Iron and the Ecology of Marine Microbes

by

Laure-Anne Ventouras

B.A., Biological Chemistry
Wellesley College, 2005

Submitted to the Department of Biological Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology
February 2013

Author: _____
Department of Biological Engineering
November 26th, 2012

Certified by: _____
Edward F. DeLong
Morton and Claire Goulder Professor in Environmental Systems
Professor of Biological Engineering & Civil and Environmental Engineering
Thesis Supervisor

Accepted by: _____
Forest M. White
Associate Professor of Biological Engineering
Chair, Biological Engineering Graduate Committee

Iron and the Ecology of Marine Microbes

by

Laure-Anne Ventouras

Submitted to the Department of Biological Engineering on November 26th, 2012
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in
Biological Engineering

Abstract

Iron is a cofactor of a number biochemical reactions that are essential for life. In the marine environment, this micronutrient is a scarce resource that limits processes of global importance such as photosynthesis and nitrogen fixation. Given that marine microorganisms play a central role in modulating such biogeochemical cycles, understanding how their distribution, diversity and activity may be affected by changes in iron availability is key. This thesis explores how the availability of iron affects the ecology of marine microbial populations and communities.

At the population level, I characterized the prevalence and diversity of iron acquisition strategies in specific populations of marine vibrios with distinct micro-habitat preferences. Using a combination of genomics and functional screens, I showed that siderophore-based iron acquisition is not conserved at the organism-level but represents a stable trait at the population level. This population-level trait further appears to play a role in driving the diversification of specific vibrio populations, especially of those that are thought to prefer particles as a micro-habitat.

At the community level, I measured whole microbial community responses to iron addition in microcosm experiments in different regions of the Pacific Ocean. Using metagenomics, I characterized the impact of iron availability on the microbial community structure of the Central Equatorial Pacific Ocean. This study showed that addition of iron to an iron-limited ecosystem triggers a phytoplankton bloom dominated by *Pseudo-nitzschia*-like diatoms, which in turn stimulate a Bacteroidetes population functionally distinct from the ambient free-living population.

In the North Pacific Subtropical Gyre, I explored how iron availability impacts microbial community gene expression dynamics. Using a metatranscriptomic approach I showed that in that environment, the impact of iron was tightly connected to the supply of other limiting macronutrients, and seems to mostly affect photosynthetic organisms. This initial study paves the way for more in depth and longer-term studies to further investigate the effects of iron on the dynamics of the microbial community in the North Pacific Subtropical Gyre. Taken together data and analyses presented in this thesis demonstrate how iron availability can shape the ecology of marine microorganisms at population, community and functional levels.

Thesis Supervisor: Edward F. DeLong
Title: Morton and Claire Goulder Professor of Environmental Systems
Professor of Biological Engineering and Civil and Environmental Engineering

This doctoral thesis has been examined by a committee of the Biological Engineering Department as follows:

Jacquin Niles
Assistant Professor of Biological Engineering
Committee Chair

Mak Saito
Associate Scientist in Marine Chemistry and Geochemistry
Woods Hole Oceanographic Institution

Acknowledgements

This thesis would not have been possible without the support of many unique individuals. Whether they guided me scientifically, shared excited a-ha! moments, or stayed by my side through ups and downs, thanks to them I grew both as a scientist and a person and owe them great gratitude for accompanying me on this journey.

First, I would like to thank my supervisor Edward DeLong for his guidance and support through the years. I thank him for believing in me and trusting me, for providing me with the academic freedom to investigate questions outside the lab's main area of expertise. I thank him for always having his door open, ready to answer questions at any time, or engage in a variety of conversations. Finally, I thank him for letting me explore other career avenues, by allowing me to take classes at the Sloan School, supporting me in attending a conference on marine policy or getting involved in the C-MORE "behind-the-scenes"– I have learned so much! I thank my committee members Jacquin Niles and Mak Saito, for the stimulating conversations and the feedback, for making sure that I kept grounded in the "iron" field.

Members (past and present) of the DeLong lab are a constant source of support –answering one-off questions, sharing protocols, or acting as a soundboard for ideas. Specifically, I'd like to thank Chon Martinez who has been invaluable in teaching me the details of microbiology and molecular biology. I also thank her for always having her door open, ready to answer questions at any time. John Eppley, Liz Ottesen and Yanmei Shi have helped tremendously in navigating the bioinformatics. Rob Young and I spent long hours discussing statistics. Adrian Sharma– you have been a rock in the later stages, making sure that I ran the river loop, took 'thoughtful' breaks, encouraged me in my writing, and kept a good humor! Jay McCarren, Julie Maresca, Virginia Rich, Gene Tyson, who were here in the early days have helped me define my scientific questions.

This thesis is the fruit of various collaborative projects. I thank Martin Polz for letting me screen his lab's collection of Vibrio strains. I also thank him for the productive conversations and the invaluable feedback on this project. Many thanks as well to Dan Repeta for sharing valuable organic matter samples and for feedback on the organic matter and transcriptomic project.

Otto Cordero– our collaboration has been one of the best experiences of my scientific career. It started around a beer at the Muddy, a simple geeky conversation about public goods and siderophores. Two years later, I think Donkey and Shrek have produced a fine piece of work- the fruit of a completely synergistic collaboration. On top of the constant scientific learning and stimulation, I thank you for the jokes, the open-ness, the David Bowie moments, the 're-thinking' the future of the world conversations.

Kathleen Munson, Mar Nieto-Cid, Erin Bertrand– thanks for embarking on a project that at first seemed mission impossible. I have learned so much from each of you- you make bubble construction fun! I also thank Anne Thompson for introducing me to the iron field and Abigail Noble for helping me refine my understanding of iron chemistry in the ocean.

Outside the lab, I have been extremely fortunate to be surrounded by a group of amazing friends who have made my time as a graduate student unique and nurturing: Arek and Sophie for the long conversations going from A to Z, Renaud for the Celtics and Amélie Nothomb!, the folks from the Club Francophone (Olivier, Damien, Yves, Blandine), the Wellesley gang (Anja, Paulina, Cristina, Cecilia, Marina, Xan, Vane, and co.), for the brunches, the NY visits, the illuminating conversations about healthcare, for being my home away from home for 10 years… and counting, Marta, Victor and Telmo, Mike and Kristin for giving me a home and stability in the last stretch, Kira and Tyler for making the Watson spirit live on, Michel for helping me in my quest to become a Québécoise– pushing me to explore the hockey love in me!

Lastly, I would like to thank my family for their unconditional love and support through the years. This would not have been possible without them. Thank you for trusting me, for letting me go explore the other side of the world, for always welcoming me back with open arms. Maman chérie– merci.

Kevin- you have been standing by my side since almost the beginning of this journey, sharing my excitements and disappointments, always patient, always encouraging, always loving, always reminding me about the joys of life. The end of this journey marks the end of one of the many adventures that await us on our longer journey- I cannot wait to tackle more together with you!

This work was supported by a gift from Adele and Gordon Binder to the Department of Biological Engineering, a Gordon and Betty Moore Foundation award to EFD, and a Center for Microbial Oceanography Research and Education NSF Science and Technology Center award EF0424599.

Table of Contents

Abstract 3

Acknowledgements 4

Table of Contents 6

List of Figures 9

List of Tables 11

List of Abbreviations 12

Chapter 1: Introduction 13
1.1. Importance of Iron in the Ocean 13
1.2. Distribution of Iron in the Ocean 15
1.3. Microbial Strategies in an Iron-poor Ocean 18
1.3.1. Iron Acquisition Strategies 18
1.3.2. Metabolic Acclimation Strategies 26
1.4. Iron's Role in Influencing Microbial Ecology 29
1.5. Thesis Goals and Overview 32
1.6. References 34

Chapter 2: Siderophore Production in the *Vibrionaceae* 41
2.1. Abstract 41
2.2. Background 42
2.2.1. Motivation 42
2.2.2. The *Vibrionaceae*: a model system for the study of siderophore production in a natural environment 42
2.3. Materials and Methods 48
2.3.1. Siderophore screen of wild isolates 48
2.3.2. Bioinformatic analysis of siderophore synthesis and transport gene clusters 50
2.3.3. Correlation Analysis 53
2.3.4. Growth Enhancement Experiments 55
2.3.5. Spotting Bioassay 55
2.4. Results 57
2.4.1. Siderophore Production is a Patchy Trait 57
2.4.2. The *Vibrionaceae* Produce a Wide Variety of Siderophores 62
2.4.3. Iron Transport Genes are Ubiquitous Across the vibrio phylogeny 68
2.4.3. Genotypic and Phenotypic Signatures of 'Cheating' 70
2.5. Discussion 77
2.5.1. Patterns of siderophore production: information of iron in the microenvironment 77
2.5.2. Patterns of siderophore production: imprints of microbe-microbe interaction 81
2.6. Conclusion 85
2.7. Acknowledgements 85
2.8. References 86
2.9. Supplementary Materials 91

2.9.1. Table S2.1 91
2.9.2. Table S2.2 111
2.9.3. Table S2.3 142
2.9.4. Published Manuscript 144

Chapter 3: Changes in Microbial Community Structure in Response to Relief of Iron Limitation in the Equatorial Pacific Ocean 156
3.1. Abstract 156
3.2. Background 157
3.3. Materials and Methods 161
3.3.1. Experimental Set-up and Sample Collection 161
3.3.2. DNA Extraction 163
3.3.3. Clone Library Construction 163
3.3.4. Phylogenetic Assignment and Tree Construction 165
3.3.5. High-Throughput Sequencing 165
3.3.6. Bioinformatic Analysis 165
3.3.7. Imaging 166
3.4. Results 168
3.4.1. Read statistics 168
3.4.2. Diatoms Bloom in Response to Iron Addition 169
3.4.3. Iron Addition Leads to a Shift in the Bacterial Community 180
3.5. Discussion 186
3.6. Conclusions and Future Directions 190
3.7. Acknowledgements 191
3.8. References 192
3.9. Supplementary Materials 196

Chapter 4: Microbial Community Dynamics in Response to Changing Iron Availability in the North Pacific Subtropical Gyre 201
4.1. Abstract 201
4.2. Background 202
4.3. Materials and Methods 206
4.3.1. Preparation of amendments 206
4.3.2. Experimental Set-up and Sample Collection 207
4.3.3. Iron Measurements 210
4.3.4. Total Organic Carbon Measurements 211
4.3.5. Flow-Cytometry Measurements 211
4.3.6. Nucleic Acids Processing and Sequencing 211
4.3.7. Bioinformatic Analysis 218
4.3.8. Statistical Analysis 224
4.4. Results 228
4.4.1. Iron and Organic Matter Affect Growth and Overall Community Structure in Different Ways 228
4.4.2. Organism Specific Patterns of Gene Expression 233
4.4.2.1. *Prochlorococcus* Displays Subtle Response to Iron Despite and Overwhelming Response to Organic Matter 235
4.4.2.2. Organism Specific Patterns of Gene Expression show no Iron-Specific Response 259
4.5. Discussion 270

4.5.1. Despite Low Iron Concentrations in the NPSG, Other Nutrients Are Limiting...270
4.5.2. Organic Matter: a poor model of iron-binding ligands...273
4.5.3. Towards a Better Understanding of the Role of Iron in the NPSG...274
4.6. Conclusions and Future Directions...276
4.7. Acknowledgements...277
4.8. References...278

Chapter 5: Conclusions and Future Directions...282
5.1. Characterizing Iron Acquisition Strategies at the the population level...282
5.1.1. Summary...282
5.1.2. Implications of Key Findings and Future Work...283
5.2. Characterizing Iron-Induced Community Changes in Microbial Communities...286
5.2.1. Metagenomic Analysis of Iron-Induced Changes in the Central Equatorial Pacific Ocean...286
5.2.2. Metatranscriptomic Analysis of Iron-Induced Changes in the North Pacific Subtropical Gyre...288
5.3. Final Thoughts...291
5.4.References...292

Appendix A: Screening of Fosmid Libraries for Siderophore Production...294
A.1. Background...294
A.2. Methods...295
A.2.1. Fosmid library construction...295
A.2.2. Heterologous host: *E.coli* BW25113 *entF trfA*...295
A.2.3. Fosmid library screening for siderophore production...297
A.2.4. Identification of Siderophore Producing Fosmid Clones...298
A.3. Results...300
A.4. Conclusion and Future Work...304
A.5. Acknowledgements...305
A.6. References...306

Appendix B: Chrome Azurol Assay...308

List of Figures

Chapter 1
Figure 1.1. Siderophore uptake in gram-negative and gram-positive bacteria....19
Figure 1.2 Schematic summarizing microbial interactions in the context of iron acquisition.31

Chapter 2
Figure 2.1. Diversity of siderophores produced by the *Vibrionaceae*....46
Figure 2.2. Bioinformatic pipeline used to infer siderophore type....52
Figure 2.3. Method of contrasts....54
Figure 2.4. Siderophore production is an intermediate frequency trait in natural *Vibrionaceae* populations....58
Figure 2.5. Frequency of siderophore production per phylogenetic node....59
Figure 2.6. Sampled environment of strains screened....60
Figure 2.7. Siderophore production correlates positively with free-living lifestyle within the *V. splendidus*- like clade....61
Figure 2.8. The *Vibrionaceae* produce a variety of siderophores....63
Figure 2.9. Distribution of iron transporters....69
Figure 2.10. Loss of siderophore synthesis genes explains patchiness of siderophore production in the *V. splendidus*-like clade....72
Figure 2.11. Growth of producers and non-producers in iron-poor conditions....74
Figure 2.12. Siderophore utilization by non-producers....75
Figure 2.13. Growth of non-producer strains is facilitated by producer strains....76

Chapter 3
Figure 3.1. Experimental Set-up....162
Figure 3.2. Visual Assessment of phytoplankton bloom....170
Figure 3.3. Shift in average percent GC content between the Fe and the control samples....171
Figure 3.4. Iron addition leads to an enrichment in eukaryotes....172
Figure 3.5. Relative abundance of genus-specific sequence reads....176
Figure 3.6. Genus-level community changes....177
Figure 3.7. Phylogenetic representation of the cloned 18S rRNA gene sequences and selected close relatives....178
Figure 3.8. Coverage of Metagenomic Reads in the Fe Sample of *Phaeodactylum tricornutum* and *Thalassiosira pseudonana* genomes....179
Figure 3.9. Genus-level Changes Specific to the Bacterial Community....184
Figure 3.10. Shift in the Functional Profile of the Bacteroidetes Population....185

Chapter 4
Figure 4.1. Experimental set-up and sampling scheme....209
Figure 4.2. Total iron concentrations show no contamination of samples....230
Figure 4.3. Addition of organic matter results in two-fold increase in total organic carbon concentration....230
Figure 4.4. Organic matter addition leads to doubling of SYBR stained cells....231
Figure 4.5. Organic matter addition leads to drastic changes in community composition....232
Figure 4.6. Organism-specific orthologous gene clusters with a significant iron effect, organic matter effect or interaction effect....234

Figure 4.7. Comparison of *Prochlorococcus* gene expression in response to iron addition and to iron limitation....242
Figure 4.8. Organic matter and iron addition relieves *Prochlorococcus* physiological stress....244
Figure 4.9. Influence of organic matter on *Prochlorococcus* is observed in gene expression pattern at T12h....245
Figure 4.10. Significant organic matter effect on *Prochlorococcus* orthologous gene clusters at T12h and T36h....248
Figure 4.11. Addition of organic matter leads to phosphate limitation after 36 hours....251
Figure 4.12. Addition of organic matter relieves nitrogen limitation....252
Figure 4.13. Comparison of the relative expression of *Prochlorococcus* clusters in OM+Fe with clusters in OM at T12h....255
Figure 4.14. *Prochlorococcus* orthologous gene clusters for which the pattern of gene expression is significantly impacted by an interaction effect at T12h....256
Figure 4.15. *Prochlorococcus* clusters of orthologous genes significantly impacted by an interaction effect at T36h....258
Figure 4.16. Fold change expression of iron-specific orthologous gene clusters at T36h for organism specific reads....264
Figure 4.17. Influence of organic matter on SAR11 is observed in gene expression pattern at T12h....267
Figure 4.18. Relative expression of specific transporters in SAR11....269

Appendix A
Figure A.1. Behavior of Heterologous *E.coli* Host on CAS Agar....302
Figure A.2. Negative Control....302
Figure A3. Functional Screen of Fosmid Libraries....302
Figure A.4. Restriction Digest Profiles of Fosmid Clones that Complemented the Heterologous Host for Siderophore Production....303

List of Tables

Chapter 1
Table 1.1. Siderophore production by marine microorganisms 22

Chapter 2
Table 2.1. Literature review of siderophore production by the *Vibrionaceae* 45
Table 2.2. Types and strengths of siderophores produced by the *Vibrionaceae* isolates investigated 79
Table 2.3. Summary of vibrio populations and the corresponding siderophore biosynthesis genes identified in this study 80
Table S2.1. Results of the siderophore screen 91
Table S2.2. Result of bioinformatic searches for siderophore biosynthetic genes in sequenced vibrio strains 111
Table S2.3. List of iron transport genes used to look for orthologs in sequenced strains 142

