^fBiological Engineering,
Massachusetts Institute of Technology, Cambridge, MA 02139

^bSynthetic Genomics, La Jolla, CA 92037

Departments of ^cBiology and ^dMarine Chemistry & Geochemistry, Woods Hole
Oceanographic Institution, Woods Hole, MA 02543

^eJoint Genome Institute, Walnut Creek, CA 94598

¹Corresponding author: delong@mit.edu

Reprinted with permission from the Proceedings of the National Academy of Sciences
© 2010 United States National Academy of Sciences

McCarren J, Becker JW, Repeta DJ, Shi Y, Young CR, Malmstrom RR, Chisholm SW,
DeLong EF (2010) Microbial community transcriptomes reveal microbes and
metabolic pathways associated with dissolved organic matter turnover in the sea.
Proc. Natl. Acad. Sci. 107 (38): 16420–16427.

Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea

Jay McCarren^{a,b}, Jamie W. Becker^{a,c}, Daniel J. Repeta^c, Yanmei Shi^a, Curtis R. Young^a, Rex R. Malmstrom^{a,d}, Sallie W. Chisholm^a, and Edward F. DeLong^{a,e,1}

Departments of ^aCivil and Environmental Engineering and ^bBiological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139; ^cDepartment of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA 02543; ^dSynthetic Genomics, La Jolla, CA 92037; and ^eJoint Genome Institute, Walnut Creek, CA 94598

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2008.

Contributed by Edward F. DeLong, August 2, 2010 (sent for review July 1, 2010)

Marine dissolved organic matter (DOM) contains as much carbon as the Earth's atmosphere, and represents a critical component of the global carbon cycle. To better define microbial processes and activities associated with marine DOM cycling, we analyzed genomic and transcriptional responses of microbial communities to high-molecular-weight DOM (HMWDOM) addition. The cell density in the unamended control remained constant, with very few transcript categories exhibiting significant differences over time. In contrast, the DOM-amended microcosm doubled in cell numbers over 27 h, and a variety of HMWDOM-stimulated transcripts from different taxa were observed at all time points measured relative to the control. Transcripts significantly enriched in the HMWDOM treatment included those associated with two-component sensor systems, phosphate and nitrogen assimilation, chemotaxis, and motility. Transcripts from *Idiomarina* and *Alteromonas* spp., the most highly represented taxa at the early time points, included those encoding TonB-associated transporters, nitrogen assimilation genes, fatty acid catabolism genes, and TCA cycle enzymes. At the final time point, *Methylophaga* rRNA and non-rRNA transcripts dominated the HMWDOM-amended microcosm, and included gene transcripts associated with both assimilatory and dissimilatory single-carbon compound utilization. The data indicated specific resource partitioning of DOM by different bacterial species, which results in a temporal succession of taxa, metabolic pathways, and chemical transformations associated with HMWDOM turnover. These findings suggest that coordinated, cooperative activities of a variety of bacterial "specialists" may be critical in the cycling of marine DOM, emphasizing the importance of microbial community dynamics in the global carbon cycle.

carbon cycle | marine | bacteria | metagenomics | metatranscriptomics

Microbial activities drive most of Earth's biogeochemical cycles. Many processes and players involved in these planetary cycles, however, remain largely uncharacterized, due to the inherent complexity of microbial community processes in the environment. Cycling of organic carbon in ocean surface waters is no exception. Though marine dissolved organic matter (DOM) is one of the largest reservoirs of organic carbon on the planet (1), microbial activities that regulate DOM turnover remain poorly resolved (2).

Marine DOM is an important substrate for heterotrophic bacterioplankton, which efficiently remineralize as much as 50% of total primary productivity through the microbial loop (3–6). Though some DOM is remineralized on short timescales of minutes to hours, a significant fraction escapes rapid removal. In marine surface waters, this semilabile DOM transiently accumulates to concentrations 2–3 times greater than are found in the deep sea (7), and represents a large inventory of dissolved carbon and nutrients that are potential substrates for marine microbes. Time-series analyses of semilabile DOM accumulation in temperate and subtropical upper ocean gyres show an annual cycle in DOC in-

ventory with net accumulation following the onset of summertime stratification, and net removal following with deep winter mixing. In addition, multiyear time-series data suggest that surface-water DOM inventories have been increasing over the past 10–20 y (8). The ecological factors behind these seasonal and decadal DOC accumulations are largely unknown. Nutrient (N, P) amendments do not appear to result in a drawdown of DOC, and other factors such as the microbial community structure and the chemical composition of semilabile DOM have been invoked to explain the dynamics of the semilabile DOC reservoir (9, 10). Whatever the cause, the balance and timing of semilabile DOM remineralization are critical factors that influence the magnitude of DOM and carbon exported to the ocean's interior through vertical mixing.

There are significant challenges associated with characterizing and quantifying complex, microbially influenced processes such as DOM cycling in the sea. These challenges include inherent phylogenetic and population diversity and variability, the complexities of microbial community metabolic properties and interactions, and those associated with measuring microbial assemblage activities and responses on appropriate temporal and spatial scales. Past approaches have included measuring the bulk response of microbial communities to nutrient addition (e.g., community substrate incorporation or respiration), following changes in total or functional group cell numbers by microscopy or flow cytometry, or monitoring changes in relative taxa abundance, typically using rRNA-based phylogenetic markers. A number of field experiments (9–13) have indicated that specific shifts in microbial community composition might be linked to surface-water carbon utilization. However, the pure compound nutrient additions (such as glucose) frequently used in such field experiments (9, 11, 14, 15) may not well approximate the environmentally relevant chemical mixtures or compound concentrations present in naturally occurring DOM.

Though complications associated with direct experimentation on natural microbial communities limit our understanding of oceanic carbon cycling to some extent, significant insight into these processes have been recently reported. For example, Carlson et al. (10) showed differences among depth-stratified microbial communities that may be related to their ability to use semilabile DOM that

Author contributions: J.M., J.W.B., D.J.R., R.R.M., and E.F.D. designed research; J.M., J.W.B., D.J.R., Y.S., and R.R.M. performed research; S.W.C. contributed new reagents/analytic tools; J.M., J.W.B., D.J.R., Y.S., C.R.Y., R.R.M., and E.F.D. analyzed data; and J.M., J.W.B., D.J.R., Y.S., C.R.Y., and E.F.D. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. SRA020733.11 and HQ012268–HQ012278).

¹To whom correspondence should be addressed. E-mail: delong@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1010732107/-DCSupplemental.

accumulates in ocean surface waters. In addition, phylogenetic analyses of time-series samples have identified some taxonomic groups that appear to be responsive to deep-water mixing events, which may be relevant to organic carbon cycling dynamics (16, 17).

To better define the processes and population dynamics associated with marine microbial DOM cycling in ocean surface waters, we performed controlled experiments using seawater microcosms amended with freshly prepared, naturally occurring DOM. High-molecular-weight DOM (HMWDOM, defined here as the size fraction >1,000 Da and <30,000 Da) was concentrated by ultrafiltration using a 1-nm membrane filter, followed by a second filtration step to remove viruses. Whole, unfiltered seawater was distributed into replicate microcosms (20 L each) that were incubated at in-situ temperatures and light intensities. The ambient concentration of dissolved organic carbon (DOC) in the unamended microcosms was 82 μ M DOC, whereas the HMWDOM-amended microcosms contained 328 μ M DOC, representing a 4-fold increase over ambient DOC concentration. Replicate control and experimental microcosms were sampled periodically over the course of a 27-h period.

The responses of microbial community members to HMWDOM addition over time were followed using flow cytometric, metagenomic, and metatranscriptomic analytical techniques. HMWDOM-induced shifts in microbial cell numbers, community composition, functional gene content, and gene expression were observed at each time point, as indicated by changes in the DOM-treated microcosms relative to an unamended control. The data indicated rapid and specific HMWDOM-induced shifts in transcription, metabolic pathway expression, and microbial growth that appear to be associated with HMWDOM turnover in ocean surface waters.

Results and Discussion

HMWDOM-Induced Cell Dynamics. Replicate microcosms were established immediately before sunrise and sampled over the course of 27 h to track the changes in microbial cell numbers, community composition, gene content, and gene expression in control vs. HMWDOM-treated microcosms. Though cell numbers in control microcosms remained constant over the time course of the experiment, the HMWDOM-treated microcosm exhibited a ~50% increase in total cells within 19 h (Fig. 1A). Assuming a 50% growth efficiency, this HMWDOM-stimulated cell growth represents consumption of less than 1% of the total added DOC. Flow cytometry indicated that the majority (> 80%) of this increase in cells was attributable to the growth of a specific population of larger, high-DNA-content cells (Fig. 1B). The distinct flow cytometric signature of the HMWDOM-responsive population at the final time point allowed us to separate these large, high-DNA-content cells for further analyses (SI Appendix, Fig. S1). Large, high-DNA-content cells were isolated and collected via fluorescence-activated cell sorting and used to generate a SSU rRNA gene amplicon library. Near full-length rRNA gene sequences from the sorted cells recovered were all affiliated with the phylum Proteobacteria, falling into one of three clades (Fig. 1C). One subset of the flow-sorted cell population contained Alphaproteobacteria, closely related to *Thalassobius* isolates within the family Rhodobacteraceae. The remaining rRNA genes from the cell-sorted population were derived from Gammaproteobacteria, with one subset most closely related to *Alteromonas* isolates, and a second subset most similar to *Methylophaga* isolates within the order Thiotrichales.

Taxon-Specific Patterns of rRNA Gene and rRNA Representation in Control vs. HMWDOM-Treated Metagenomic and Metatranscriptomic Datasets. Community genomic DNA samples from T₀ and T_{27hrs} were pyrosequenced on the Roche 454 FLX platform, yielding \approx 500,000 reads per sample (Table 1). Though SSU rDNA genes represent a small fraction (~1%) of the total genomic pyrosequencing reads, sufficient data (~500–750 individual reads) was available for phylogenetic analyses, which avoids PCR bias, and other artifacts associated with PCR amplicon “pyrotag” libraries (18–20). Classification of these of rRNA genes (Methods) provided

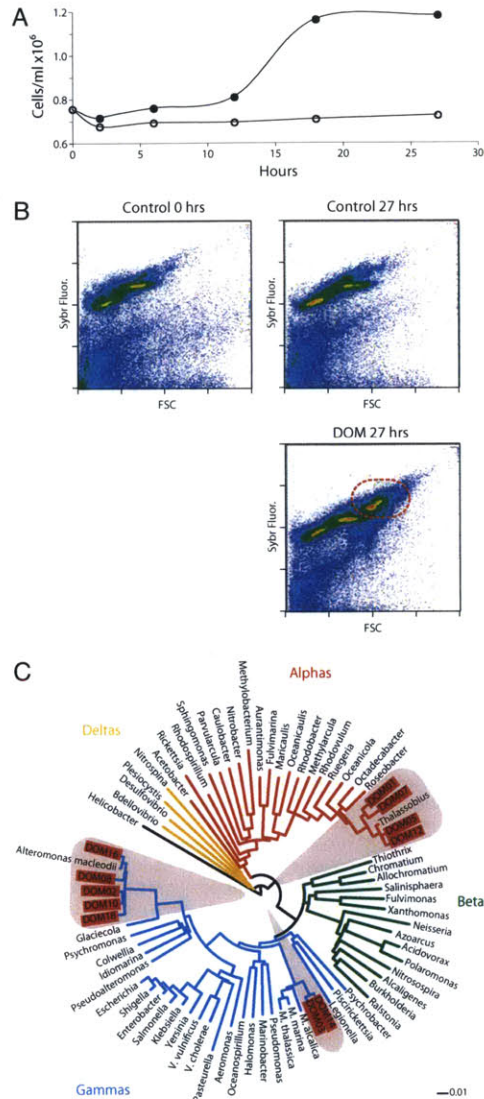


Fig. 1. Dynamics of microbial populations during 27-h microcosm incubations. (A) Flow cytometric counts of microbial cells from control (○) and DOM-amended (●) treatments. Samples displayed in B highlighted in red. (B) Flow cytometry scatterplots from selected samples show little change in the distribution of cell size [as measured by forward scatter (FSC)] and DNA content (SYBR fluorescence) of control samples from beginning to end of the experiment, whereas most of the increase in cell numbers observed in the DOM-amended treatment can be attributed to the appearance of larger, high-DNA-content cells (circled in red). (C) Weighted neighbor-joining tree of selected SSU rDNA sequences from proteobacterial type strains and the sequences obtained from flow cytometric sorting of the larger, higher-DNA-content population of cells present after DOM amendment. The sequences obtained from the flow-sorted population are restricted to three specific taxonomic clades: Rhodobacteraceae, Methylophaga, and *Alteromonas*.

Table 1. Number of pyrosequences analyzed in control and treatment DNA and cDNA libraries

Treatment	Sample	0 h	2 h	12 h	27 h
Control	DNA	557,099	NA	NA	422,666
	cDNA (non rRNA)	505,075 (18,345)	221,751 (12,658)	470,578* (12,934)	514,670 (18,078)
+DOM	DNA	NA	NA	NA	526,681
	cDNA (non rRNA)	NA	230,376 (14,762)	251,690 (15,748)	751,284 (42,689)

*One of two technical replicate sequencing runs for this sample contained a spuriously high representation of a single sequence (~4.2% of reads) not present in the other replicate sequencing run. These nearly perfect duplicate reads (>99% nucleotide identity and read-length difference of <5 bp) were removed before subsequent analysis.

an overview of microbial community composition over the course of the experiment (Fig. 2A, inner rings). As expected, typically abundant planktonic bacterial taxa such as *Pelagibacter* (Rickettsiales) and *Prochlorococcus* (Cyanobacteria) were highly represented (Fig. 2A and *SI Appendix*, Fig. S2). The community

composition of the control microcosm did not change substantially from the beginning to the end of the experiment. In contrast, the representation of several taxonomic groups increased in the HMWDOM-amended microcosm over the 27-h incubation. Three specific gammaproteobacterial groups—the families Idioma-

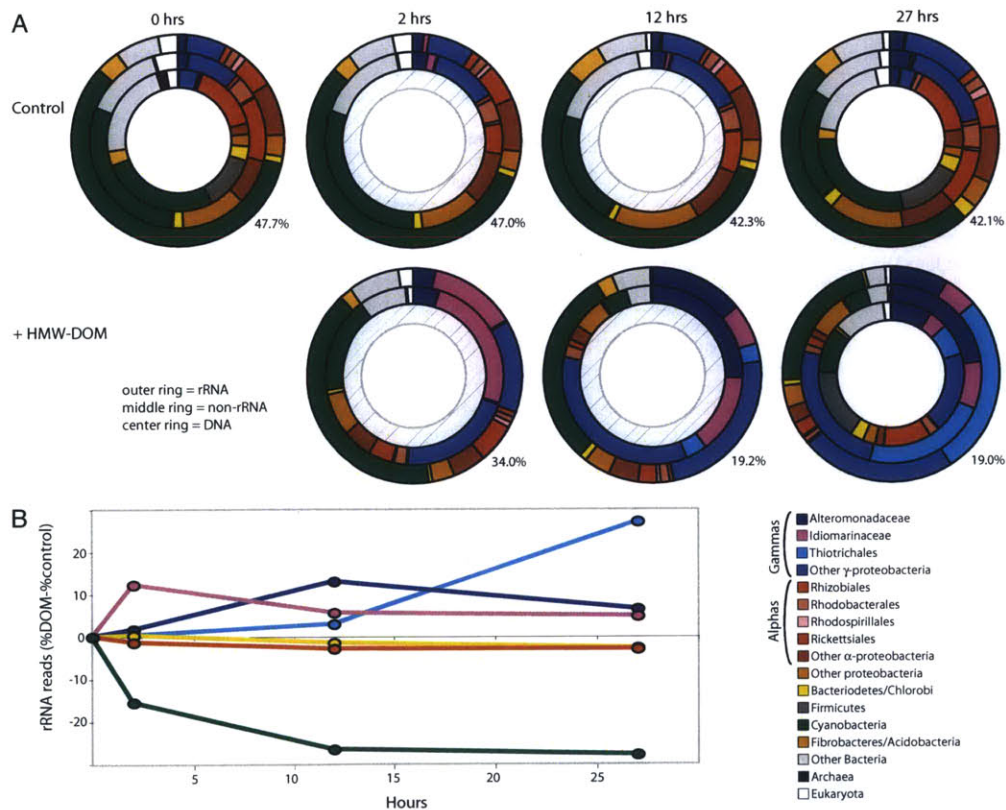


Fig. 2. Microbial community composition assessed by taxonomic classification of metagenomic and metatranscriptomic sequence reads. (A) SSU rRNA reads (outer ring) and non-rRNA reads (middle ring) from metatranscriptomic datasets as well as those reads from metagenomic datasets identified as SSU rDNA reads (center ring). Only taxonomic groups that represent >1% of total reads in at least one dataset have been included with all other groups binned together with unassigned reads. In some instances, reads can only be confidently assigned to broad class- and order-level taxonomic groups and are labeled as such. For mRNA datasets, some reads have no significant blast hits, the percentage of which is noted beside each sample. (B) Tracking the changes in community composition by comparing the difference between the DOM-amended treatment and control reveals distinct taxonomic groups responding at each time point. Only taxonomic groups showing more than $\pm 2\%$ change are plotted.

inaceae and Alteromonadaceae (both of which fall in the order Alteromonadales) and the order Thiotrichales—all increased in rRNA gene representation following HMWDOM amendment (Fig. 2*A* and *B* and *SI Appendix*, Fig. S2). Two of these HMWDOM-stimulated groups (Alteromonadaceae and Thiotrichales) corresponded to the same dominant groups found in the FACS-sorted, high-DNA-containing cell populations (Fig. 1). The Rhodobacteraceae group that was recovered in the flow-sorted population did not, however, show a corresponding rRNA enrichment in the HMWDOM-treated metagenomic or metatranscriptomic datasets. These alphaproteobacteria may simply represent a background population of cells that were sorted along with the DOM-stimulated gammaproteobacteria because their flow cytometric signal overlapped with the large, high-DNA-content cell fraction.

Analyses of metagenomic sequence reads yields information on the relative representation of taxonomic groups, but not absolute cell numbers. Though cyanobacteria represented more than a quarter of all SSU rRNA genes throughout the time course of the experiment in the control microcosm, in the HMWDOM treatment they comprised only 10% of the rRNA sequence reads by 27 h. Enumeration of *Prochlorococcus* cells via flow cytometry indicated, however, that absolute *Prochlorococcus* cell numbers changed by less than 1% in the HMWDOM-amended microcosm. The changes in community composition observed in the metagenomic datasets therefore appear due to the growth of specific population members (in particular, Alteromonadaceae and Thiotrichales) and not to the disappearance of other dominant groups.

Compared with SSU rDNA reads from metagenomic DNA datasets, pyrosequencing of total community cDNA yielded orders of magnitude more total rRNA sequences that could be similarly classified taxonomically (Fig. 2*A*, outer rings). [The cDNAs in this study were not subjected to upstream rRNA subtraction procedures that have been reported in other metatranscriptomic studies (21–23).] In contrast to rRNA gene abundance in the DNA, rRNA in the cDNA pool reflects the cellular abundance of specific phylogenetic groups, as well as their cellular rRNA copy numbers. For example, the rRNAs of several groups (e.g., Rickettsiales, Firmicutes, and Archaea) were less abundant in the cDNA datasets in comparison with their corresponding genes in the genomic DNA dataset (Fig. 2 and *SI Appendix*, Fig. S2). Conversely, cyanobacterial rRNAs were more highly represented in the cDNA than the corresponding rRNA genes in the DNA (Fig. 2 and *SI Appendix*, Fig. S2). Similarly, in the 27 h post-HMWDOM amendment, the Thiotrichales comprised nearly one-third of all SSU rRNA sequences in the cDNA, but represented less than 8% of all SSU rRNA genes in the DNA of the same sample.

Taxon-Specific Responses to HMWDOM Addition Inferred from Functional Gene Transcript Abundance. Taxonomic classification of non-rRNA transcripts from cDNA datasets (Fig. 2*A*, middle ring; *Methods*) generally paralleled the trends observed for rRNA taxon abundance, indicating parallel responses in both functional gene transcript and rRNAs (Fig. 2). Two exceptions to this correspondence were observed: cyanobacterial rRNA sequences were present in much greater abundance than non-rRNA cyanobacterial transcripts at all time points in both the control and the HMWDOM treatment. Conversely, Idiomarinaceae and Alteromonadaceae were underrepresented in rRNAs, relative to non-rRNA transcripts present in the HMWDOM-treated microcosm cDNAs.

Distinct shifts in the cDNAs of specific subpopulations occurred in response to HMWDOM addition. Though the control remained virtually unchanged throughout the experiment, at each time point following HMWDOM addition, a different taxonomic group dominated the cDNA pool for both rRNA and non-rRNA transcripts (Fig. 2*A* and *B*). Two hours post-HMWDOM amendment, Idiomarinaceae sequences represented nearly 13% of all rRNA sequences in the cDNAs from the HMWDOM treatment, though they remained less than 1% of the total rRNA sequences in all

control cDNAs. By 12 h, the abundance of Idiomarinaceae rRNA sequences in the HMWDOM treatment receded closer to control values, whereas Alteromonadaceae rRNA sequences in the transcript pool rose to 15% of the total rRNAs relative to the control (Fig. 2*B*). Similarly, by the end of the experiment, Alteromonadaceae rRNA sequences decreased in relative abundance compared with earlier time points, when Thiotrichales-like rRNA represented the most abundant rRNAs. Strikingly, though Thiotrichales-like rRNAs represented approximately one-third of the total rRNA sequences in cDNA at the final HMWDOM-treated time point, Thiotrichales never represented more than 0.04% of in any of the controls at all time points.

Idiomarinaceae and Alteromonadaceae are closely related families within the order Alteromonadales (24). Because these closely related taxa were differentially represented at two different time points in the HMWDOM treatment, we searched for potential differences in their functional gene transcript representation at different times. All sequence reads having a best match to the full genome sequence of these two dominant taxa [*Idiomarina loihiensis* (25) and *Alteromonas macleodii* (26)] were analyzed separately for each taxonomic bin (*SI Appendix*, Tables S1 and S2). There were many similarities in the distribution of cDNA reads of functional gene categories between the two taxa. Examination of the 2-h and 12-h HMWDOM microcosm time points for Idiomarinaceae and Alteromonadaceae, respectively, indicated that transcript representation for many nutrient acquisition genes were similarly abundant within both taxonomic groups at the two different time points. An outer membrane receptor for a TonB-associated iron transporter was among the most abundant transcripts for both Idiomarinaceae and Alteromonadaceae. Similarly, the three genes require for the glutamine synthase cycle involved in nitrogen assimilation were abundant in each taxonomic bin. Genes involved in fatty acid catabolism were abundant in both Idiomarinaceae and Alteromonadaceae bins (*SI Appendix*, Tables S1 and S2). Additionally, the two enzymes specific for the glyoxylate cycle (isocitrate lyase and malate synthase), which could use acetyl-CoA output by the β -oxidation of fatty acids, were abundant in both bins. One striking difference between the two different Alteromonadales cDNA bins was the high representation of one gene, triacylglycerol lipase (10-fold more abundant in treatment than control), found only among Idiomarinaceae-like reads. Interestingly, triacylglycerol lipase reads were virtually absent from reads assignable to the Alteromonadaceae bin.

The taxonomic groups that appeared most responsive to HMWDOM addition comprised only a small fraction of the starting microbial community. In contrast, transcripts from typically more dominant taxa such as *Pelagibacter* and *Prochlorococcus* decreased in relative abundance in the HMWDOM treatment over time. Additionally, because the differences in transcript abundance between control and treatment were small for *Prochlorococcus* and *Pelagibacter*, our sequencing depth allowed the detection of only a few significantly different transcripts between controls and treatments (*SI Appendix*, Figs. S3 and S4). Only seven *Pelagibacter* ORFs were identified as having statistically significant changes in transcript abundance ($P < 0.001$; *Methods*) in the HMWDOM-treated sample vs. the control (*SI Appendix*, Fig. S3). This small number of transcriptionally responsive ORFs (within our detection limits) was consistent with the hypothesis that *Pelagibacter* has a relatively small genome and streamlined regulatory network (27) and so may be less responsive to large fluctuations in ambient nutrient concentrations. The absolute *Pelagibacter* cell numbers appear to have increased slightly over the course of incubation in the treatment relative to the control, as evidenced by its higher gene abundances in the treatment relative to *Prochlorococcus* (whose absolute cell numbers remained constant as determined by flow cytometry; Fig. 2). The enrichment of transcripts encoding DNA-directed RNA polymerase and methionine biosynthesis protein (*SI Appendix*, Fig. S3) may indicate some utilization of some fraction of HMWDOM by *Pelagibacter*

cells to obtain reduced sulfur for the biosynthesis of sulfur-containing amino acids (28). The depletion of proteorhodopsin transcripts in the treatment at the final time point (*SI Appendix, Fig. S3*) suggested a potentially diminished requirement for proteorhodopsin phototrophy, with the increase in carbon availability. For *Prochlorococcus*, most of the significantly different transcripts were depleted in the treatment relative to the control at the earlier time points, whereas a few transcripts were enriched at the final time point. Several of these treatment-stimulated *Prochlorococcus* transcripts appeared to be involved with cellular repair processes, including oxidative damage protection and protein folding (*SI Appendix, Fig. S4*).

Small RNAs. Thirty putative sRNA (psRNA) clusters comprising >100 reads were identified, 20 of which showed statistically significant differences in abundance between the treatment and control for one or more time points (*SI Appendix, Fig. S5*). Based on the Rfam 10.0 database (<http://rfam.sanger.ac.uk/>), five clusters were identified as transfer-messenger RNA (tmRNA), and one was RNaseP RNA. Notably, all but one tmRNA cluster was overrepresented in the treatment, in part reflecting increases in specific taxa in the treatment vs. control (Fig. 1). For instance, cluster 7 tmRNA, which was overrepresented at 2 h, was most closely related to *Idiomarinaceae*, whereas *Methylophaga*-like cluster 9 tmRNA was enriched at later time points. Several psRNA clusters mapped into previously reported abundant psRNA groups found in microbial community transcripts sampled from the water column at Station ALOHA (29) (*SI Appendix, Fig. S5*). Five apparently different psRNA clusters (cluster 2, 3, 4, 8, and 14) were adjacent to genes encoding class II fumarate hydratase, an enzyme that catalyzes the reversible hydration/dehydration of fumarate to S-malate in the tricarboxylic acid cycle. To test the possibility that these clusters belonged to the same group but did not merge due to stringent clustering method, we performed pairwise alignment analysis among representative sequences of these five clusters (*SI Appendix, Fig. S6*). Only cluster 3 and cluster 14 merged (based on high sequence identity in the alignment at the end of both sequences), confirming that several divergent psRNA species, all adjacent to fumarate hydratase genes, were enriched in response to HMWDOM addition.