Chapter 3
Table 3.1. Read statistics 168
Table 3.2. Summary statistics of diatom-specific reads 175
Table 3.3. Sequence reads recruiting to Bacteroidetes genomes 181

Chapter 4
Table 4.1. Coordinates of stations where organic matter was obtained 206
Table 4.2. Set of primers used for processing RNA samples 215
Table 4.3. Summary of probe sets used for the sample-specific rRNA subtraction 216
Table 4.4. Summary of amounts of DNA/RNA retrieved or used at each step 217
Table 4.5. Summary of organism-centric generation of clusters of orthologous genes 221
Table 4.6. Summary of sequencing statistics for gDNA and cDNA samples 223
Table 4.7. Models of differential expression for which posterior probabilities were computed with baySeq 225
Table 4.8. Summing of models 227
Table 4.9. List of *Prochlorococcus* orthologous gene clusters for which the pattern of gene expression is significantly impacted by iron at T12h 236
Table 4.10. List of *Prochlorococcus* orthologous gene clusters for which the pattern of gene expression is significantly impacted by iron at T36h 239
Table 4.11. List of orthologous gene clusters significantly impacted by an iron effect for the most abundant heterotrophs 262

Appendix A
Table A.1. Summary of Results 301

List of Abbreviations

BLAST	Basic Local Alignment Search Tool
DFB	Desferrioxamine B
DTPA	Diethyltriaminepentaacetic acid
EDDA	Ethylenediamine di(*o*-hydroxy-phenyl-acetic acid)
EDTA	Ethylenediamine tetra-acetic acid
HDPE	High density polyethylene
HOT	Hawaii Ocean Time Series
HNLC	High-nitrate low chlorophyll
KEGG	Kyoto Encyclopedia of Genes and Genomes
LB	Luria Bertani
LDPE	Low density polyethylene
NCBI	National Center for Biotechnology Information
NPSG	North Pacific Subtropical Gyre
NTA	Nitriloacetate
PIPES	1,4-Piperazinediethanesulfonic acid
PVDF	Polyvinyl difluoride

Chapter 1

Introduction

Iron is essential to almost all life on earth. Because iron is a cofactor of a number of enzymes involved in key biochemical reactions, its availability to marine microorganisms impacts global biogeochemical cycles (Morel & Price 2003; Jickells et al. 2005; Moore et al. 2009). However, despite being the fourth most abundant metal in the Earth's crust, biologically available iron is scarce in a number of environments, including the ocean (Bruland & Lohan 2003). Given this limited resource, recent evidence suggests that marine microorganisms have evolved complex iron-acquisition and metabolic adjustment strategies to efficiently compete in this environment. In order to accurately model and understand global ocean processes, in depth knowledge of the interplay between iron and biology in the marine environment is needed. To this end, combining the knowledge of iron chemical speciation (an active area of research) with an understanding of iron-related pathways possessed by marine microbes is essential. Novel genomic and metagenomic technologies enable us to probe the biological component of the interplay between iron and microbial communities at an unprecedented level of detail. Building on these new technologies, this thesis contributes to furthering our understanding of how iron affects the ecology of marine microbes.

1.1. Importance of Iron in the Ocean

Microbial biochemical processes play a critical role in ensuring the recycling of elements necessary for life (Falkowski et al. 2008). In the ocean, these processes have been shown to impact the carbon, nitrogen, phosphorus, silicon and sulfur cycles with important consequences for global biogeochemistry (Arrigo 2005). For instance, marine photosynthetic microorganisms are thought to account for about half of total photosynthetic carbon fixation on the planet (Morel & Price 2003). Because iron acts as a cofactor to a number of enzymatic reactions underlying these processes, its distribution and availability in the ocean

can have global consequences. Indeed, iron is needed to all forms of life on Earth (recent findings suggest that *Borrelia burgendorfii* may not need iron). Iron's extensive range of redox potentials makes it an ideal cofactor for enzymes that mediate electron-transfer reactions, including but not limited to: Fe-S proteins and cytochromes of the respiratory and photosynthetic electron transport chains, enzymes involved in pigment (including chlorophyll) biosynthesis, ribonucleotide reductases involved in DNA biosynthesis, some dioxygenases involved in the degradation of aromatic compounds, the nitrogenase enzyme complex, nitrate reductase and glutamate-synthase all involved in nitrogen metabolism, and superoxide dismutase needed during oxidative stress (Crichton 2001; Raven 1988). In the early 1990's, Martin hypothesized and shortly thereafter confirmed that primary production in large areas of the World's oceans is in fact limited by low iron inputs (Martin 1990; Martin et al. 1994). Spurred by the observation that high iron atmospheric inputs to the surface ocean coincided with the last glacial maximum, Martin's 'iron hypothesis' underlines how the distribution and availability of iron in the ocean can potentially have important global repercussions (Martin 1990).

In recent decades, the 'iron hypothesis' has been tested in various oceanic basins. It was initially verified shortly after Martin's seminal paper with the observation that addition of iron to surface seawaters of the Equatorial Pacific Ocean resulted in a massive phytoplankton bloom during the IronEx I and IronEx II expeditions (Martin et al. 1994; Coale et al. 1996). A number of subsequent mesocosm experiments showed that the growth of photosynthetic microorganisms, especially that of large eukaryotic phytoplankton (such as diatoms) was limited by low iron concentrations in a number of high-nitrate low-chlorophyll (HNLC) seascapes similar to the Equatorial Pacific, as in the Southern Ocean (Boyd et al. 2000; Hoffmann et al. 2006; Assmy et al. 2007) and the Subarctic Pacific Ocean (Tsuda et al. 2005). In addition to limiting photosynthetic microorganisms, it has been hypothesized that iron –because of its need in the enzyme nitrogenase –could limit nitrogen fixation in environments where both iron and nitrogen-sources (nitrate+nitrite) are scarce, as in the low-nitrate low-chlorophyll (LNLC) regions of the ocean (Karl & Letelier 2008). In the eastern low-latitude North Atlantic Ocean, the finding that high abundances of the diazotroph *Trichodesmium* spp. are correlated to high iron inputs from Saharan-sand-laden winds supports this hypothesis (Tyrrell et al. 2003). This hypothesis was subsequently confirmed in the North Atlantic

Ocean, where addition of both iron and phosphate to surface waters enhanced nitrogen fixation rates (Mills et al. 2004). While most efforts to date have focused on understanding the role of iron in limiting the growth and activity of photosynthetic and N_2-fixing microorganisms (Boyd et al. 2007), a few studies have suggested that iron availability could also affect the growth and activity of marine microbial heterotrophs (Tortell et al. 1996), which as the main actors of nutrient remineralization (Azam et al. 1983) are central to sustaining the ocean's ecosystem services. Therefore, understanding how iron availability in the ocean impacts biological processes is important to further our understanding of the global ocean ecosystem.

1.2. Distribution of Iron in the Ocean

Given the demonstrated importance of iron in regulating key biogeochemical processes, understanding the distribution of iron and its bioavailable forms in the marine environment is essential.

Despite the need of iron in biological processes, subnanomolar concentrations of dissolved iron (averaging 0.07 nmol kg^{-1} in open ocean environments) have routinely been measured in surface seawaters (Johnson et al. 1997). The scarcity of iron in the marine environment can be explained by the complex interaction of physical, chemical and biological processes that govern its distribution. The main geochemical sources that supply iron to the oceans include aerial dust deposition (Sedwick 2005; Mahowald et al. 2005), riverine discharge, iron plumes originating at hydrothermal input (Sedwick et al. 1992; Boyle et al. 2005), upwelling of sediments (Bruland et al. 2005) and horizontal advection of metals from the continental margins (Lam et al. 2006; Lam & Bishop 2008). Rapid biological uptake and sorption of iron onto particulate matter counterbalance these as the main sinks of iron in the marine environment (Bruland & Lohan 2003). Governed by competing abiotic and biotic processes, the vertical distribution of iron is typical of hybrid-type metals: at times nutrient-like with low surface concentrations, at times scavenging-like with concentrations decreasing with depth. As other hybrid-type metals, the distribution of iron is highly variable both across spatial and time scales (Bruland & Lohan 2003; Saito & Moffett 2002; Noble et al. 2008). Given this complexity, determining which of the different forms of iron may be biologically available is essential.

The unique redox properties of iron further impact its availability in the marine environment. Except for certain regions of the ocean such as the Southern Ocean where Fe(II) can be stable for weeks (Croot et al., 2001; Croot et al., 2005), the more biologically useful form of iron, ferrous iron, Fe(II) is generally not thermodynamically stable at circumneutral pH in the presence of oxygen as it is readily oxidized to the ferric form, Fe(III) (Bruland & Lohan 2003). In turn, Fe(III) reacts with hydroxyl anions and forms insoluble ferric-oxyhydroxides and other refractory mineral phases such as hematite and goethite (Bruland & Lohan 2003). These refractory phases precipitate and are scavenged by particulate matter, which sinks to the bottom of the ocean, hence decreasing the dissolved iron pool in surface waters (operational definition of <0.4 μm; most recently defined as <0.2 μm by GEOTRACES), hypothesized to be the more labile and biologically accessible pool (Bruland & Lohan 2003).

Recent advances in the field have demonstrated that as much as 99% of the dissolved iron pool in the ocean is bound by organic ligands of unknown structure, function and origin (Rue & Bruland 1995; Macrellis et al. 2001). Detected directly in seawater samples via competitive ligand exchange voltammetric techniques, two pools of iron-ligands, L_1 and L_2, have been characterized, operationally defined by their distinct conditional stability constants to ferric iron (Rue & Bruland 1997). L_1, the strongest pool of ligands with a conditional stability constant of $\sim 10^{12}$ M^{-1} is usually found in the upper layers of the water column, while L_2, the weaker pool of ligands with a stability constant of $\sim 10^{11}$ M^{-1} is more abundant at depth (Rue & Bruland 1997). To date, these iron-binding ligands have been detected in a variety of ocean biomes, including the North Pacific Subtropical Gyre (Rue & Bruland 1995), the Western Mediterranean (van den Berg 1995), the Equatorial Pacific (Rue & Bruland 1997), the North Sea (Gledhill et al. 1998), the North, South and Equatorial Atlantic (Witter & Luther 1998; Powell & Donat 2001), Southern Pacific (Nolting et al. 1998; Ibisanmi et al. 2011), the Arabian Sea (Witter, Lewis, and Luther 2000) as well as in the California upwelling region (Macrellis et al. 2001). Photoreactivity of natural iron-binding ligands was also detected, indicating that iron-cycling and hence bio-availability in surface seawater may be impacted by light (Johnson et al. 1994). While the exact chemical composition of these ligands has yet to be determined, the finding that ligand concentrations correlate well with dissolved iron concentrations (Buck & Bruland 2007)

indicates that these may play a key role in maintaining iron in the soluble pool, suggesting that iron-binding ligands in the ocean may be one of the main sources of bio-available iron. **We aim to test this hypothesis in Chapter 4.**

1.3. Microbial Strategies in an Iron-Poor Ocean

In order to better understand how iron affects microbial communities in the ocean, it is key to identify not only the different iron species present in the water column (Section 1.2), but also to determine what types of iron-related metabolism is harbored by marine microorganisms. While investigations to unravel these strategies are still under way, a conceptual picture is emerging suggesting these can be split into iron acquisition strategies on one hand, and metabolic acclimation strategies on the other. While high-affinity iron acquisition strategies ensure a sufficient supply of iron to the cell, metabolic acclimation strategies tend to reduce the absolute iron requirement of the cell.

1.3.1 Iron Acquisition Strategies

The main challenge faced by microorganisms in acquiring iron in the marine environment is the fact that the dominant form, Fe(III) is either insoluble or complexed by organic ligands whose role in influencing bio-availability is as of yet unclear (see Section 1.2). Hence, strategies for iron acquisition would require 1) the excretion of strong iron chelators (siderophores) that could compete with the ambient organic ligands and shuttle Fe(III) back to the cell, 2) the direct uptake of the ambient Fe(III)-organic ligand complex, or 3) a solubilization step potentially via reduction to the more soluble ferrous iron form.

1.3.1.1 Siderophore – strong iron chelators

Siderophores are small molecules with extremely high affinity constants for iron that are produced by microorganisms under iron-limiting conditions. Because they have been studied extensively in the context of microbial pathogenesis due to their link with virulence, much is known about siderophore-mediated iron acquisition in pathogenic and terrestrial microorganisms (Andrews et al 2003). When iron is scarce, siderophore biosynthesis is upregulated, usually via *fur* (ferric uptake regulator). Siderophores are excreted into the cell's surrounding medium, where they strongly chelate ferric iron. Because they have extremely high complex formation constants to ferric iron Fe(III), siderophores are considered to be specific ferric-iron chelators and are able to compete with the organic ligands that bind iron in the extracellular medium. In gram-negative bacteria, the ferric-iron-siderophore complex is recognized by

specific outer-membrane receptors, which shuttle the complex into the periplasm via an active TonB-dependent transport mechanism (see Figure 1.1). The complex is then brought into the cytoplasm via ABC transporters (Braun et al. 1998). In gram-positive bacteria, both membrane-bound substrate-binding proteins and ABC transporters mediate the transport of the ferric-iron-siderophore complex inside the cell (Heinrichs et al. 2004). Release of Fe(III) from the siderophore is achieved in one of two ways 1) destructive hydrolysis of the siderophore once the Fe(III)-siderophore complex is in the cytoplasm, or 2) intracellular reduction of Fe(III) to Fe(II) by intracellular ferric iron reductases. Because siderophores chelate ferrous iron much less strongly than ferric iron, release via the reductive mechanism leaves the siderophore intact for potential re-use. (Figure 1.1).

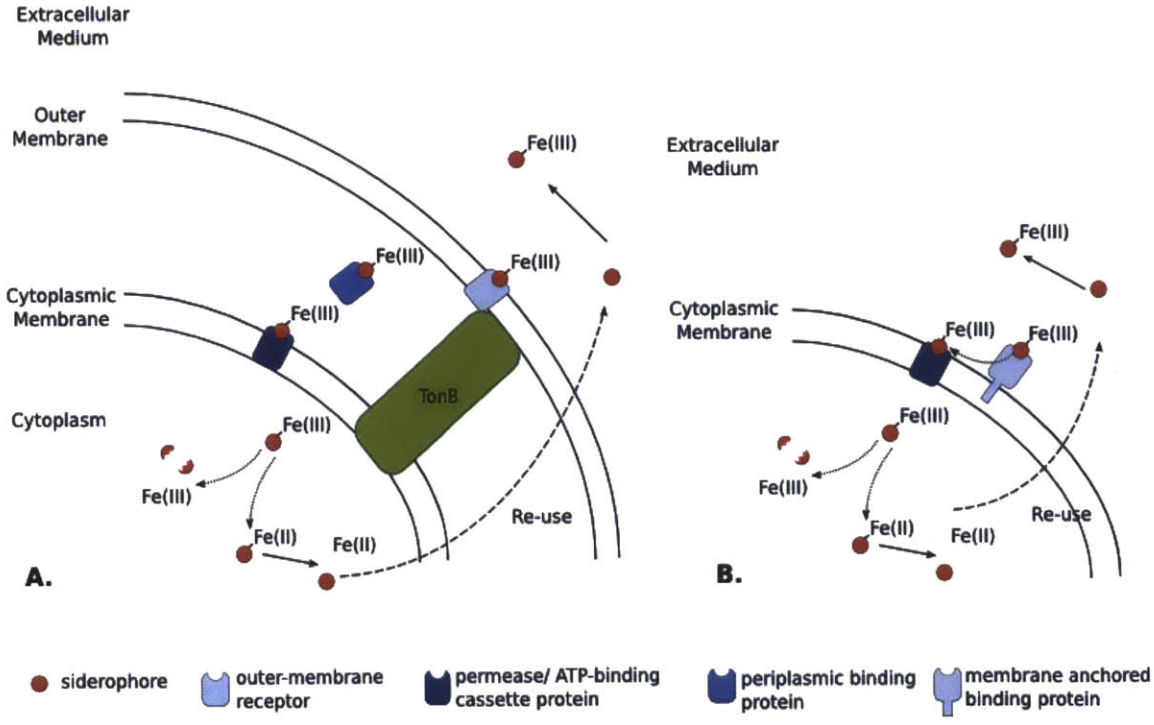


Figure 1.1: Siderophore uptake in Gram-Negative (A) and Gram-Positive (B) Bacteria. Adapted from Andrews et al. 2003. Siderophores are excreted into the extracellular medium. It is recognized by the cell's cognate outer-membrane receptor. Inside the cell, Fe(III) is released from the siderophore either via hydrolysis of the siderophore or via reduction of the ferric iron.

While the chemical characterization, synthesis and regulation of siderophores produced by pathogenic and terrestrial microorganisms are well understood, studies on siderophores produced by marine microorganisms are in their infancy. For a while, the usefulness of siderophores in a dilute environment such as the ocean was questioned: what would be the incentive to produce and secrete a costly molecule if it was to be lost in the extracellular environment (Volker & Wolf-gladrow 1999)? Yet, with their high affinity constant to iron, siderophores could present useful ways to remove Fe(III) from the uncharacterized pool of iron-binding organic ligands found in seawater. In fact, it was even suggested that part of these organic ligands might be siderophores themselves. Indeed, voltammetric measurements to determine the Fe(III) stability constant of these ligands show that these are extremely similar to the stability constants of siderophores (Rue & Bruland 1995). Another study revealed that organic ligands obtained from seawater in the California upwelling possess hydroxamate and catecholate chemical groups, typical iron-chelating moieties of siderophores (Macrellis et al. 2001). Hence, if some of the iron-binding ligands found in seawater (described in section 1.2) are actually siderophores, these molecules could play a more important and broader role in making iron available to the microbial community than what was originally thought.

Siderophore production and uptake in the marine environment has in fact been reported for a number of microorganisms. Back in the 1970's, bioassays revealed the presence of siderophores in homogenized mud and cyanobacterial mats sampled in coastal marine environments (Estep et al. 1975). This was followed by the isolation of schizokinen from cultures of iron-limited *Anabaena* sp. (Simpson & Neilands 1976). Since then, as summarized in Table 1.1, siderophore production has been detected for a variety of marine bacterial (Vraspir & Butler 2009), including a few cyanobacterial (Wilhelm & Trick 1994) as well as fungal isolates (Holinsworth & Martin 2010). Early reports also show that cultures of marine heterotrophs were able to access iron provided via siderophores (Granger & Price 1999). Confirming the initial findings that some marine isolates produce siderophores, synthesis and transport genes have been identified in the genomes and metagenomes of marine microorganisms (Hopkinson & Morel 2009). Corroborating these results, siderophores have recently been directly detected in seawater samples obtained across the North Atlantic Ocean (Mawji et al. 2008).

Efforts targeted towards the structural characterization of marine siderophores have revealed that marine microorganisms produce the three major types of siderophores: hydroxamates, catecholates and carboxylates (Table 1.1). These studies have also showed that some siderophores produced by marine microorganisms have unique chemical properties, amphiphilicity (Martinez 2000) and photoreactivity (Barbeau et al. 2001). With additional hydrophobic fatty acid side chains, amphiphilic siderophores can partition with the membrane of the microorganisms that produce them, potentially facilitating their uptake in a dilute environment (Martinez 2000). While a few amphiphilic siderophores have been described for pathogenic bacteria, including *Mycobacteria* (Snow & White 1969) and *Acinetobacteria* (Luo et al. 2005), most have been described for marine isolates, further indicating that this property seems of particular advantage in the marine environment.

Table 1.1: Siderophore production by marine microorganisms. [1]

Producer	Comments	Siderophores	Characteristics	Reference
Alteromonas	- γ-proteobacteria	Alterobactin A and B	-catecholate -carboxylate	(Reid & Alison Butler 1991; Reid et al. 1993; Holt et al. 2005)
Anabaena	-cyanobacteria -freshwater strain	Schizokinen	-hydroxamate	(Jeanjean et al. 2008) Simpson and Neilands, 1976
		Anachelin	-catecholate	(Ito et al. 2001)
Halomonas	- γ-proteobacteria	Aquachelin	-photoreactive	(Martinez 2000)
		Loihichelins	-photoreactive -amphiphilic	
Marinobacter	- γ-proteobacteria	Marinobactin A-E	-amphiphilic	(Martinez 2000; Martinez et al 2007)
		Vibrioferrin	-carboxylate -photoreactive	(Amin et al. 2007; Amin et al. 2012)
		Desferrioxamine E	-hydroxamate	(Amin et al. 2012)
		Petrobactin	- photoreactive	(Hickford et al. 2004)
Ochrobactrum	- α-proteobacteria	Ochrobactin	-amphiphilic -photoreactive	(Martin et al. 2006)
		Aerobactin	-carboxylate -photoreactive	(Haygood et al. 1993)
Photobacterium	- γ-proteobacteria	Piscibactin		(Souto et al. 2012)
Pseudoalteromonas	- γ-proteobacteria	Pseudoalterobactin A and B	-carboxylate	(Kanoh et al. 2003)
Synechococcus	-cyanobacteria -coastal strain	Synechobactin	-amphiphilic -photoreactive	(Ito et al. 2005)
Vibrio	- γ-proteobacteria	Aerobactin	-photoreactive	(Okujo et al. 1994)
		Anguibatin	-mixed hydroxamate and catecholate	(Actis et al. 1986; Sandy et al. 2010)
		Amphibactin B-I	-hydroxamate -amphiphilic	(Martinez et al. 2003)
		Bisucaberin	-hydroxamate	(Takahashi et al. 1987)
		Fluvibactin	-catecholate -hydroxyphenyl-oxazolone	(Yamamoto et al. 1993)
		Ochrobactin Ochrobactin-OH	-amphiphilic -photoreactive?	(Vraspir et al. 2011; Gauglitz et al. 2012)
		Vanchrobactin, (di- and tri-)	-catecholate	(Soengas et al. 2006; Sandy et al. 2010)
		Vibriobactin	- catecholate	(Griffiths et al. 1984)
		Vibrioferrin	- hydroxamate	(Yamamoto et al. 1994)
		Vulnibactin	-mixed hydroxamate and catecholate	(Biosca et al. 1996)
Aureobasidium	Ascomycetes	fusigen	-hydroxamate	(Wang et al. 2009)
Penicillium bilaii		pistillarin	-catecholate	(Capon et al. 2007)
Cunninghamella elegans		rhizoferrin	-carboxylate	(Martin et al. 2006)
	Zygomycetes		-carboxylate	(Vala et al. 2000)
	Ascomycota		-hydroxamates	(Baakza et al. 2004)
	Aspergilla		- hydroxamates	(Vala et al. 2006)

[1] photoreactive siderophores have the characteristic α-hydroxy carboxylic acid group (Barbeau et al. 2003).

The other special property of marine siderophores is photoreactivity. In the presence of a carboxylate side-chain, a ligand to metal charge transfer (LMCT) reaction is initiated by the absorption of a photon, resulting in the reduction of Fe(III) to Fe(II) and the concomitant oxidation of the carboxylate moiety (release of CO_2) (Faust & Zepp 1993; Abrahamson et al. 1994). First described for a marine siderophore with aquachelin (Barbeau et al. 2001) isolated from *Halomonas aquamarina* (Martinez et al., 2000), this process was thereafter experimentally confirmed for a number of siderophores produced by marine microorganisms, including aerobactin (Küpper et al. 2006), vibrioferrin (Amin et al. 2009), petrobactin (Barbeau et al. 2002), synechobactin (Ito & A Butler 2005), ochrobactin (Martin et al. 2006). Many of the known siderophores produced by marine microbes have a carboxylate moiety suggesting that light driven processes might play a particularly important role in siderophore-mediated iron uptake (Gauglitz et al. 2012; Butler & Theisen 2010).

As discussed earlier siderophore production by marine bacteria is mostly described for isolates belonging to the alpha or gamma proteobacteria class (Table 1.1) and production by a eukaryotic phytoplankton species has yet to be discovered. While this is potentially an artifact of the 'culturability' of gamma-proteobacteria, this also raises the question of the universality of a siderophore-based iron acquisition strategy within the marine environment. Indeed, from an evolutionary point of view, the production and secretion of costly molecules, which are likely to be lost in the dilute environment does not make sense. In fact, comparative modeling of the kinetics of siderophore uptake and diffusion verses the kinetics of surface-bound iron reductases suggests that siderophore production only makes thermodynamic sense when cells have reached a certain density, as would be the case in blooms or in biofilms formed on particles (Volker & Wolf-gladrow 1999). **In this thesis (chapter 2), we explore the hypothesis that siderophore production does not only depend on microbial phylogeny but also on habitat.**

1.3.1.2 Direct uptake of ferric-iron complexes

Another strategy utilized by microorganisms to transport iron is the direct uptake of ferric-iron complexes present in their environment without having to first synthesize costly molecules. Such ferric-iron complexes include siderophores produced by other microorganisms (Braun & Hantke 2011), as demonstrated with cultures of marine bacteria to which exogenous siderophores were added (Granger et al. 1999). Furthermore, a number of molecules that are found in the extracellular medium have the ability to chelate iron (ie. phosphate, albeit weak, is an iron chelator). Some of these are byproducts of central metabolism, such as citrate, an intermediate in the TCA cycle. Its abundance and role as an iron-chelator has previously been examined (Pierre et al. 2000), and has been recognized to play a role in the iron nutrition of plant-associated bacteria and fungi (Guerinot et al. 1990). The negative correlation between iron availability and the prevalence of *FecA* (ferric-dicitrate outer-membrane receptor) homologs in metagenomes from the Global Ocean Survey suggests that citrate or similar molecules might also play an important role in iron nutrition of marine microorganisms (Toulza et al. 2012).

Pathogens that invade mammalian hosts are known to recognize and take up heme, an abundant iron-chelating protein found in the host's environment (Genco & Dixon 2001). Surprisingly, heme-specific membrane transporters were detected in the genomes of a number of marine bacteria usually found to be associated with particulate matter, including some *Vibrionaceae*, *Pseudoalteromonas* and *Pseudomonas*, to cite a few (Hopkinson et al. 2008). Further investigation showed that addition of exogenous heme to cultures of *Microscilla*, led to an upregulation of heme-specific transporters as well as the uptake of Fe bound to heme (Hopkinson et al. 2008). While the exact nature of the iron-chelating organic ligands found in the marine environment is not known, we could hypothesize that marine microbes carry transporters that could specifically recognize them and potentially directly internalize them.

1.3.1.3. Extracellular Iron Reduction.

The third general iron acquisition mechanism utilized by marine microorganisms is the extracellular reduction of iron. This can be mediated by membrane-bound iron reductases, which is a well-documented form of iron acquisition by fungi such as yeast (Askwith & Kaplan 1997), higher plants (Bienfait 1987) and to a lesser extent by some bacteria (Cowart 2002). Culture-based experiments (Maldonado & Price 2000) as well as recent analyses of genomes (Armbrust et al., 2004; Kustka et al. 2007) have demonstrated that some eukaryotic photosynthetic microorganisms such as diatoms also possess these membrane-bound ferric-iron reductases. In addition to surface-bound reductases, reactive oxygen species (ROS), including hydrogen peroxide and superoxide can also mediate the reduction of ferric iron bound to organic matter (Rose & Waite 2005). These ROS, which are produced either via photolysis of colored dissolve organic matter or leaked by phytoplankton (Marshall et al. 2005) have been linked to iron acquisition in cyanobacteria (Rose et al. 2005). Finally, as described in section 1.2.2.1, reduction of Fe(III) bound to organic ligands can be mediated by the photoreactivity of the organic ligands that bind iron (Barbeau 2006).

Because eukaryotic phytoplankton generally do not possess siderophore-specific transporters, it is widely accepted (and verified experimentally) that extracellular iron reduction plays a key role in the way in which they acquire iron (Shaked & Lis 2012). As experimental observations improve, the models of iron acquisition by marine phytoplankton are constantly re-evaluated (Morel et al 2008). Points of disagreement between these models include the iron species taken up: the Fe' model proposes that soluble Fe(III)' (referring to the sum of inorganic species) is taken up via ferric iron ABC transporters (Hudson & Morel 1990; Hudson & Morel 1993; Sunda & Huntsman 1995), while both the Fe(II)s (Shaked et al 2005) and the FeL models (Salmon et al. 2006) suggest that soluble ferrous iron is taken up via ferrous iron ABC transporters. Another point of disagreement is the form of ferric iron prior to reduction: in the Fe(II)s model both transiently soluble Fe(III) (Fe(III)') and Fe(III) bound to organic ligands are subject to reduction, whereas in FeL model, only Fe(III) bound to organic ligands is reduced (Morel et al. 2008). Regardless of these differences, all models propose that the main type of iron available for uptake by eukaryotic phytoplankton is unchelated iron, which is generated in one way or another by an initial extracellular reduction step. A recent bioinformatics study shows that the most common iron transporter

found in bacterial genomes and metagenomes is the Fe^{3+} ABC transporter (Hopkinson et al. 2012), suggesting that unchelated iron generated after extracellular iron reduction might not only be important for eukaryotic phytoplankton but for bacterial species as well.

1.3.2 Metabolic Acclimation Strategies

In addition to highly specific iron transport mechanisms used to enhance the supply of iron into the cell, marine microorganisms have evolved a variety of metabolic acclimation strategies. Uncovering these various acclimation strategies is an active area of research, but it would seem that most of them effectively decrease the cell's absolute iron requirement.

Because microorganisms differ in their requirement for iron, they display a variety of acclimation strategies to iron scarcity. Effects of iron limitation are especially dramatic for marine photo-autotrophs (including cyanobacteria and eukaryotic phytoplankton) and nitrogen fixers. A fully functional photosynthetic apparatus requires 22-23 atoms of iron (Ferreira & Straus 1994), while the nitrogenase enzyme complex, which catalyzes nitrogen fixation requires 30 atoms of iron (Rees et al. 2005). This high iron requirement makes photosynthesizers and diazotrophs especially vulnerable to iron stress, illustrated by decreased photosynthetic efficiency of both natural phytoplankton assemblages (Behrenfeld & Kolber 1999) and cultures (Berman-Frank et al. 2001), and a 50% decrease in rates of nitrogen fixation in iron-limited cultures of some diazotrophs such as *Trichodesmium* (Berman-Frank et al. 2001). Numerous studies have focused on understanding metabolic adjustment strategies of phytoplankton both at the physiological and genomic/transcriptomic level, (reviewed for cyanobacteria and eukaryotic algae in Morrissey & Bowler 2012).

Iron stress in photo-autotrophs and diazotrophs leads to a number of physiological and molecular changes. One of these strategies is the decrease in the iron quota of the cell as has been demonstrated in cultures of *Prochlorococcus* and *Trichodesmium* (Block et al. 1997; Berman-Frank et al. 2001). At the physiological level, this is achieved by the down-regulation of iron-rich proteins, such as chlorophyll (a process also known as chlorosis) (Oquist 1971; 1974), photosystems I and II, phycobilisomes, Fe-S cluster

containing proteins (Ferreira & Straus 1994). At the molecular level, the reduction in the iron quota of the cell is achieved by substituting iron-containing enzymes with iron-free enzymes, such as the replacement of the Fe-S cluster containing electron carrier ferredoxin by the iron-free flavodoxin (Ferreira & Straus 1994; Laroche et al. 1996). A recent metagenomic analysis of *Prochlorococcus* clades in iron-limited waters revealed the importance of the latter strategy in the wild (Rusch et al. 2010): these clades possessed a much lower number of iron-containing proteins (ie. loss of nitrate reductase, plastoquinol terminal oxidase, ferredoxin and cytochrome C_m) when compared to clades from more iron-replete environments (Rusch et al. 2010). Other interesting strategies include the recycling of iron between metalloenzymes involved in distinct processes; the unicellular photosynthetic diazotroph *Crocosphaera* shuttles iron between photosynthetic enzymes, which are utilized during the day and nitrogen fixation enzymes, active at night (Saito et al. 2010). One of the consequences of iron-stress at the cellular level is a decrease in the cell-division rate, as observed for wild populations of *Prochlorococcus* (Mann & Chisholm 2000). This presumably results in cells that are smaller than their counterparts found in iron-replete regions, a worthwhile acclimation strategy as a larger surface-to-volume ratio allows smaller cells to take better advantage of the kinetics of diffusion (Sunda & Huntsman 1995).

In comparison to photo-autotrophs and diazotrophs, acclimation strategies utilized by marine heterotrophs have not been as widely studied. Yet, iron-containing enzymes are an integral part of the respiratory electron transport chain, DNA and nucleotide biosynthesis, suggesting that iron deprivation would impact heterotrophic bacteria as well. Indeed, in some regions of the ocean, heterotrophic bacteria account for ~40% (Tortell et al. 1996) to ~80% (Schmidt & Hutchins 1999) of the biogenic iron pool. Iron-enrichment experiments in the Southern Ocean revealing that heterotrophic bacteria can also be limited by iron (Pakulski et al. 1996) and the finding that marine bacteria can account for up to 50% of the total planktonic iron uptake (Tortell et al. 1999) further support the notion that lack of iron impacts marine heterotrophs. More recently, it has been demonstrated that the bacterial community in the Ross Sea is not only limited by the lack of dissolved organic matter but also by iron (Bertrand et al., 2011). Hence, we can expect that these bacteria have also evolved special mechanisms to cope with iron scarcity. In fact, some of

the acclimation strategies cited above, such as the replacement of iron-containing ferredoxin with iron-free flavodoxin could be used by heterotrophs as well (Sancho 2006).