Global trends in functional gene transcript abundances in the HMWDOM treatment vs. control. All non-rRNA cDNA sequences were compared with NCBI-nr, KEGG (30), and GOS protein clusters databases (31) using BLASTX (32). We focused in particular on quantifying KEGG ortholog abundances in the HWM DOM-treated microcosm relative to the unamended controls across all time points (*SI Appendix, Tables S3–S6*).

Among all of the controls (0 h, 2 h, 12 h, and 27 h), only a few orthologs exhibited significant changes between time points ($n = 43$; *SI Appendix, Table S3*). Among these significantly different orthologs, about half were due to differences between the initial time point (0 h) and the other controls. In contrast, a larger number of orthologs exhibited differences in abundance between the pooled controls and the HMWDOM treatment (*SI Appendix, Tables S4–S6*). At 2 h post-HMWDOM addition, 67 KEGG orthologs exhibited differences from the control, with 58 of those enriched in the treatment vs. pooled controls (detectable effect sizes of enriched orthologs: 2.0- to 550-fold change; *SI Appendix, Table S4*). At 12 h, 221 differences were apparent, and 200 of those were enriched in the treatment vs. controls (detectable effect sizes of enriched orthologs: 2.3- to 2,200-fold change; *SI Appendix, Table S5*). At 27 h, 390 differences were detected, and 311 of those orthologs were enriched in the treatment (detectable effect sizes of enriched orthologs: 1.6- to 1,100-fold change; *SI Appendix, Table S6*).

Significantly enriched transcripts in the HMWDOM treatment included those encoding enzymes in KEGG pathways for carbohydrate, nitrogen, methane, sulfur, and fatty acid metabolic genes. Numerous transcripts associated with signal transduction and membrane transport pathways were also enriched in the

HMWDOM treatment. Amino acid and nucleotide metabolism were also enriched in the HMWDOM addition microcosms, as were transcripts encoding enzymes involved in transcription and translation. The effect for all of these categories, however, was much more pronounced for the 12- and 27-h post-HMWDOM treatments than for the 2-h treatment. This is apparently due to the fact that the predominant DOM-responsive taxa were initially low in numbers, but increased in both cell density and transcriptional activity over the time course of the experiment.

At 12 h in the HMWDOM microcosm a variety of two-component sensor systems and several transporters were overrepresented. Particularly abundant were genes involved in nutrient acquisition. Specifically, both the components of the phosphate two-component sensor system (phoB, phoR, phoA, and OmpR phoB) as well as all components of the ABC transporter for phosphate (pstS, pstC, pstA, and pstB) were overrepresented at 12 and 27 h post-HMWDOM addition. At 27 h post-HMWDOM addition, members of several two-component sensor systems are enriched, including those associated with glucose (BarA, UvrY, CsrA), glucose-6-P (UhpB), nitrogen (GlnL, GlnG), C4-dicarboxylate (YfhK, YfhA), redox state of the quinone pool (ArcA), misfolded proteins (CpxR), carbon storage (BarA, UvrY, CsrA), and bacterial flagellar chemotaxis (CheA, CheV, CheY). Flagellar biosynthesis-associated transcripts were also similarly enriched, with 18 of 42 KO associated with flagellar biosynthesis more the 4-fold more abundant in the amended microcosm relative to controls.

Transcripts encoding components of the GS/GOGAT pathway (glutamine and glutamate synthesis) were also significantly enriched in the HMWDOM treatment. Nitrogen two-component systems enriched in the DOM treatment transcript pool (GlnL, GlnG) typically sense nitrogen limitation via the intracellular glutamine pool and respond to nitrogen limitation by activating glutamate metabolism (33), which is consistent with the observed elevated GS/GOGAT transcript levels. Other enzymes in the nitrogen pathway, however, appeared relatively unchanged except for aminomethyltransferase (involved in glycine synthesis), which was less prevalent in the HMWDOM treatment. [Transcripts for one specific family of Amt family ammonium transporters from *Prochlorococcus* were significantly depleted in the HMWDOM treatment (*SI Appendix, Fig. S4*)]. Similar to the signatures of nitrogen limitation, the prevalence of the OmpR family phosphate two-component system, and the enrichment of a PIT family inorganic phosphate transporter, suggested that over the course of the experiment, the HMWDOM microcosm community was experiencing nitrogen and phosphate limitation as a consequence of the elevated DOC levels relative to the control.

Transcripts associated with sulfur-metabolizing enzymes were enriched in the HMWDOM treatment at the final time point and included enzymes associated with sulfate metabolism, and serine metabolism. Serine metabolism produces acetate that potentially could be shunted into the reductive carboxylate cycle, also enriched in the DOM treatment. Transcripts encoding three enzymes of the fatty acid metabolism pathway were also enriched in the HMWDOM treatment, as well as those encoding a short-chain fatty acid transporter. Furthermore, fatty acid biosynthesis pathway transcripts were significantly depleted in the HMWDOM treatment, suggesting a potential shift to catabolic metabolism of fatty acid-like molecules in the HMWDOM treatment. At the first time point, just 2 h postamendment, the two most enriched transcripts that corresponded to KEGG orthologs were triacylglycerol lipase and acyl-CoA dehydrogenase (50-fold and 109-fold, respectively). These enzymes catalyze two early steps in the catabolism of triacylglycerols (TAGs). These signals may be the result of cell wall material copartitioning in the HMWDOM concentrate, or the tendency of lipid compounds to associate with HMWDOM concentrates (34).

Methylophaga species were the most highly represented single taxon in both rRNA and functional gene transcripts in the

HMWDOM microcosm at the final time point. Consistent with this observation, two key enzymes involved in the ribulose mono-phosphate (RuMP) pathway, hexulose-6-phosphate synthase and 6-phospho-3-hexuloisomerase, were also highly abundant in the amended microcosm (eighth and second most abundant, respectively) while remaining undetected in the control. The cyclical RuMP pathway is an assimilatory pathway that is widespread in bacteria, functioning as a pathway for formaldehyde fixation and detoxification. In the first two reactions in this pathway, formaldehyde is condensed with ribulose-5 phosphate, which is then isomerized to fructose-6-phosphate. Moreover, gene transcripts for the enzymes encoding many of the steps in this pathway were enriched by the end of this experiment (Fig. 3) and increased over the time course of the experiment (Poisson ANOVA; *SI Appendix, Table S7*). Though a large variety of one-carbon compounds are processed through the RuMP pathway, all methyltrophic pathways share formaldehyde as a common entry point. Formaldehyde can also be oxidized to CO₂ via several routes, and several of the enzymes involved in these dissimilatory pathways were also abundant in the amended treatment (Fig. 3), particularly those associated with the tetrahydromethanopterin-dependent pathway. In total, the data reflected the enrichment of pathways for both assimilatory and dissimilatory single-carbon compound utilization, which coincided with the appearance of an actively

growing *Methylophaga* population in the HMWDOM treatment (Figs. 2 and 3).

Conclusions

Semilabile DOM may support up to 40% of marine bacterial carbon demand (35, 36), yet little is known about the specific microorganisms and metabolic pathways responsible for its degradation and transformation in the ocean's water column. There is growing evidence that microbial transformation of semilabile DOM renders DOM less and less labile, further increasing accumulation in oligotrophic gyres and ultimately leading to export as refractory DOM (36). Microbial population dynamics and metabolic processes are therefore central to understanding the cycling of DOM in the sea.

In this study, short-term incubation of bacterial populations from surface seawater with naturally occurring HMWDOM from the same environment revealed specific shifts in microbial cells, rRNAs, and DOM-responsive gene transcripts relative to unamended controls. Cell numbers nearly doubled specifically in response to HMWDOM. Flow sorting and rRNA gene and transcript abundances consistently indicated the stimulation of several phylogenetic groups within the Alteromonadales (*Idiomarina* and *Alteromonas* sp.) and Thiotrichales (*Methylophaga* sp.). Analysis of microbial cDNA abundances over time via pyrosequencing revealed that 2 h after DOM addition, close relatives of *Idioma-*

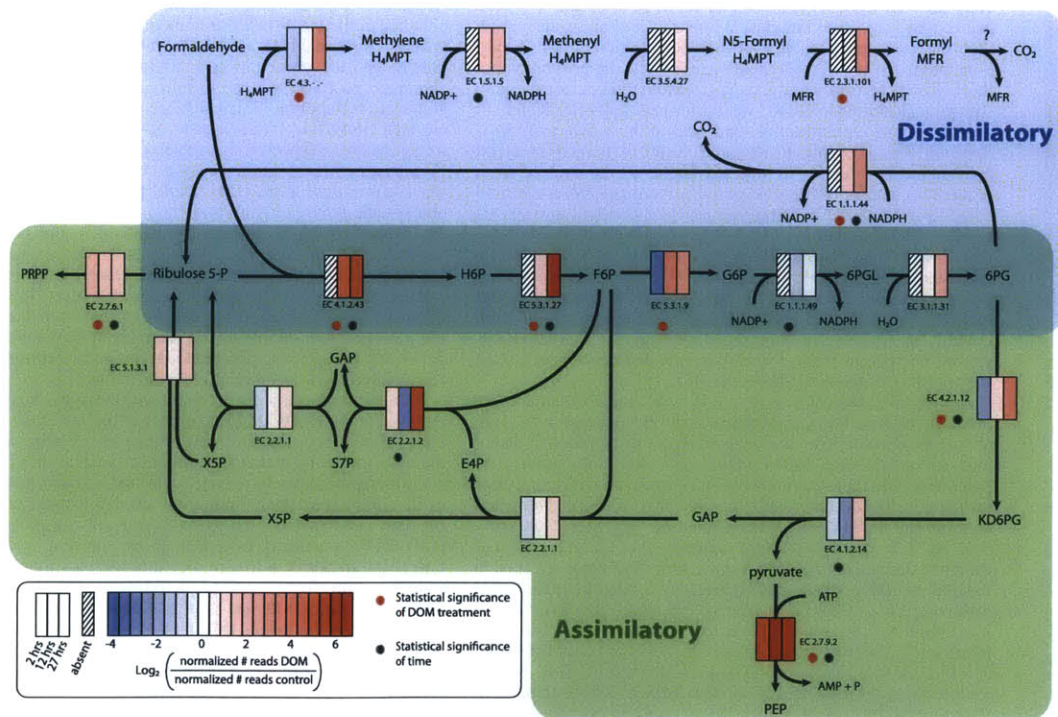


Fig. 3. Diagram of representative dissimilatory and assimilatory methyltrophic pathways and enzymes that show increased transcript abundance following DOM amendment. A KEGG ortholog-based expression ratio comparing normalized abundances of reads present in the DOM-amended treatment with those from an untreated control at 2, 12, and 27 h following DOM addition. Asterisks mark those enzymes showing statistically significant differences in transcript abundance relative to time and/or unamended control (*SI Appendix, Table S7*). H₄MPT, tetrahydromethanopterin; MFR, formylmethanofuran; H6P, hexulose-6-phosphate; F6P, fructose-6-phosphate; 6PGL, 6-phosphogluconolactone; 6PG, 6-phosphogluconate; KD, ketodeoxy; PEP, phosphoenolpyruvate; GAP, glyceraldehyde phosphate; E4P, erythrose-4-phosphate; X5P, xylulose-5-phosphate; S7P, sedoheptulose-5-phosphate; PRPP, phosphoribosyl diphosphate.

rina sp. were stimulated by HMWDOM. In apparent microbial succession, a few hours later, *Alteromonas macleodii*-like rRNAs and mRNAs increased dramatically relative to the unamended control. After 27 h, the same indicators showed that *Methylophaga* sp. (order Thiotrichales) predominated. We interpret this succession as a specific metabolic sequence and successional cascade that reflects sequential processing and degradation of specific components within HMWDOM. Analyses also indicated that 27 h post-DOM addition, both the dissimilatory and assimilatory single-carbon compound utilization pathways were highly expressed, coincident with the appearance and high abundance of *Methylophaga* sp. at the final time point.

The data indicate several specific groups of bacteria that appear to operate in succession and synergy to catalyze the turnover of naturally occurring HMWDOM in the marine environment. These findings may reflect regular (and predictable) metabolic cascades and community succession patterns that in part regulate the transformation and turnover of naturally occurring semilabile DOM. Furthermore, our findings are suggestive of some of the chemical attributes and degradation patterns of naturally occurring DOM. In previous chemical analyses, about 15% of DOM carbohydrate has been shown to consist of methyl sugars (37, 38). Our present findings suggest that *Alteromonadales* (specifically, *Idiomarina* spp. and *Alteromonas macleodii*) might be metabolizing semilabile DOM methyl sugars to methanol or formaldehyde, and carbon dioxide, among other products. The methanol and/or formaldehyde produced could be further oxidized and incorporated by *Methylophaga* sp. in the terminal portion of this aerobic food chain. Such a specific carbon compound-driven syntrophy has rarely been observed in aerobic microbial consortia. Although confirmation awaits further experimentation and chemical analyses, if correct, DOM methyl sugar metabolism might provide a partial explanation for the ubiquitous presence of methylotrophs in open-ocean and coastal environments (12, 39–42).

In summary, the experimental metatranscriptomic approach described here is beginning to reveal metabolic pathways and microbial taxa involved in the chemical transformation and turnover of naturally occurring marine DOM. These techniques can be used to track a variety of microbial processes in the environment, and set the stage for future inquiries on the nature and details of microbial community environmental responses and dynamics in situ. In this study, we gained detailed perspective on microbial community dynamics and metabolism associated with the ocean carbon cycle in marine surface waters. The apparent resource partitioning of DOM by different bacterial species that was suggested by the data supports the significance of microbial community dynamics in the ocean's carbon cycle. The findings also underscore the importance of describing microbial synergistic interactions and population dynamics occurring on relatively short time-scales of hours to days.

Methods

Microcosm Setup and Biomass Sampling. Seawater for microcosm incubation experiments was collected (23°12.88' N, 159°8.17' W) from 75-m depth, predawn, on August 16, 2007, during the Center for Microbial Oceanography: Research and Education (C-MORE) BLOOMER Cruise. See *SI Appendix* for further details on the seawater collection and microcosm preparation.

HMW DOM Preparation. Surface seawater obtained from the uncontaminated underway system of the *R/V Kilo Moana* was filtered to remove microbes and small particles using a clean (10% HCl overnight soak), 0.2- μ m Whatman Polycap TC polyether sulfone capsule filter. HMWDOM was concentrated using a custom-built ultrafiltration apparatus equipped with a stainless-steel membrane housing and centripetal pump along with a fluorinated high-density polyethylene reservoir. The system was plumbed with Teflon tubing and PVDF valves, and fitted with a dual thin-film ultrafiltration membrane element (Separation Engineering). The membrane has a 1-nm pore size that nominally retains organic matter of a molecular weight greater than 1,000 Da (>98% rejection of vitamin B₁₂). Membranes were precleaned with 0.01 mol L⁻¹ hydrochloric acid (overnight wash) and 0.01 mol L⁻¹ sodium hydroxide (over-

night wash), and rinsed with copious amounts of distilled water until the pH returned to neutral. Membranes were flushed with 100 L of seawater for 45 min just before sample collection. Surface seawater (2,000 L) was concentrated 100-fold over a period of 24 h. Samples were taken for DOC quantification from the inflow and permeate during ultrafiltration, and of the concentrate upon completion. A 2-L subsample of the concentrate was prefiltered using a 0.2- μ m Polycap TC filter (Whatman) before filtration through a preirradiated 30-kDa Ultracel regenerated cellulose membrane loaded in a high-output stirred cell (Millipore) to remove viral particles.

Dissolved Organic Carbon. DOC samples of 30 mL were transferred into combusted (450 °C for 8 h) glass vials and acidified with 150 μ L of a 25% phosphoric acid solution before sealing with acid-washed Teflon septa and storage at 4 °C until processing. Analysis was performed using the high-temperature combustion method on a Shimadzu TOC-VCSH with platinumized alumina catalyst. Sample concentrations were determined alongside potassium hydrogen phthalate standards and consensus reference materials (CRM) provided by the DOC-CRM program (<http://www.rsmas.miami.edu/groups/biogeochem/CRM.html>).

Flow Cytometry and Cell Sorting. At each time point, 1 mL of seawater was preserved with 0.125% glutaraldehyde (final concentration), frozen in liquid nitrogen, and stored at -80 °C for subsequent flow cytometric analysis and cell sorting using an Influx (Becton Dickinson). Before counting and sorting, samples were stained with SYBR Green (Invitrogen) for 15 min, and DNA-containing cells were identified based on fluorescence and scatter signals (43). See *SI Appendix* for further details on cell sorting and rRNA amplicon sequencing from the sorted population.

RNA Amplification and cDNA Synthesis. Metatranscriptome analyses were performed as previously described (44) with minor modifications. Briefly, 100 ng of total RNA was amplified using MessageAmp II (Ambion) following the manufacturer's instructions and substituting the T7-Bpml-(dT)₁₆VN oligo (44) in place of that supplied with the kit. Amplified RNA was then reverse transcribed into cDNA using SuperScript Double-Stranded cDNA Synthesis kit (Invitrogen) and random hexamer priming. Last, the cDNA was digested with Bpml and used for pyrosequencing. See *SI Appendix* for further details on pyrosequencing.

Bioinformatic Analyses. Full-length SSU rDNA amplicon sequences from flow-sorted cells were classified using both the Greengenes (45) NAST aligner and the Ribosomal Database Project (RDP) naïve Bayesian classifier (46). Resulting alignments were compared with the SILVA (47) databases using ARB (48). RDP classifier results were compared also with type strains using tools available at the RDP (49) and Interactive Tree of Life web sites (50).

cDNA datasets were parsed to separate rRNA sequences from the remaining non-rRNA sequences. rRNA sequences were identified as previously described (44) using a bit-score cutoff of 40 for BLASTN (32) searches against a custom SS, SSU, 18S, 23S, and 28S rRNA databases. Non-rRNA sequences were compared with NCBI-nr, KEGG, and GOS protein clusters databases using BLASTX (32) for functional gene analyses as previously described (29, 44). See *SI Appendix* for further details.

Statistical Analyses. Statistical analyses were conducted on KEGG ortholog groups using the packages DegSeq (51) and ShotgunFunctionalizer (52) in the R Statistical Package (53). In all statistical analyses, we assumed that the data (counts for a particular KEGG ortholog group) followed a Poisson sampling distribution. Analyses were conducted at the individual gene level as well as at the pathway level. See *SI Appendix* for further details on statistical analyses.

Accession Numbers. All 454 FLX pyrosequencing .sff files have been deposited in the GenBank database under accession no. SRA020733.11. Full-length SSU SSU rRNA sequences obtained from flow-sorted cells have been deposited to the GenBank/EMBL/DDJB databases under accession nos. HQ012268-HQ012278.

ACKNOWLEDGMENTS. We thank the captain and crew of the *R/V Kilo Moana* for facilitating sample collection, Chief Scientist Ricardo Letelier and all participants of the C-MORE BLOOMER cruise for help and encouragement, and Rachel Barry for pyrosequence library production and sequencing. This work was supported by the Gordon and Betty Moore Foundation (E.F.D., S.W.C., and D.J.R.), the Office of Science-Biological and Environmental Research, US Department of Energy (E.F.D. and S.W.C.), the National Science Foundation (D.J.R.), and National Science Foundation Science and Technology Center Award EF0424599 (to E.F.D. and S.W.C.). This article is a contribution from the National Science Foundation Science and Technology Center for Microbial Oceanography: Research and Education (C-MORE).

- Hedges JJ (1992) Global biogeochemical cycles: Progress and problems. *Mar Chem* 39: 67–93.
- Ogawa H, Amagai Y, Koike I, Kaiser K, Benner R (2001) Production of refractory dissolved organic matter by bacteria. *Science* 292:917–920.
- Pomeroy LR (1974) Oceans food web, a changing paradigm. *Bioscience* 24:499–504.
- Azam F, et al. (1983) The ecological role of water-column microbes in the sea. *Mar Ecol Prog Ser* 10:257–263.
- Azam F (1998) Microbial control of oceanic carbon flux: The plot thickens. *Science* 280: 694–696.
- Ducklow H (1999) The bacterial component of the oceanic euphotic zone. *FEMS Microbiol Ecol* 30:1–10.
- Benner R, Pakulski JD, McCarthy M, Hedges JJ, Hatcher PG (1992) Bulk chemical characteristics of dissolved organic matter in the ocean. *Science* 255:1561–1564.
- Church MJ, Ducklow HW, Karl DM (2002) Multiyear increases in dissolved organic matter inventories at station ALOHA in the North Pacific Subtropical Gyre. *Limnol Oceanogr* 47:1–10.
- Carlson C, et al. (2002) Effect of nutrient amendments on bacterioplankton production, community structure, and DOC utilization in the northwestern Sargasso Sea. *Aquat Microb Ecol* 30:19–36.
- Carlson C, et al. (2004) Interactions among dissolved organic carbon, microbial processes, and community structure in the mesopelagic zone of the northwestern Sargasso Sea. *Limnol Oceanogr* 49:1073–1083.
- Pinhassi J, Berman T (2003) Differential growth response of colony-forming alpha- and gamma-proteobacteria in dilution culture and nutrient addition experiments from Lake Kinneret (Israel), the Eastern Mediterranean Sea, and the Gulf of Eilat. *Appl Environ Microbiol* 69:199–211.
- Pinhassi J, et al. (2004) Changes in bacterioplankton composition under different phytoplankton regimes. *Appl Environ Microbiol* 70:6753–6766.
- Schafer H, et al. (2001) Microbial community dynamics in Mediterranean nutrient-enriched seawater mesocosms: Changes in the genetic diversity of bacterial populations. *FEMS Microbiol Ecol* 34:243–253.
- Allers E, et al. (2007) Response of Alteromonadaceae and Rhodobacteriaceae to glucose and phosphorus manipulation in marine mesocosms. *Environ Microbiol* 9:2417–2429.
- Cecilia A, Jakob P (2006) *Roseobacter* and SAR11 dominate microbial glucose uptake in coastal North Sea waters. *Environ Microbiol* 8:2022–2030.
- Hansell DA, Carlson CA (2001) Biogeochemistry of total organic carbon and nitrogen in the Sargasso Sea: Control by convective overturn. *Deep Sea Res Part II Top Stud Oceanogr* 48:1649–1667.
- Morris RM, et al. (2005) Temporal and spatial response of bacterioplankton lineages to annual convective overturn at the Bermuda Atlantic Time-Series Study site. *Limnol Oceanogr* 50:1687–1696.
- Quince C, et al. (2009) Accurate determination of microbial diversity from 454 pyrosequencing data. *Nat Methods* 6:639–641.
- Kunin V, Engelbrekton A, Ochman H, Hugenholtz P (2010) Wrinkles in the rare biosphere: Pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ Microbiol* 12:118–123.
- Turnbaugh PJ, et al. (2010) Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins. *Proc Natl Acad Sci USA* 107: 7503–7508.
- Stewart FJ, Ottesen EA, DeLong EF (2010) Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. *ISME J* 4: 896–907.
- Vila-Costa M (2010) Transcriptomic analysis of a marine bacterial community enriched with dimethylsulfoniopropionate. *ISME J*, 10.1038/ismej.2010.62.
- Gilbert JA, et al. (2008) Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities. *PLoS ONE* 3:e3042.
- Ivanova EP, Flavier S, Christen R (2004) Phylogenetic relationships among marine Alteromonas-like proteobacteria: Emended description of the family Alteromonadaceae and proposal of Pseudoalteromonadaceae fam. nov., Colwelliaceae fam. nov., Shewanellaceae fam. nov., Moritellaceae fam. nov., Ferrimonadaceae fam. nov., Idiomarinaceae fam. nov. and Psychromonadaceae fam. nov. *Int J Syst Evol Microbiol* 54:1773–1788.
- Hou S, et al. (2004) Genome sequence of the deep-sea gamma-proteobacterium *Idiomarina loihiensis* reveals amino acid fermentation as a source of carbon and energy. *Proc Natl Acad Sci USA* 101:18036–18041.
- Ivars-Martinez E, et al. (2008) Comparative genomics of two ecotypes of the marine planktonic copiotroph *Alteromonas macleodii* suggests alternative lifestyles associated with different kinds of particulate organic matter. *ISME J* 2:1194–1212.
- Giovannoni SJ, et al. (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* 309:1242–1245.
- Tripp HJ, et al. (2008) SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* 452:741–744.
- Shi Y, Tyson GW, DeLong EF (2009) Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column. *Nature* 459:266–269.
- Kanehisa M, et al. (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 36:D480–D484.
- Yooseph S, et al. (2007) The Sorcerer II Global Ocean Sampling expedition: Expanding the universe of protein families. *PLoS Biol* 5:e16.
- Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402.
- Zimmer DP, et al. (2000) Nitrogen regulatory protein C-controlled genes of *Escherichia coli*: Scavenging as a defense against nitrogen limitation. *Proc Natl Acad Sci USA* 97:14674–14679.
- Mannino A, Harvey HR (1999) Lipid composition in particulate and dissolved organic matter in the Delaware Estuary: Sources and diagenetic patterns. *Geochim Cosmochim Acta* 63:2219–2235.
- Repeta DJ, Aluwihare LI (2006) High molecular weight dissolved organic carbon cycling as determined by natural abundance radiocarbon measurements of neutral sugars. *Limnol Oceanogr* 51:1045–1053.
- Lou Y-W, Friedrichs MAM, Doney SC, Church MJ, Ducklow HW (2010) Oceanic heterotrophic bacterial nutrition by semilabile DOM as revealed by data assimilative modeling. *Aquat Microb Ecol* 60:273–287.
- Panagiotopoulos C, Repeta DJ, Johnson CG (2007) Characterization of methyl sugars, 3-deoxysugars and methyl deoxysugars in marine high molecular weight dissolved organic matter. *Org Geochem* 38:884–896.
- Quan TM, Repeta DJ (2007) Characterization of high molecular weight dissolved organic carbon using periodate over-oxidation. *Mar Chem* 105:183–193.
- Lidstrom ME (2006) Aerobic methylotrophic prokaryotes. *The Prokaryotes*, ed Dworkin M (Springer, New York), 3rd Ed, pp 618–634.
- Neufeld JD, Boden R, Moussard H, Schafer H, Murrell JC (2008) Substrate-specific clades of active marine methylotrophs associated with a phytoplankton bloom in a temperate coastal environment. *Appl Environ Microbiol* 74:7321–7328.
- Neufeld JD, Chen Y, Dumont MG, Murrell JC (2008) Marine methylotrophs revealed by stable-isotope probing, multiple displacement amplification and metagenomics. *Environ Microbiol* 10:1526–1535.
- Neufeld JD, et al. (2007) Stable-isotope probing implicates *Methylophaga* spp and novel Gammaproteobacteria in marine methanol and methylamine metabolism. *ISME J* 1:480–491.
- Marie D, Partensky F, Jacquet S, Vaulot D (1997) Enumeration and cell cycle analysis of natural populations of marine picoplankton by flow cytometry using the nucleic acid stain SYBR Green I. *Appl Environ Microbiol* 63:186–193.
- Frias-Lopez J, et al. (2008) Microbial community gene expression in ocean surface waters. *Proc Natl Acad Sci USA* 105:3805–3810.
- DeSantis TZ, et al. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72:5069–5072.
- Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73:5261–5267.
- Pruesse E, et al. (2007) SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 35: 7188–7196.
- Ludwig W, et al. (2004) ARB: A software environment for sequence data. *Nucleic Acids Res* 32:1363–1371.
- Cole JR, et al. (2009) The Ribosomal Database Project: Improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* 37:D141–D145.
- Letunic I, Bork P (2007) Interactive Tree of Life (ITOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics* 23:127–128.
- Wang L, Feng Z, Wang X, Wang X, Zhang X (2010) DEGseq: An R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26: 136–138.
- Kristiansson E, Hugenholtz P, Dalevi D (2009) ShotgunFunctionalizer: An R-package for functional comparisons of metagenomes. *Bioinformatics* 25:2737–2738.
- R Development Core Team (2010) *R: A language and environment for statistical computing, version 2.11.1* (R Foundation for Statistical Computing, Vienna). Available at <http://www.r-project.org/>.