Despite the limited number of studies aimed at understanding the response of marine heterotrophs to iron limitation, based on studies in pathogenic and terrestrial microorganisms, we can expect that iron stress would also affect global cellular processes of marine heterotrophs. Indeed, most microorganisms regulate their iron metabolism in response to iron availability via a global regulator such as *fur* (ferric uptake regulator) (Andrews et al., 2003). When complexed with Fe^{2+}, Fur binds to the fur recognition sequence, the 'Fur box', composed of GATAAT repeats upstream of the gene it targets and of which it blocks transcription (Andrews et al., 2003; Baichoo et al., 2002). In *E.coli*, over 100 genes are regulated by *fur*, including genes involved in iron transport, respiration, the TCA cycle, DNA synthesis and redox stress tolerance, to cite a few (Andrews et al., 2003). Confirming the global impact of iron stress on marine heterotrophs, one recent study shows that over 23 genes have at least a 50% higher expression level in iron-depleted cultures of the widely distributed *Candidatus Pelagibacter ubique* compared to the control (Smith et al. 2010). While the *Candidatus P. ubique* genome lacks a siderophore biosynthetic gene cluster (Giovannoni et al. 2005), one of the most strongly upregulated genes was *sfuC*, coding for the permease of the of the Iron(III) ABC transporter (Smith et al. 2010). In fact, one of the most widely studied microbial responses to iron stress is the upregulation of high-affinity iron transport systems such as siderophores, discussed previously (Andrews et al., 2003). Another common acclimation strategy is the upregulation of iron storage proteins, such as ferritins, under iron-replete conditions (Andrews et al. 2003).

1.4. Iron's role in influencing microbial ecology

The variety of iron transport systems and metabolic acclimation strategies discussed in Sections 1.2 and 1.3 indicate the diversity of coping mechanisms utilized by various microbial species, under various conditions of iron availability. Beyond the context of the individual organism, iron availability (or the lack thereof) also shapes the ecology of microbial assemblages, not only by spawning competition between species but also by mediating the development of inter-species synergistic interactions.

Together with light, pH, salinity and temperature, nutrient availability is one of the main factors that controls the growth of phytoplankton and bacterioplankton in the ocean. As nutrient supplies to the surface ocean are limited, resource competition amongst microbial species has been shown to impact species succession and the composition of microbial communities (Dugdale 1967; Tilman et al. 1982). Indeed, in environments with low nutrient inputs, thriving microbial species are those that can best access scarce resources or that have evolved alternative metabolic strategies reducing the need for a limiting nutrient. For example, under conditions of nitrate and nitrite limitation (and provided all other macro- and micro-nutrients are present in sufficient concentrations), diazotrophs that can instead utilize atmospheric N_2 as a source of nitrogen thrive. As a necessary but scarce micronutrient, iron is emerging as one of the factors shaping microbial community composition in the marine environment. Recent experiments in the HNLC region of the Equatorial Pacific Ocean and Southern Ocean confirm this. Indeed, upon relief of iron limitation, the community shifts from one dominated by small picophytoplankton (with high surface-to-volume ratios facilitating the uptake of scarce nutrients and lower absolute iron-requirements) to one dominated by large eukaryotic phytoplankton (Cavender-Bares et al. 1999; Hoffmann et al. 2006).

While siderophore biosynthesis seems to be restricted to a subset of marine bacteria, siderophore-specific transporters are harbored by a much wider group of microorganisms (Hopkinson & Morel 2009; Hopkinson & Barbeau 2012), suggesting that despite not producing siderophores many microorganisms can nonetheless transport them. The findings that natural phytoplankton assemblages (Maldonado & Price 1999; Hutchins et al. 1999) as well as cultures of eukaryotic photosynthetic microbes including *Chlorella*

vulgaris (Allnutt & Bonner 1987), *Thalassiosira oceanica* (Maldonado & Price 2001), and *Phaeodactylum tricornutum* (Horstmann & Soria-Dengg 1995; Soria-Dengg et al. 2001) can utilize (mostly via extracellular iron reduction of the siderophore-bound Fe(III)) the iron bound to siderophores supports this hypothesis. By benefitting from the siderophore without incurring the energetic costs associated with its production, these microorganisms are potentially likely to reach high abundances given the presence of a few siderophore producers.

In addition to generating competitive interactions, the lack of iron in the marine environment can potentially foster synergistic inter-species and even inter-kingdom interactions. Indeed, a recent study suggests that the vibrioferrin siderophore produced by a dinoflagellate-associated *Marinobacter* is at the center of a mutualistic relationship in which the bacterium produces a photoreactive siderophore, which upon photoreduction provides bioavailable iron to the dinoflagellate in turn providing dissolved organic matter to the bacterium (Amin et al., 2009). This interplay between extracellular reduction of iron, accepted as a valid mechanism for iron acquisition by eukaryotic phytoplankton and siderophores, likely produced by bacteria points to the complexity of iron acquisition in the marine environment and to how it may mediate interesting microbe-microbe interactions. Further considering the recent findings that marine fungi also produce a suite of siderophores, reviewed by Holinsworth et al. (Holinsworth & Martin 2010), we can expect to discover many more interesting microbe-microbe interactions generated by the race to acquire iron in the iron-scarce marine environment (see Figure 1.2).

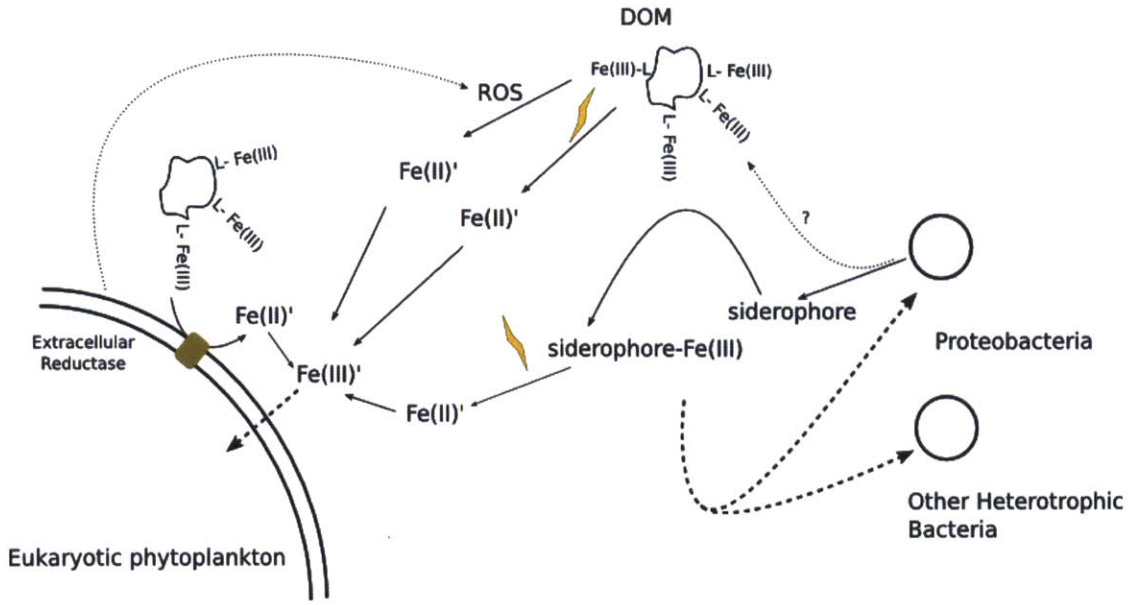


Figure 1.2: Schematic summarizing microbial interactions in the context of iron acquisition Eukaryotic phytoplankton and bacteria compete for iron in iron-scarce regions but can also benefit each other by engaging in mutualistic processes.

1.5. Thesis goals and overview

The overall goal of this thesis is to better understand how iron shapes the ecology of marine microbes that are central to the functioning of the ocean ecosystem, at the population and community levels. In this thesis we aim to explore the underlying transport and metabolic adjustment processes (at the genomic and transcriptomic level) that impact community structure and function in response to iron availability.

In Chapter 2, I explore how patterns of siderophore production are related to the structure of a population of closely-related *Vibrionaceae* isolates. By performing a functional screen of ~1,300 *Vibrionaceae* isolates combined with the genome analysis of 69 of these we can ask: **how widespread and diverse is siderophore production amongst the** *Vibrionaceae*? Because the *Vibrionaceae* isolates studied were obtained from a variety of micro-habitats, I ask: **is siderophore production correlated to habitat?** Finally, given the high-throughput nature of the data combined with genomic information, I can ask: **Is there genomic evidence for iron-mediated microbe-microbe interactions?**

In Chapter 3, I investigate how iron affects the composition, and hence ecosystem function of a natural microbial assemblage in the Equatorial Pacific (Chapter 3). By performing a microcosm experiment in which iron is provided to the natural microbial assemblage combined with a high-throughput sequencing/ metagenomic approach, I can ask **how does iron affect microbial community structure-both in terms of diversity and functional gene content?**

In Chapter 4, I explore the influence of iron-availability on microbial community dynamics by investigating the gene expression response of a natural microbial assemblage in the North Pacific Subtropical Gyre to different forms of iron. Using an experimental metatranscriptomic approach I ask: **how does iron affect community composition and gene expression patterns of a microbial community indigenous to a low-nitrate low-chlorophyll region?** In light of the potential role of iron-binding ligands in increasing iron-bioavailability (Section 1.2), I also ask: **does the concurrent addition of iron and organic matter enhance the iron specific signal?**

By exploring patterns of siderophore production at the population level and investigating the impact of iron availability on microbial community composition and dynamics *in situ*, this thesis provides a unique overview of how iron availability influences the ecology of microbial microbes in the wild.

1.6. References

Abrahamson, H.B., Rezvani, A.B. & Brushmiller, J.G., 1994. Photochemical and spectroscopic studies of complexes of iron (III) with citric acid and other carboxylic acids. Inorganica Chimica Acta, 226, pp. 117-127.

Actis, L. a et al., 1986. Characterization of anguibactin, a novel siderophore from Vibrio anguillarum 775(pJM1). Journal of Bacteriology, 167(1), pp.57-65.

Allnutt, F.C. & Bonner, W.D., 1987. Characterization of Iron Uptake from Ferrioxamine B by Chlorella vulgaris. Plant physiology, 85(3), pp.746-50.

Amin, S. a et al., 2007. Boron binding by a siderophore isolated from marine bacteria associated with the toxic dinoflagellate Gymnodinium catenatum. Journal of the American Chemical Society, 129(3), pp. 478-9.

Amin, S. a et al., 2012. Iron transport in the genus Marinobacter. Biometals, 25(1), pp.135-47.

Amin, S. a et al., 2009. Vibrioferrin, an unusual marine siderophore: iron binding, photochemistry, and biological implications. Inorganic chemistry, 48(23), pp.11451-8.

Andrews, S.C., Robinson, A.K. & RodrÃguez-QuiÃ±ones, F., 2003. Bacterial iron homeostasis. FEMS Microbiology Reviews, 27(2-3), pp.215-237.

Arrigo, K.R., 2005. Marine microorganisms and global nutrient cycles. Nature, 437(7057), pp.349-55.

Askwith, C. & Kaplan, J., 1997. An oxidase-permease-based iron transport system in Schizosaccharomyces pombe and its expression in Saccharomyces cerevisiae. The Journal of biological chemistry, 272(1), pp.401-5.

Assmy, P. et al., 2007. Mechanisms determining species dominance in a phytoplankton bloom induced by the iron fertilization experiment EisenEx in the Southern Ocean. Deep Sea Research Part I: Oceanographic Research Papers, 54(3), pp.340-362.

Azam, F. et al., 1983. The Ecological Role of Water-Column Microbes in the Sea. Marine Ecology Progress Series, 10, pp.257-263.

Baakza, A. et al., 2004. A comparative study of siderophore production by fungi from marine and terrestrial habitats. Journal of Experimental Marine Biology and Ecology, 311, pp.1-9.

Baichoo, N., Helmann, JD. 2002. Recognition of DNA by Fur: a reinterpretation of the Fur box consensus sequence. Journal of Bacteriology, 184(21), pp. 5826-32.

Barbeau, K et al., 2001. Photochemical cycling of iron in the surface ocean mediated by microbial iron(III)-binding ligands. Nature, 413(6854), pp.409-13.

Barbeau, Katherine et al., 2002. Petrobactin, a photoreactive siderophore produced by the oil-degrading marine bacterium Marinobacter hydrocarbonoclasticus. Journal of the American Chemical Society, 124(3), pp.378-9.

Barbeau, Katherine et al., 2003. Photochemical reactivity of siderophores produced by marine heterotrophic bacteria and cyanobacteria, based on characteristic Fe(III) binding groups. Limnology and Oceanography, 48(3), pp.1069-1078.

Barbeau, Katherine, 2006. Photochemistry of organic iron(III) complexing ligands in oceanic systems. Photochemistry and photobiology, 82(6), pp.1505-16.

Behrenfeld, M. & Kolber, Z., 1999. Widespread iron limitation of phytoplankton in the south pacific ocean. Science (New York, N.Y.), 283(5403), pp.840-3.

van den Berg, C.M.G., 1995. Evidence for organic complexation of iron in seawater. Marine Chemistry, 50(1-4), pp.139-157.

Berman-Frank, I. et al., 2001. Iron availability, cellular iron quotas, and nitrogen fixation in Trichodesmium. Limnology and Oceanography, 46(6), pp.1249-1260.

Bertrand, E.M., Saito, M.A., Lee, P.A., Dunbar, R.B., Sedwick, P.N., DiTullio, G.R., 2011. Iron limitation of a springtime bacterial and phytoplankton community in the Ross Sea: implications for Vitamin B12 nutrition. Frontiers in Microbiology. 2, 160.

Bienfait, H.F., 1987. Biochemical basis of iron efficiency reactions in plants. In G. Winkelmann, D. van der Helm, & J. B. Neilands, eds. Iron transport in microbes, plants and animals. Weinheim, pp. 339-349.

Biosca, E.G. et al., 1996. Siderophore-mediated iron acquisition mechanisms in Vibrio vulnificus biotype 2. Applied and environmental microbiology, 62(3), pp.928-35.

Block, H.A., Sunda, W.G. & Huntsman, S.A., 1997. Interrelated influence of iron , light and cell size on marine phytoplankton growth. Nature, 2051(1977), pp.389-392.
Boyd, P.W. et al., 2000. A mesoscale phytoplankton bloom in the polar Southern Ocean stimulated by iron fertilization. Nature, 407(6805), pp.695-702.
Boyd, P.W. et al., 2007. Mesoscale iron enrichment experiments 1993-2005: synthesis and future directions. Science (New York, N.Y.), 315(5812), pp.612-7.
Boyle, E. a. et al., 2005. Iron, manganese, and lead at Hawaii Ocean Time-series station ALOHA: Temporal variability and an intermediate water hydrothermal plume. Geochimica et Cosmochimica Acta, 69(4), pp.933-952.
Braun, V. & Hantke, K., 2011. Recent insights into iron import by bacteria. Current opinion in chemical biology, 15(2), pp.328-34.
Braun, V., Hantke, K. & Koster, W., 1998. Bacterial iron transport: mechanisms, genetics and regulation. Metal Ions in Biological Systems, 35, pp.67-145.
Bruland, K W & Lohan, M.C., 2003. Controls of Trace Metals in Seawater. In H. Holland & K. Turekian, eds. Treatise on Geochemistry. Elsevier, pp. 23-47.
Bruland, Kenneth W. et al., 2005. Iron, macronutrients and diatom blooms in the Peru upwelling regime: brown and blue waters of Peru. Marine Chemistry, 93(2-4), pp.81-103.
Buck, K.N. & Bruland, Kenneth W., 2007. The physicochemical speciation of dissolved iron in the Bering Sea, Alaska. Limnology and Oceanography, 52(5), pp.1800-1808.
Butler, Alison & Theisen, R.M., 2010. Iron(III)-siderophore coordination chemistry: reactivity of marine siderophores. Coord. Chem. Rev., 254(Iii), pp.288-296.
Capon, R.J. et al., 2007. Citromycetins and bilains A-C: new aromatic polyketides and diketopiperazines from Australian marine-derived and terrestrial Penicillium spp. Journal of natural products, 70(11), pp. 1746-52.
Cavender-Bares, K.K. et al., 1999. Differential response of equatorial Pacific phytoplankton to iron fertilization. Limnology and Oceanography, 44(2), pp.237-246.
Coale, K.H. et al., 1996. A massive phytoplankton bloom induced by an ecosystem-scale iron fertilization experiment in the equatorial Pacific Ocean. Nature, 383, pp.495-501.
Cowart, R.E., 2002. Reduction of iron by extracellular iron reductases: implications for microbial iron acquisition. Archives of biochemistry and biophysics, 400(2), pp.273-81.
Crichton, R., 2001. The importance of iron for biological systems. In R. Crichton, ed. Inorganic Biochemistry of Iron Metabolism: From Molecular Mechanisms to Clinical Consequences. Chichester, NY: John Wiley and Sons Ltd, pp. 17-48.
Croot, P.L., Bowie, A.R., Frew, R.D., Maldonado, M.T., Hall, J.A., Safi, K.A., La Roche, J., Boyd, P.W., Law, C.S., 2001. Retention of dissolved iron and Fe(II) in an iron induced Southern Ocean phytoplankton bloom. Geophysical Research Letters. 28, pp. 3425-3428.
Croot, P.L., Laan, P., Nishioka, J., Strass, V., Cisewski, B., Boye, M., Timmermans, K.R., Bellerby, R.G., Goldson, L., Nightingale, P., de Baar, H.J.W. 2005. Spatial and temporal distribution of Fe(II) and H_2O_2 during EisenEx, an open ocean mesoscale iron enrichment. Marine Chemistry, 95, pp. 65-88.
Dugdale, R.C., 1967. Nutrient Limitation in the Sea: Dynamics, Identification and Significance. Limnology and Oceanography, 12, pp.685-695.
Estep, M., Armstrong, J.E. & Van Baalen, C., 1975. Evidence for the occurrence of specific iron (III)-binding compounds in near-shore marine ecosystems. Applied microbiology, 30(2), pp.186-8.
Falkowski, P.G., Fenchel, T. & Delong, E.F., 2008. The microbial engines that drive Earth's biogeochemical cycles. Science (New York, N.Y.), 320(5879), pp.1034-9.
Faust, B. & Zepp, R., 1993. Photochemistry of Aqueous Iron(III)-Polycarboxylate Complexes: Roles in the Chemistry of Atmospheric and Surface Waters. Environmental science & technology, 27(12), pp. 2517-2522.
Ferreira, F. & Straus, N.A., 1994. Iron deprivation in cyanobacteria. Journal of Applied Phycology, 199(Ps Ii), pp.199-210.
Gauglitz, J.M., Zhou, H. & Butler, Alison, 2012. A suite of citrate-derived siderophores from a marine Vibrio species isolated following the Deepwater Horizon oil spill. Journal of inorganic biochemistry, 107(1), pp.90-5.

Genco, C.A. & Dixon, D.W., 2001. Emerging strategies in microbial haem capture. Molecular microbiology, 39, pp.1-11.
Giovannoni, S.J. et al., 2005. Genome streamlining in a cosmopolitan oceanic bacterium. Science (New York, N.Y.), 309(5738), pp.1242-5.
Gledhill, M. et al., 1998. Variability in the speciation of iron in the northern North Sea. Marine Chemistry, 59(3-4), pp.283-300.
Granger, J. & Price, Neil M., 1999. The importance of siderophores in iron nutrition of heterotrophic marine bacteria. Limnology and Oceanography, 44(3), pp.541-555.
Griffiths, G.L. et al., 1984. Vibriobactin , a Siderophore from Vibrio cholerae. J. Biol. Chem., 259(1), pp. 383-385.
Guerinot, M.L., Meidl, E.J. & Plessner, O., 1990. Citrate as a siderophore in Bradyrhizobium japonicum. Journal of bacteriology, 172(6), pp.3298-303.
Haygood, M.G., Holt, P.D. & Butler, Alison, 1993. Aerobactin production by a planktonic marine Vibrio sp. Limnology and Oceanography, 38(5), pp.1091-1097.
Heinrichs, D.E. et al., 2004. Iron Transport Systems in Pathogenic Bacteria: Staphylococcus, Streptococcus, Bacillus. In J. H. Crosa, A. R. Mey, & S. M. Payne, eds. Iron Transport in Bacteria. Washington, DC: ASM Press, pp. 387-401.
Hickford, S.J.H. et al., 2004. Petrobactin sulfonate, a new siderophore produced by the marine bacterium Marinobacter hydrocarbonoclasticus. Journal of natural products, 67(11), pp.1897-9.
Hoffmann, L.J. et al., 2006. Different reactions of Southern Ocean phytoplankton size classes to iron fertilization. Limnology and Oceanography, 51(3), pp.1217-1229.
Holinsworth, B. & Martin, J.D., 2010. Siderophore production by marine-derived fungi. Biometals, 22(4), pp.625-632.
Holt, P.D. et al., 2005. Iron(III) coordination chemistry of alterobactin A: a siderophore from the marine bacterium Alteromonas luteoviolacea. Inorganic chemistry, 44(21), pp.7671-7.
Hopkinson, B.M. & Barbeau, K. a., 2012. Iron transporters in marine prokaryotic genomes and metagenomes. Environmental Microbiology, 14, pp.114-128.
Hopkinson, B.M. & Morel, François M M, 2009. The role of siderophores in iron acquisition by photosynthetic marine microorganisms. Biometals, 22(4), pp.659-69.
Hopkinson, B.M., Roe, K.L. & Barbeau, Katherine a, 2008. Heme uptake by Microscilla marina and evidence for heme uptake systems in the genomes of diverse marine bacteria. Applied and environmental microbiology, 74(20), pp.6263-70.
Horstmann, U. & Soria-Dengg, S., 1995. Ferrioxamines B and E as iron sources for the marine diatom Phaeodactylum tricornutum. Marine Ecology Progress Series, 127, pp.269-277.
Hudson, R.J.M. & Morel, Francois M M, 1990. Iron Transport in Marine Phytoplankton: Kinetics of Cellular and Medium Coordination Reactions. Limnology and Oceanography, 35(5), pp.1002-1020.
Hudson, R.J.M. & Morel, François M. M., 1993. Trace metal transport by marine microorganisms: implications of metal coordination kinetics. Deep Sea Research Part I: Oceanographic Research Papers, 40(1), pp.129-150.
Hutchins, David A et al., 1999. Competition among marine phytoplankton for different chelated iron species. Nature, 400, pp.1-4.
Ibisanmi, E. et al., 2011. Vertical distributions of iron-(III) complexing ligands in the Southern Ocean. Deep Sea Research Part II: Topical Studies in Oceanography, 58(21-22), pp.2113-2125.
Ito, Y. & Butler, A, 2005. Structure of synechobactins, new siderophores of the marine cyanobacterium, Synechococcus sp. PCC 7002. Limnology and Oceanography, 50, pp.1918-1923.
Ito, Y., Okada, S. & Murakami, M., 2001. Two structural isomeric siderophores from the freshwater cyanobacterium Anabaena cylindrica. Tetrahedron, 57, pp.9093-9099.
Jeanjean, R. et al., 2008. A large gene cluster encoding peptide synthetases and polyketide synthases is involved in production of siderohores and oxidative stress response in the cyanobacterium Anabaena sp. strain PCC 7120. Environmental microbiology, 10(10), pp.2574-2585.
Jickells, T.D. et al., 2005. Global iron connections between desert dust, ocean biogeochemistry, and climate. Science (New York, N.Y.), 308(5718), pp.67-71.

Johnson, K.S. et al., 1994. Iron photochemistry in seawater from the equatorial Pacific. Marine Chemistry, 46, pp.319-334.
Johnson, K.S., Gordon, R.M. & Coale, K.H., 1997. What controls dissolved iron concentrations in the world ocean? Marine Chemistry, 57(3-4), pp.137-161.
Kanoh, K. et al., 2003. Pseudoalterobactin A and B, New Siderophores Excreted by Marine Bacterium Pseudoalteromonas sp. KP20-4. Journal of Antibiotics, 56(10), pp.871-875.
Karl, D. & Letelier, R., 2008. Nitrogen fixation-enhanced carbon sequestration in low nitrate, low chlorophyll seascapes. Marine Ecology Progress Series, 364, pp.257-268.
Kustka, A.B., Allen, A.E. & Morel, François M. M., 2007. Sequence Analysis and Transcriptional Regulation of Iron Acquisition Genes in Two Marine Diatoms 1. Journal of Phycology, 43(4), pp.715-729.
Küpper, F.C. et al., 2006. Photoreactivity of iron(III)-aerobactin: photoproduct structure and iron(III) coordination. Inorganic chemistry, 45(15), pp.6028-33.
Lam, P.J. et al., 2006. Wintertime phytoplankton bloom in the subarctic Pacific supported by continental margin iron. Global Biogeochemical Cycles, 20(1), pp.1-12.
Lam, P.J. & Bishop, J.K.B., 2008. The continental margin is a key source of iron to the HNLC North Pacific Ocean. Geophysical Research Letters, 35(7), pp.1-5.
Laroche, J. et al., 1996. Flavodoxin as an in situ marker for iron stress in phytoplankton. Nature, 382, pp. 802-805.
Luo, M., Fadeev, E. a & Groves, J.T., 2005. Membrane dynamics of the amphiphilic siderophore, acinetoferrin. Journal of the American Chemical Society, 127(6), pp.1726-36.
Macrellis, H.M. et al., 2001. Collection and detection of natural iron-binding ligands from seawater. Marine Chemistry, 76(3), pp.175-187.
Mahowald, N.M. et al., 2005. Atmospheric global dust cycle and iron inputs to the ocean. Global Biogeochemical Cycles, 19(4).
Maldonado, M.T. & Price, Neil M, 2000. Nitrate regulation of Fe reduction and transport by Fe-limited Thalassiosira oceanica. Limnology and Oceanography, 45(3), pp.814-826.
Maldonado, M.T. & Price, Neil M, 2001. Reduction and Transport of organically bound iron by Thalassiosira oceanica (Bacillariophyceae). Journal of Phycology, 309, pp.298-309.
Maldonado, M.T. & Price, Neil M, 1999. Utilization of iron bound to strong organic ligands by plankton communities in the subarctic Pacific Ocean. Deep Sea Research Part II: Topical Studies in Oceanography, 46, pp.2447-2473.
Mann, E.L. & Chisholm, S.W., 2000. Iron limits the cell division rate of Prochlorococcus in the eastern equatorial Pacific. Limnology and Oceanography, 45(5), pp.1067-1076.
Marshall, J.-A. et al., 2005. Superoxide production by marine microalgae. Marine Biology, 147(2), pp. 533-540.
Martin, J.D. et al., 2006. Structure and membrane affinity of new amphiphilic siderophores produced by Ochrobactrum sp. SP18. Journal of biological inorganic chemistry: JBIC: a publication of the Society of Biological Inorganic Chemistry, 11(5), pp.633-41.
Martin, JH, 1990. Glacial-Interglacial CO2 Change: The Iron Hypothesis. Paleoceanography, 5(1), pp.1-13.
Martin, John et al., 1994. Testing the iron hypothesis in ecosystems of the equatorial Pacific Ocean. , pp. 123-129.
Martinez, J. S., 2000. Self-Assembling Amphiphilic Siderophores from Marine Bacteria. Science, 287(5456), pp.1245-1247.
Martinez, Jennifer S et al., 2003. Structure and membrane affinity of a suite of amphiphilic siderophores produced by a marine bacterium. Proceedings of the National Academy of Sciences of the United States of America, 100(7), pp.3754-9.
Martinez, Jennifer S & Butler, Alison, 2007. Marine amphiphilic siderophores: marinobactin structure, uptake, and microbial partitioning. Journal of inorganic biochemistry, 101(11-12), pp.1692-8.
Mawji, E. et al., 2008. Hydroxamate siderophores: occurrence and importance in the Atlantic Ocean. Environmental science & technology, 42(23), pp.8675-80.
Mills, M.M. et al., 2004. Iron and phosphorus co-limit nitrogen fixation in the eastern tropical North Atlantic. Nature, 429(6989), pp.292-4.

Moore, M. et al., 2009. Large-scale distribution of Atlantic nitrogen fixation controlled by iron availability. Nature Geoscience, 2(12), pp.867-871.
Morel, F M M & Price, N M, 2003. The biogeochemical cycles of trace metals in the oceans. Science (New York, N.Y.), 300(5621), pp.944-7.
Morel, Francois, Kustka, A, Shaked, Y., 2008. The role of unchelated Fe in the iron nutrition of phytoplankton. Limnology and Oceanography, 53(1), pp.400-404.
Morrissey, J. & Bowler, C., 2012. Iron utilization in marine cyanobacteria and eukaryotic algae. Frontiers in microbiology, 3(March), p.43.
Noble, A.E. et al., 2008. Cobalt, manganese, and iron near the Hawaiian Islands: A potential concentrating mechanism for cobalt within a cyclonic eddy and implications for the hybrid-type trace metals. Deep Sea Research Part II: Topical Studies in Oceanography, 55(10-13), pp.1473-1490.
Nolting, R.F. et al., 1998. Fe(III) speciation in the high nutrient , low chlorophyll Pacific region of the Southern Ocean. Marine Chemistry, 62, pp.335-352.
Okujo, N. & Yamamoto, S, 1994. Identification of the siderophores from Vibrio hollisae and Vibrio mimicus as aerobactin. FEMS microbiology letters, 118(1-2), pp.187-92.
Oquist, G., 1971. Changes in Pigment Composition and Photosynthesis Induced by Iron-Deficiency in the Blue-Green Alga Anacystis nidulans. Physiologia Plantarum, 25(2), pp.188-191.
Oquist, G., 1974. Iron Deficiency in the Blue-Green Alga Anacystis nidulans: Changes in Pigmentation and Photosynthesis. Physiologia Plantarum, 30(1), pp.30-37.
Pakulski, J. et al., 1996. Iron stimulation of Antarctic bacteria. Nature, 383, pp.133-134.
Pierre, J.L. & Gautier-Luneau, I., 2000. Iron and citric acid: a fuzzy chemistry of ubiquitous biological relevance. Biometals, 13(1), pp.91-6.
Powell, R.T. & Donat, J.R., 2001. Organic complexation and speciation of iron in the South and Equatorial Atlantic. Deep Sea Research Part II: Topical Studies in Oceanography, 48(13), pp.2877-2893.
Raven, J. a., 1988. The iron and molybdenum use efficiencies of plant growth with different energy, carbon and nitrogen sources. New Phytologist, 109(3), pp.279-287.
Rees, D.C. et al., 2005. Structural basis of biological nitrogen fixation. Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, 363(1829), pp.971-84; discussion 1035-40.
Reid, R.T. et al., 1993. A siderophore from a marine bacterium with an exceptional ferric ion affinity constant. Nature, 366, pp.455-458.
Reid, R.T. & Butler, Alison, 1991. Investigation of the mechanism of iron acquisition by the marine bacterium Alteromonas luteoviolaceus: Characterization of siderophore production. Limnology and Oceanography, 36, pp.1783-1792.
Richier, S. et al., 2012. Abundances of Iron-Binding Photosynthetic and Nitrogen-Fixing Proteins of Trichodesmium Both in Culture and In Situ from the North Atlantic L. J. Stal, ed. PLoS ONE, 7(5), p.e35571.
Rose, A.L. et al., 2005. Use of superoxide as an electron shuttle for iron acquisition by the marine cyanobacterium Lyngbya majuscula. Environmental science & technology, 39(10), pp.3708-15.
Rose, A.L. & Waite, T.D., 2005. Reduction of organically complexed ferric iron by superoxide in a simulated natural water. Environmental science & technology, 39(8), pp.2645-50.
Rue, E.L. & Bruland, Kenneth W, 1995. Complexation of iron (III) by natural organic ligands in the Central North Pacific as determined by a new competitive ligand equilibration / adsorptive cathodic stripping voltammetric method. Marine Chemistry, 50, pp.117-138.
Rue, E.L. & Bruland, Kenneth W, 1997. The role of organic complexation on ambient iron chemistry in the equatorial Pacific Ocean and the response of a mesoscale iron addition experiment. Limnology and Oceanography, 42, pp.901-910.
Rusch, D.B. et al., 2010. Characterization of Prochlorococcus clades from iron-depleted oceanic regions. PNAS.
Saito, M.A. et al., 2010. Iron conservation by reduction of metalloenzyme inventories in the marine diazotroph Crocosphaera watsonii. Proceedings of the National Academy of Sciences.
Saito, M.A. & Moffett, J.W., 2002. Temporal and spatial variability of cobalt in the Atlantic Ocean. Geochimica et Cosmochimica Acta, 66(11), pp.1943-1953.
Salmon, T.P. et al., 2006. The FeL model of iron acquisition: Nondissociative reduction of ferric complexes in the marine environment. Limnology and Oceanography, 51(4), pp.1744-1754.