APPENDIX B

Distinct dissolved organic matter sources induce rapid transcriptional responses in sympatric populations of *Prochlorococcus*, *Pelagibacter* and the OM60 clade

Adrian K. Sharma^{a,*}

Jamie W. Becker^{b,*}

Elizabeth A. Ottesen^c

Jessica A. Bryant^a

Solange Duhamel^d

David M. Karl^e

Otto X. Cordero^a

Daniel J. Repeta^f

Edward F. DeLong^{a,g,1}

Departments of ^aCivil and Environmental Engineering and ^gBiological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

Departments of ^bBiology and ^fMarine Chemistry & Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA 02543

^cDepartment of Microbiology, University of Georgia, Athens, GA 30602

^dLamont Doherty Earth Observatory, Columbia University, Palisades, NY 10964

^eDepartment of Oceanography, University of Hawaii at Manoa, Honolulu, HI 96822

¹Corresponding author: delong@mit.edu

*These authors contributed equally to this work

In prep: Environmental Microbiology

Abstract

Marine dissolved organic matter (DOM) contains nearly as much carbon as the Earth's atmosphere. To investigate the roles of planktonic marine microbes in the biogeochemical cycling of DOM, we followed the responses of a surface water microbial assemblage to DOM derived from an axenic culture of *Prochlorococcus marinus* and high-molecular weight DOM concentrated from nearby surface waters.

Metatranscriptomic analyses indicated these treatments induced similar microbial community response dynamics due to the selective utilization of DOM components common to both treatments, including utilization of nitrogenous compounds. The rapid transcriptional response demonstrated by the *Prochlorococcus* population was consistent with utilization of proteinaceous material in the DOM. *Pelagibacter* also exhibited a rapid response to DOM additions that included increased transcript abundance for genes associated with utilization of methylated organic nitrogen compounds including methylamine- and glycine betaine-like compounds. Whereas *Prochlorococcus* and *Pelagibacter* responded rapidly to utilize DOM, opportunistic members of the NOR5/OM60 clade required more time to exhibit an observable response, which included elevated expression of genes associated with polysaccharide uptake and oxidation. In total the results indicated that nitrogen and DOM resource partitioning play a key role in structuring turnover of dissolved organic carbon pools by marine planktonic microbial communities.

Introduction

About one half of global primary production occurs in the ocean (Field et al., 1998) where a diverse group of phytoplankton fix carbon and nutrients into particulate organic matter (Azam, 1998). Exudation of metabolic waste products, viral lysis and predation all release a portion of microbial production into the water column as dissolved organic matter (DOM), a complex mixture of biochemicals of varying biological availability (lability) (Carlson, 2002) that changes in time and space (Aluwihare et al., 1999; Kujawinski et al., 2009; Mopper et al., 2007). DOM supports secondary production and microbial respiration (Hansell et al., 2009; del Giorgio and Duarte, 2002), with heterotrophic picoplankton representing the main consumers. Understanding how picoplankton interact with this dynamic DOM reservoir is complicated by the inherent phylogenetic and population diversity of microbial communities, the complexities of their collective metabolic properties and interactions, and by our ability to measure microbial assemblage activities and responses on appropriate temporal and spatial scales. For these reasons, characterizing and quantifying microbial DOM cycling in the sea is a significant challenge.

Several recent studies using experimental microcosms have explored the consumption of phytoplankton-derived isotopically labeled DOM sources by examining uptake patterns and changes in community composition among diverse taxonomic groups (Nelson and Carlson, 2012; Sarmiento and Gasol, 2012). These studies indicate that organisms with different taxonomic affiliations and varying ecological growth strategies exhibit preferences in both the phytoplankton-derived origin and compositional

properties of DOM. Other recent studies have combined meta-omics approaches with temporal field observations or experimental perturbations in coastal systems to gain insight into taxon-specific microbial activity in response to changes in naturally derived sources of DOM (Poretsky et al., 2010; Rinta-Kanto et al., 2012; Teeling et al., 2012). These studies highlight patterns of taxon-specific resource partitioning of DOM, community strategies for energy scavenging under carbon limitation, and temporal successions of microbial populations in response to dynamic changes in DOM during a natural phytoplankton bloom. These studies also demonstrate the utility of pairing DOM uptake experiments with meta-omics methodologies as a means of uncovering metabolic strategies involved in microbial DOM consumption.

The details of the functional and metabolic roles of specific microbial taxa in DOM degradation however, still remain largely unknown. Community response dynamics to DOM perturbations across short time scales are only poorly understood, as most methods lack the necessary sensitivity to track transient responses. Such resolution would provide insight into the complex response mechanisms of microbial communities that result from both ecological variables and DOM resource partitioning.

Here we report a microcosm-based DOM perturbation experiment in an oligotrophic region of the ocean focused on measuring rapid temporal response dynamics and functional roles of oligotrophic taxa that are ubiquitous in the open ocean (Lauro et al., 2009; Yooseph et al., 2010; Nelson and Carlson, 2012). Incubations were conducted at Station ALOHA in the North Pacific Subtropical Gyre (NPSG), an oceanic region where the infrequency of deep-water mixing events results in low inorganic nutrient

concentrations and limitation of primary production (Karl and Lukas, 1996; Karl et al., 2008). To study microbial communities under extreme oligotrophic conditions, perturbation experiments were conducted in late spring, a time of the year when the water column at Station ALOHA is highly stratified (Karl et al., 2012) and inorganic nutrient levels are most depleted (see supplementary figure S1).

In order to study differences in utilization of DOM from different sources, we examined the response of surface water microbes to two distinct DOM types in our microcosms. To examine breakdown of compounds in the standing DOM pools, we concentrated naturally occurring high-molecular weight dissolved organic matter (HMWDOM) on site from Station ALOHA surface seawater using an approach similar to that of McCarren *et al.* (2010). This size-fractionated DOM pool is considered to be “semi-labile”, rich in polysaccharides (Aluwihare et al., 2005) and other high-molecular weight compounds that might be preferred by copiotrophic taxa (McCarren et al., 2010). In order to examine breakdown of newly produced “labile” DOM, a second DOM source was prepared by concentrating the hydrophobic fraction of exudate from an axenic culture of *Prochlorococcus* strain MIT9313 (ProDOM). *Prochlorococcus* is the dominant photoautotroph in nutrient poor ocean gyres and heterotrophic taxa in these regions are likely adapted to utilizing substrates derived from their photosynthate (Partensky et al., 1999; Bertlisson et al., 2005). Both DOM amendments were used to enrich surface seawater microbial communities in controlled microcosm experiments. We monitored microbial community growth, composition, and gene expression along with exoenzyme

activity over a 36-hour period, generating a detailed look at the temporal and functional responses of different microbial taxa to changes in ambient DOM quantity and quality.

Materials and Methods

Experimental setup and sample collection

The experimental setup, including RNA and DNA sampling, is described briefly below. For details, see Supplementary File 1. Seawater for microcosm incubation experiments was collected from the bottom of the mixed layer (35 m) and transferred to 20 L carboys. 2 L of DOM concentrate (HMWDOM or ProDOM) was added to 18 L of seawater. Carboys were incubated in a deck-board incubator that simulated surface seawater temperature and light level intensity at 25-45 m. At selected time points, bacterioplankton biomass from ~2 L samples was rapidly collected for RNA extraction by filtration. At both the beginning and the end of the experiment, biomass was similarly collected for DNA extraction.

DNA and RNA isolation

Total DNA was extracted and purified using the Quick-Gene 610 l system (Fujifilm, Tokyo, Japan) and DNA Tissue Kit L with a modified lysis protocol. Total RNA was extracted using the *mirVana*TM miRNA Isolation kit (Ambion) as described in (Frias-Lopez et al., 2008), with modifications to account for the recovery of RNA from Sterivex filters (see Supplementary File 1 for further details on DNA and RNA isolation).

rRNA subtraction, RNA amplification, cDNA synthesis and pyrosequencing

Bacterial and archaeal ribosomal RNA transcripts were removed from total RNA extracts using a subtractive hybridization protocol described in (Stewart et al., 2010). Sample-specific rRNA probes for this study were based on amplicons obtained from DNA sampled from 0h and 36h microcosm communities. First round cDNA synthesis and RNA amplification were performed using the MessageAmp™ II-Bacteria kit (Ambion) and second round cDNA synthesis using the SuperScript® III First-Strand Synthesis System (Invitrogen) as described previously in (Frias-Lopez et al., 2008) and (Shi et al., 2009). Pyrosequencing was performed using Titanium series chemistry on a Roche Genome Sequencer FLX instrument (half-plate run per sample) as indicated in (Stewart et al., 2012).

Flow cytometry

Flow cytometry was performed as described in (McCarren et al., 2010). *Prochlorococcus* cells were identified based on fluorescence under a 692 nm laser (red fluorescence, 692_40) and vertical forward scatter. Prior to total cell counts, samples were stained with SYBR Green (Invitrogen, Carlsbad CA) for 10 min, and cells were identified based on fluorescence and scatter signals (Marie et al., 1997).

Exoenzyme assay

β -glucosidase activity was measured as an increase in fluorescence of the product 4-Methylumbelliferone (MUF) released after enzymatic hydrolysis of the non-fluorescent

4-Methylumbelliferyl- β -D-glucopyranoside (MUF-Glc; Sigma-Aldrich) substrate (see Supplementary File 1 for details).

Preparation of DOM amendments

High-molecular weight DOM was isolated and concentrated from surface seawater as described in McCarren *et al.* (2010) with a few modifications as indicated in Supplemental File 1. Samples were taken for TOC quantification, cell counts, and viral particle counts from the raw seawater, 0.2 μ m pre-filtered, and permeate water periodically during ultrafiltration and from the concentrate upon completion. This sample suite was also taken after serial filtration of the concentrate through a 0.1 μ m Polycap TC prefilter (Whatman) followed by a 30-kDa polyethersulfone membrane (Millipore) to remove viral particles as described in McCarren *et al.* (2010), before adding 2L of concentrated material to the 18L microcosm. Cell and viral counts determined pre- and post-30kDa filtration using flow cytometry and fluorescence microscopy indicated the removal of cells, cell debris, and the reduction of virus particles below ambient concentrations of seawater from the mixed-layer at Station ALOHA (data not shown). *Prochlorococcus*-derived DOM was isolated and concentrated from an axenic culture of *Prochlorococcus* strain MIT9313 grown in 20 L of Pro99 medium prepared according to existing protocols (Moore *et al.*, 2007) in sterile Sargasso seawater. Details regarding the growth of MIT9313 and solid-phase extraction of DOM can be found in Supplemental File 1. Dried ProDOM material was stored in a combusted amber vial in the dark prior to resuspension at sea in 1.6 ml of filtered (0.1 μ m Polycap

TC; Whatman) seawater collected from 35 m at Station ALOHA. Half of this resuspension was added to 2L of sterile 35m seawater, and subsequently added to the 18L microcosm.

Quantification of organic carbon and dissolved nitrogen

Total organic carbon (TOC) and total dissolved nitrogen (TDN) were measured using the high temperature combustion method on a Shimadzu TOC-V_{CSH} with platinized aluminum catalyst coupled to a TNM-1 total nitrogen detector, while particulate organic carbon (POC) was measured at the University of California Davis Stable Isotope Facility. Details regarding sample handling and processing are provided in Supplementary File 1.

Chromatographic separation and detection of MIT9313 metabolites

Chromatographic separation and detection of metabolites derived from *Prochlorococcus* strain MIT9313 was achieved using an Agilent 1200 series liquid chromatograph coupled to an Agilent 6130 mass spectrometer with an atmospheric electrospray ionization source. Mass spectral data was acquired from 100-2000 Da in the positive mode and Ions with a minimum signal intensity 5-fold greater than the maximum noise level were included in analysis. Details regarding run conditions and feature detection are provided in Supplementary File 1.

Bioinformatics

Parameters for *in silico* removal of reads containing rRNA, the removal of

artificially duplicated sequences generated during pyrosequencing, BLASTX searches, and the assignment of higher order taxonomy in MEGAN (Huson et al., 2007) can be found in Supplementary File 1. Statistical comparison of metatranscriptomes from treatment and control time points were also performed using baySeq, a method that uses an empirical Bayes approach to detect patterns of differential expression (DE) among genes within a set of samples (Hardcastle and Kelly, 2010). The differential expression of reference genes was determined at the whole community level for each time point individually and the taxonomic results of this analysis are represented in Figure 2C (see Supplementary File 1 for further details).

Taxon-specific ortholog sequence clusters were generated separately for *Prochlorococcus*, *Pelagibacter*, and the OM60 clade using sequenced genome representatives from NCBI. Within each taxonomic bin, sequence counts for genes present in multiple reference genomes were compiled to generate “ortholog cluster”-based transcript abundances. This approach was implemented to avoid artificial division of transcript pools from environmental organisms amongst multiple, imperfectly matched reference sequences (see Ottesen et al., 2013). Specific details regarding the construction of these ortholog clusters for our study and the associated bioinformatic parameters are presented in Supplementary File 1. Taxon-specific ortholog count files were used in baySeq in pairwise differential gene expression tests to identify orthologs with a posterior probability greater than or equal to 0.9 of being significantly enriched or under-represented in a single treatment relative to the control at each time point. As outlined in the results and discussion, DE orthologs were further binned into central or auxiliary

pathways (for specific details see Supplementary File 1). Changes in KEGG pathway abundances supported by DE KOs between treatment and control time points for *Prochlorococcus* and *Pelagibacter* are displayed in heatmaps (Figure 3) which were generated in R using the heatmap.2 function in gplots (Warnes et al., 2009) (<http://hosho.ees.hokudai.ac.jp/~kubo/Rdoc/library/gplots/html/00Index.html>).

Results and discussion

Microbial community growth and exoenzyme activity

Both amendments induced similar increases in total cell numbers relative to the control (Figure 1B and C), despite the fact that the two DOM sources differed not only in quality but also in quantity of organic carbon and nitrogen added (Figure 1A). The HMWDOM amendment increased concentrations of both dissolved organic carbon (DOC) and dissolved organic nitrogen (DON) by approximately 140%, while the ProDOM amendment increased DOC and DON by only 7%. DOC concentrations were 191 $\mu\text{M C}$ for HMWDOM and 85.5 $\mu\text{M C}$ for ProDOM, approximately 2.4x and 1.1x the ambient value of 79.9 $\mu\text{M C}$ in the control microcosm. Both amendments increased DOC and DON concentrations in a manner that was proportional to the ratio of DOC:DON in the control. Despite these substantial differences in substrate quantity, both treatments induced an approximate 6% increase in cell density relative to the control (Figure 1B and 1C), suggesting that the ProDOM treatment contained a higher proportion of labile DOM that could be converted into cellular biomass. Whereas the HMWDOM addition concentrated a fraction of the DOM already present in the sample, chemical analysis of

the ProDOM material using high performance liquid chromatography-electrospray ionization mass spectrometry revealed the presence of 1,491 distinct low-molecular weight (<1.2 kDa) compounds (data not shown; see experimental procedures), and this amendment could have introduced new DOM components to the microcosm.

Flow cytometric analysis indicated two stages of diauxic-like of growth in the microbial community over the 36 h time course in both treatments relative to the control. The first and largest increase in cell numbers occurred between 12 and 19 h in both treatments relative to the control, where *Prochlorococcus* cells accounted for approximately 70% of the total cell growth observed in treatments between these time points (Figures 1B-1G). The second, less pronounced growth stage in DOM amended microcosms occurred between 19 and 36 h. Here, *Prochlorococcus* comprised a much smaller fraction of the total cell growth in the HMWDOM and ProDOM microcosms, indicating that heterotrophs could be responsible for the later increases in cell numbers.

The final time point in both treatments was characterized by an increase in β -glucosidase exoenzyme activity (Figure 1H and 1I) consistent with heterotrophic growth at later time points. β -glucosidase is an enzyme produced by heterotrophic picoplankton that catalyzes the selective cleavage of glucosidic bonds in order to break down oligosaccharides into smaller sugars that can be transported into the cell. At the 36-hour time point, assays indicated a 130% increase in activity in the HMWDOM treatment and a 46% increase in activity in the ProDOM treatment relative to the control. These findings indicate the presence of polysaccharides in both treatments, and the level of β -glucosidase activity per unit carbon added in each treatment suggests that labile

polysaccharides likely comprised a substantial proportion of the DOC in the ProDOM amendment. The observed increase in β -glucosidase activity at the final time point in both treatments was likely related to heterotrophic growth and activity in the latter stages of the experiment.

Meta-genomic and -transcriptomic profiling of microcosm community structure

To further examine microbial community response dynamics to DOM additions from both a taxonomic and metabolic perspective, metagenomic and metatranscriptomic data was obtained before perturbation as well as 2, 12, 27, and 36 h after amendment. Table 1 outlines read numbers and database statistics for community DNA and cDNA samples sequenced from each of the different microcosms. Microbial community composition and transcriptional activity in the 35 m seawater used for our microcosm experiments (0 h DNA and cDNA) was dominated by *Prochlorococcus* and *Pelagibacter* (Figure 2A). Surface ecotypes of *Prochlorococcus* (Johnson et al., 2006) comprised approximately 50% of meta-genomic and 36% of metatranscriptomic reads with a significant BLASTx hit to the NCBI-nr database. The next most abundant group in our starting community was *Pelagibacter*, which accounted for 10% of assignable reads in 0 h DNA and cDNA samples. The vast majority of *Pelagibacter*-like sequences shared highest similarity to the *Pelagibacter ubiquus* HTCC7211 genome, a strain cultivated from the oligotrophic Sargasso Sea (Stingl et al., 2007).

Metagenomic samples revealed small increases in the relative abundance of a variety of heterotrophic groups in both HMWDOM and ProDOM 36 h DNA when

compared to the control 36 h DNA (Figure 2B). These groups were OM60, Alteromonadales, Rhodobacterales, SAR116 and Flavobacteriales. DNA sequences from the SAR11 group increased in the ProDOM treatment by nearly 50% over the control, suggesting a strong growth response to this DOM source. Metatranscriptomic samples indicated subtle changes in taxonomic composition over time in response to DOM enrichments, with increased representation of cDNA reads from *Prochlorococcus* and *Pelagibacter* in treatments at multiple time points. The 36 h time point from both treatments also displayed increases in the transcriptomic activity of OM60 and Alteromonadales.

While these temporal sequence data indicated only minor taxonomic shifts in response to DOM perturbations on this short timescale, they did indicate changes in gene expression patterns within the ambient microbial community at Station ALOHA. To identify NCBI-nr reference genes with statistically significant read counts between metatranscriptomes at each time point, we used baySeq, a Bayesian method for identifying differential gene expression between samples (Hardcastle and Kelly, 2010). Any reference gene with a posterior probability of ≥ 0.9 of being differentially expressed between a treatment and the control was considered significant. Figure 2C shows the taxonomic association of only those NCBI-nr genes that were significantly enriched in cDNA samples from the treatments relative to the control at each time point where significantly underrepresented genes were excluded. From a taxonomic perspective, both treatments exhibited a similar temporal trend in differential gene expression. At early time points the majority of differentially expressed transcripts was associated with

Prochlorococcus and *Pelagibacter*, with increasing transcript abundances from different heterotrophic taxa at the later time points. *Pelagibacter* also exhibited rapid differential gene expression responses in both DOM treatments. These taxonomic trends in differential gene expression mirrored the microbial community growth and abundance patterns (Figure 1B-G), where *Prochlorococcus*-like cells represented the majority of growth at earlier time points, but not later. Combined with high β -glucosidase activity at 36 h (Figure 1H and I), these independent methods of analysis support the hypothesis that a microbial succession in growth and activity occurred in our treatments in response to DOM perturbations. Abundant, oligotrophic taxa (*Prochlorococcus* and *Pelagibacter*) were observed to rapidly respond to changes in the ambient DOM pool. In contrast, the opportunistic heterotrophic taxa (OM60, Alteromonadales, etc.) gradually increased in numbers, and their transcriptional responses became more apparent, in the latter stages of the incubation.

Taxon-specific responses to DOM perturbations inferred from genome-centric transcriptomic analyses

To gain additional insight into microbial community dynamics and the DOM substrate utilization patterns driving the microbial successions previously described, we performed differential gene expression analyses on those taxa with the greatest percentages of differentially expressed genes (relative to the control; Figure 2C). These taxa were represented by *Prochlorococcus*, *Pelagibacter*, and the OM60 clade of gammaproteobacteria. The latter group is comprised of photoheterotrophic organisms

capable of aerobic anoxygenic photosynthesis (Fuchs et al., 2007; Cho et al., 2007) that have a cosmopolitan distribution in marine surface waters, but tend to be more abundant in nutrient rich coastal systems relative to the oligotrophic open ocean (Pernthaler and Pernthaler, 2005; Yan et al., 2009). Within each taxon bin, transcript counts for genes shared between multiple reference genomes of the same taxa were combined into ortholog counts annotated with KEGG pathway information [see Ottesen *et. al.*, (2013, experimental procedures and Dataset S1)]. Analyses of transcriptional dynamics focused on changes in relative transcript abundance within each specific taxonomic bin. Significant differential expression of KEGG annotated orthologs (Datasets S2-S5) was used to direct and support comparisons of taxon-specific pathway abundances through time. In some cases, non-significant orthologs are discussed in taxon specific analyses as supporting information.

Prochlorococcus

DOM enrichments induced specific, rapid changes in *Prochlorococcus* gene expression. A set of three pathways had differential expression (DE) of KEGG orthologs (KOs) that were enriched in treatments at early time points, but underrepresented at later ones (Figure 3A). At the 2 h time point, these included the protein biosynthesis pathways Ribosome and Translation factors, which together accounted for the DE of 31 KOs in HMWDOM (dataset S2) and 39 in ProDOM (dataset S3). The third pathway, RNA degradation, was only enriched at 2 h due to DE of DnaK and GroEL. These chaperones help polypeptide chains emerging from the ribosome reach their native conformation.

These observations suggest that DOM additions quickly stimulated a burst of protein biosynthesis in *Prochlorococcus*.

Accompanying this presumptive burst in protein biosynthesis, *Prochlorococcus* exhibited an immediate and sustained increase in transcripts mapping to KOs involved in genome replication and cell division (Figure 3A). Differentially expressed KOs included 11 different KO's in HMWDOM (Dataset S2) and 24 in ProDOM (Dataset S3) associated with KEGG pathways for Chromosome, DNA replication, and DNA recombination and repair. In addition, KOs from lipid and starch biosynthesis pathways were significantly more abundant in DOM-enriched microcosms, suggesting increased levels of membrane biogenesis and carbon storage (Figure 3A). These observations are further supported by the enrichment of various orthologs from carbon fixation and carbon concentrating mechanisms in treatments at early time points (Dataset S2 and S3). The greater number of differentially expressed KO's for protein, cell division and lipid biosynthesis pathways in ProDOM, combined with their higher percent increase relative to the HMWDOM treatment indicates that the magnitude of response by *Prochlorococcus* at the level of gene expression was more pronounced for ProDOM. These observations support the hypothesis (also supported by the flow cytometry data) that ProDOM contained a greater amount of labile DOM, despite the higher quantity of carbon supplied in the HMWDOM treatment (Figure 1A). The fact that genome replication and cell division pathways were underrepresented in the treatments at 27 h appears to be the result of slower *Prochlorococcus* growth in the control (Figure 1D and 1E).

Prochlorococcus populations in the control microcosm appeared to dedicate a greater fraction of its transcriptome to nitrogen acquisition and assimilation relative to the treatments. KOs involved in nitrogen metabolism and transport pathways were under-represented in treatment samples (Figure 3A), including but not limited to the ammonium assimilation protein glutamine synthetase, and both the permease and substrate binding subunits of the Urea ABC transporter. Additionally, transcripts encoding an ammonium transporter ortholog unassigned in KEGG, represented the only *Prochlorococcus* ortholog that was significantly underrepresented in both treatments at every time point (see Cluster 287, Datasets S2 and S3). The genes encoding glutamine synthetase, and urea and ammonium transporters are highly expressed by *Prochlorococcus* as a nitrogen scavenging response to nitrogen stress in culture (Tolonen et al., 2006). The higher expression of these nitrogen acquisition orthologs in *Prochlorococcus* transcriptomes in the control microcosm, combined with the growth and biosynthetic signals in treatments, suggests that both DOM enrichments provided a source of DON.