Sancho, J., 2006. Flavodoxins: sequence, folding, binding, function and beyond. Cellular and molecular life sciences: CMLS, 63(7-8), pp.855-64.
Sandy, M. et al., 2010. Vanchrobactin and anguibactin siderophores produced by Vibrio sp. DS40M4. Journal of natural products, 73(6), pp.1038-43.
Schmidt, M.A. & Hutchins, D A, 1999. Size-fractionated biological iron and carbon uptake along a coastal to offshore transect in the NE Pacific. Deep Sea Research Part II: Topical Studies in Oceanography, 46, pp.2487-2503.
Sedwick, P.N, McMurtry, G.. & Macdougall, J.., 1992. Chemistry of hydrothermal solutions from Pele's Vents, Loihi Seamount, Hawaii. Geochimica et Cosmochimica Acta, 56(10), pp.3643-3667.
Sedwick, P. N., 2005. Iron in the Sargasso Sea (Bermuda Atlantic Time-series Study region) during summer: Eolian imprint, spatiotemporal variability, and ecological implications. Global Biogeochemical Cycles, 19(4).
Shaked, Y., Kustka, A.B. & Morel, François M. M., 2005. A general kinetic model for iron acquisition by eukaryotic phytoplankton. Limnology and Oceanography, 50(3), pp.872-882.
Shaked, Y. & Lis, H., 2012. Disassembling Iron Availability to Phytoplankton. Frontiers in Microbiology, 3(April), pp.1-26.
Simpson, F. & Neilands, J., 1976. Siderochromes in cyanophyceae: Isolation and Characterization of Schizokinen from Anabaena sp. Journal of Phycology, 12(1), pp.44-48.
Smith, D.P. et al., 2010. Transcriptional and translational regulatory responses to iron limitation in the globally distributed marine bacterium Candidatus pelagibacter ubique. PloS one, 5(5), p.e10487.
Snow, G. a & White, a J., 1969. Chemical and biological properties of mycobactins isolated from various mycobacteria. The Biochemical journal, 115(5), pp.1031-50.
Soengas, R.G. et al., 2006. Structural characterization of vanchrobactin, a new catechol siderophore produced by the fish pathogen Vibrio anguillarum serotype O2. Tetrahedron Letters, 47(39), pp. 7113-7116.
Soria-Dengg, S., Reissbrodt, R. & Horstmann, U., 2001. Siderophores in marine coastal waters and their relevance for iron uptake by phytoplankton: experiments with the diatom Phaeodactylum tricornutum. Marine Ecology Progress Series, 220, pp.73-82.
Souto, A. et al., 2012. Structure and Biosynthetic Assembly of Piscibactin, a Siderophore from Photobacterium damselae subsp. piscicida , Predicted from Genome Analysis. European Journal of Organic Chemistry, 2012(29), pp.5693-5700.
Sunda, W.G. & Huntsman, S.A., 1995. Iron uptake and growth limitation in oceanic and coastal phytoplankton. Marine Chemistry, 50, pp.189-206.
Takahashi, A. et al., 1987. Bisucaberin, a New Siderophore, Sensitizing Tumor Cells to Macrophage-Mediated Cytolysis. Journal of Antibiotics, 40, pp.1671-1676.
Tilman, D., Kilham, S.S. & Kilham, P., 1982. Phytoplankton Community Ecology: The Role of Limiting Nutrients. Annual Review of Ecology and Systematics, 13(1), pp.349-372.
Tortell, P. et al., 1999. Marine bacteria and biogeochemical cycling of iron in the oceans. FEMS Microbiology Ecology, 29(1), pp.1-11.
Tortell, P., Maldonado, M.T. & Price, Neil M, 1996. The role of heterotrophic bacteria in iron-limited ocean ecosystems. Nature, 383, pp.330-332.
Toulza, E. et al., 2012. Analysis of the Global Ocean Sampling (GOS) Project for Trends in Iron Uptake by Surface Ocean Microbes F. Rodriguez-Valera, ed. PLoS ONE, 7(2), p.e30931.
Tsuda, A. et al., 2005. Responses of diatoms to iron-enrichment (SEEDS) in the western subarctic Pacific, temporal and spatial comparisons. Progress In Oceanography, 64(2-4), pp.189-205.
Tyrrell, T. et al., 2003. Large-scale latitudinal distribution of Trichodesmium spp . in the Atlantic Ocean. Journal of Plankton Research, 25(4), pp.405-416.
Vala, A., Vaidya, S. & Dube, H., 2000. Siderophore production by facultative marine fungi. Indian Journal of Marine Sciences, 29, pp.339-340.
Vala, A.K., Dave, B.P. & Dube, H.C., 2006. Chemical characterization and quantification of siderophores produced by marine and terrestrial aspergilli. Canadian Journal of Microbiology, 52, pp.603-607.
Volker, C. & Wolf-gladrow, D.A., 1999. Physical limits on iron uptake mediated by siderophores or surface reductases. Marine Chemistry, 65, pp.227-244.

Vraspir, J.M. & Butler, Alison, 2009. Chemistry of Marine Ligands and Siderophores. Annual Review of Marine Science, 1, pp.43-63.
Vraspir, J.M., Holt, P.D. & Butler, Alison, 2011. Identification of new members within suites of amphiphilic marine siderophores. Biometals, 24(1), pp.85-92.
Wang, W. et al., 2009. Chemical and biological characterization of siderophore produced by the marine-derived Aureobasidium pullulans HN6.2 and its antibacterial activity. Biometals, 22, pp.965-972.
Wilhelm, S.W. & Trick, C.G., 1994. Iron-limited growth of cyanobacteria: production is a common response Multiple siderophore. Limnology and Oceanography, 39(8), pp.1979-1984.
Witter, A.E., Lewis, B.L. & Luther III, G.W., 2000. Iron speciation in the Arabian Sea. Deep Sea Research Part II: Topical Studies in Oceanography, 47(7-8), pp.1517-1539.
Witter, A.E. & Luther, G.W., 1998. Variation in Fe-organic complexation with depth in the Northwestern Atlantic Ocean as determined using a kinetic approach. Marine Chemistry, 62, pp.241-258.
Yamamoto, Shigeo et al., 1994. Structure and Iron Transport Activity of Vibrioferrin, a New Siderophore of Vibrio parahaemolyticus. Journal of Biochemistry, 115, pp.868-874.
Yamamoto, Shigeo et al., 1993. Structures of 2 Polyamine Containing Catecholate Siderophores From Vibrio Fluvialis. Journal of Biochemistry, 113(5), pp.538-544.

Chapter 2

Siderophore Production in the *Vibrionaceae*

2.1. Abstract

Iron is scarce in many environments, including the ocean. Under iron limiting conditions, pathogenic and terrestrial microorganisms excrete small peptides, siderophores, to acquire iron from their environment. In the ocean, it remains unclear whether siderophores are a widely used strategy for iron acquisition. Indeed, free-living microorganisms may have less of an incentive than particle-associated ones to produce energetically expensive siderophores, which upon excretion may be lost in the dilute environment. Here we ask whether siderophore production amongst a group of closely related *Vibrionaceae* isolates obtained from a variety of micro-niches is linked to habitat. A large phenotypic screen of ~1,700 isolates revealed that siderophore production is a patchy trait across the *Vibrionaceae* phylogeny and positively correlated with a free-living lifestyle. Analysis of 69 genomes showed that in some cases siderophore synthesis genes had specifically been excised from the genomes of non-producers, while corresponding transport genes had been retained. We hypothesized and experimentally confirmed that these non-producers were *'cheaters'* and could effectively obtain their iron by using the siderophores produced by closely related producing strains. These findings suggest that for the group of isolates investigated, interactions between producers and non-producers might be a greater determinant than kinetics in predicting siderophore production: on particles, where encounter rates between microorganisms are greater, there is more opportunity for 'cheating' in benefitting from the siderophore produced by a close relative. This study contributes to our understanding of iron acquisition in the marine environment, suggesting that the structure of the marine micro-environment influences siderophore production in a predictable way, with ecological factors, such as encounter frequency determining the frequency of non-producers. Our results further provide evidence that the incentive to cheat in microbial interactions can drive genomic diversity.

With Otto X. Cordero[1], Edward F. DeLong[1,2], Martin F. Polz[1].
[1]Department of Civil and Environmental Engineering, MIT, Cambridge, MA, USA
[2]Department of Biological Engineering, MIT, Cambridge, MA, USA

The published manuscript can be found in Section 2.9.4:

Cordero, O.X.*, Ventouras, L.A.*, DeLong, E.F., Polz, M.F. (2012) Public good dynamics drive evolution of iron acquisition strategies in natural bacterioplankton populations. *Proc. Natl. Acad. Sci. USA,* doi:10.1073/pnas.1213344109

2.2. Background

2.2.1 Motivation

All microorganisms (save for some *Borrelia*) require iron for growth. However, at circumneutral pH, the dominant iron species consists of insoluble Fe(III)-oxyhydroxides. Hence, unlike for other essential metal ions, Fe(III) first needs to be solubilized before being imported by the cell. To this end, microorganisms secrete small molecular weight (<1.5kDa) ferric-iron specific chelators called siderophores (Wandersman & Delepelaire 2004). In order to deal with iron scarcity in the marine environment, a number of marine microorganisms (marine fungi, bacteria as well as some archaea) are known to produce siderophores, see Chapter 1, Table 1 (Vraspir & Butler 2009; Sandy & Butler 2009). Yet, the usefulness of siderophores as a widely utilized strategy to acquire iron in the marine environment is regularly being questioned: the most abundant marine bacteria *Pelagibacter ubique* (Giovannoni et al. 2005) and *Prochlorococcus marinus* (Rocap et al. 2003) carry no siderophore biosynthetic gene cluster. Indeed, why would microorganisms produce a costly molecule if the chances of losing it by diffusion to the surrounding dilute marine environment are great? A model comparing the kinetics of siderophore diffusion and uptake with the kinetics of surface-bound iron reductases, suggests that siderophore production is thermodynamically beneficial only under specific conditions: when cells have reached a high enough density to ensure that the ambient siderophore concentration balances out the chances of loss by diffusion (Voelker & Wolf-Gladrow 1999). This condition is likely to be fulfilled at hotspots of microbial activity such as during blooms of a specific microbial species, around plumes of dissolved organic matter, or on particulate matter that is colonized by microbial biofilms (Azam & Malfatti 2007). Here, using a collection of *Vibrionaceae* isolates collected from various micro-niches, we explore whether siderophore production is in fact correlated with specific habitats.

2.2.2 The Vibrionaceae: a model system for the study of siderophore production in a natural environment

The *Vibrionaceae* is a family of usually motile, gram-negative γ-proteobacteria endemic to aquatic systems. These heterotrophs are ubiquitous across marine environments from the open-ocean (DeLong et

al. 2006) to freshwater ecosystems, estuaries and coastal zones as well as the deep-sea (Thompson et al. 2004). Usually representing less than 1% of the bacterioplankton community (Eilers et al. 2000; Suzuki et al. 2004; DeLong et al. 2006; Rusch et al. 2007; Feingersch et al. 2010; Gilbert et al. 2010), they can nonetheless thrive in many of the distinct niches that are characteristic of the marine micro-environment: free-living in the water column, in association with algae (Thompson et al. 2004; Reen et al. 2006), with zooplankton (Huq et al. 1983), with chitin-rich marine snow particles (Huq et al. 1983), or as mutualists (Nyholm & McFall-Ngai 2004) or pathogens of a number of eukaryotic hosts (Thompson et al. 2004). Some vibrio species are not specialized to a single niche, but instead display a variety of lifestyles: *V. fischeri* spends part of its life as a free-living microorganism, and the other part as a symbiont of the squid *Euprymna scolopes* (Ruby 1998). Because vibrios are found in such different niches, they provide for an interesting model system to explore whether siderophore production in the marine environment may be linked to habitat.

In addition to inhabiting diverse micro-niches, the *Vibrionaceae* are known to produce a variety of siderophores. Because of the significance of vibrio-mediated infections both in humans and in seafood, and the connection between siderophore production and virulence, most studies of siderophore production have focused on the pathogenic members of the *Vibrionaceae* (Table 2.1). These systems have been studied in detail, from the determination of the chemical structure of the siderophore (Griffiths et al. 1984; Actis et al. 1986; Yamamoto, et al. 1994; Soengas et al. 2006) to the description of plausible biosynthetic pathways and identification of the genes responsible for both synthesis and transport (Crosa & Walsh 2002; Tanabe, et al. 2003a; Wang et al. 2007; Balado et al. 2006). More recent efforts targeted at identifying siderophores produced by marine bacteria, have led to the discovery and structural elucidation of a number of novel siderophores (Table 1.1). Because members of the *Vibrionaceae* are easily brought into culture (Eilers et al. 2000), quite a few of these newly-characterized marine siderophores are in fact produced by vibrio isolates, which may or may not be pathogenic. Together, these efforts reveal that the members of the *Vibrionaceae* family produce a very diverse set of siderophores, with either hydroxamate, catecholate (and phenolate) or carboxyl iron-chelating moeities (Figure 2.1). They also suggest that siderophores are produced by a variety of species, not necessarily restricted to one branch of the *Vibrionaceae* phylogeny. Apart for the

siderophores produced by pathogens associated with specific hosts, these studies lack information regarding the distinct micro-niche of the siderophore producing vibrio strain. These studies are also heavily biased towards the identification of siderophores produced by pathogenic vibrios. To address these two issues, we chose to screen a set of *Vibrionaaceae* isolates obtained from a wide variety of micro-niches, including both pathogenic and non-pathogenic strains and for which a distinct habitat association was available.

Table 2.1: Literature review of siderophore production by the *Vibrionaceae*.

Organism	Siderophore Name	Siderophore type	Ecology	References
Vibrio anguillarum 775(pJM1)	Anguibactin	catechol	Fish pathogen, responsible for vibriosis.	(Crosa 1980; Actis et al. 1986)
	Vanchrobactin	catechol	Anguibactin is plasmid encoded and increases virulence	(Soengas et al. 2006)
Vibrio alginolyticus	Vibrioferrin	hydroxy-carboxylate		(Wang et al. 2007; Wang et al. 2008)
Vibrio parahaemolyticus	Vibrioferrin	hydroxy-carboxylate.	Associated with oysters, causes gastroenteritis in humans.	(Yamamoto et al. 1992; Yamamoto et al. 1994; Yamamoto al. 1994a; Tanabe et al. 2003)
Vibrio cholerae	vibriobactin	catechol	Human pathogen. Siderophore production is not linked to increased virulence.	(Griffiths et al. 1984; Sigel et al. 1985)
Vibrio vulnificus	vulnibactin	hydroxamate and catechol	Eel pathogen, opportunistic human pathogen. Vulnibactin production is linked to virulence	(Wright et al. 1981; Noriyuki Okujo et al. 1996; Biosca et al. 1996)
Vibrio hollisae	aerobactin	hydroxamate	Human pathogen	(Okujo et al 1994)
Vibrio mimicus	aerobactin	hydroxamate	Human pathogen	(Okujo et al 1994)
Vibrio salmonicida	bisucaberin	hydroxamate	Fish pathogen	
Vibrio splendidus	aerobactin	hydroxamate	Aerobactin was also produced by close relatives *V. orientalis, V. nereis, V. tubiashi*	(Murakami et al. 1998)
V. fluvialis, V.furnissii	fluvibactin.	cathecol	Enteric pathogen	(Shigeo Yamamoto et al. 1993)
Vibrio sp. S4BW	ochrobactin-OH	amphiphilic, photoreactive (citrate-based)	Isolated from the Gulf of Mexico, after the Deepwater Horizon oil spill	(Gauglitz et al. 2012)
Vibrio sp. HC0601C5	Amphibactin	amphiphilic, secreted into culture medium	Isolated from the Santa Barbara basin	(Vraspir et al. 2011)
Vibrio sp. DS40M4	vanchrobactin anguibactin		Isolated from an open ocean water sample, on the continental slope between Cape Verde and the Canary Islands	(Sandy et al. 2010)
Vibrio sp. R10	Amphibactin	amphiphilic, not secreted into culture medium	Isolated from near-shore waters at Roatan, Honduras	(Martinez et al. 2003)
Vibrio sp. DS40M5	ochrobactin aerobactin	amphiphilic, photoreactive	Isolated from an open ocean water sample, on the continental slope between Cape Verde and the Canary Islands	(Haygood et al. 1993; Martin et al. 2006)

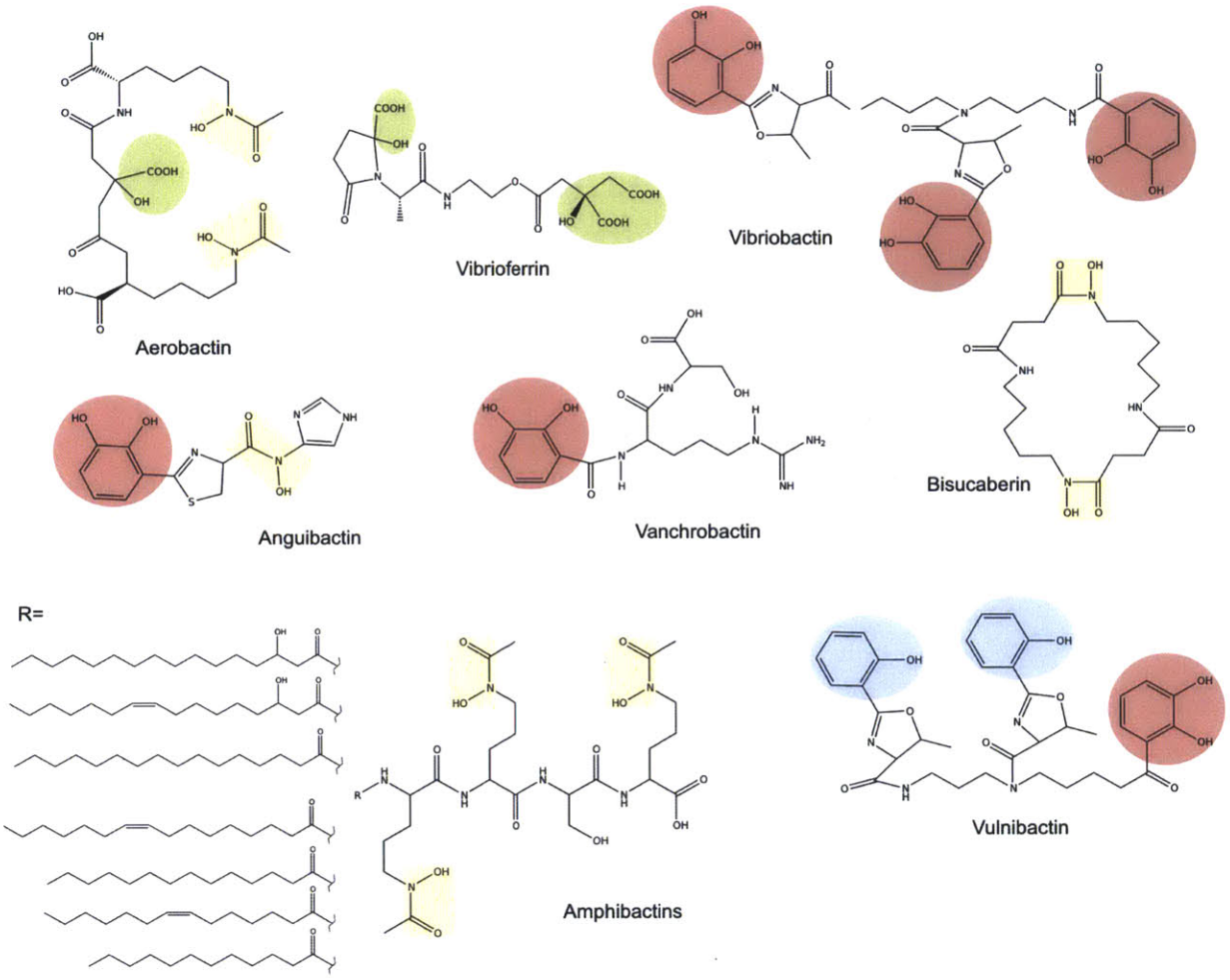


Figure 2.1: Diversity of siderophores produced by the *Vibrionaceae.* Examples of siderophores produced by some vibrio isolates. Iron-chelating groups are highlighted: hydroxamate moieties in yellow, catecholate in red, carboxylate in green and phenolate in blue.