To obtain a better understanding of the DOM utilization patterns that contributed to the physiological responses observed in *Prochlorococcus*, the DE of orthologs belonging to auxiliary KEGG pathways (as opposed to core pathways involved in central metabolic processes) was examined in detail. The number of DE *Prochlorococcus* orthologs binned into these two pathway categories (auxiliary vs. core) is presented in table 2 and details regarding KEGG pathway assignments as core or auxiliary are presented in Supplementary Files 1 and 2. At 2 h, *Prochlorococcus* transcripts mapping to the KO D-amino acid oxidase exhibited an ~8-fold increase in both treatments relative

to the control (Tables S1 and S2). D-amino acid oxidase catalyzes the breakdown of D-amino acids into their corresponding oxo-acids and ammonium and its increased activity here suggests a source of proteinaceous material common to both treatments, perhaps in the form of peptidoglycan. Breakdown of D-amino acids should directly correspond to an increase in glutamine and glutamate via GS-GOGAT ammonium assimilation (Muro-Pastor et al., 2005). An increase in cellular concentrations of glutamate is supported by the DE of a number of KOs involved in the biosynthesis of aspartate, proline, and arginine - amino acids that use glutamate as a metabolic precursor (Tables S1 and S2). In addition to peptidoglycan as a possible source of organic nitrogen, the enrichment of various DE orthologs encoding peptidases and proteases was observed in treatment transcriptomes, suggesting *Prochlorococcus* was capable of the uptake and degradation of oligopeptides present in the DOM additions. Transcripts encoding proline iminopeptidase were enriched in both treatments at 27 h, and these enzymes are known to function in the proteolytic systems of lactic acid bacteria growing on milk proteins (Kunji et al., 1996). Transcripts for five different *Prochlorococcus* proteases were exclusively enriched in the ProDOM treatment, and four of them were enriched at 2 h, indicating a rapid response to an influx of protein. These observations indicate that the *Prochlorococcus*-derived DOM fraction we used was richer in labile protein material than HMWDOM derived from seawater, despite the large discrepancy in organic carbon quantity (Figure 1A).

Overall, these results indicate that DOM enrichments provided *Prochlorococcus* with a source of DON substrates in the form of proteinaceous material. Studies have

shown that *Prochlorococcus* utilizes various nutrient acquisition strategies to circumvent nitrogen and phosphorus depletion (Martiny et al., 2006; Martiny et al., 2009), and highlight the ability of this cyanobacterium to utilize organic nutrients for growth (Martínez et al., 2012; Gómez-Pereira et al., 2013). Some studies (Zubkov et al., 2004; Mary et al., 2008) have suggested that the ecological success of *Prochlorococcus* in oligotrophic regions of the ocean is due in part to their high uptake rate of amino acids, which can account for 33% of the total bacterioplankton turnover of amino acids in oligotrophic parts of the Arabian Sea (Zubkov et al., 2003). These authors note that the classical distinction between auto- and hetero-trophic organisms in the marine environment is blurred in oligotrophic waters where photosynthetic cyanobacteria often demonstrate mixotrophic tendencies by utilizing organic nutrients. An increasing body of evidence demonstrates the ability of phytoplankton to utilize organic matter to fulfill nutrient requirements (Palenik and Morel, 1990; Berg et al., 2003; Bronk et al., 2007; Mulholland and Lee, 2009; Baran et al., 2011; Muñoz-Marín et al., 2013). Our study strongly supports these views given the rapid response of the *in situ* *Prochlorococcus* population to DOM enrichment.

Pelagibacter

Pelagibacter also exhibited rapid changes in gene expression in response to both HMWDOM and ProDOM enrichment, particularly in protein biosynthesis pathways (Figure 3B). The majority of DE KO's from the pathways ribosome (ribosomal proteins) and chaperones (e.g. DnaK and GroEL) occurred early (2 and 12 h) in both treatments

(Datasets S4 and S5) suggesting that, like *Prochlorococcus*, *Pelagibacter* responded to the influx of organic matter by rapidly modifying its proteome. *Pelagibacter* also demonstrated transcriptional growth signals in response to DOM perturbation (Figure 3A and 3B). The DE KOs falling into the categories of DNA replication proteins and chromosome were represented by a total of 6 different KOs in HMWDOM but only a single KO in ProDOM (Dataset S4 and S5). Sowell et al. (2008) found that *Pelagibacter* cells remained physiologically poised to quickly resume growth from stationary phase when conditions improve, which is supported by the rapid response of *Pelagibacter* observed here.

Like *Prochlorococcus*, *Pelagibacter* cells responded to DOM addition by decreasing the expression of orthologs involved in nitrogen acquisition (Figure 3B). The ammonium assimilation protein glutamine synthetase was detected as significantly underrepresented in *Pelagibacter* reads from HMWDOM (2 h) and ProDOM (2 h, 36 h) and tended to be more highly expressed in control microcosms at all time points with the exception of 27 h. Two other DE nitrogen metabolism KOs, homologs of the glycine cleavage system T (GcvT) protein, were also underrepresented in both treatments at various time points. The canonical GcvT protein degrades the amino acid glycine by removing a methyl group for C1 oxidation via the tetrahydrofolate (THF) pathway and producing ammonium that can be assimilated by the cell. Interestingly, both GcvT-like aminomethyltransferase homologs are linked to genes potentially involved in either the uptake or degradation of methylated compounds in multiple *Pelagibacter* genomes (Sun et al., 2011). This suggests a role for these enzymes in nitrogen acquisition via the

breakdown of methylated organic nitrogen substrates.

A net decrease in transporter gene expression by *Pelagibacter* was observed in both DOM treatments after two hours (Figure 3B), suggesting that *Pelagibacter* cells were reallocating cellular resources away from nutrient acquisition and diverting them to biosynthetic processes. In a study conducted in the nutrient-deplete Sargasso Sea, transport functions were shown to dominate the metaproteome of *Pelagibacter*, indicating that these organisms depend on a variety of high affinity transporters to fulfill their nutrient requirements in oligotrophic environments (Sowell et al., 2009). *Pelagibacter* orthologs involved in a variety of transport functions were underrepresented in treatments (Tables S3 and S4), including ABC transporter substrate-binding proteins for sugars, phosphate and periplasmic C₄-dicarboxylate-binding receptors. However, the majority of DE transport orthologs (13/23 in HMWDOM and 10/18 in ProDOM) function in the uptake of organic and inorganic nitrogen containing compounds.

Whereas transporters in general were underrepresented in *Pelagibacter* transcripts from our organic nutrient-rich treatments, a small number of transport orthologs were stimulated by the DOM additions (Tables S5 and S6), suggesting selective up-regulation of transporters for specific substrates. These transporters included a gene encoding a homolog of an ABC transporter-like substrate binding protein for polar amino acids (PheC). The gene encoding this protein is linked to a sarcosine oxidase operon in multiple *Pelagibacter* genomes, suggesting this transporter could be involved in the uptake of methylated organic nitrogen compounds. Sarcosine oxidase and its homologs are known to function in the degradation of substrates such as glycine betaine and

methylamines (Latypova et al., 2010; Sun et al., 2011). *pheC* was enriched in the HMWDOM treatment relative to the control microcosm at all time points, but only significantly DE at 2 h and 36 h. None of the genes encoding the four sarcosine oxidase subunits were detected as differentially expressed under this condition, however transcripts encoding the SoxG subunit were 5.6-fold higher in the HMWDOM treatment at 36 h (posterior probability = 0.73). A later response was observed for *pheC* in ProDOM (36 h), along with the enrichment of other homologs of transporters for glycine betaine (27 h) and a general amino acid substrate-binding protein (36 h). Increased expression of these transporters at later time points may indicate that endogenous microbial community activity, stimulated early on by the ProDOM amendment, such as increased primary production caused by growth of *Prochlorococcus* cells, could have caused the synthesis of labile DON substrates in treatment microcosms over time.

To gain further insight into *Pelagibacter* DOM substrate utilization we examined differentially expressed KOs belonging to auxiliary pathways outside of those core pathways involved in central metabolic processes. Pathways designated as core or auxiliary are listed in Supplementary file 2 and the breakdown of differentially expressed KOs among these categories is shown in table 2. The increased expression of the gene encoding formate-tetrahydrofolate ligase (*fhs*) in *Pelagibacter* transcripts from the HMWDOM treatment (Table S7) suggested an influx of C1 groups entering the THF oxidation pathway. The FHS protein catalyzes the second-to-last step of the C1 THF oxidation pathway, and couples substrate level phosphorylation to the oxidation of formyl-tetrahydrofolate. Transcripts encoding this protein were on average 62% higher in

the HMWDOM treatment relative to the control at every time point, but only detected as significantly DE at 27 h. This observation, combined with enrichment of transcripts for the sarcosine oxidase-linked transporter *pheC* (Table S5) in the HMWDOM treatment, potentially links the uptake of methylated organic nitrogen compounds to C1 oxidation in *Pelagibacter*, providing a mechanism for nitrogen acquisition and the production of energy needed to fuel the growth signals observed in KEGG pathways (Figure 3B).

At the 2 h time point in the ProDOM treatment, two *Pelagibacter* orthologs involved in homocysteine biosynthesis were enriched (Table S8). These were homoserine O-acetyltransferase (*metX*) and O-acetylhomoserine (thiol)-lyase (*metY*). Homocysteine is required by the enzyme betaine-homocysteine methyltransferase (BHMT) for the first demethylation step in glycine betaine degradation (Barra et al., 2006). Although the gene encoding BHMT was not detected as DE over the course of the experiment (likely due to its low representation among SAR11 transcripts), its expression was 4-fold higher in ProDOM relative to the control at 2 h (posterior probability = 0.63). A gene encoding a homolog of γ -butyrobetaine dioxygenase (*γ -bbh*) was enriched in ProDOM at 12 h and this protein catalyzes the first step in the degradation of γ -butyrobetaine, a substance whose degradation proceeds through the L-carnitine degradation pathway resulting in the production of glycine betaine (Kleber, 1997). In the *Pelagibacter* HTCC7211 genome, homologs encoding subunits of a L-proline/glycine betaine ABC transporter are linked to the *γ -bbh* gene. Together these observations suggest that the ProDOM treatment provided a source of γ -butyrobetaine- and glycine betaine-like substrates, thereby supplying *Pelagibacter* with a source of both nitrogen and energy.

A study by Sun *et al.* (2011) used both genomic data and physiological tests to show that *Pelagibacter* is capable of the uptake and degradation of a wide variety of one-carbon compounds including methyl functional groups from methylated compounds, which provide a source of energy via the C1 tetrahydrofolate (THF) oxidation pathway. Our results suggested that in addition to a source of energy, *Pelagibacter* utilizes methylated organic nitrogen compounds such as methylamines and glycine betaine from the ambient DOM pool to acquire nitrogen. A similar scenario has been observed for *Pelagibacter* and organophosphorus molecules in the Sargasso Sea, where phosphate concentrations during summer stratification can be one to two orders of magnitude lower than at Station ALOHA (Wu *et al.*, 2000; Cavender-Bares *et al.*, 2001). A metaproteomic study conducted in the euphotic zone of the Sargasso Sea found that the periplasmic substrate-binding protein for phosphonate acquisition was among the most frequently detected *Pelagibacter* proteins (Sowell *et al.*, 2009), indicating that these organisms rely on the ambient DOM pool for survival under nutrient poor conditions.

OM60 Clade

Analysis at the whole community level found that orthologs from genomes within the OM60 clade of gammaproteobacteria were highly represented among a collection of DE orthologs from heterotrophs that displayed increased abundance and activity at later time points in both treatments (27 and 36 h; Figure 2C), relative to the control. Group-specific ortholog clusters were generated for seven OM60 genomes available at NCBI. Of the 3,666 OM60 clade ortholog clusters identified among DNA and cDNA samples

(Dataset S1), 37 were detected as differentially expressed in cDNA samples from the HMWDOM treatment and 70 from the ProDOM treatment. The vast majority of these DE orthologs occurred at the 27 and 36-hour time points, and those enriched in treatments included orthologs from KEGG pathways involved in protein, nucleotide and peptidoglycan biosynthesis (Tables S9 and S10). Interestingly, these later time points also coincide with the growth of non-*Prochlorococcus* cell types in treatments as determined by flow cytometry data and increasing activity of β -glucosidase (Figure 1).

Differentially expressed orthologs reveal OM60 polysaccharide utilization in both DOM treatments. For example, cluster 4450 encodes a predicted beta-glucoside-specific TonB-receptor family protein that was enriched in both treatments at 27 and 36 h (Tables S9 and S10). In the HTCC2148 genome, this TonB receptor homolog is located downstream of a glycosyl hydrolase gene, which was significantly enriched in OM60 transcripts in the HMWDOM treatments at 27 h, and displayed a 3.6-fold increase in ProDOM (posterior probability = 0.65). Glycosyl hydrolases (GHs) are enzymes that catalyze the hydrolysis of the glycosidic linkages in polysaccharides and are often secreted outside the cell. Further annotation of the protein-encoding sequence of this GH homolog from the HTCC2148 genome using DBCAN, an HMM-based database for carbohydrate-active enzymes (<http://csbl.bmb.uga.edu/dbCAN/>), indicates that the first 300 residues share strong similarity to GH16 family hydrolases. Enzymes of the GH16 family are known to cleave β -1,4 or β -1,3 linked glucans and galactans (Baumann et al., 2007; Hehemann et al., 2010), and these latter types of linkages are abundant in HMWDOM (Aluwihare et al., 1997). β -glucanases of the GH16 family share a signature

catalytic motif EXDXXE in the active site (http://www.cazypedia.org/index.php/Glycoside_Hydrolase_Family_16) (Michel et al., 2001), which is present in the GH16-encoding homolog from the HTCC2148 genome and among the transcripts mapping to this region of the gene in both treatments at the 27 h time point. The differential expression of this GH16 homolog provides a potential link between the OM60 clade and the degradation of polysaccharides with β -glycosidic linkages. The increased expression of this GH16 homolog could also contribute to the increased β -glucosidase exoenzyme activity observed in the treatment microcosms. However, GH16 enzymes preferentially hydrolyze longer-chain substrates, generating shorter oligosaccharides more similar to the analogs used in our assay and preferred by β -glucosidase enzymes.

In accordance with signals for polysaccharide utilization and uptake, OM60 transcripts from the ProDOM treatment (Table S10) showed enrichment in genes encoding glycolysis (enolase and pyruvate kinase at 27 h) and citric acid cycle enzymes (isocitrate dehydrogenase and succinyl-CoA synthetase at 36 h). That similar signals were not observed in HMWDOM (Table S9) may indicate a slower response time in this treatment compared to ProDOM. This hypothesis is supported by the higher proportion of OM60 transcripts relative to the entire community in ProDOM cDNA samples at 27 and 36 h (Figure 2A) and the overall increase in DE orthologs in this treatment. These observations again suggest that the ProDOM amendment contained a higher concentration of more labile substrates, which may have accelerated the rate at which the heterotrophic population was able to respond.

Like *Pelagibacter*, OM60 cells in the control microcosm appeared to allocate a larger fraction of their transcriptome to the expression of orthologs involved in nutrient acquisition (Tables S11 and S12), which may suggest a more universal response of marine heterotrophs to acquire a broad range of substrates in oligotrophic conditions. While the expression of a few TonB-dependent transporters (TBDT) appeared to be stimulated in response to DOM enrichments, many were underrepresented in treatments and on average these 12 orthologs exhibit a ~8-fold decrease relative to the control. TBDT are an extremely diverse group of proteins that can facilitate the transport of a variety of large molecules across the outer membrane, including rare metals or vitamins, proteins, organic acids and also polymeric substances such as polysaccharides (Schauer et al., 2008). Organisms of the OM60 clade contain anywhere from 27 to 37 TBDT per genome, indicating that these types of transporters play a large role in complementing the nutritional requirements of the cell (Tang et al., 2012).

The OM60 group was a low abundance heterotrophic taxa that demonstrated increasing transcriptional activity over time in response to both DOM enrichments. Using fluorescence *in-situ* hybridization, Yan et al. (2009) found organisms of the OM60 clade were much less abundant in open ocean samples (0.1-0.5% of cells) than in coastal systems (3-7% of cells), suggesting these organisms thrive in more nutrient rich systems. Increased abundance of OM60 cells has also been shown to co-occur with algal blooms (Eilers et al., 2001; Alonso-Sáez et al., 2007; Yan et al., 2009) and many studies demonstrate that diverse low abundance taxa bloom under increasing concentrations of organic nutrients (Cottrell and Kirchman, 2000; Eilers et al., 2000; McCarren et al.,

2010; Romera-Castillo et al., 2011; Tada et al., 2011; Nelson and Carlson, 2012). These opportunistic taxa exhibit a “feast or famine” lifestyle (Nissen, 1987; Flårdh et al., 1992; Srinivasan and Kjelleberg, 1998) and are often referred to as copiotrophs (Lauro et al., 2009; Yooseph et al., 2010). Teeling *et al.* (2012) observed successions in low abundance copiotrophs in response to a naturally occurring phytoplankton bloom, accompanied by metaproteomic signals for polysaccharide utilization. These signals were based on increases in expression of glycosyl hydrolases and TonB transport enzymes, which were similar to our observations for OM60 transcripts at later time points in this study. Although our amendments differed greatly in the quantity and quality of the DOM added, similar to organic nutrient rich niches generated by natural phytoplankton blooms, they both stimulated the subsequent activity of lower abundance heterotrophs adapted to ephemeral nutrient enrichment in oligotrophic waters.

Conclusions

We investigated the response of an oligotrophic microbial community to two environmentally relevant sources of organic carbon using controlled microcosm experiments in the North Pacific Subtropical Gyre. Our findings provide insight into potential response dynamics of oligotrophic bacteria to elevated DOM concentrations due to seasonal accumulation related to stratification, as well as ephemeral increases connected to episodic cyanobacterial blooms. Both treatments produced similar patterns in cell growth, taxonomic response, and exoenzyme activity despite differences in the quantity and quality of the carbon added. Combined these results suggest the ProDOM

treatment contained a much greater proportion of labile carbon, relative to HMWDOM, which represents a standing stock of semi-labile and refractory organic carbon subjected to persistent heterotrophic activity.

The temporal patterns observed over the course of our incubations appear to be the result of oligotroph/copiotroph response dynamics structured by nitrogen acquisition and DOM resource partitioning. The genome-wide transcriptional patterns of *Prochlorococcus* and *Pelagibacter* we report here suggest that the success of these oligotrophic taxa relies not only on strategies that economize cellular nutrient requirements, but also on regulatory and functional adaptations that allow these organisms to quickly respond to influxes of labile organic nutrients, such as DON, from the ambient DOM pool. Oligotrophic conditions at Station ALOHA almost certainly explain the low abundance and activity of copiotrophs like members of the OM60 clade (Lauro et al., 2009; Yooseph et al., 2010; Nelson and Carlson, 2012). An increase in the activity of these heterotrophs was detectable in DOM amended microcosms only in the latter stages of incubation, fueled by the increased DOC and DON concentrations. This is consistent with previous studies that have shown that utilization of carbon-rich organic material, such as sugars, depends on availability of nitrogen and phosphorus from other sources (Bianchi et al., 1998; Skoog et al., 2002; Piontek et al., 2011).

McCarren *et al.* (2010) conducted a similar HMWDOM microcosm experiment within the mixed layer of the North Pacific Subtropical Gyre, however results from that study showed strong responses in copiotrophic taxa, particularly among organisms of the Alteromonadales. Differences in response dynamics between that study and the one

presented here are likely the result of multiple variables. The McCarren *et al.* treatment had a DOC concentration that was 300% greater than ambient conditions, more than double the amount of HMWDOM used here, and the percent abundance of cDNA reads from Alteromonadales was 300% greater in the starting community of that experiment relative to that observed in our experiments. These two variables implicate a potential founder effect for an organism known to have rapid growth kinetics under increasing substrate concentrations (Eilers *et al.*, 2000). Differences between the starting community composition and activity of the McCarren *et al.* study and the experiments described here were potentially influenced by inorganic nitrogen availability, as in-situ concentrations of nitrate and nitrite were approximately 400% lower and the mixed layer was 40m shallower in this study (data obtained from low level nitrogen measurements at <http://hahana.soest.hawaii.edu/hot/hot-dogs/>).

A major challenge in microbial oceanography is to understand the mechanisms driving changes in community composition and activity across temporal and seasonal time scales (Fuhrman *et al.*, 2006; Giovannoni and Vergin, 2012; Ottesen *et al.*, 2013). In areas where seasonal stratification of the water column regularly occurs (Karl *et al.*, 2012), an extremely oligotrophic microbial assemblage may result due to inorganic nutrient depletion. Amendment of such an assemblage with naturally derived DOM sources resulted in rapid responses from oligotrophic taxa to acquire nitrogen from the most labile compounds, but later appeared to favor transcriptional activity and growth of opportunistic copiotrophs utilizing polysaccharides. These observations suggest that

nitrogen acquisition via DOM resource partitioning plays a key role in structuring turnover of dissolved organic carbon pools by marine planktonic microbial communities.

Acknowledgements

The authors would like to thank the cruise chief scientist S. Wilson and the captain and crew of the R/V *Ka`imikai-o-Kanaloa* for their help in collecting samples for this study. Our gratitude goes to R. Barry and T. Palden for all their hard work in preparing libraries for pyrosequencing and to J. Eppley for his great work on the DeLong lab sequencing pipeline line. Many thanks to the Chisholm lab: P. Berube for providing a large volume of spent *Prochlorococcus* media for ProDOM extraction and A. Coe, J. Thompson and S. Roggensack for flow cytometry training. We thank C. Johnson for assistance with chromatographic and spectrometric data acquisition, M. Nieto-Cid for help with TOC/TDN analysis, and J-H. Hehemann for annotation of the OM60 glycosyl hydrolase and informative discussion. Thanks to L.A. Ventouras, C. Young, A. Martinez, F. Stewart and Y. Shi for valuable scientific discussions.

This work is a contribution from the National Science Foundation Science and Technology Center for Microbial Oceanography: Research and Education (C-MORE). This work was supported by a National Science Foundation Science and Technology Center Award to E.F.D., D.J.R., and D.M.K. (EF0424599), grants to D.M.K., D.J.R and E.F.D from the Gordon and Betty Moore Foundation, a gift from the Agouon Institute

(to E.F.D.), and a fellowship (202180) to A.K.S from the Canadian Institutes of Health Research (CIHR).

Table 1. Read numbers and statistics

cDNA sample	Total reads ¹	Non-rRNA reads ²	% rRNA ³	NCBI hits ⁴
0h	618918	369456	38.2	249185
Control 2h	711088	411428	38.6	298271
Control 12h	616224	431332	28.4	294160
Control 27h	593531	406165	30.2	261611
Control 36h	602581	417565	29.3	255126
HMWDOM 2h	556200	308146	42.4	194588
HMWDOM 12h	564333	401107	27.6	278836
HMWDOM 27h	534137	456035	14.0	314063
HMWDOM 36h	573227	462651	18.4	297466
ProDOM 2h	654338	472796	26.3	357097
ProDOM 12h	647187	255872	57.8	176581
ProDOM 27h	638776	432165	30.5	323866
ProDOM 36h	585229	439791	23.1	289130
DNA sample	Total reads	Non-rRNA reads	% rRNA	NCBI hits
0h	594218	591180	0.49	389357
Control 36	638559	635073	0.50	430104
HMWDOM 36	696659	692255	0.51	467678
ProDOM 36	618145	614493	0.52	452396

¹Total number of sequence reads per run.

² Number of sequence reads after removal of rRNA sequences.

³ The percentage of the total number of sequenced reads that had a best BLASTN hit to rRNA.

⁴Non-replicate, non-rRNA reads with a significant BLASTX hits to proteins in the NCBI non-redundant database.

Table 2. Number of *Prochlorococcus* and *Pelagibacter* ortholog clusters detected as differentially expressed in cDNA samples¹

Organism and Treatment	Total DE + ²	Core DE + ³	Auxiliary DE + ⁴	Unassigned DE + ⁵	Total DE -	Core DE -	Auxiliary DE -	Unassigned DE -
<i>Prochlorococcus</i> HMWDOM	301/2743	114 (38%)	23 (8%)	164 (54%)	130/2743	72 (55%)	11 (9%)	47 (36%)
<i>Prochlorococcus</i> ProDOM	444 /2743	174 (39%)	35 (8%)	235 (53%)	211/2743	113 (54%)	22 (10%)	76 (36%)
<i>Pelagibacter</i> HMWDOM	100/2094	50 (50%)	17 (17%)	33 (33%)	46/2094	25 (55%)	2 (4%)	19 (41%)
<i>Pelagibacter</i> ProDOM	63/2094	32 (51%)	9 (14%)	22 (35%)	44/2094	23 (52%)	3 (7%)	18 (41%)

¹The total number of ortholog clusters detected in all cDNA samples was 2734 for *Prochlorococcus* and 2,094 for *Pelagibacter*. Those orthologs that were enriched in treatments relative to controls are indicated by + and those that were underrepresented as -. The number of ortholog clusters detected as DE in *Pelagibacter* is lower due to decreased sequencing depth compared to *Prochlorococcus*.

²Total DE refers to the total number of orthologs detected as DE as either enriched (+) in treatments relative to controls, or underrepresented (-)

³Core DE refers to the total number of orthologs detected as DE in a core pathway and the number in brackets represents their fraction of total DE orthologs

⁴Auxiliary DE refers to the total number of orthologs detected as DE in an auxiliary pathway and the number in brackets represents their fraction of total DE orthologs

⁵Unassigned DE refers to the total number of DE orthologs that were not assigned to KEGG level 3 pathways and the number in brackets represents their fraction of total DE orthologs

Figure 1. Microbial community dynamics determined by flow cytometry and β -glucosidase exoenzyme assays. A) Shows the percent increase over ambient seawater concentration of DOC and DON in microcosm perturbation experiments. Panels B) HMWDOM and C) ProDOM show total community cell counts, while panels D) HMWDOM and E) ProDOM show *Prochlorococcus marinus* cell counts. Panels F) HMWDOM and G) ProDOM show *Prochlorococcus* growth as a percentage of total community growth that occurred within each of these treatments between consecutive time points. Panels H) and I) plot the cell specific β -glucosidase activity for control and treatment microcosm communities through time.

Figure 1.