The *Vibrionaceae* strains investigated in this study were originally isolated from a variety of micro-niches in the Northwestern Atlantic coastal environment and phylogenetically characterized as part of two previous studies (Hunt et al. 2008; Preheim et al. 2011). These isolates are organized in phylogenetic clusters (determined by the highly conserved *hsp60* gene) of closely related but non-clonal genotypes. An unsupervised machine learning approach was used to determine the propensity of each isolate to associate with a specific sampled environment (Hunt et al. 2008). Because phylogenetic groups co-clustered with similar habitat, these were hypothesized to constitute samples of wild microbial, ecologically cohesive populations (Hunt et al. 2008). Hence, the *Vibrionaceae* isolates investigated in this study belong to different populations, for a which a clear resource preference (micro-habitat) is known, including: zooplankton exoskeleton, macro-algae, gut of invertebrates (Preheim, et al. 2011; Preheim 2010), as well as four distinct filter-size fractions each corresponding to a different lifestyle (Hunt et al. 2008). The smallest fraction (<1μm) encompasses free-living cells which are likely to exploit dissolved nutrients or highly labile aggregates, whereas the largest size-fraction (>64μm) includes isolates that are likely to live in association with chitinous or algal particles where bacteria can attach and form biofilms.

Taking advantage of the fact that each isolate is assigned to a distinct population, we can specifically ask: **(1) What is the pattern of siderophore production amongst a group of closely related *Vibrionaceae*? (2) Is this pattern correlated to population habitat? (3) What is the diversity of siderophore produced by the *Vibrionaceae*? (4) Can we find evidence of microbe-microbe interactions in these samples of wild isolates?**

2.3. Materials and Methods

2.3.1. Siderophore screen of wild isolates

We screened a total of 1,710 *Vibrionaceae* isolates obtained from seawater collected at the Plum Island Estuary, MA as part of two previous studies. In study 1, isolates were obtained from size-fractionated seawater samples taken in the fall and in the spring (Hunt et al. 2008). In study 2, isolates were obtained from different body parts of invertebrates and from macro-algae found in seawater (Preheim et al. 2011). Roughly 56% of the isolates we screened for siderophore production originated from study 1. The isolates were chosen such that each ecologically cohesive population previously identified as part of these two studies was represented. Siderophore production was assessed using the well-established Chrome-Azurol S (CAS) assay (Schwyn & Neilands 1987). Biological replicate measurements for both the liquid and solid versions of the assay were performed in high-throughput, using either 96-well 2mL culture blocks (Corning, Corning, NY, USA) or 96-well microtiter plates for absorbance measurements (Corning).

For the solid version of the assay, strains were grown in duplicates overnight in 2216 marine broth (Difco, Franklin Lakes NJ, USA) at room temperature and stamped directly onto CAS agar plates. When tightly bound to Fe(III), the Chrome Azurol S dye present in the CAS agar plates is blue. When the CAS dye is not bound to Fe(III) it turns to orange (Schwyn & Neilands 1987). Hence, siderophore production could be assessed visually as colonies that produce a siderophore are surrounded by an orange halo. We scored siderophore production based on the size of the halo as "+++", "++", "+", "+-" or "-".

CAS agar plates were prepared by mixing the CAS-Fe-HDTMA dye with growth media appropriate for *Vibrionaceae*. For 1 L of CAS-agar, 100 mL of CAS-Fe-HDTMA dye was mixed with 900 mL of freshly prepared growth media. The CAS-Fe-HDTMA dye was prepared in advance as follows, for 1L: 10 mL of a 10 mM ferric chloride ($FeCl_3$) in 100 mM hydrochloric acid (HCl) solution was mixed with 590 mL of a 1 mM aqueous solution of chrome azurol S (CAS). The Fe-CAS solution was then added to 400 mL of a 2 mM aqueous solution of hexadecyl-trimethyl-ammonium bromide (HDTMA). The resulting CAS-Fe-HDTMA solution was autoclaved for 25 min in a polycarbonate bottle that had previously been

soaked overnight in 10% HCl then rinsed 5 times with MilliQ water. The CAS-Fe-HDTMA dye was stored at room temperature covered from light until usage.

The growth media was prepared as follows, for 1 L of CAS-agar: 30.24g of 1,4-piperazine-diethanesulfonic acid (PIPES), together with 1 g of ammonium chloride (NH_4Cl), 3 g potassium phosphate (KH_2PO_4), 20 g sodium chloride (NaCl), was dissolved into MilliQ water by adjusting the pH with 10 M NaOH to 6.8. Note that the commonly used phosphate buffer disodium phosphate ($Na_2HPO_4 \cdot 7H_2O$) was omitted as phosphate can chelate iron and lead to a discoloration of the CAS dye. As a solidifying agent, 9 g of agar noble (Difco) were added to the solution. We found that the more commonly used solidifying agents, agarose and agar also led to a discoloration of the CAS dye, likely due to higher phosphate content. The volume was adjusted to 860 mL and the solution was autoclaved. After cooling, 30 mL of a sterile 10% (w/v) Casamino acids (Difco) aqueous solution and 10 mL of a sterile 20% (w/v) glucose aqueous solutions were added. Finally, the 100 mL of CAS-Fe-HDTMA were added to the growth media. The final concentrations of the CAS-agar components are: 100 mM PIPES, 18 mM NH_4Cl, 22 mM KH_2PO_4, 2% (w/v) NaCl, 0.3% casamino acids, 0.2 % glucose, 10 μM $FeCl_3$, 58 μM CAS, 80 μM HDTMA. Unless otherwise noted, all chemicals were obtained from Sigma-Aldrich (St. Louis, MO, USA). The detailed protocol can be found in Appendix B1 of this thesis.

For the liquid version of the assay, strains were first grown overnight in duplicates in 2216 marine broth (Difco) at room temperature, then transferred into the *Vibrionaceae* growth media described above (without the solidifying agent nor the CAS-Fe-HDTMA dye). After overnight growth at room temperature, strains were further transferred into iron-poor media, consisting of the *Vibrionaceae* growth media described above amended with 100 μM 2,2'-bipyridyl (Sigma-Aldrich) and 10 nM $FeCl_3$ to induce siderophore production. Cultures were centrifuged and 99 μL of supernatant were mixed in a 1:1 ratio with liquid CAS dye and 2 μL shuttle solution (Schwyn & Neilands 1987). The mixture was incubated in the dark for 15 min, after which absorbance at 630 nm was measured on a Synergy2 filter-based multi-mode plate reader (Biotek, Winooski, VT). In the presence of a siderophore, the absorbance of the dye at 630 nm is

quenched. The liquid CAS dye and shuttle solution were prepared as previously described (Schwyn & Neilands 1987).

The outcome of both the liquid and the solid versions of the CAS assay were taken into consideration when establishing whether a strain produced a siderophore or not. To combine the results from both screens, only the strains that grew –either in the iron poor media (with an absorbance value at 600 nm greater than 0.15) or in the solid CAS assay– were considered. Production of a siderophore was established when at least two out of four (biological duplicates for each the liquid and the solid version of the assay) measurements were positive. Siderophore production was considered positive for all absorbances at 630 nm < 0.3 and for all solid CAS assay outcomes rated as "+++", "++", "+", or "+-". The raw results and the production data (resolved between liquid and solid versions of the assay) for the 1,012 strains that grew can be found in the supplementary material of this chapter, Section 2.9.1.

2.3.2. Bioinformatic Analysis of Siderophore Synthesis and Transport Gene Clusters

Draft genomes were available for 69 of the 1,710 isolates screened. Gene finding and annotations had previously been performed using the SEED subsystems (Overbeek et al. 2005) and the RAST server (Aziz et al. 2008). Siderophore biosynthesis clusters were identified using a combination of annotation text searches and the software AntiSMASH developed to identify biosynthetic clusters of secondary metabolites (Medema et al. 2011). AntiSMASH was run on all the contigs for all 69 genomes using default parameters. For the purpose of this study, we focused on the 'nrps' and 'siderophore' metabolite clusters identified by AntiSMASH. These matched well with the annotation text searches that were concurrently performed. Text searches were performed for the following expressions: siderophore, actin (encompassing many of the names of known siderophores), ferrin, pyoverdin, chelin, hydroxamate, catechol. Genes in the detected siderophore synthesis clusters were further characterized by searching for conserved domains in PFAM Protein Families Database (Punta et al. 2012). The specificity of adenylation domains identified by PFAM was determined using NRPS predictor 2 (Röttig et al. 2011). The different types of siderophores produced were inferred by manual inspection of the siderophore synthesis clusters, taking into account the specificity

of neighboring transporters, the types of adenylation-domain substrates and BLAST homology searches against NCBI's non-redundant (nr) protein database (Altschul et al. 1997). Table S2.2, section 2.9.2 in supplementary materials of this chapter displays the AntiSMASH, PFAM and NRPS Predictor 2 search results. Figure 2.2 summarizes this bioinformatic approach.

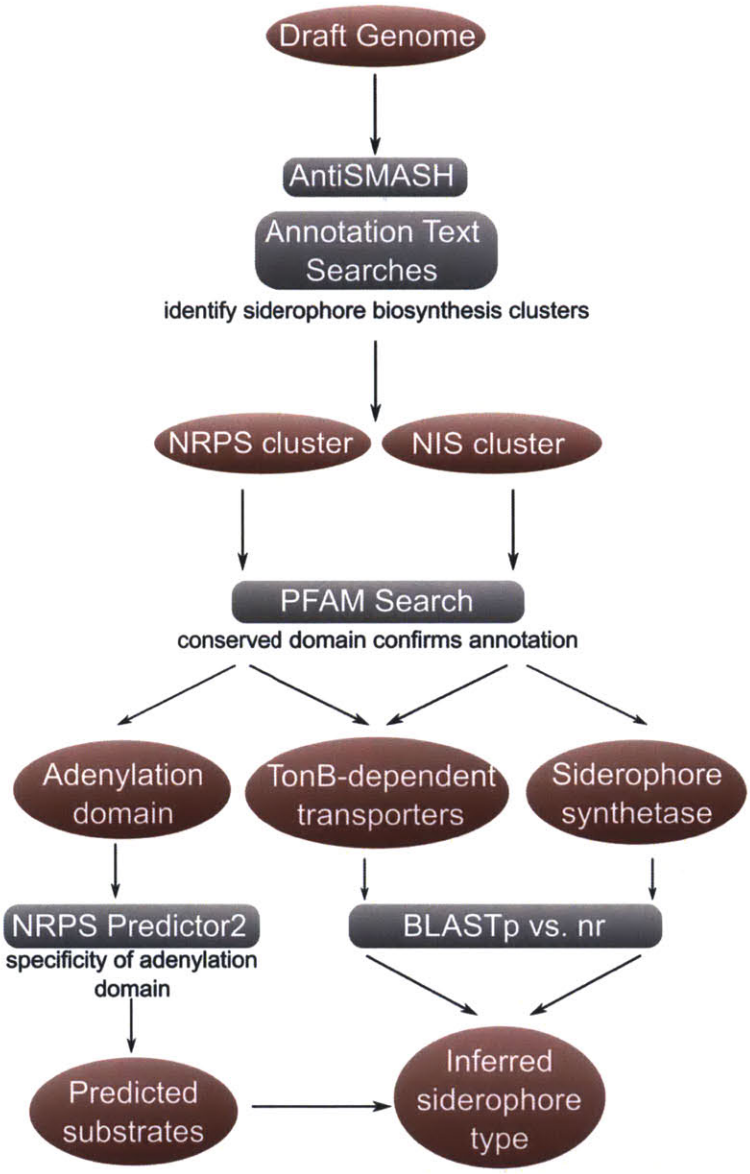


Figure 2.2. Bioinformatic pipeline used to infer siderophore type. Bioinformatic pipeline used to infer the type of siderophore produced from the available draft genomes. Bioinformatic software used in the analysis is highlighted in grey and references to each software appear in the text.

To map the presence or absence of different iron-transporters across the 69 genomes, the following approach was used. We identified transporters involved in general iron acquisition by performing annotation text searches for the terms 'siderophore', 'ferrichrome', 'heme', 'hydroxamate', 'catechol', and 'ferrous' followed by manual inspection of the results for the retrieved genes' potential role in transport. We identified specific siderophore transporters by manually inspecting genes neighboring the siderophore biosynthetic clusters identified above and retaining those that had significant matches to PFAM ABC transporters (Pfam00005) or TonB-dependent transporter (Pfam00593) families. From these, we picked one representative for each type of transporter examined and searched all 69 genomes for the corresponding orthologous genes. Orthologous genes across all 69 genomes were determined by using the software OrthoMCL using the following parameters: 50% sequence identity at the amino acid level (Chen et al. 2006). The list of representative genes used to search for orthologous genes can be found as part of Table S2.2 (Section 2.9).

2.3.3. Correlation Analysis

For this analysis, we only considered strains that grew on CAS-agar belonging to the *Vibrio splendidus*-like clade. This clade consists of distinct subpopulations of species, including *V. splendidus, V. crassostreae, V. tasmaniensis,* and *V. cyclotrophicus,* which are thought to specialize in exploiting different marine particulate nutrient patches (Hunt et al. 2008). The dependence of the variables tested was corrected based on Felsenstein's method of contrasts (Felsenstein 1985). Spearman rank correlations were then calculated between the phylogenetic contrasts of the 'siderophore production' trait and the different sampling size fractions, which include: <1 μm, between 1 μm and 5 μm, between 5 μm and 64 μm, and >64 μm. The phylogeny is based on DNA sequence of the well-conserved *hsp60* gene. We plotted the distributions of the correlation values calculated for 100 bootstrap trees to correct for uncertainty in the phylogeny. Distributions were smoothed using kernel density estimation as implemented in the "density" function in the statistical software R.

The approach used to calculate independent phylogenetic contrasts can be briefly summarized as follows. Siderophore production was determined at each internal nodes as a percentage of producers under that node: for example, at the root there is 40% production, whereas under the *V. ordalii* ancestral node, there is ~90% of production. The same approach is used to assign a habitat to each internal node. Contrasts (differences) are then computed between independent sister branches: to each pair of independent sister branches we assign a transition value for both siderophore production and habitat (see Figure 2.3). This matrix of independent values is then used to compute the Spearman rank correlation. This approach assumes a Brownian motion of evolutionary traits as well as equal phylogenetic branch lengths. Nodes for which both transistion values were 0 were removed. To remove the noise, evolutionary transitions between leaves or between leaves and internal nodes were not considered.