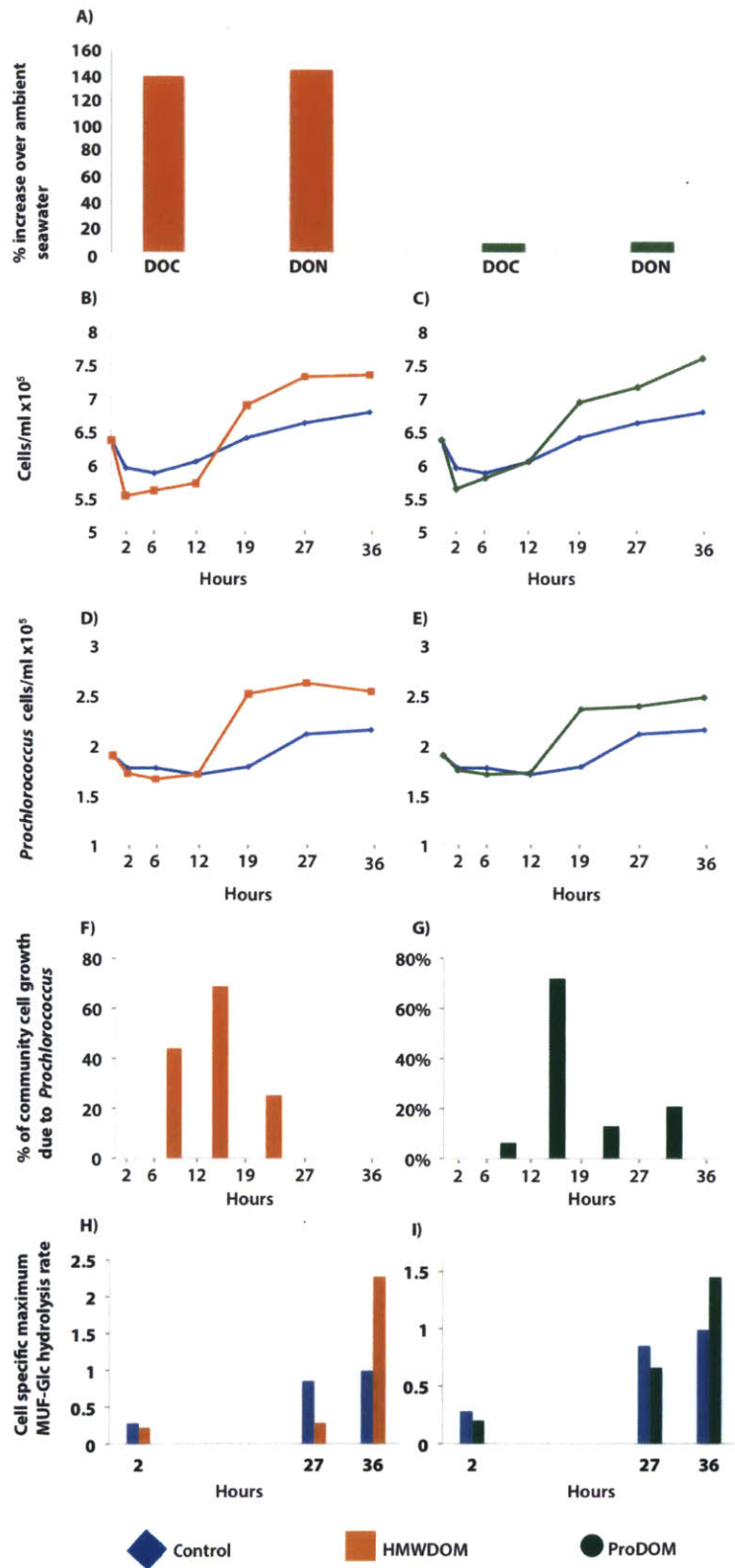


Figure 2. NCBI order level taxonomy of DNA and cDNA reads through time. A) Proportion of the total number of assignable reads represented by a taxonomic order in DNA and cDNA libraries. Only those groups that represented >3% of assignable reads are shown B) The same as in A but only for selected heterotrophic taxa C) The taxonomic association of those NCBI-nr genes detected as enriched in cDNA from treatments relative to controls at each time point show which taxa exhibited changes in gene expression in response to DOM addition. The numbers in brackets along the x-axis denote the total number of NCBI-nr genes detected as significantly enriched in the treatment at each time point. All taxa shown in A) are present in C) with the exception of Caudovirales. Taxonomic representation of reads at the order level was chosen to visually reduce the number of taxa represented on the plots, while simultaneously representing genomes from key divisions.

Figure 2.

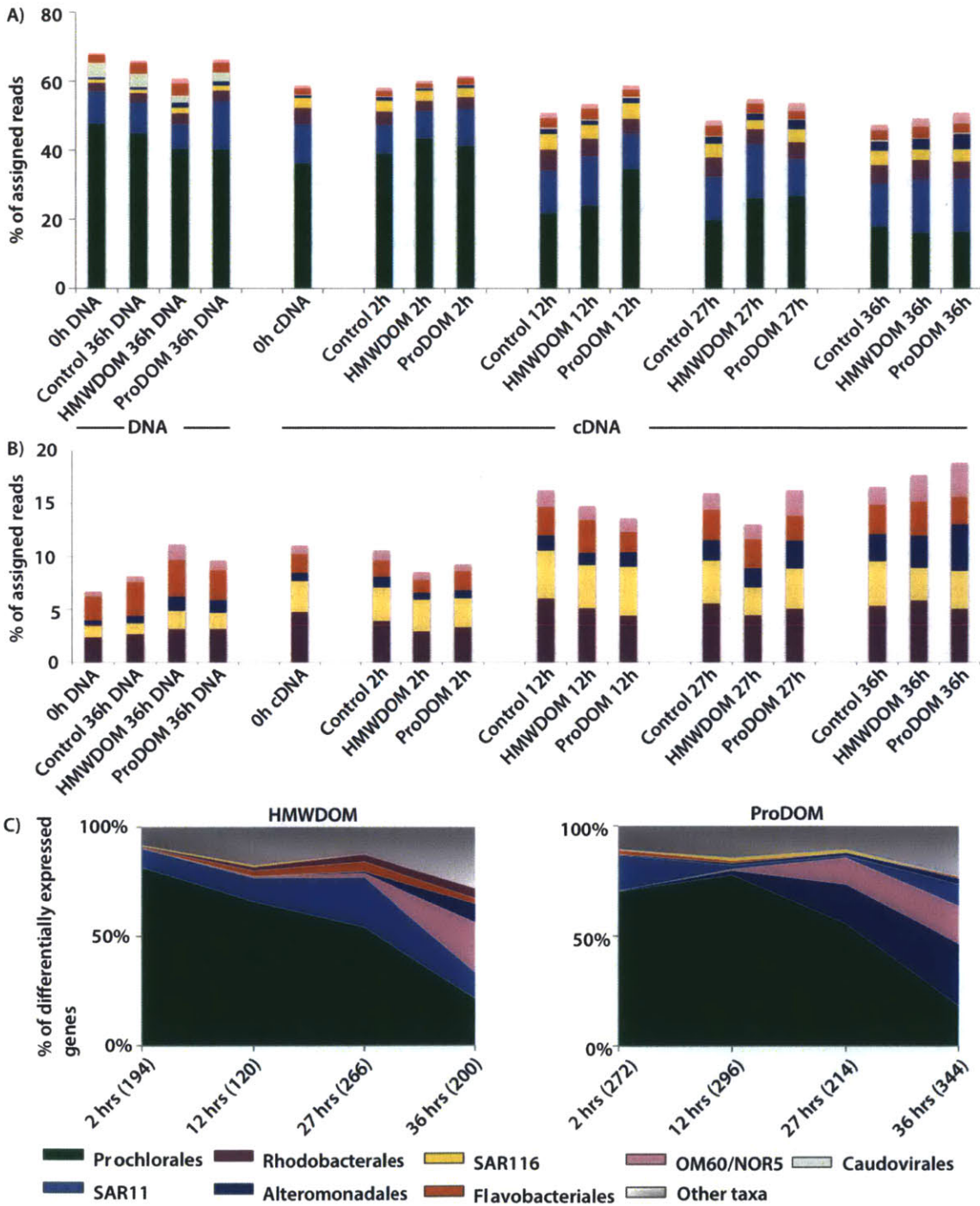
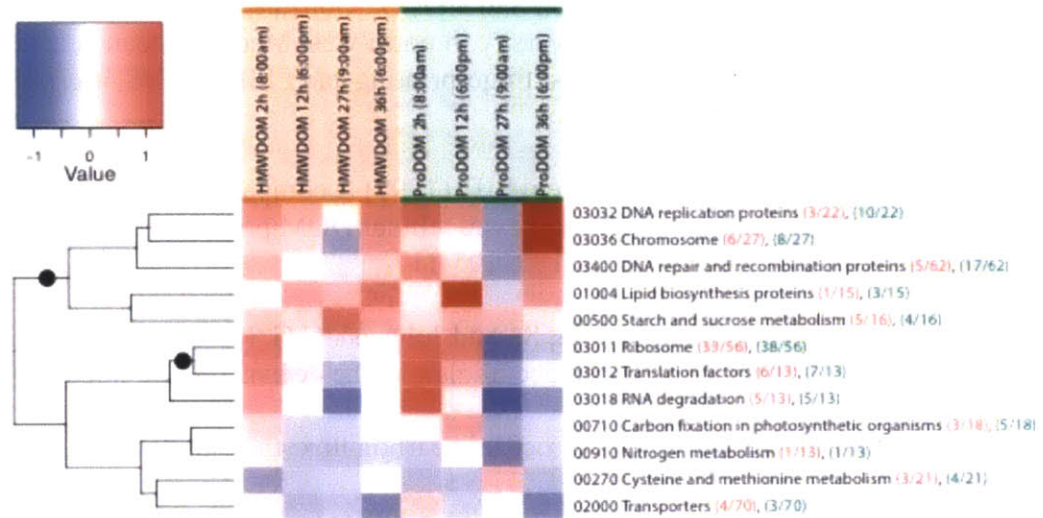


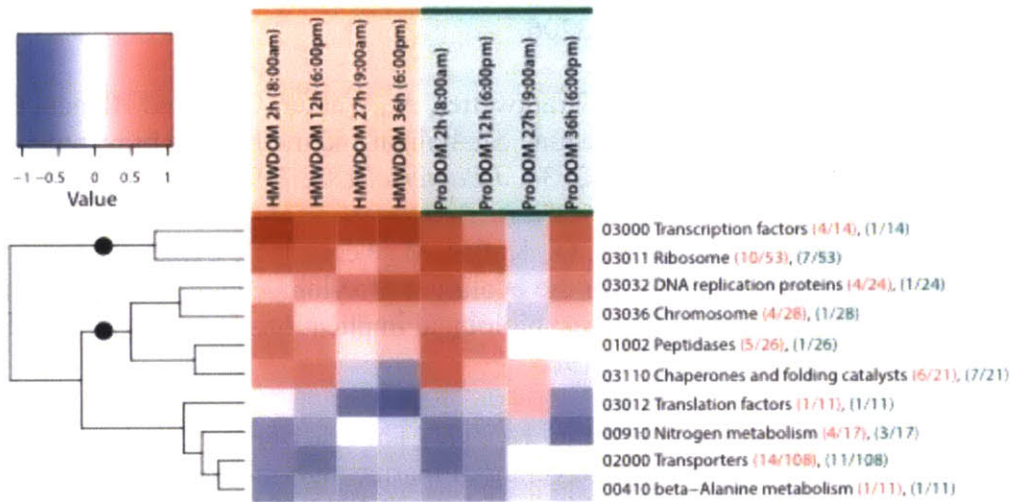
Figure 3. Heatmaps depicting the relative abundance of metabolic pathways in the metatranscriptome of various taxonomic groups over time for both DOM treatments. Pathway abundances for cDNA reads from each sample were calculated as a fraction of sequences mapping to a pathway over the total number of cDNA hits for a particular taxon with a significant match in KEGG. Level 3 pathway abundance is calculated as the percent change in the treatment relative to the control ($\% \text{ treatment} - \% \text{ control} / \% \text{ control}$), such that a value of 1 in the heatmap represents a 100% increase in the pathway in the treatment relative to the control. The dendrogram clusters pathways by similar mean abundance values. A) *Prochlorococcus* transcriptome. Only those pathways that have significantly differentially expressed KOs that were either enriched or underrepresented in both treatments relative to the control in at least 3 of 4 the time points are shown. B) *Pelagibacter* transcriptome. The criteria for choosing pathways to display for *Pelagibacter* was slightly different from *Prochlorococcus* in that only one treatment had to have DE KOs at 3 of 4 time points, rather than both and this was due to the decreased sequencing depth in *Pelagibacter*. Conflicting pathways that contained differentially expressed KOs that are both enriched and underrepresented in treatments at 3 of 4 time points are excluded. The clusters indicated by the black dots represent pathways enriched in the treatment by these criteria, and the remainder those underrepresented. The numbers in brackets next to the pathway names (x/y) indicate the total sum of KOs detected as differentially expressed (x) over the total number of KOs detected in that pathway (y) across all time points for HMWDOM (orange) and ProDOM (green). Note that DE KOs are sometimes assigned to multiple pathways encoding similar metabolic functions (e.g. DNA replication proteins and Chromosome). Therefore, combining the number of DE KOs from two pathways does not always equal the sum of the numbers in brackets.

Figure 3.

A) *Prochlorococcus*



B) *Pelagibacter*



References

- Alonso-Sáez, L., Balagué, V., Sà, E.L., Sánchez, O., González, J.M., Pinhassi, J., et al. (2007) Seasonality in bacterial diversity in north-west Mediterranean coastal waters: assessment through clone libraries, fingerprinting and FISH. *FEMS Microbiol Ecol* 60: 98-112
- Aluwihare, L.I., Repeta, D.J., and Chen, R.F. (1999) A comparison of the chemical characteristics of oceanic DOM and extracellular DOM produced by marine algae. *Marine Ecology Progress Series* 186: 105-117
- Aluwihare, L.I., Repeta, D.J., Pantoja, S., and Johnson, C.G. (2005) Two chemically distinct pools of organic nitrogen accumulate in the ocean. *Science* 308: 1007-1010
- Azam, F. (1998) Microbial control of oceanic carbon flux: the plot thickens. *Science*: 694-695
- Baran, R., Bowen, B.P., and Northen, T.R. (2011) Untargeted metabolic footprinting reveals a surprising breadth of metabolite uptake and release by *Synechococcus* sp. PCC 7002. *Mol Biosyst* 7: 3200-3206
- Barra, L., Fontenelle, C., Ermel, G., Trautwetter, A., Walker, G.C., and Blanco, C. (2006) Interrelations between glycine betaine catabolism and methionine biosynthesis in *Sinorhizobium meliloti* strain 102F34. *J Bacteriol* 188: 7195-7204
- Baumann, M.J., Eklöf, J.M., Michel, G., Kallas, M., Teeri, T.T., Czjzek, M., and Brumer, H. (2007) Structural evidence for the evolution of xyloglucanase activity from xyloglucan endo-transglycosylases: biological implications for cell wall metabolism. *The Plant Cell Online* 19: 1947-1963
- Berg, G.M., Repeta, D.J., and LaRoche, J. (2003) The role of the picoeukaryote *Aureococcus anophagefferens* in cycling of marine high-molecular weight dissolved organic nitrogen. *Limnology and oceanography* : 1825-1830
- Bertlison, S., Berglund, O., Pullin, M.J., and Chisholm, S.W. (2005) Release of dissolved organic matter by *Prochlorococcus*. *Vie et Milieu* 55: 225-232
- Bianchi, A., Van Wambeke, F., and Garcin, J. (1998) Bacterial utilization of glucose in the water column from eutrophic to oligotrophic pelagic areas in the eastern North Atlantic Ocean. *Journal of marine systems* 14: 45-55
- Bronk, D.A., See, J.H., Bradley, P., and Killberg, L. (2007) DON as a source of bioavailable nitrogen for phytoplankton. *Biogeosciences* 4: 283-296

- Carlson, C.A. (2002) Production and removal processes, In *Biogeochemistry of marine dissolved organic matter*. Hansell, D.A. (eds). Academic Press, pp. 91-152
- Carlson, C.A., Giovannoni, S.J., Hansell, D.A., Goldberg, S.J., Parsons, R., and Vergin, K. (2004) Interactions among dissolved organic carbon, microbial processes, and community structure in the mesopelagic zone of the northwestern Sargasso Sea. *Limnology and Oceanography* : 1073-1083
- Carlson, C.A., Morris, R., Parsons, R., Treusch, A.H., Giovannoni, S.J., and Vergin, K. (2009) Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the northwestern Sargasso Sea. *ISME J* 3: 283-295
- Cho, J.C., Stapels, M.D., Morris, R.M., Vergin, K.L., Schwalbach, M.S., Givan, S.A., et al. (2007) Polyphyletic photosynthetic reaction centre genes in oligotrophic marine Gammaproteobacteria. *Environ Microbiol* 9: 1456-1463
- 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
- del Giorgio, P.A., and Duarte, C.M. (2002) Respiration in the open ocean. *Nature* 420: 379-384
- Eiler, A., Hayakawa, D.H., Church, M.J., Karl, D.M., and Rappé, M.S. (2009) Dynamics of the SAR11 bacterioplankton lineage in relation to environmental conditions in the oligotrophic North Pacific subtropical gyre. *Environ Microbiol* 11: 2291-2300
- Eilers, H., Pernthaler, J., and Amann, R. (2000) Succession of pelagic marine bacteria during enrichment: a close look at cultivation-induced shifts. *Applied and environmental microbiology* 66: 4634-4640
- Eilers, H., Pernthaler, J., Peplies, J., Glöckner, F.O., Gerdts, G., and Amann, R. (2001) Isolation of novel pelagic bacteria from the German bight and their seasonal contributions to surface picoplankton. *Appl Environ Microbiol* 67: 5134-5142
- Field, C.B., Behrenfeld, M.J., Randerson, J.T., and Falkowski, P. (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281: 237-240
- Flårdh, K., Cohen, P.S., and Kjelleberg, S. (1992) Ribosomes exist in large excess over the apparent demand for protein synthesis during carbon starvation in marine *Vibrio* sp. strain CCUG 15956. *J Bacteriol* 174: 6780-6788

- Frias-Lopez, J., Shi, Y., Tyson, G.W., Coleman, M.L., Schuster, S.C., Chisholm, S.W., and DeLong, E.F. (2008) Microbial community gene expression in ocean surface waters. *Proceedings of the National Academy of Sciences* 105: 3805-3810
- Fuchs, B.M., Spring, S., Teeling, H., Quast, C., Wulf, J., Schattener, M., et al. (2007) Characterization of a marine gammaproteobacterium capable of aerobic anoxygenic photosynthesis. *Proc Natl Acad Sci U S A* 104: 2891-2896
- Fuhrman, J.A., Hewson, I., Schwalbach, M.S., Steele, J.A., Brown, M.V., and Naeem, S. (2006) Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc Natl Acad Sci U S A* 103: 13104-13109
- Giovannoni, S.J., and Vergin, K.L. (2012) Seasonality in ocean microbial communities. *Science* 335: 671-676
- Hansell, D.A., Carlson, C.A., Repeta, D.J., and Schlitzer, R. (2009) Dissolved organic matter in the ocean: a controversy stimulates new insights
- Hardcastle, T.J., and Kelly, K.A. (2010) baySeq: empirical Bayesian methods for identifying differential expression in sequence count data. *BMC Bioinformatics* 11: 422
- Hehemann, J.H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., and Michel, G. (2010) Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* 464: 908-912
- Huson, D.H., Auch, A.F., Qi, J., and Schuster, S.C. (2007) MEGAN analysis of metagenomic data. *Genome research* 17: 377-386
- Johnson, Z.I., Zinser, E.R., Coe, A., McNulty, N.P., Woodward, E.M., and Chisholm, S.W. (2006) Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* 311: 1737-1740
- Karl, D.M., and Lukas, R. (1996) The Hawaii Ocean Time-series (HOT) program: Background, rationale and field implementation. *Deep Sea Research Part II: Topical Studies in Oceanography* 43: 129-156
- Karl, D.M., Bidigare, R.R., Church, M.J., Dore, J.E., Letelier, R.M., Mahaffey, C., and Zehr, J. (2008) The nitrogen cycle in the North Pacific Trades biome: an evolving paradigm. *Nitrogen in the Marine Environment* : 705-769
- Karl, D.M., Björkman, K.M., Dore, J.E., Fujieki, L., Hebel, D.V., Houlihan, T., et al. (2001) Ecological nitrogen-to-phosphorus stoichiometry at station ALOHA. *Deep Sea Research Part II: Topical Studies in Oceanography* 48: 1529-1566

- Karl, D.M., Church, M.J., Dore, J.E., Letelier, R.M., and Mahaffey, C. (2012) Predictable and efficient carbon sequestration in the North Pacific Ocean supported by symbiotic nitrogen fixation. *Proc Natl Acad Sci U S A* 109: 1842-1849
- Kleber, H.P. (1997) Bacterial carnitine metabolism. *FEMS microbiology letters* 147: 1-9
- Kujawinski, E.B., Longnecker, K., Blough, N.V., Vecchio, R.D., Finlay, L., Kitner, J.B., and Giovannoni, S.J. (2009) Identification of possible source markers in marine dissolved organic matter using ultrahigh resolution mass spectrometry. *Geochimica et Cosmochimica Acta* 73: 4384-4399
- Kunji, E.R.S., Mierau, I., Hagting, A., Poolman, B., and Konings, W.N. (1996) The proteolytic systems of lactic acid bacteria. *Antonie van Leeuwenhoek* 70: 187-221
- Latypova, E., Yang, S., Wang, Y.S., Wang, T., Chavkin, T.A., Hackett, M., et al. (2010) Genetics of the glutamate-mediated methylamine utilization pathway in the facultative methylotrophic beta-proteobacterium *Methyloversatilis universalis* FAM5. *Mol Microbiol* 75: 426-439
- Lauro, F.M., McDougald, D., Thomas, T., Williams, T.J., Egan, S., Rice, S., et al. (2009) The genomic basis of trophic strategy in marine bacteria. *Proc Natl Acad Sci U S A* 106: 15527-15533
- Marie, D., Partensky, F., Jacquet, S., and Vaulot, D. (1997) Enumeration and Cell Cycle Analysis of Natural Populations of Marine Picoplankton by Flow Cytometry Using the Nucleic Acid Stain SYBR Green I. *Appl Environ Microbiol* 63: 186-193
- Martiny, A.C., Coleman, M.L., and Chisholm, S.W. (2006) Phosphate acquisition genes in *Prochlorococcus* ecotypes: evidence for genome-wide adaptation. *Proc Natl Acad Sci U S A* 103: 12552-12557
- Martiny, A.C., Kathuria, S., and Berube, P.M. (2009) Widespread metabolic potential for nitrite and nitrate assimilation among *Prochlorococcus* ecotypes. *Proc Natl Acad Sci U S A* 106: 10787-10792
- Martínez, A., Osburne, M.S., Sharma, A.K., DeLong, E.F., and Chisholm, S.W. (2012) Phosphite utilization by the marine picocyanobacterium *Prochlorococcus* MIT9301. *Environ Microbiol* 14: 1363-1377
- Mary, I., Garczarek, L., Tarran, G.A., Kolowrat, C., Terry, M.J., Scanlan, D.J., et al. (2008) Diel rhythmicity in amino acid uptake by *Prochlorococcus*. *Environ Microbiol* 10: 2124-2131

- McCarren, J., Becker, J.W., Repeta, D.J., Shi, Y., Young, C.R., Malmstrom, R.R., et al. (2010) Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. *Proc Natl Acad Sci U S A* 107: 16420-16427
- Michel, G., Chantalat, L., Duee, E., Barbeyron, T., Henrissat, B., Kloareg, B., and Dideberg, O. (2001) The kappa-carrageenase of *P. carrageenovora* features a tunnel-shaped active site: a novel insight in the evolution of Clan-B glycoside hydrolases. *Structure* 9: 513-525
- Moore, L.R., Coe, A., Zinser, E.R., Saito, M.A., Sullivan, M.B., Lindell, D., et al. (2007) Culturing the marine cyanobacterium *Prochlorococcus*. *Limnol. Oceanogr. Methods* 5: 353-362
- Mopper, K., Stubbins, A., Ritchie, J.D., Bialk, H.M., and Hatcher, P.G. (2007) Advanced instrumental approaches for characterization of marine dissolved organic matter: extraction techniques, mass spectrometry, and nuclear magnetic resonance spectroscopy. *Chem Rev* 107: 419-442
- Mulholland, M.R., and Lee, C. (2009) Peptide hydrolysis and the uptake of dipeptides by phytoplankton. *Limnol. Oceanogr* 54: 856-868
- Muro-Pastor, M.I., Reyes, J.C., and Florencio, F.J. (2005) Ammonium assimilation in cyanobacteria. *Photosynthesis Research* 83: 135-150
- Nelson, C.E., and Carlson, C.A. (2012) Tracking differential incorporation of dissolved organic carbon types among diverse lineages of Sargasso Sea bacterioplankton. *Environmental Microbiology* : 1500-1516
- Nissen, H. (1987) Long term starvation of a marine bacterium, *Alteromonas denitrificans*, isolated from a Norwegian fjord. *FEMS Microbiology Letters* 45: 173-183
- Ottesen, E.A., Young, C.R., Eppley, J.M., Ryan, J.P., Chavez, F.P., Scholin, C.A., and Delong, E.F. (2013) Pattern and synchrony of gene expression among sympatric marine microbial populations. *Proc Natl Acad Sci U S A* 110: E488-E497
- Palenik, B., and Morel, F.M.M. (1990) Comparison of cell-surface L-amino acid oxidases from several marine phytoplankton. *Mar. Ecol. Prog. Ser* 59: 195-201
- Partensky, F., Hess, W.R., and Vaulot, D. (1999) *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol Mol Biol Rev* 63: 106-127

- Pernthaler, A., and Pernthaler, J. (2005) Diurnal variation of cell proliferation in three bacterial taxa from coastal North Sea waters. *Appl Environ Microbiol* 71: 4638-4644
- Piontek, J., Händel, N., De Bodt, C., Harlay, J., Chou, L., and Engel, A. (2011) The utilization of polysaccharides by heterotrophic bacterioplankton in the Bay of Biscay (North Atlantic Ocean). *Journal of Plankton Research* 33: 1719-1735
- Poretsky, R.S., Sun, S., Mou, X., and Moran, M.A. (2010) Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon. *Environ Microbiol* 12: 616-627
- Rinta-Kanto, J.M., Sun, S., Sharma, S., Kiene, R.P., and Moran, M.A. (2012) Bacterial community transcription patterns during a marine phytoplankton bloom. *Environ Microbiol* 14: 228-239
- Romera-Castillo, C., Sarmiento, H., Alvarez-Salgado, X.A., Gasol, J.M., and Marrasé, C. (2011) Net production and consumption of fluorescent colored dissolved organic matter by natural bacterial assemblages growing on marine phytoplankton exudates. *Appl Environ Microbiol* 77: 7490-7498
- Sarmiento, H., and Gasol, J.M. (2012) Use of phytoplankton-derived dissolved organic carbon by different types of bacterioplankton. *Environ Microbiol* 14: 2348-2360
- Schauer, K., Rodionov, D.A., and de Reuse, H. (2008) New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'? *Trends Biochem Sci* 33: 330-338
- Shi, Y., Tyson, G.W., and DeLong, E.F. (2009) Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column. *Nature* 459: 266-269
- Skoog, A., Whitehead, K., Sperling, F., and Junge, K. (2002) Microbial glucose uptake and growth along a horizontal nutrient gradient in the North Pacific. *Limnology and oceanography* : 1676-1683
- Sowell, S.M., Wilhelm, L.J., Norbeck, A.D., Lipton, M.S., Nicora, C.D., Barofsky, D.F., et al. (2009) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J* 3: 93-105
- Srinivasan, S., and Kjelleberg, S. (1998) Cycles of famine and feast: the starvation and outgrowth strategies of a marine Vibrio. *Journal of biosciences* 23: 501-511
- Stewart, F.J., Dalsgaard, T., Young, C.R., Thamdrup, B., Revsbech, N.P., Ulloa, O., et al. (2012) Experimental incubations elicit profound changes in community transcription in OMZ bacterioplankton. *PLoS One* 7: e37118

- Stewart, F.J., Ottesen, E.A., and DeLong, E.F. (2010) Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. *ISME J* 4: 896-907
- Stingl, U., Tripp, H.J., and Giovannoni, S.J. (2007) Improvements of high-throughput culturing yielded novel SAR11 strains and other abundant marine bacteria from the Oregon coast and the Bermuda Atlantic Time Series study site. *ISME J* 1: 361-371
- Sun, J., Steindler, L., Thrash, J.C., Halsey, K.H., Smith, D.P., Carter, A.E., et al. (2011) One carbon metabolism in SAR11 pelagic marine bacteria. *PLoS One* 6: e23973
- Tada, Y., Taniguchi, A., Nagao, I., Miki, T., Uematsu, M., Tsuda, A., and Hamasaki, K. (2011) Differing growth responses of major phylogenetic groups of marine bacteria to natural phytoplankton blooms in the western North Pacific Ocean. *App. Environ. Microbiol.* 77: 4055-4065
- Tang, K., Jiao, N., Liu, K., Zhang, Y., and Li, S. (2012) Distribution and functions of TonB-dependent transporters in marine bacteria and environments: implications for dissolved organic matter utilization. *PLoS One* 7: e41204
- Teeling, H., Fuchs, B.M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C.M., et al. (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* 336: 608-611
- Tolonen, A.C., Aach, J., Lindell, D., Johnson, Z.I., Rector, T., Steen, R., et al. (2006) Global gene expression of *Prochlorococcus* ecotypes in response to changes in nitrogen availability. *Molecular Systems Biology* 2
- Warnes, G.R., Bolker, B., and Lumley, T. (2009) gplots: Various R programming tools for plotting data. *R package version 2*
- Wu, J., Sunda, W., Boyle, E.A., and Karl, D.M. (2000) Phosphate depletion in the western North Atlantic Ocean. *Science* 289: 759-762
- Yan, S., Fuchs, B.M., Lenk, S., Harder, J., Wulf, J., Jiao, N.Z., and Amann, R. (2009) Biogeography and phylogeny of the NOR5/OM60 clade of Gammaproteobacteria. *Syst Appl Microbiol* 32: 124-139
- Yooseph, S., Nealson, K.H., Rusch, D.B., McCrow, J.P., Dupont, C.L., Kim, M., et al. (2010) Genomic and functional adaptation in surface ocean planktonic prokaryotes. *Nature* 468: 60-66

Zubkov, M.V., Fuchs, B.M., Tarran, G.A., Burkill, P.H., and Amann, R. (2003) High rate of uptake of organic nitrogen compounds by *Prochlorococcus* cyanobacteria as a key to their dominance in oligotrophic oceanic waters. *Appl Environ Microbiol* 69: 1299-1304

Zubkov, M.V., Tarran, G.A., and Fuchs, B.M. (2004) Depth related amino acid uptake by *Prochlorococcus* cyanobacteria in the Southern Atlantic tropical gyre. *FEMS Microbiol Ecol* 50: 153-161

Supplementary Information

Includes: detailed experimental procedures, Figure S1 and Tables S1-S12.

Experimental Procedures

Experimental setup

Seawater for microcosm incubation experiments was collected (22° 45'N, 158° 00'W) from the bottom of the mixed layer (35 m) at dawn, on May 27, 2010. Hydrocasts for sampling were conducted using a conductivity-temperature-depth (CTD) rosette sampler aboard the R/V *Ka'imikai-o-Kanaloa*. Water was transferred to acid-washed, sample-water rinsed 20 L polycarbonate bottles. The deck-board incubator was a blue light type, which simulated light levels at ca. 10m (roughly 35% surface irradiance). Twenty-liter carboys were wrapped in a single layer of black fiberglass screen, to further decrease the light level inside the carboys to 14% surface irradiance, the *in situ* light intensity at 25-45m. Carboys were incubated in the deck-board incubators supplied with flow-through surface seawater to maintain near *in situ* temperatures. 2 L of DOM concentrate (HMWDOM or ProDOM) was added to 18L of water obtained from 35 m depth for a total initial volume of 20 L. Final DOC concentrations were 191 $\mu\text{M C}$ for HMWDOM, approximately 2.4x the ambient value of 79.9 $\mu\text{M C}$, and 85.5 $\mu\text{M C}$ for ProDOM, approximately 1.07x the ambient value of 79.9 $\mu\text{M C}$.

RNA and DNA sampling

At selected time points, bacterioplankton biomass from ~2 L sample was rapidly

collected for RNA extraction by first pre-filtering through a 1.6 μm glass fiber filter and then harvesting cells onto 0.2 μm Sterivex (Millipore) filters. Filtration was limited to less than 10 minutes. 1.8 ml of RNAlater[®] (Applied Biosystems, Foster City CA) was added to the filter, which was subsequently capped and flash frozen at -80 °C. Samples were transported frozen to the laboratory in a dry shipper and stored at -80 °C until RNA extraction procedures. At both the beginning and the end of the experiment, biomass was similarly collected for DNA samples. 18 L of seawater for T0 DNA sample collection were directly taken from the CTD bottle (not from the microcosms) and 5-6 L of water were filtered from the end of the experiment for DNA extractions. Filter units for DNA extraction were filled with lysis buffer (50 mM Tris-HCl, 40 mM EDTA, and 0.75 M sucrose), capped and frozen at -20 °C until extraction.

RNA extraction

Total RNA was extracted using the *mirVana*[™] miRNA Isolation kit (Ambion) as described in (Frias-Lopez et al., 2008), with modification to account for the recovery of RNA from Sterivex filters. Filters were thawed on ice, and then RNAlater[®] was expelled via syringe and discarded. The standard amounts of Lysis/Binding Buffer and miRNA Homogenate Additive were doubled, and following vortexing and incubation steps, the resulting lysate removed with a syringe and divided into two 2ml tubes, which were processed separately through the remainder of the standard *mirVana*[™] miRNA Isolation kit protocol. Following purification of the total RNA from the *mirVana*[™] columns, samples resulting from a single filter were combined back together for genomic DNA

removal with TURBO DNA-free™, then purified and concentrated using the RNeasy MinElute Cleanup kit (Qiagen).

DNA Isolation

Total DNA was extracted and purified using the Quick-Gene 610 l system (Fujifilm, Tokyo, Japan) and DNA Tissue Kit L with a modified lysis protocol. 2 mg of lysozyme in lysis buffer (described above) was added to thawed Sterivex filters, which were incubated with rotation to mix at 37 °C for 45 min. In total, 100 µl each of the kit buffers EDT and MDT were added and incubated at 55 °C for 2 h with rotation. The lysate was decanted from the filter, 2 ml LDT solution was added to the lysate, and incubated at 55 °C for a further 15 min without rotation. Finally, 2.7 ml EtOH was added, mixed using a vortex, then sample loaded onto the QuickGene instrument for purification according to the manufacturer's DNA Tissue protocol.

Exoenzyme assay

β-glucosidase activity was measured as an increase in fluorescence of the product 4-Methylumbelliferone (MUF) released after enzymatic hydrolysis of the non-fluorescent 4-Methylumbelliferyl-β-D-glucopyranoside (MUF-Glc; Sigma-Aldrich) substrate. The kinetic parameters of β-glucosidase activity were measured in a series of eight different MUF-Glc concentrations, ranging from 0.05 to 100 µM (final concentration). The highest concentrations were saturating. The kinetic parameters were determined using the Hanes–Woolf plot graphical representation of the rearrangement of the Michaelis–Menten

equation as follows: $S:V = K_m :V_m + S:V_m$, with the MUF-Glc concentration (S), the hydrolysis rate (V), the maximum hydrolysis rate (V_m), and the half-saturation constant (K_m). All samples were incubated in the dark at in situ temperature in an incubator. Hydrolysis of MUF-Glc to MUF (excitation and emission: 359 and 449 nm, respectively) was measured on a Kontron SFM25 spectrofluorometer. At least four measurements were obtained within 18 h to verify the linearity of the assay. A standard curve using MUF (Sigma-Aldrich) from 0 to 500 nM in 0.2- μ m-filtered and boiled seawater was used to calculate hydrolysis rates. Blanks (i.e., ultrapure water) and killed controls (i.e., sample fixed with 0.2% paraformaldehyde, final concentration) were run periodically at saturating concentration and indicated no significant autohydrolysis of the substrate.

Preparation of DOM amendments

For preparation of HMWDOM, 434 L of surface seawater was obtained using acid-cleaned Teflon tubing connected to a compressed air-driven diaphragm pump (Wilden) and concentrated 100-fold over a period of 36 h using a single thin-film ultrafiltration membrane element (1 kDa; Spearation Engineering) in a custom-built polycarbonate membrane housing as previously described (McCarren et al., 2010). For preparation of ProDOM, a 20L culture of *Prochlorococcus* strain MIT9313 was maintained at 22 °C and ca. 20 μ E/m²/s and monitored for growth using bulk fluorescence. Upon reaching stationary phase (ca. 14 days), cell biomass was removed by centrifugation, followed by gentle 0.1 μ m filtration (Whatman Polycap 36 TC capsule filter). The filtrate was then stored briefly in the dark at 4 °C until solid-phase extraction

of metabolites. Filtrate was removed from the fridge and acidified to pH 2-3 by adding trace metal grade hydrochloric acid before loading onto a custom packed column loaded with soxhlet purified octadecyl (C₁₈) functionalized silica gel (Sigma-Aldrich) at a rate of 2 ml/min. The column was then washed with 500 ml of acidified ultrapure water (pH 2-3) at a flow rate of 1 ml/min to remove salts before eluting with 10 column volumes of acidified HPLC-grade methanol (pH 2-3) at a rate of 1 ml/min. Salt removal was confirmed using a silver nitrate solution and the methanol elution was collected in a combusted flask and concentrated down to 1.5 ml using a rotary evaporator. A 50 µl subsample was placed onto a combusted 25 mm 0.7 µm glass fiber filter (Whatman GF/F) for POC analysis to quantify the amount of organic material recovered via solid-phase extraction. A 100 µl subsample was also taken for HPLC/MS analysis. The remaining sample was dried using filtered high purity nitrogen gas, rinsed with ultrapure water to remove residual methanol, and then dried again before resuspension in sterile seawater.

Quantification of organic carbon and dissolved nitrogen

Combusted glassware (450 °C for 8 h) was used for all sampling. Sub-samples of 30 ml for total organic carbon (TOC) and total dissolved nitrogen (TDN) were transferred into glass vials and acidified with 150 µl of a 25% phosphoric acid solution before sealing with acid-washed Teflon lined septa and storage in the dark at 4 °C until processing. Sample concentrations were determined using the high temperature combustion method alongside potassium hydrogen phthalate and potassium nitrate

standards and consensus reference materials provided by the DOC-CRM program (www.rsmas.miami.edu/groups/biogeochem/CRM.html). Particulate organic carbon (POC) analysis of solid-phase extracted *Prochlorococcus*-derived material was measured by placing 50 μl of sample onto a combusted 25 mm 0.7 μm glass fiber filter (Whatman GF/F) and allowing methanol to evaporate in a chemical hood. Filters were then placed inside a combusted glass petri dish, wrapped in foil, and immediately frozen. This process was repeated to obtain duplicate samples and blank filters with 50 μl of pure methanol added were also prepared. Filters were later thawed and put in a drying oven (60 °C) overnight to ensure they were thoroughly dried before encapsulation into 9x10 mm tin capsules and shipped to the University of California Davis Stable Isotope Facility for quantification.

Preparation of DOM amendments

High-molecular weight DOM was isolated and concentrated from surface seawater as previously described (McCarren et al., 2010) with the following modifications. 434 L of surface seawater was obtained using acid-cleaned Teflon tubing connected to a compressed air-driven diaphragm pump (Wilden) and concentrated 100-fold over a period of 36 h using a single thin-film ultrafiltration membrane element (Spearation Engineering) in a custom-built polycarbonate membrane housing. Samples were taken for TOC quantification, cell counts, and viral particle counts from the raw, 0.2 μm pre-filtered, and permeate water periodically during ultrafiltration and from the concentrate upon completion. This sample suite was also taken after serial filtration of

the concentrate through a 0.1 μm Polycap TC prefilter (Whatman) followed by 30-kDa to remove viral particles as previously described (McCarren et al., 2010).

Prochlorococcus-derived DOM was isolated and concentrated from an axenic culture of *Prochlorococcus* strain MIT9313 grown in 20 L of Pro99 medium prepared according to existing protocols (Moore et al., 2007) in sterile Sargasso seawater. The culture was maintained at 22 °C and ca. 20 $\mu\text{E}/\text{m}^2/\text{s}$ and monitored for growth using bulk fluorescence. Upon reaching stationary phase (ca. 14 days), cell biomass was removed by centrifugation, followed by gentle 0.1 μm filtration (Whatman Polycap 36 TC capsule filter). The filtrate was then stored briefly in the dark at 4 °C until solid-phase extraction of metabolites. Filtrate was removed from the fridge and acidified to pH 2-3 by adding trace metal grade hydrochloric acid before loading onto a custom packed column loaded with soxhlet purified octadecyl (C_{18}) functionalized silica gel (Sigma-Aldrich) at a rate of 2 ml/min. The column was then washed with 500 ml of acidified ultrapure water (pH 2-3) at a flow rate of 1 ml/min to remove salts before eluting with 10 column volumes of acidified HPLC-grade methanol (pH 2-3) at a rate of 1 ml/min. Salt removal was confirmed using a silver nitrate solution and the methanol elution was collected in a combusted flask and concentrated down to 1.5 ml using a rotary evaporator. A 50 μl subsample was placed onto a combusted 25 mm 0.7 μm glass fiber filter (Whatman GF/F) for POC analysis to quantify the amount of organic material recovered via solid-phase extraction. A 100 μl subsample was also taken for HPLC/MS analysis. The remaining sample was dried using filtered high purity nitrogen gas, rinsed with ultrapure water to remove residual methanol, and then dried again. The dried material was stored

in a combusted amber vial in the dark prior to resuspension at sea in 1.6 ml of filtered (0.1 μm Polycap TC; Whatman) seawater collected from 35 m at Station ALOHA. 800 μl of the resuspension was added to 2L of sterile seawater, and subsequently added to the 18L microcosm.

Chromatographic separation and detection of MIT9313 metabolites

Chemical analysis of ProDOM material by high performance liquid chromatography–electrospray ionization mass spectrometry (HPLC-ESI-MS) revealed the production and release of 1,491 low-molecular weight features with signal intensity at least 5-fold greater than the maximum noise level. These features ranged in size from 100.2-1,111.5 m/z, with the majority falling between 150-450 m/z. Most features were detected between 20-40 min when the mobile phase went from 64 to 100% methanol, indicating the presence of predominantly nonpolar material. A control sample of Pro99 medium incubated without inoculation and processed as described above for the ProDOM amendment confirmed that these features were absent from the background medium.

Chromatographic separation was performed using an Agilent 1200 series liquid chromatograph comprised of a G1379B degasser, G1312A binary pump, G1367C automatic liquid sampler and F1315C diode array detector. The mobile phases were aqueous (A) formic acid (0.1%) and methanolic (B) formic acid (0.1%). 25 μl of the *Prochlorococcus*-derived sample was injected onto a ZORBAX SB-C18 column (Agilent; 3.5 μm 4.6x150 mm) at a flow of 1 ml/min (starting with 100% A, ramping to

80% B at 25 min, ramping to 100% B at 35 min and holding until 55 min, ramping to 0% B at 65 min and holding until 75 min). Full scan absorbance data were acquired from 210 to 800 nm with a 2.0 nm step and 4 nm slit width. Mass spectrometry was performed in-line using an Agilent 6130 (single quadrupole) mass spectrometer with an atmospheric electrospray ionization source. Source conditions were as follows: drying gas at 11.5 L/min, nebulizer at 60 psig, drying gas temperature at 300 °C, capillary voltages at + or - 4000 V. Acquisition ranges were from 100-2000 Da in the positive mode and used a fragmentor at 4.0, threshold at 150 and a step size of 0.1. Data was processed using the molecular profiling software MZmine 2 (Pluskal et al., 2010). Ions with a minimum signal intensity at least 5-fold greater than the maximum noise level were detected using a centroid mass detector. Chromatograms were then built from the raw data using the same minimum signal intensity, a retention time tolerance of +/-5 s, and a mass tolerance of +/-0.3 *m/z*.

Bioinformatics

Table 1 outlines read numbers and database statistics for community DNA and cDNA samples sequenced from each of the three different microcosms; Control, HMWDOM and ProDOM. Metagenomic samples were sequenced from the 35 m seawater used in the microcosms (0 h) and additionally from the end point of each microcosm (36h). Metatranscriptomic samples were sequenced from 0 h, and from four different time points in each microcosm; 2h, 12h, 27h and 36h. DNA and cDNA samples generated an average of 614,000 reads per sample. Sequences derived from rRNA were

identified using BLASTN with a bit score cutoff of 50 against a database composed of 5S, 16S, 18S, 23S and 28S rRNA sequences from microbial genomes and the SILVA LSU and SSU databases (<http://www.arb-silva.de>). Reads with best BLASTN hits to rRNA averaged 0.5% and 31% in DNA and cDNA libraries respectively, and these sequences were excluded from further analyses. Non-rRNA sequences with identical start sites (first 3 bp), 99% identity and ≤ 1 -bp length difference were identified as probable artificially duplicated sequences (Stewart et al., 2010) and removed using the cd-hit program (Li and Godzik, 2006) and scripts developed by (Gomez-Alvarez et al., 2009). Non-rRNA sequences were compared with the 31 May 2010 version of NCBI's non-redundant (nr) protein reference database using BLASTX, and a bit score cutoff of 50 was used to identify significant matches. An average of 68% of the non-rRNA reads in each library had a best BLASTX hit in the NCBI-nr database (bit score >50), yielding an average of 434,000 and 276,000 NCBI-nr hits in DNA and cDNA libraries respectively. The MEGAN program (Huson et al., 2007) was used to assign sequences to a higher-order taxonomy. All analyses used a bit score cutoff of 50 and database matches with bit scores within 3% of the top-scoring hit. The number of reads with significant matches to different taxonomic orders was normalized according to the total number of all significant hits to the NCBI nr database for an individual sample.

Statistical analysis

For statistical comparisons of metatranscriptomic samples from treatment and control time points, each sequence within a sample was assigned to a single reference

gene in the NCBI-nr database based on BLASTX alignment bit score. When a single sequence aligned equally well to multiple potential reference genes, it was assigned to the reference gene that was most frequently identified in the dataset. Whole community reference gene hit counts were normalized to the total reads that matched the database for an individual sample. Reference gene abundances between samples were compared using baySeq, a method that uses an empirical Bayes approach to detect patterns of differential gene expression within a set of samples (Hardcastle and Kelly, 2010). BaySeq can perform pairwise sample comparisons, but is also capable of more complex comparisons to account for experimental designs involving multiple sample groups, such as the two treatments used here. We took advantage of the ability of baySeq to extract information from multiple sample groups, and simultaneously evaluated five different differential gene expression models to categorize the differential expression of reference genes at the whole community level for each time point individually.

The first model (DE.1) examined the reference gene counts from the control microcosm relative to both treatments, to identify the differential expression of genes that were common to both treatments. The second model (DE. 2) identified only those reference genes that were significantly differentially expressed in the HMWDOM treatment relative to both the control and the ProDOM treatment. The third model (DE.3) identified only those reference genes that were significantly differentially expressed in the ProDOM treatment relative to both the control and the HMWDOM treatment. Models DE.2 and DE.3 were used to identify biological signals specific to the degradation of each DOM source. The fourth model, DE.all identified those reference genes that were

differentially expressed in the Control, HMWDOM, and ProDOM data and accounted for varying levels of gene expression within a single reference gene across all three conditions. Finally, the fifth model, Non-differentially expressed (NDE) assessed the probability of the expression of a reference gene being unaffected by the treatments. Bayseq estimates a posterior probability of each of the models that define patterns of differential or non-differential expression for each reference gene, such that the sum of all probabilities for each of the five models for the count data for a single reference gene equals one.

To detect the significance of the affect of a single treatment or both treatments, the posterior probabilities of certain models could be summed. For example, a reference gene was considered differentially expressed in the HMWDOM treatment if the summed posterior probabilities from models DE.2 and DE.4 were greater than 0.9, because both of these models account for an affect specifically due to this treatment. Similarly, a reference gene was considered differentially expressed in the ProDOM treatment if the summed posterior probabilities from models DE.3 and DE.4 were greater than 0.9. If the posterior probabilities from models DE.1 and DE.4 were greater than 0.9, then a reference gene was considered differentially expressed in both treatments at a similar level. Additionally, differentially expressed reference genes were counted as enriched or underrepresented in a treatment based on their fold change between treatment and control. Differentially expressed reference genes from each of these five models provided a preliminary framework with which to understand similarities and differences that occurred in the treatments through time and helped guide and refine subsequent analyses

at the organism level (particularly categorizing pathways as central versus auxiliary). The data used to generate the plot in figure 2C was obtained by the following analysis: 1) Enumerate the number of differentially expressed reference genes enriched in a single treatment for each individual species name 2) Assign each species name and its differentially expressed reference gene counts to a taxonomic order.

Taxon-specific ortholog clustering

Pairwise reciprocal best BLAST hits between translated coding sequences of reference genomes within a single group were compiled to generate ortholog cluster assignments. For *Prochlorococcus*, the 13 genomes were: *P. marinus* str. AS9601; *P. marinus* str. MIT 9202; *P. marinus* str. MIT 9211; *P. marinus* str. MIT 9215; *P. marinus* str. MIT 9301; *P. marinus* str. MIT 9303; *P. marinus* str. MIT 9312; *P. marinus* str. MIT 9313; *P. marinus* str. MIT 9515; *P. marinus* str. NATL1A; *P. marinus* str. NATL2A; *P. marinus* subsp. *marinus* str. CCMP1375; *P. marinus* subsp. *pastoris* str. CCMP1986/MED4. Four *Pelagibacter* genomes were included in our analyses and these were: *P. Ubique* HTCC 7211, *P. Ubique* HTCC1062; *P. Ubique* HTCC1002; *alpha proteobacterium* HIMB114. Seven OM60 genomes were included: *gamma proteobacterium* HIMB55, *gamma proteobacterium* NOR5-3, *gamma proteobacterium* NOR51-B, *Congregibacter litoralis* KT71, *marine gamma proteobacterium* HTCC2080, *gamma proteobacterium* IMCC3088, *marine gamma proteobacterium* HTCC2148. Identification of shared genes in each of these taxon specific groups used an e-value cutoff of 10^{-5} and required 30% alignment identity over 80% of the longer sequence.

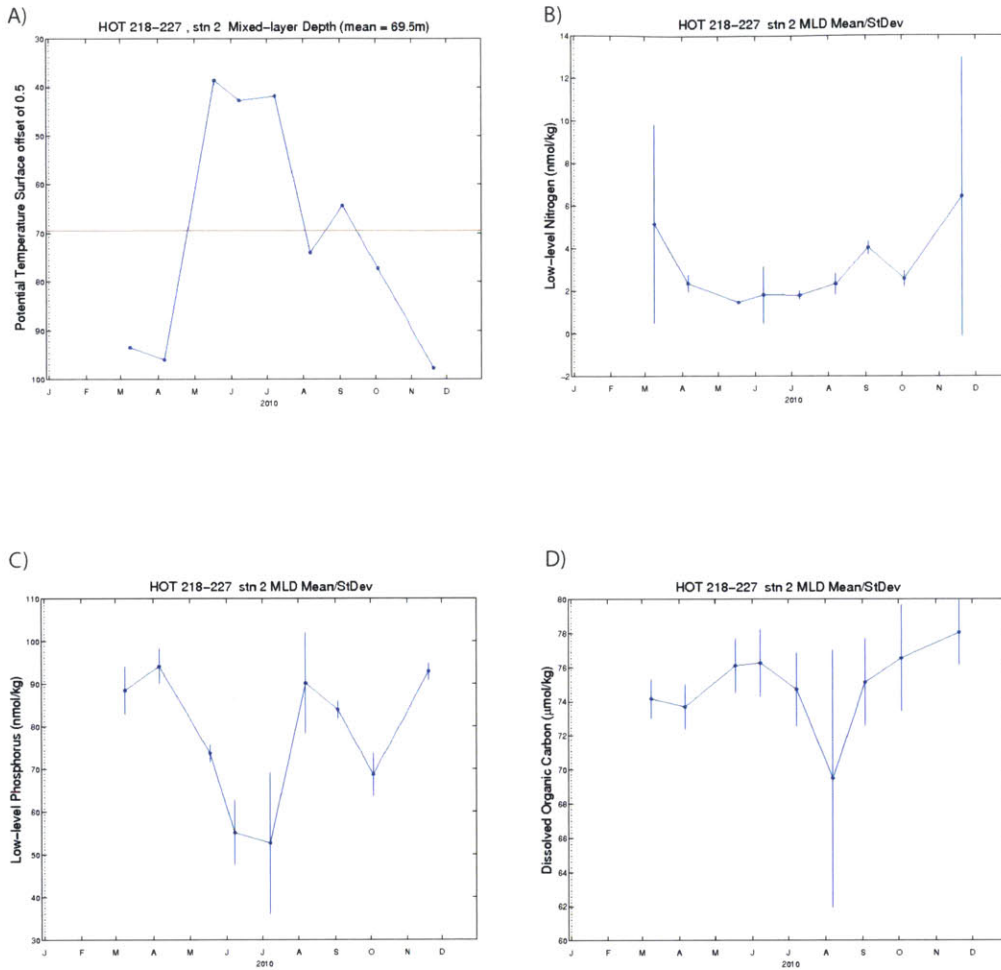
Functional annotation of ortholog clusters used KEGG Genomes annotations where available (Ogata et al., 1999). Genomes from these groups lacking curated annotations were analyzed using the KEGG automated annotation pipeline (KAAS) (Moriya et al., 2007).

In some cases, metatranscriptomic sequences were mapped to reference genes that were not derived from sequenced genomes (i.e. environmental clones). Where possible, these references were assigned to ortholog clusters based on single-directional peptide BLAST (significance cutoffs as above). cDNA reads from our experiment with top BLASTx hits to a reference gene belonging to one of these three taxon bins were then mapped to their respective ortholog cluster (Dataset S1). For sequences matching equally well to multiple genes within the database (i.e. to multiple taxa), all matches were required to fall within the Cyanobacteria for assignment to *Prochlorococcus*, within the SAR11 cluster for *Pelagibacter* and within the OM60 clade itself for OM60 reads. Taxon-specific ortholog count files were used in baySeq in pairwise differential gene expression tests to identify those orthologs that were either enriched or under-represented in a single treatment relative to the control at each time point. For taxon-specific analyses, we opted to do pairwise comparisons between a single treatment and sample at each time point for organism specific bins and tease apart the differences and similarities between treatments by examining differentially expressed orthologs in central and auxiliary metabolic pathways as outlined in the results and discussion. The structure for this analysis pipeline was informed and guided by preliminary results from the whole community baySeq analysis, which indicated that many differentially expressed reference

genes shared between both treatments were due to growth signals such as ribosomal proteins. The differential gene expression of individual KEGG annotated orthologs was used to direct and support comparisons of taxon-specific pathway abundances between treatment and control cDNA samples (Figure 3, Datasets S2-S5). Changes in pathway abundances supported by DE KOs between treatment and control time points for *Prochlorococcus* and *Pelagibacter* were displayed in heatmaps (Figure 3) which were generated in R using the heatmap.2 function in gplots (Warnes et al., 2009) (<http://hosho.ees.hokudai.ac.jp/~kubo/Rdoc/library/gplots/html/00Index.html>).

Many DE orthologs from *Prochlorococcus* and *Pelagibacter* were from central metabolic pathways involved in growth, biosynthetic, or photosynthetic responses, providing information about the physiological state of the cell. To more efficiently examine the differential expression of orthologs involved in the degradation of specific DOM compounds, DE KOs involved in central metabolic pathways were filtered from the complete list of all DE orthologs detected from these organisms. The complete list of central metabolic pathways used to filter DE orthologs for each organism can be found in Supplementary file 2, along with the resulting list of auxiliary pathways. To eliminate redundancy, this list of central metabolic pathways also includes any pathway represented in the heatmaps in Figure 3, and often includes pathways that had no representation among the cDNA reads (but were included in that organism's complete list of pathways because they were present among DNA reads). DE orthologs from auxiliary pathways for *Prochlorococcus* are in tables S1-2 and *Pelagibacter* in tables S7-8.

In the case of the OM60 clade, the majority of DE occurred in the final two time points, and all DE orthologs detected as enriched or underrepresented are displayed in Tables S9-S12.



Supplementary Figure 1. Physiochemical characteristics of the mixed surface layer at Station ALOHA in 2010 plotted from the Hawaii Ocean Time-series Data Organization & Graphical System (HOT-DOGS) website (<http://hahana.soest.hawaii.edu/hot/hot-dogs/>) A) Monthly mixed layer depth calculated using potential temperature. B) Mean and standard deviation of monthly Low-level Nitrogen within the mixed layer depth. C) Mean and standard deviation of monthly Low-level Phosphate within the mixed layer depth. D) Mean and standard deviation of monthly Dissolved Organic Carbon within the mixed layer depth.

Table S1: Differentially expressed Prochlorococcus Orthologs enriched in HMWDOM belonging to auxiliary KEGG level 3 pathways¹

Ortholog	KO	KEGG Level ²	Avg. fold change ³	Time point ⁴
Cluster 1909	K00540 K00273, DAO; D-amino-acid oxidase [EC:1.4.3.3]	00260 Glycine, serine and threonine metabolism 00472 D-Arginine and D-ornithine metabolism	7.96	2h (8:00 AM)
Cluster 729	K00812, aspB; aspartate aminotransferase [EC:2.6.1.1]	00250 Alanine, aspartate and glutamate metabolism 00330 Arginine and proline metabolism	3.76	2h (8:00 AM)
Cluster 172	K00297, metF; methylenetetrahydrofolate reductase (NADPH) [EC:1.5.1.20]	00680 Methane metabolism 00670 One carbon pool by folate	3.63	2h (8:00 AM)
Cluster 25	K00946, thiL; thiamine-monophosphate kinase [EC:2.7.4.16]	00730 Thiamine metabolism	3.04	2h (8:00 AM)
Cluster 1458	K00611, OTC, argF, argI; ornithine carbamoyltransferase [EC:2.1.3.3]	00330 Arginine and proline metabolism	2.41	2h (8:00 AM)
Cluster 53	K00620 argJ; glutamate N-acetyltransferase / amino-acid N-acetyltransferase [EC:2.3.1.35 2.3.1.1]	00330 Arginine and proline metabolism	3.98	12h (6:00 PM),36h (6:00 PM)
Cluster 449	K00286, proC; pyrroline-5-carboxylate reductase [EC:1.5.1.2]	00330 Arginine and proline metabolism	2.38	12h (6:00 PM),36h (6:00 PM)
Cluster 1180	K08479 sasA; two-component system, OmpR family, clock-associated histidine kinase SasA [EC:2.7.13.3]	02020 Two-component system 02022 Two-component system	5.70	27h (9:00 AM)
Cluster 2296	K03073 secE; preprotein translocase subunit SecE	02044 Secretion system 03060 Protein export	5.07	27h (9:00 AM)
Cluster 1344	rhomboid family membrane serine protease	01002 Peptidases	4.94	27h (9:00 AM)
Cluster 393	K01259 pip; proline iminopeptidase [EC:3.4.11.5]	00330 Arginine and proline metabolism 01002 Peptidases	4.65	27h (9:00 AM)
Cluster 1861	K01358 clpP, CLPP; ATP-dependent Clp protease, protease subunit [EC:3.4.21.92]	01002 Peptidases 04112 Cell cycle - Caulobacter	4.16	27h (9:00 AM)
Cluster 1490	putative signal peptidase; Signal peptidase I;	01002 Peptidases	2.94	27h (9:00 AM)
Cluster 227	ATP-dependent ClpB protease Hsp 100	01002 Peptidases	2.60	27h (9:00 AM),36h (6:00 PM)
Cluster 1812	K00820, K00820, glmS; glucosamine--fructose-6-phosphate aminotransferase (isomerizing) [EC:2.6.1.16]	00250 Alanine, aspartate and glutamate metabolism 00520 Amino sugar and nucleotide sugar	2.33	27h (9:00 AM),36h (6:00 PM)
Cluster 1858	K01714 dapA; dihydrodipicolinate synthase [EC:4.2.1.52]	00300 Lysine biosynthesis	1.73	27h (9:00 AM)
Cluster 147	K10697 rpaA; two-component system, OmpR family, response regulator RpaA	02020 Two-component system 02022 Two-component system	1.66	27h (9:00 AM)
Cluster 3062	K02473, wbpP; UDP-N-acetylglucosamine 4-epimerase [EC:5.1.3.7]	00520 Amino sugar and nucleotide sugar metabolism	inf	36h (6:00 PM)
Cluster 5393	K03116 tatA; sec-independent protein translocase protein Tata	03060 Protein export 03070 Bacterial secretion system	inf	36h (6:00 PM)
Cluster 1643	K00969 nadD; nicotinate-nucleotide adenyltransferase [EC:2.7.7.18]	00760 Nicotinate and nicotinamide metabolism	10.53	36h (6:00 PM)
Cluster 623	K00383, GSR, gor; glutathione reductase (NADPH) [EC:1.8.1.7]	00480 Glutathione metabolism	5.26	36h (6:00 PM)
Cluster 339	K07738 nrdR; transcriptional repressor NrdR	03000 Transcription factors	2.98	36h (6:00 PM)
Cluster 1168	K01649, leuA; 2-isopropylmalate synthase [EC:2.3.3.13]	00290 Valine, leucine and isoleucine biosynthesis 00620 Pyruvate metabolism	2.87	36h (6:00 PM)

¹ Prochlorococcus orthologs detected among all cDNA and DNA samples in this experiment represented 140 KEGG level 3 pathways that were sorted into 50 core pathways and 90 auxiliary pathways. Core pathways were defined as those involved in DNA replication, cell growth, biosynthesis and photosynthesis. Auxiliary pathways represent the remainder of pathways. Only those Prochlorococcus orthologs that were detected as DE and belonged to an auxiliary KEGG level 3 pathway are represented. For each KO, only two pathways are shown. Sometimes a Prochlorococcus ortholog had two KO numbers, and in those cases, only a single functional annotation is represented. Peptidase orthologs without a KO number were manually assigned to a pathway based on NCBI annotation.

² Amino acid metabolism was included among the auxiliary pathways as it should be directly affected by increased nitrogen availability.

³ Indicates the average fold change if the ortholog was differentially expressed at multiple time points

⁴ Time point(s) that the ortholog was detected as differentially expressed

Table S2: Differentially expressed *Prochlorococcus* Orthologs enriched in ProDOM belonging to auxiliary KEGG level 3 pathways¹

Ortholog	KO	KEGG Levels ²	Avg. fold change ³	Time point ⁴
Cluster 1883	metal-dependent protease; conserved hypothetical protein ;	01002 Peptidases	inf	2h (8:00 AM)
Cluster 1909	K00540 K00273, DAO; D-amino-acid oxidase [EC:1.4.3.3]	00260 Glycine, serine and threonine metabolism 00472 D-Arginine and D-ornithine metabolism	9.35	2h (8:00 AM),12h (6:00 PM)
Cluster 5067	K01625 eda; 2-dehydro-3-deoxyphosphogluconate aldolase / 4-hydroxy-2-oxoglutarate aldolase [EC:4.1.2.14 4.1.3.16]	00030 Pentose phosphate pathway 00330 Arginine and proline metabolism	9.32	2h (8:00 AM)
Cluster 158	putative metal-dependent protease; putative molecular chaperone	01002 Peptidases	8.93	2h (8:00 AM)
Cluster 48	K01585, speA; arginine decarboxylase [EC:4.1.1.19]	00330 Arginine and proline metabolism	5.96	2h (8:00 AM)
Cluster 83	Dipeptidyl aminopeptidases/acylaminoacyl-peptidases	01002 Peptidases	4.86	2h (8:00 AM)
Cluster 623	K00383, GSR, gor; glutathione reductase (NADPH) [EC:1.8.1.7]	00480 Glutathione metabolism	4.40	2h (8:00 AM)
Cluster 25	K00946, thiL; thiamine-monophosphate kinase [EC:2.7.4.16]	00730 Thiamine metabolism	4.07	2h (8:00 AM)
Cluster 1528	K01255 CARP, pepA; leucyl aminopeptidase [EC:3.4.11.1]	01002 Peptidases 00480 Glutathione metabolism	3.07	2h (8:00 AM)
Cluster 1458	K00611, OTC, argF, argI; ornithine carbamoyltransferase [EC:2.1.3.3]	00330 Arginine and proline metabolism	2.79	2h (8:00 AM)
Cluster 172	K00297, metF; methylenetetrahydrofolate reductase (NADPH) [EC:1.5.1.20]	00680 Methane metabolism 00670 One carbon pool by folate	2.78	2h (8:00 AM)
Cluster 153	K11329 rpaB; two-component system, OmpR family, response regulator RpaB	02020 Two-component system 02022 Two-component system	1.77	2h (8:00 AM)
Cluster 1699	K10206, LL-diaminopimelate aminotransferase [EC:2.6.1.83]	00300 Lysine biosynthesis	1.77	2h (8:00 AM)
Cluster 1743	K03076 secY; preprotein translocase subunit SecY	03060 Protein export 03070 Bacterial secretion system	1.70	2h (8:00 AM)
Cluster 1344	rhomboid family membrane serine protease	01002 Peptidases	15.92	12h (6:00 PM),27h (9:00 AM)
Cluster 1848	K00794 ribH; riboflavin synthase beta chain [EC:2.5.1.-]	00740 Riboflavin metabolism	4.54	12h (6:00 PM)
Cluster 522	K03118 tatC; sec-independent protein translocase protein TatC	03060 Protein export 03070 Bacterial secretion system	3.92	12h (6:00 PM)
Cluster 1023	K03568 tidD; TidD protein	01002 Peptidases	3.53	12h (6:00 PM)
Cluster 1893	K00605, gcvT; aminomethyltransferase [EC:2.1.2.10]	00260 Glycine, serine and threonine metabolism 00910 Nitrogen metabolism	2.06	12h (6:00 PM)
Cluster 1812	K00820, glmS; glucosamine-fructose-6-phosphate aminotransferase (isomerizing) [EC:2.6.1.16]	00250 Alanine, aspartate and glutamate metabolism 00520 Amino sugar and nucleotide sugar	2.58	12h (6:00 PM)
Cluster 1830	K01077, phaA, phaB; alkaline phosphatase [EC:3.1.3.1]	00361 gamma-Hexachlorocyclohexane degradation 02020 Two-component system	2.37	12h (6:00 PM)
Cluster 504	K02259 COX15; cytochrome c oxidase subunit XV assembly protein	00860 Porphyrin and chlorophyll metabolism 00190 Oxidative phosphorylation	1.81	12h (6:00 PM)
Cluster 393	K01259 pip; proline iminopeptidase [EC:3.4.11.5]	00330 Arginine and proline metabolism 01002 Peptidases	4.06	27h (9:00 AM)
Cluster 1690	trypsin-like serine protease	01002 Peptidases	2.58	27h (9:00 AM)
Cluster 1508	K01358 clpP, CLPP; ATP-dependent Clp protease, protease subunit [EC:3.4.21.92]	01002 Peptidases 04112 Cell cycle - Caulobacter	2.22	27h (9:00 AM)
Cluster 1509	K01358 clpP, CLPP; ATP-dependent Clp protease, protease subunit [EC:3.4.21.92]	01002 Peptidases 04112 Cell cycle - Caulobacter	2.18	27h (9:00 AM)
Cluster 818	K00392, sir; sulfite reductase (ferredoxin) [EC:1.8.7.1]	00450 Selenoamino acid metabolism 00920 Sulfur metabolism	2.06	27h (9:00 AM)
Cluster 1643	K00969 nadD; nicotinate-nucleotide adenyltransferase [EC:2.7.7.18]	00760 Nicotinate and nicotinamide metabolism	17.05	36h (6:00 PM)
Cluster 1187	K01583, K01582, lysine decarboxylase [EC:4.1.1.19]	00330 Arginine and proline metabolism 00310 Lysine degradation	5.07	36h (6:00 PM)
Cluster 449	K00286, proC; pyrroline-5-carboxylate reductase [EC:1.5.1.2]	00330 Arginine and proline metabolism	4.75	36h (6:00 PM)
Cluster 53	K00620 argI; glutamate N-acetyltransferase / amino-acid N-acetyltransferase [EC:2.3.1.35 2.3.1.1]	00330 Arginine and proline metabolism	4.51	36h (6:00 PM)
Cluster 13	K01755, argH; argininosuccinate lyase [EC:4.3.2.1]	00250 Alanine, aspartate and glutamate metabolism 00330 Arginine and proline metabolism	4.03	36h (6:00 PM)
Cluster 339	K07738 nrdR; transcriptional repressor NrdR	03000 Transcription factors	3.34	36h (6:00 PM)
Cluster 36	K05912 K00436 E1.12.1.2; hydrogen dehydrogenase [EC:1.12.1.2]	00630 Glyoxylate and dicarboxylate metabolism 00680 Methane metabolism	2.12	36h (6:00 PM)
Cluster 350	K03797, prc, ctpA; carboxyl-terminal processing protease [EC:3.4.21.102]	01002 Peptidases	1.98	36h (6:00 PM)

¹ *Prochlorococcus* orthologs detected among all cDNA and DNA samples in this experiment represented 140 KEGG level 3 pathways that were sorted into 50 core pathways and 90 auxiliary pathways. Core pathways were defined as those involved in DNA replication, cell growth, biosynthesis and photosynthesis. Auxiliary pathways represent the remainder of pathways. Only those *Prochlorococcus* orthologs that were detected as DE and belonged to an auxiliary KEGG level 3 pathway are represented. For each KO, only two pathways are shown. Sometimes a *Prochlorococcus* ortholog had two KO numbers, and in those cases, only a single functional annotation is represented. Peptide orthologs without a KO number were manually assigned to a pathway based on NCBI annotation.

² Amino acid metabolism was included among the auxiliary pathways as it should be directly affected by increased nitrogen availability.

³ Indicates the average fold change if the ortholog was differentially expressed at multiple time points

⁴ Time point(s) that the ortholog was detected as differentially expressed

Table S3: Differentially expressed Pelagibacter transport Orthologs underrepresented in HMWDOM¹

Ortholog	Annotation	Avg. fold change ²	Time point ³
Cluster 1009	K02029 ABC-type amino acid transport system, permease component	0.00	2h (8:00 AM)
Cluster 973	Probable ammonium transporter, marine subtype	0.43	2h (8:00 AM),12h (6:00 PM),27h (9:00 AM),36h (6:00 PM)
Cluster 643	K02002 Glycine betaine/proline transport system substrate-binding protein (proX)	0.44	2h (8:00 AM),12h (6:00 PM)
Cluster 600	K02051 Sulfonate/nitrate/taurine transport system substrate-binding protein (ssuA, tauA)	0.45	2h (8:00 AM)
Cluster 1323	K01999 Branched-chain amino acid transport system substrate-binding protein (livK)	0.57	2h (8:00 AM),12h (6:00 PM),27h (9:00 AM),36h (6:00 PM)
Cluster 1203	K09969 General L-amino acid transport system substrate-binding protein (aapJ, bztA)	0.59	2h (8:00 AM),12h (6:00 PM),27h (9:00 AM),36h (6:00 PM)
Cluster 971	TRAP-type bacterial extracellular solute-binding protein, family 7	0	12h (6:00 PM)
Cluster 1289	TRAP dicarboxylate transporter, dctp subunit	0.19	12h (6:00 PM)
Cluster 1830	TRAP-type extracellular solute-binding protein	0.22	12h (6:00 PM)
Cluster 753	K02002 Glycine betaine/proline transport system substrate-binding protein (proX)	0.30	12h (6:00 PM)
Cluster 2267	Ammonium transporter	0.36	12h (6:00 PM),27h (9:00 AM),36h (6:00 PM)
Cluster 1286	Ammonium transporter	0.50	12h (6:00 PM),36h (6:00 PM)
Cluster 1145	K06901 Xanthine/uracil/vitamin C permease family protein	0.53	12h (6:00 PM),27h (9:00 AM),36h (6:00 PM)
Cluster 688	K10018 Octopine/nopaline transport system substrate-binding protein (occt, nocT)	0.54	12h (6:00 PM)
Cluster 1786	K02027 ABC-type sugar transport system, periplasmic	0.55	12h (6:00 PM),36h (6:00 PM)
Cluster 297	TRAP dicarboxylate transporter - Dctp subunit (mannitol/chloroaromatic compounds)	0.57	12h (6:00 PM)
Cluster 130	K01999 Branched-chain amino acid transport system substrate-binding protein (livK)	0.60	12h (6:00 PM),36h (6:00 PM)
Cluster 696	K02055 Spermidine/putrescine-binding periplasmic protein	0.70	12h (6:00 PM)
Cluster 1189	K02040 Phosphate transport system substrate-binding protein (pstS)	0.27	27h (9:00 AM)
Cluster 924	K02027 ABC-type sugar transport system, periplasmic	0.41	27h (9:00 AM),36h (6:00 PM)
Cluster 456	Arabinose efflux permease	0.42	27h (9:00 AM)
Cluster 557	K02195 Heme exporter protein C (ccmC)	0.44	36h (6:00 PM)
Cluster 462	TRAP-type bacterial extracellular solute-binding protein, family 7	0.58	36h (6:00 PM)

Table S4: Differentially expressed Pelagibacter transport Orthologs underrepresented in ProDOM¹

Ortholog	Annotation	Avg. fold change ²	Time point ³
Cluster 971	bacterial extracellular solute-binding protein, family 7	0.04	2h (8:00 AM),12h (6:00 PM)
Cluster 1289	TRAP dicarboxylate transporter, dctp subunit	0.11	2h (8:00 AM)
Cluster 2267	Ammonium transporter	0.28	2h (8:00 AM),36h (6:00 PM)
Cluster 973	Probable ammonium transporter, marine subtype	0.40	2h (8:00 AM),12h (6:00 PM),36h (6:00 PM)
Cluster 1254	K02012 Iron(III) transport system substrate-binding protein (afuA, fbpA)	0.41	2h (8:00 AM)
Cluster 1286	Ammonium transporter	0.45	2h (8:00 AM),12h (6:00 PM)
Cluster 1145	K06901 Xanthine/uracil/vitamin C permease family protein	0.48	2h (8:00 AM),36h (6:00 PM)
Cluster 600	K02051 Sulfonate/nitrate/taurine transport system substrate-binding protein (ssuA, tauA)	0.48	2h (8:00 AM)
Cluster 1323	K01999 Branched-chain amino acid transport system substrate-binding protein (livK)	0.49	2h (8:00 AM),12h (6:00 PM)
Cluster 462	TRAP-type bacterial extracellular solute-binding protein, family 7	0.50	2h (8:00 AM),36h (6:00 PM)
Cluster 643	K02002 Glycine betaine/proline transport system substrate-binding protein (proX)	0.50	2h (8:00 AM),12h (6:00 PM)
Cluster 924	K02027 ABC-type sugar transport system, periplasmic	0.52	2h (8:00 AM)
Cluster 1203	K09969 General L-amino acid transport system substrate-binding protein (aapJ, bztA)	0.54	2h (8:00 AM),12h (6:00 PM)
Cluster 889	TRAP dicarboxylate transporter- dctp subunit	0.54	2h (8:00 AM)
Cluster 130	K01999 Branched-chain amino acid transport system substrate-binding protein (livK)	0.58	2h (8:00 AM),12h (6:00 PM),36h (6:00 PM)
Cluster 696	K02055 Spermidine/putrescine-binding periplasmic protein	0.68	2h (8:00 AM)
Cluster 1786	K02027 ABC-type sugar transport system, periplasmic	0.53	12h (6:00 PM)
Cluster 1189	K02040 Phosphate transport system substrate-binding protein (pstS)	0.23	27h (9:00 AM)

¹ DE Orthologs represented are either annotated with KEGG level 3 pathway 02000 Transporters or identified as transporters from ortholog annotations unassigned in KEGG²Indicates the average fold change if the ortholog was differentially expressed at multiple time points³Time point(s) that the ortholog was detected as differentially expressed

Table S5: Differentially expressed Pelagibacter transport Orthologs enriched in HMWDOM¹

Ortholog	Annotation	Avg. fold change ²	Time point ³
Cluster 1389	K02196 Heme exporter protein D (ccmD)	inf	2h (8:00 AM)
Cluster 703	K02030 ABC-type amino acid transport substrate-binding protein (PheC)	2.54	2h (8:00 AM),36h (6:00 PM)
Cluster 3	K11720 Predicted permease YjgP/YjgQ family protein	inf	27h (9:00 AM)
Cluster 646	K02023 ABC sugar transporter, ATP-binding protein;	inf	27h (9:00 AM)
Cluster 1423	K07003 Resistance-Nodulation-Cell Division Superfamily transporter	2.47	27h (9:00 AM),36h (6:00 PM)
Cluster 698	K02010 Iron(III) transport system ATP-binding protein [EC:3.6.3.30]	14.19	36h (6:00 PM)
Cluster 865	K03499 Potassium transporter peripheral membrane componen (trka)	6.21	36h (6:00 PM)

Table S6: Differentially expressed Pelagibacter transport Orthologs enriched in ProDOM¹

Ortholog	Annotation	Avg. fold change ²	Time point ³
Cluster 1389	K02196 Heme exporter protein D (ccmD)	inf	2h (8:00 AM)
Cluster 1492	K09013 FeS assembly ATPase SufC;	3.50	2h (8:00 AM)
Cluster 1423	K07003 Resistance-Nodulation-Cell Division Superfamily transporter	3.15	2h (8:00 AM)
Cluster 643	K02002 Glycine betaine/proline transport system substrate-binding protein (proX)	2.00	27h (9:00 AM)
Cluster 250	Lysine exporter protein; transporter;	3.92	36h (6:00 PM)
Cluster 703	K02030 ABC-type amino acid transport substrate-binding protein (PheC)	2.19	36h (6:00 PM)
Cluster 1203	K09969 General L-amino acid transport system substrate-binding protein (aapJ, bztA)	1.80	36h (6:00 PM)

¹ DE Orthologs represented are either annotated with KEGG level 3 pathway 02000 Transporters or identified as transporters from ortholog annotations unassigned in KEGG

²Indicates the average fold change if the ortholog was differentially expressed at multiple time points

³Time point(s) that the ortholog was detected as differentially expressed

Table S7: Differentially expressed *Pelagibacter* Orthologs enriched in HMWDOM belonging to auxiliary KEGG level 3 pathways¹

Ortholog	KO	KEGG Level ²	Avg. fold change ³	Time point ⁴
Cluster 1393	K01586 lysA; diaminopimelate decarboxylase	00300 Lysine biosynthesis	6.13	12h (6:00 PM)
Cluster 577	K00013 hisD; histidinol dehydrogenase	00340 Histidine metabolism	5.82	12h (6:00 PM)
Cluster 647	K03072 secD; preprotein translocase subunit SecD	02044 Secretion system 03070 Bacterial secretion system	2.18	12h (6:00 PM)
Cluster 324	K03074 secF; preprotein translocase subunit SecF	02044 Secretion system 03070 Bacterial secretion system	2.90	12h (6:00 PM)
Cluster 896	K01755 argH, ASL; argininosuccinate lyase	00250 Alanine, aspartate and glutamate metabolism 00330 Arginine and proline metabolism	3.87	27h (9:00 AM)
Cluster 300	K01938 fhs; formate-tetrahydrofolate ligase	00630 Glyoxylate and dicarboxylate metabolism 00670 One carbon pool by folate	2.14	27h (9:00 AM)
Cluster 479	K07638 envZ; two-component system, OmpR family, osmolarity sensor histidine kinase EnvZ	02020 Two-component system	16.39	27h (9:00 AM)
Cluster 1220	K00820 E2.6.1.16, glmS; glucosamine-fructose-6-phosphate aminotransferase (isomerizing)	00250 Alanine, aspartate and glutamate metabolism 00520 Amino sugar and nucleotide sugar metabolism	6.88	2h (8:00 AM)
Cluster 14	K00240 sdhB; succinate dehydrogenase iron-sulfur protein	00020 Citrate cycle (TCA cycle) 00190 Oxidative phosphorylation	5.14	36h (6:00 PM)
Cluster 1279	K01653 E2.2.1.65, ilvH, ilvN; acetylactate synthase I/III small subunit	00290 Valine, leucine and isoleucine biosynthesis 00650 Butanoate metabolism	inf	36h (6:00 PM)
Cluster 85	K00620 argJ; glutamate N-acetyltransferase / amino-acid N-acetyltransferase	00330 Arginine and proline metabolism	4.68	36h (6:00 PM)
Cluster 1084	K00075 murB; UDP-N-acetylmuramate dehydrogenase	00520 Amino sugar and nucleotide sugar metabolism 00550 Peptidoglycan biosynthesis	inf	36h (6:00 PM)
Cluster 612	K01069 E3.1.2.6, gloB; hydroxyacylglutathione hydrolase	00620 Pyruvate metabolism	inf	36h (6:00 PM)
Cluster 1363	K11175 purN; phosphoribosylglycinamide formyltransferase 1	00670 One carbon pool by folate 00230 Purine metabolism	inf	36h (6:00 PM)
Cluster 1587	K02221 yggT family protein	02044 Secretion system	11.35	36h (6:00 PM)
Cluster 1442	K02653 Type II Secretion System PilC;	02044 Secretion system 02035 Bacterial motility proteins	3.28	36h (6:00 PM)
Cluster 576	K03217 yidC, spolIII, OXA1; preprotein translocase subunit YidC	02044 Secretion system 03070 Bacterial secretion system	2.26	36h (6:00 PM)

Table S8: Differentially expressed *Pelagibacter* Orthologs enriched in ProDOM belonging to auxiliary KEGG level 3 pathways¹

Ortholog	KO	KEGG Level ²	Avg. fold change ³	Time point ⁴
Cluster 404	K01903 malate-CoA ligase subunit beta	00020 Citrate cycle (TCA cycle) 00640 Propanoate metabolism	10.96	2h (8:00 AM)
Cluster 85	K00620 argJ; glutamate N-acetyltransferase / amino-acid N-acetyltransferase	00330 Arginine and proline metabolism	7.09	2h (8:00 AM)
Cluster 752	K00641 E2.3.1.31, metX; homoserine O-acetyltransferase	00920 Sulfur metabolism 00270 Cysteine and methionine metabolism	3.03	2h (8:00 AM)
Cluster 1217	K01740 E2.5.1.49, metY; O-acetylhomoserine (thiol)-lyase	00270 Cysteine and methionine metabolism	2.52	2h (8:00 AM)
Cluster 837	K00471 E1.14.11.1; gamma-butyrobetaine dioxygenase	00310 Lysine degradation	4.09	12h (6:00 PM)
Cluster 1279	K01653 E2.2.1.65, ilvH, ilvN; acetylactate synthase I/III small subunit	00290 Valine, leucine and isoleucine biosynthesis 00650 Butanoate metabolism	inf	36h (6:00 PM)
Cluster 664	K00858 E2.7.1.23; NAD+ kinase	00760 Nicotinate and nicotinamide metabolism	6.99	36h (6:00 PM)
Cluster 371	K04042 glmU; bifunctional UDP-N-acetylglucosamine pyrophosphorylase / Glucosamine-1-phosphate N-acetyltransferase	00520 Amino sugar and nucleotide sugar metabolism	4.78	36h (6:00 PM)
Cluster 450	K03210 yajC; preprotein translocase subunit YajC	02044 Secretion system 03070 Bacterial secretion system	3.19	36h (6:00 PM)

¹*Pelagibacter* orthologs among all cDNA and DNA samples in this experiment represented 141 KEGG level 3 pathways that were sorted into 79 core pathways and 62 auxiliary pathways. Core pathways were defined as those generally involved in DNA replication, cell growth and biosynthetic processes. Auxiliary pathways represent non-core pathways. Only those *Pelagibacter* orthologs that were detected as DE and belonged to an auxiliary KEGG level 3 pathway are represented here. A single KO can sometimes belong to multiple pathways and in these instances only two pathways are shown. Sometimes a *Pelagibacter* ortholog had two KO numbers, and in those cases, only a single functional annotation is represented.

² Amino acid metabolism was included among the auxiliary pathways as it should be directly affected by increased nitrogen availability.

³Indicates the average fold change if the ortholog was differentially expressed at multiple time points

⁴Time point(s) that the ortholog was detected as differentially expressed

Table S9: Differentially expressed OM60 Orthologs enriched in HMWDOM¹

Hierarchy	Annotation	KEGG Level3	Avg. fold change ²	Time point ³
Cluster 258	Cytochrome C ¹ superfamily protein cytochrome c, class II protein cytochrome c556	unassigned	inf	2h (8:00 AM)
Cluster 277	K00939 E2.7.4.3, adk adenylate kinase (ATP-AMP transphosphorylase) [EC:2.7.4.3]	00230 Purine metabolism	inf	12h (6:00 PM)
Cluster 2332	K03551 ruvB holliday junction DNA helicase	03440 Homologous recombination 03400 DNA repair and recombination proteins	inf	27h (9:00 AM)
Cluster 3076	K00615 E2.2.1.1, tktA, tktB transketolase [EC:2.2.1.1]	00030 Pentose phosphate pathway 00710 Carbon fixation in photosynthetic organisms	inf	27h (9:00 AM)
Cluster 2148	K00820 E2.6.1.16, glmS glucosamine-fructose-6-phosphate aminotransferase (isomerizing) [EC:2.6.1.16]	00250 Alanine, aspartate and glutamate metabolism 00520 Amino sugar and nucleotide sugar metabolism	inf	27h (9:00 AM)
Cluster 2139	K02110 ATPFOC, atpE F-type H ⁺ -transporting ATPase subunit c [EC:3.6.3.14]	00190 Oxidative phosphorylation	inf	27h (9:00 AM)
Cluster 382	K00798 E2.5.1.17, cobO, btuR cob(II)alamin adenosyltransferase [EC:2.5.1.17]	00860 Porphyrin and chlorophyll metabolism	inf	27h (9:00 AM)
Cluster 1333	K02970 RP-S21, rpsU small subunit ribosomal protein S21	03011 Ribosome	inf	27h (9:00 AM), 36h (6:00 PM)
Cluster 4450	TonB-dependent receptor subfamily protein TonB-dependent receptor, plug	unassigned	inf	27h (9:00 AM), 36h (6:00 PM)
Cluster 9889	Glycosyl hydrolases family 16	unassigned	4.86	27h (9:00 AM)
Cluster 2112	K01878 glyQ glycyL-tRNA synthetase alpha chain [EC:6.1.1.14]	00970 Aminoacyl-tRNA biosynthesis	inf	36h (6:00 PM)
Cluster 8320	hypothetical protein MGP2080_01411	unassigned	inf	36h (6:00 PM)
Cluster 1937	K01887 RARS, argS arginyl-tRNA synthetase [EC:6.1.1.19]	00970 Aminoacyl-tRNA biosynthesis	inf	36h (6:00 PM)
Cluster 3300	K09903 pyrH Uridylate kinase UMP kinase [EC:2.7.4.22]	00240 Pyrimidine metabolism	12.04	36h (6:00 PM)
Cluster 1226	K03666 hfq RNA chaperone Hfq Host factor Hfq	03036 Chromosome 03018 RNA degradation	10.90	36h (6:00 PM)
Cluster 2758	K02916 RP-L35, rpmI large subunit ribosomal protein L35	03011 Ribosome	10.90	36h (6:00 PM)
Cluster 3107	K03561 transporter, MotA/TolQ/ExbB proton channel family protein TonB system biopolymer transport component	unassigned	3.31	36h (6:00 PM)

¹ Includes all OM60 orthologs detected as differentially expressed regardless of KEGG annotation² Indicates the average fold change if the ortholog was differentially expressed at multiple time points³ Time point(s) that the ortholog was detected as differentially expressed

Table S10: Differentially expressed OM60 Orthologs enriched in ProDOM¹

Hierarchy	Annotation	KEGG Level3	Avg. fold change ²	Time point ³
Cluster 268	K03317 putative Na ⁺ dependent nucleoside transporter NupC protein	unassigned	inf	2h (8:00 AM)
Cluster 1965	K00324 pntA NAD(P) transhydrogenase subunit alpha [EC:1.6.1.2]	00760 Nicotinate and nicotinamide metabolism	inf	12h (6:00 PM)
Cluster 2133	K03495 tRNA uridine 5-carboxymethylaminomethyl modification enzyme GidA glucose-inhibited division protein A	03036 Chromosome	inf	12h (6:00 PM)
Cluster 1800	K06149 universal stress protein UspA-like protein universal stress protein family, putative	unassigned	inf	12h (6:00 PM)
Cluster 4573	K00099 dxr 1-deoxy-D-xylulose-5-phosphate reductoisomerase [EC:1.1.1.267]	00900 Terpenoid backbone biosynthesis	inf	12h (6:00 PM)
Cluster 5505	Predicted Fe-S oxidoreductase putative Fe-S oxidoreductase	unassigned	inf	12h (6:00 PM)
Cluster 1297	pentapeptide repeat domain protein	unassigned	inf	27h (9:00 AM)
Cluster 2148	K00820 E2.6.1.16, glmS glucosamine-fructose-6-phosphate aminotransferase (isomerizing) [EC:2.6.1.16]	00250 Alanine, aspartate and glutamate metabolism 00520 Amino sugar and nucleotide sugar metabolism	inf	27h (9:00 AM)
Cluster 2180	K01142 E3.1.11.2, xthA exodeoxyribonuclease III [EC:3.1.11.2]	03400 DNA repair and recombination proteins 03410 Base excision repair	inf	27h (9:00 AM)
Cluster 2539	K01689 ENO, eno enolase [EC:4.2.1.11] phosphopyruvate hydratase Enolase	00010 Glycolysis / Gluconeogenesis	inf	27h (9:00 AM)
Cluster 4450	TonB-dependent receptor subfamily protein TonB-dependent receptor, plug	unassigned	inf	27h (9:00 AM),36h (6:00 PM)
Cluster 3118	K00873 PK, pyk pyruvate kinase [EC:2.7.1.40] pyruvate kinase II	00620 Pyruvate metabolism 00010 Glycolysis / Gluconeogenesis	11.88	27h (9:00 AM)
Cluster 192	K00382 DLD, lpd, pdhD dihydroliipoamide dehydrogenase [EC:1.8.1.4]	00020 Citrate cycle (TCA cycle) 00620 Pyruvate metabolism	5.62	27h (9:00 AM)
Cluster 193	K00658 DLS, sucB 2-oxoglutarate dehydrogenase E2 component (dihydroliipoamide succinyltransferase) [EC:2.3.1.61]	00020 Citrate cycle (TCA cycle) 00310 Lysine degradation	4.86	27h (9:00 AM)
Cluster 4655	K00526 E1.17.4.1B, nrdB, nrdF ribonucleoside-diphosphate reductase beta chain [EC:1.17.4.1]	00240 Pyrimidine metabolism 00230 Purine metabolism	4.73	27h (9:00 AM)
Cluster 2298	K00525 E1.17.4.1A, nrdA, nrdE ribonucleoside-diphosphate reductase alpha chain [EC:1.17.4.1]	00240 Pyrimidine metabolism 00230 Purine metabolism	2.58	27h (9:00 AM)
Cluster 1333	K02970 RP-S21, rpsU small subunit ribosomal protein S21	03010 Ribosome 03011 Ribosome	inf	36h (6:00 PM)
Cluster 1042	hypothetical protein MGP2080_03820 conserved hypothetical protein FlgN protein	unassigned	inf	36h (6:00 PM)
Cluster 3589	K00031 IDH1, IDH2, icd isocitrate dehydrogenase [EC:1.1.1.42]	00020 Citrate cycle (TCA cycle)	inf	36h (6:00 PM)
Cluster 3636	K03106 SRP54, ffh signal recognition particle subunit SRP54 (TC 3.A.5.1.1)	03060 Protein export 03070 Bacterial secretion system	inf	36h (6:00 PM)
Cluster 1937	K01887 RARS, argS arginyl-tRNA synthetase [EC:6.1.1.19]	00970 Aminoacyl-tRNA biosynthesis	inf	36h (6:00 PM)
Cluster 2234	K03671 Thioredoxin	03110 Chaperones and folding catalysts	inf	36h (6:00 PM)
Cluster 1040	K02386 flgA flagella basal body P-ring formation protein	02040 Flagellar assembly 02035 Bacterial motility proteins	inf	36h (6:00 PM)
Cluster 1701	TonB-dependent outer membrane receptor	unassigned	inf	36h (6:00 PM)
Cluster 2112	K01878 glyQ glycyL-tRNA synthetase alpha chain [EC:6.1.1.14]	00970 Aminoacyl-tRNA biosynthesis	inf	36h (6:00 PM)
Cluster 190	K01902 sucD succinyl-CoA synthetase alpha subunit [EC:6.2.1.5]	00020 Citrate cycle (TCA cycle)	inf	36h (6:00 PM)
Cluster 3627	K01733 E4.2.3.1, thrC threonine synthase [EC:4.2.3.1] Threonine synthase	00260 Glycine, serine and threonine metabolism 00750 Vitamin B6 metabolism	inf	36h (6:00 PM)
Cluster 7065	K01130 E3.1.6.1, asIA arylsulfatase [EC:3.1.6.1] Sulfatase, secreted	00140 Steroid hormone biosynthesis 00600 Sphingolipid metabolism	inf	36h (6:00 PM)
Cluster 3167	penicillin-binding protein, beta-lactamase class C	unassigned	inf	36h (6:00 PM)
Cluster 766	K00919 ispE 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase [EC:2.7.1.148]	00900 Terpenoid backbone biosynthesis	inf	36h (6:00 PM)
Cluster 1021	K10941 flrA sigma-54 specific transcriptional regulator, flagellar regulatory protein A	03000 Transcription factors	inf	36h (6:00 PM)
Cluster 2762	K04764 integration host factor, alpha subunit	03036 Chromosome 03032 DNA replication proteins	inf	36h (6:00 PM)
Cluster 7376	K03071 secB preprotein translocase subunit SecB Protein export cytoplasm chaperone protein	03070 Bacterial secretion system 03110 Chaperones and folding catalysts	15.43	36h (6:00 PM)
Cluster 3117	K03704 putative 'Cold-shock' DNA-binding domain protein protein CspE	03000 Transcription factors	10.29	36h (6:00 PM)
Cluster 1226	K03666 hfq RNA chaperone Hfq Host factor Hfq	03036 Chromosome 03018 RNA degradation	9.82	36h (6:00 PM)
Cluster 3465	K06142 hypothetical protein MGP2080_08019 Outer membrane protein (OmpH-like)	unassigned	5.14	36h (6:00 PM)
Cluster 1084	K02952 RP-S13, rpsM small subunit ribosomal protein S13	03010 Ribosome 03011 Ribosome	3.68	36h (6:00 PM)
Cluster 3107	K03561 transporter, MotA/TolQ/ExbB proton channel family protein TonB system biopolymer transport component	unassigned	3.36	36h (6:00 PM)

¹ Includes all OM60 orthologs detected as differentially expressed regardless of KEGG annotation² Indicates the average fold change if the ortholog was differentially expressed at multiple time points³ Time point(s) that the ortholog was detected as differentially expressed

Table S11: Differentially expressed OM60 Orthologs underrepresented in HMWDOM¹

Hierarchy	Annotation	KEGG Level3	Avg fold change ²	Time point ³
Cluster 1453	aerobic-type carbon monoxide dehydrogenase, large subunit CoxL/CutL-like protein	unassigned	0.00	2h (8:00 AM)
Cluster 2338	K03640 18K peptidoglycan-associated outer membrane lipoprotein, secreted OmpA/MotB	unassigned	0.00	12h (6:00 PM)
Cluster 6892	K09516 RETSAT all-trans-retinol 13,14-reductase [EC:1.3.99.23] FAD dependent oxidoreductase domain protein	00830 Retinol metabolism	0.24	12h (6:00 PM), 36h (6:00 PM)
Cluster 7809	TonB-dependent receptor	unassigned	0.00	27h (9:00 AM)
Cluster 9740	K02275 coxB cytochrome c oxidase subunit II [EC:1.9.3.1]	00190 Oxidative phosphorylation	0.00	27h (9:00 AM)
Cluster 8518	TonB-dependent receptor	unassigned	0.00	27h (9:00 AM)
Cluster 4025	K02014 TonB-dependent receptor domain protein outer membrane receptor protein	unassigned	0.23	27h (9:00 AM), 36h (6:00 PM)
Cluster 7577	putative hexachlorocyclohexane dehydrochlorinase 1	unassigned	0.00	36h (6:00 PM)
Cluster 3451	K01474 hyuB N-methylhydantoinase B/acetone carboxylase, alpha subunit [EC:3.5.2.14]	00330 Arginine and proline metabolism	0.00	36h (6:00 PM)
Cluster 387	TonB-dependent receptor plug domain protein	unassigned	0.00	36h (6:00 PM)
Cluster 8853	TonB-dependent receptor	unassigned	0.00	36h (6:00 PM)
Cluster 64	K13482 xdhB xanthine dehydrogenase large subunit [EC:1.17.1.4]	00230 Purine metabolism	0.00	36h (6:00 PM)
Cluster 2921	K13643 iron-sulfur cluster assembly transcription factor IscR transcriptional regulator ⁴	03000 Transcription factors	0.06	36h (6:00 PM)
Cluster 3712	K03088 SIG3.2, rpoE RNA polymerase sigma-70 factor	03020 RNA polymerase	0.10	36h (6:00 PM)
Cluster 4700	hypothetical protein OMB55_00019360 hypothetical protein MGP2080_08414	unassigned	0.16	36h (6:00 PM)
Cluster 8788	TonB-dependent receptor	unassigned	0.18	36h (6:00 PM)
Cluster 7324	K16089 K02014 putative TonB-dependent receptor hypothetical protein NOR51B_2110 Outer membrane protein	unassigned	0.26	36h (6:00 PM)
Cluster 5367	hypothetical protein MGP2080_00560 hypothetical protein IMCC3088_1417	unassigned	0.26	36h (6:00 PM)
Cluster 3689	hypothetical protein OMB55_00014460 conserved hypothetical protein secreted protein	unassigned	0.42	36h (6:00 PM)
Cluster 3082	outer membrane cobalamin receptor protein TonB-dependent receptor domain protein	unassigned	0.43	36h (6:00 PM)

¹ Includes all OM60 orthologs detected as differentially expressed regardless of KEGG annotation² Indicates the average fold change if the ortholog was differentially expressed at multiple time points³ Time point(s) that the ortholog was detected as differentially expressed

Table S12: Differentially expressed OM60 Orthologs underrepresented in ProDOM¹

Hierarchy	Annotation	KEGG Level3	Avg. fold change ²	Time point ³
Cluster 7809	TonB-dependent receptor	unassigned	0.00	12h (6:00 PM)
Cluster 4882	K02404 GTP-binding signal recognition particle SRP54 flagellar biosynthetic protein FlhF	02035 Bacterial motility proteins	0.00	12h (6:00 PM)
Cluster 27	K03314 nhaB Na ⁺ /H ⁺ antiporter	unassigned	0.00	12h (6:00 PM)
Cluster 3548	K01690 edd phosphogluconate dehydratase [EC:4.2.1.12] 6-phosphogluconate dehydratase	00030 Pentose phosphate pathway	0.00	12h (6:00 PM)
Cluster 4025	K02014 TonB-dependent receptor domain protein outer membrane receptor protein	unassigned	0.23	12h (6:00 PM),27h (9:00 AM),36h (6:00 PM)
Cluster 1229	K01448 N-acetylmuramoyl-L-alanine amidase domain protein AmiC	03036 Chromosome	0.00	27h (9:00 AM)
Cluster 3615	DNA integration/recombination/inversion protein phage integrase family	unassigned	0.00	27h (9:00 AM)
Cluster 8325	hypothetical protein MGP2080_01466	unassigned	0.06	27h (9:00 AM)
Cluster 6468	K02014 TonB-dependent receptor domain protein Outer membrane protein	unassigned	0.06	27h (9:00 AM)
Cluster 7992	K00257 E1.3.99.- [EC:1.3.99.-] Butyryl-CoA dehydrogenase	00281 Geraniol degradation 00624 1- and 2-Methylnaphthalene degradation	0.07	27h (9:00 AM)
Cluster 5678	K00174 korA 2-oxoglutarate ferredoxin oxidoreductase subunit alpha [EC:1.2.7.3]	00020 Citrate cycle (TCA cycle) 00720 Reductive carboxylate cycle (CO2 fixation)	0.10	27h (9:00 AM)
Cluster 1127	K03073 secE preprotein translocase subunit SecE [TC 3.A.5.1.1]	03060 Protein export 03070 Bacterial secretion system	0.12	27h (9:00 AM)
Cluster 4700	hypothetical protein OMB55_00019360 hypothetical protein MGP2080_08414	unassigned	0.12	27h (9:00 AM)
Cluster 392	K01637 E4.1.3.1, aceA isocitrate lyase [EC:4.1.3.1]	00630 Glyoxylate and dicarboxylate metabolism	0.22	27h (9:00 AM)
Cluster 2316	K02014 TonB-dependent receptor domain protein outer membrane receptor protein	unassigned	0.27	27h (9:00 AM)
Cluster 3801	hypothetical protein NOR51B_2319 TonB-dependent receptor, plug	unassigned	0.25	27h (9:00 AM),36h (6:00 PM)
Cluster 6892	K09516 RETSAT all-trans-retinol 13,14-reductase [EC:1.3.99.23] FAD dependent oxidoreductase domain protein	00830 Retinol metabolism	0.25	27h (9:00 AM),36h (6:00 PM)
Cluster 4041	K00134 GAPDH, gapA glyceraldehyde 3-phosphate dehydrogenase [EC:1.2.1.12]	00010 Glycolysis / Gluconeogenesis	0.00	36h (6:00 PM)
Cluster 1317	K00966 GMP manose-1-phosphate guanylyltransferase [EC:2.7.7.13]	00051 Fructose and mannose metabolism 00520 Amino sugar and nucleotide sugar metabolism	0.00	36h (6:00 PM)
Cluster 6134	conserved hypothetical protein	unassigned	0.04	36h (6:00 PM)
Cluster 2974	K03320 Ammonium Transporter Family subfamily protein	unassigned	0.06	36h (6:00 PM)
Cluster 2503	RND transporter, HAE1/HME family, permease protein	unassigned	0.17	36h (6:00 PM)
Cluster 8788	TonB-dependent receptor	unassigned	0.18	36h (6:00 PM)
Cluster 5367	hypothetical protein MGP2080_00560 hypothetical protein IMCC3088_1417	unassigned	0.18	36h (6:00 PM)
Cluster 3712	K03088 SIG3.2, rpoE RNA polymerase sigma-70 factor	03020 RNA polymerase	0.19	36h (6:00 PM)
Cluster 4786	glycine/D-amino acid oxidase, deaminating; putative monomeric sarcosine oxidase	unassigned	0.22	36h (6:00 PM)
Cluster 3689	hypothetical protein OMB55_00014460 conserved hypothetical protein secreted protein	unassigned	0.28	36h (6:00 PM)
Cluster 7264	TonB-dependent receptor domain protein hypothetical protein NOR51B_558	unassigned	0.32	36h (6:00 PM)
Cluster 930	K00257 E1.3.99.- [EC:1.3.99.-] Acyl-CoA dehydrogenase	00281 Geraniol degradation 00624 1- and 2-Methylnaphthalene degradation	0.33	36h (6:00 PM)
Cluster 1528	K15987 V-type H(+)-translocating pyrophosphatase pyrophosphate-energized proton pump	unassigned	0.35	36h (6:00 PM)
Cluster 3082	outer membrane cobalamin receptor protein TonB-dependent receptor domain protein	unassigned	0.41	36h (6:00 PM)
Cluster 3201	Oar-like outer membrane protein protein, OmpA family	unassigned	0.42	36h (6:00 PM)

¹ Includes all OM60 orthologs detected as differentially expressed regardless of KEGG annotation² Indicates the average fold change if the ortholog was differentially expressed at multiple time points³ Time point(s) that the ortholog was detected as differentially expressed

References

- Frias-Lopez, J., Shi, Y., Tyson, G.W., Coleman, M.L., Schuster, S.C., Chisholm, S.W., and DeLong, E.F. (2008) Microbial community gene expression in ocean surface waters. *Proceedings of the National Academy of Sciences* 105: 3805-3810
- Gomez-Alvarez, V., Teal, T.K., and Schmidt, T.M. (2009) Systematic artifacts in metagenomes from complex microbial communities. *ISME J* 3: 1314-1317
- Hardcastle, T.J., and Kelly, K.A. (2010) baySeq: empirical Bayesian methods for identifying differential expression in sequence count data. *BMC Bioinformatics* 11: 422
- Huson, D.H., Auch, A.F., Qi, J., and Schuster, S.C. (2007) MEGAN analysis of metagenomic data. *Genome research* 17: 377-386
- Li, W., and Godzik, A. (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22: 1658-1659
- McCarren, J., Becker, J.W., Repeta, D.J., Shi, Y., Young, C.R., Malmstrom, R.R., et al. (2010) Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. *Proc Natl Acad Sci U S A* 107: 16420-16427
- Moore, L.R., Coe, A., Zinser, E.R., Saito, M.A., Sullivan, M.B., Lindell, D., et al. (2007) Culturing the marine cyanobacterium *Prochlorococcus*. *Limnol. Oceanogr. Methods* 5: 353-362
- Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C., and Kanehisa, M. (2007) KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* 35: W182-W185
- Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa, M. (1999) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic acids research* 27: 29-34
- Pluskal, T., Castillo, S., Villar-Briones, A., and Oresic, M. (2010) MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 11: 395
- Stewart, F.J., Ottesen, E.A., and DeLong, E.F. (2010) Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. *ISME J* 4: 896-907
- Warnes, G.R., Bolker, B., and Lumley, T. (2009) gplots: Various R programming tools for plotting data. *R package version 2*