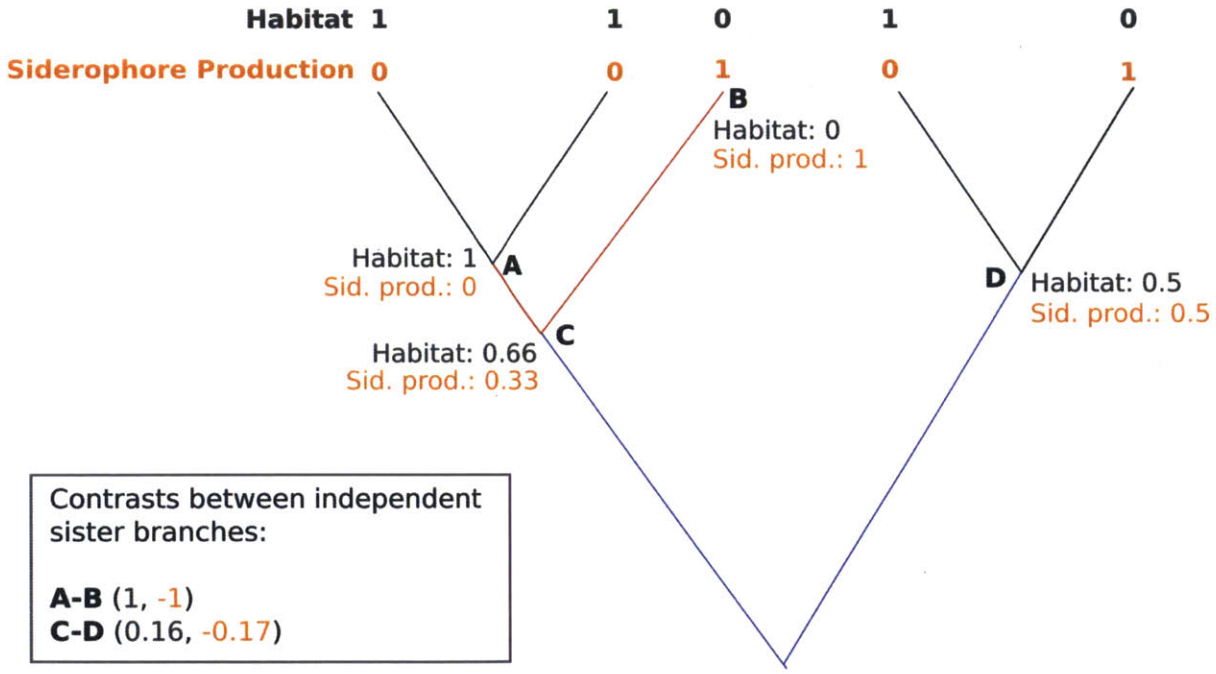


Figure 2.3. Method of contrasts. Nodes are denoted by the letters A through D. Pairs of independent sister branches are color-coded red and blue. For each node, the frequency of habitat and siderophore production under that node is determined.

2.3.4. Growth Enhancement Experiments

Growth enhancement experiments were performed for the two non-producer strains: ZF90 and ZF264. Both strains carry siderophore specific transporter (ZF-90: aerobactin transporters; ZF264: vibrioferrin transporters). Strains were grown overnight at room temperature in 2216 marine broth (Difco). Cultures were pelleted (2 min at 10,000 rpm in a microcentrifuge), the supernatant was discarded and pellets were washed once in modified M9 salts (42.2 mM dibasic sodium phosphate (Na_2HPO_4), 22 mM monobasic potassium phosphate (KH_2PO_4), 18.7 mM ammonium chloride (NH_4Cl), and 2% *w/v* sodium chloride (NaCl)) to remove any carry-over 2216 media. Washed cells were pelleted and resuspended in modified M9 salts. Cells were then inoculated (1:1,000) into growth media (modified M9 salts, 0.3% casamino acids, 200mM magnesium sulfate ($MgSO_4^{2-}$), 20μM calcium chloride ($CaCl_2$), and vitamins: 0.1 $\mu g.L^{-1}$ vitamin B12, 2 $\mu g.L^{-1}$ biotin, 5 $\mu g.L^{-1}$ calcium pantothenate, 2 $\mu g.L^{-1}$ folic acid, 5 $\mu g.L^{-1}$ nicotinamide, 10 $\mu g.L^{-1}$ pyridoxin hydrochloride, 5 $\mu g.L^{-1}$ riboflavin, 5 $mg.L^{-1}$ thiamin hydrochloride) with (iron-poor) or without (iron-replete) the iron-specific chelator EDDA, ethylenediamine-*N,N'*-diacetic acid (Sigma-Aldrich). Cells were incubated for 24 hours at room temperature. Absorbance at 600 nm was measured every 3 hours. Purified ferric-aerobactin (EMC microcollections, Tübingen, Germany) and vibrioferrin (generously provided by Dr. Carl Carrano, San Diego State University) were added at specified concentrations. Cell-free supernatant from the aerobactin producer 5F-79 and the vibrioferrin producer 12B01 grown under iron-poor conditions (modified M9 + 100 μM EDDA) were obtained by filtering the supernatant of stationary phase cultures through a 3kDalton membrane using an Amicon Ultra® centrifugal unit (Millipore, Billerica, MA, USA).

2.3.5. Spotting Bioassay

Solid media containing 42.2 mM Na_2HPO_4, 22 mM KH_2PO_4, 18.7 mM NH_4Cl, 2% *w/v* NaCl, 0.3% casamino acids, 200 mM $MgSO_4$, 20 μM $CaCl_2$, and vitamins (see previous section) was prepared with or without addition of 120 μM of iron-specific chelator EDDA. Overnight cultures of non-producer strains (ZS139, ZF90) grown in 2216, were rinsed and resuspended in modified M9 salts (see above) and spread on the solid media. The aerobactin producer strain, 5F79 was grown overnight in modified M9,

washed in 2% salt water to remove residual media and spotted onto the solid media. Plates were incubated at room temperature and presence of colonies was checked after 24 hours.

2.4. Results

2.4.1. Siderophore Production is a Patchy Trait

Of the 1,710 *Vibrionaceae* isolates, 1,012 (~60%) grew on the CAS-agar and were retained for further analysis. Roughly ~40% (394) of these were identified as siderophore producers per the assay pipeline we used (Figures 2.4 and 2.5). Siderophore production, while spread across the *Vibrionaceae* phylogenetic tree has a patchy distribution: all lineages seem to have evolved to include both producers and non-producers. Furthermore, siderophore production does not seem to be a species-specific trait. For example the *Vibrio splendidus* group, marked "3" in Figure 2.4, is clearly composed of both producers and non-producers. Some groups however are characterized by particularly high frequency of producers, as is the case for *Allivibrio fischerii*, *Vibrio ordalii*, *Vibrio sp. F12* (Figure 2.5). All three populations display siderophore production frequencies nearing 100%, far above the overall average of 40%. *Allivibrio fischerii* and *Vibrio ordalii* are known to be host-associated strains (Nyholm et al. 2009; Austin 2011), suggesting that the evolutionary pressure to produce siderophores is especially high in those environments. In contrast, a handful of populations display average frequencies of production less than 20%. Despite these observations, most lineages center around 40% of siderophore production (Figures 2.4 and 2.5).

As described in the methods, we tested for correlations between the 'siderophore production' trait and the 'sampling-size fraction' trait. The correlation test was performed only for the *Vibrio splendidus*-like strains to discount potential biases introduced by populations that are known to be host associated. Figure 2.6 shows the sampling size fraction associated with each of these strains (the raw data can be found in Table S2.1, section 2.9.1). We found a positive non-zero correlation for siderophore production and a <1 μm sampling size fraction (Figure 2.7). This means that we are more likely to find *Vibrio splendidus*-like isolates that produce siderophores in the free-living fraction. On the other hand, we found a negative non-zero correlation between siderophore production and isolates sampled from a > 64 μm size fraction, meaning that a higher proportion of non-producers could be found on large particles. Interestingly, the frequency of siderophore production correlates with particle size in a predictable fashion: the smaller the particle, the more positive the correlation.

Figure 2.4. Siderophore production is an intermediate frequency trait in natural *Vibrionaceae* populations. The distribution of producers on the phylogeny, as found by the CAS screen. The tree is based on the genetic marker *hsp60* and comprises different genotypic clusters previously found to have cohesive ecology and hypothesized to represent samples from natural ecotypes. The numbers indicate : 1. *Vibrio ordalii*, 2. *Allivibrio fischerii*. The clade descending from node 3 corresponds to close relatives of *V. splendidus* (the *V. splendidus*-like clade), and includes: 4. *V. crassostreae*, 5, *V. cyclotrophicus*, 6. *V. splendidus*. The tree and associated assay data are plotted using the imaging software 'interactive Tree Of Life' (iTOL) (Letunic & Bork 2007). Raw data can be found in the supplementary materials for this chapter, section 2.9.1.

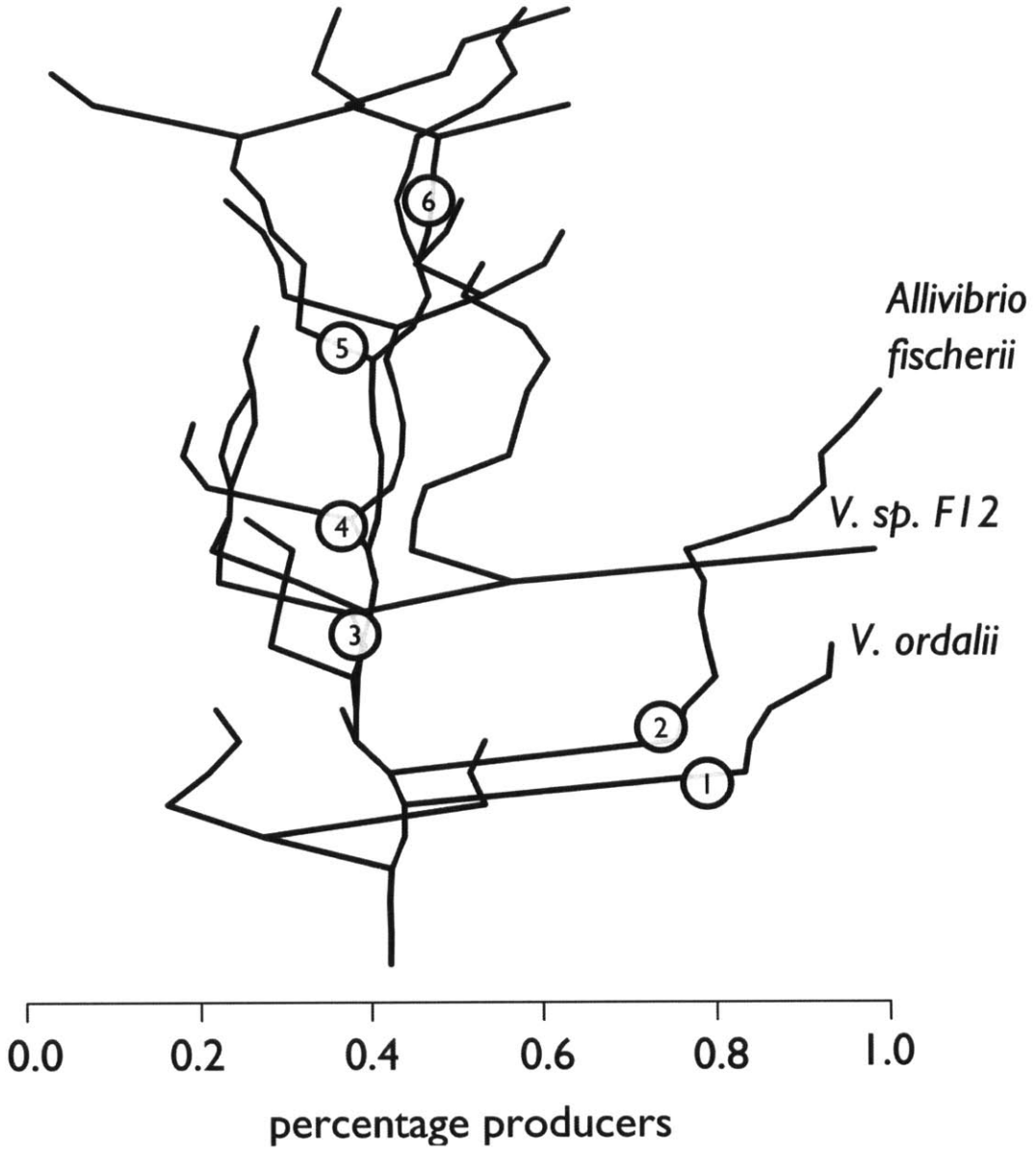


Figure 2.5. Frequency of siderophore production per phylogenetic node. This figure represents the data shown in Figure 2.4 in terms of percentage of producers descending of each internal node of the phylogeny. Some nodes have been marked to help guide the comparison with Figure 2.4. Among the populations with a high incidence of producers are animal host associated *V. ordalii* and *A. fischerii*. The clade descending from node 3 corresponds to close relatives of the *V. splendidus* species. The patchy distribution of producers suggests that non-siderophore producers are either utilizing exogenously produced siderophores or have evolved other iron acquisition systems.

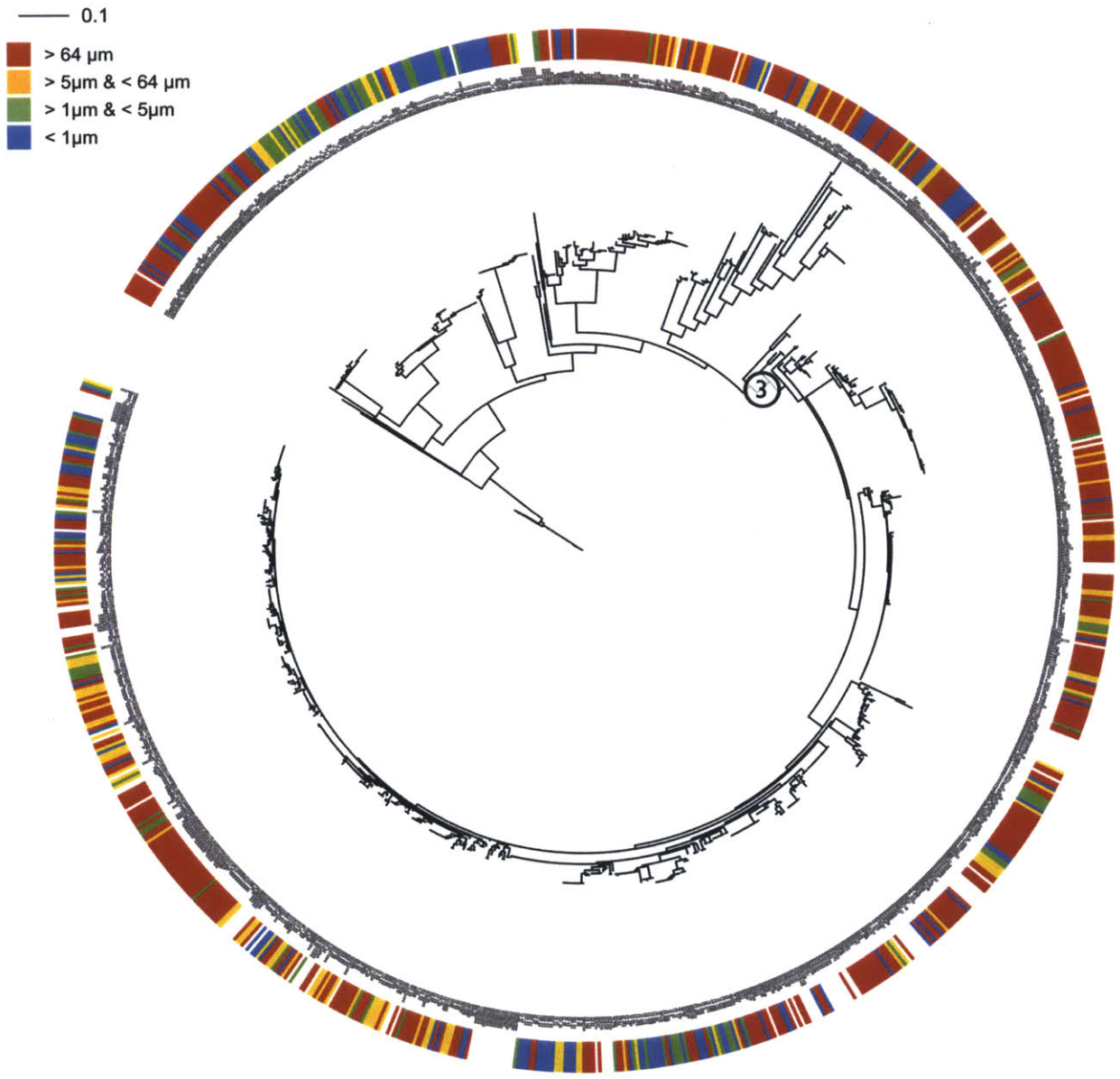


Figure 2.6. Sampled environment of strains screened. This figure displays the different sampled environments (size fractions) for the vibrio strains screened for siderophore production. We only represent the vibrio strains that grew. The tree is the same as the one displayed in Figure 2.4. It is based on the conserved *hsp60* gene. Sampled environments are color-coded. The clade descending from node 3 is the *V. splendidus*-like clade, for which siderophore production data represented in Figure 2.4 and sampled environment shown here are used to calculate the correlations depicted in Figure 2.7.

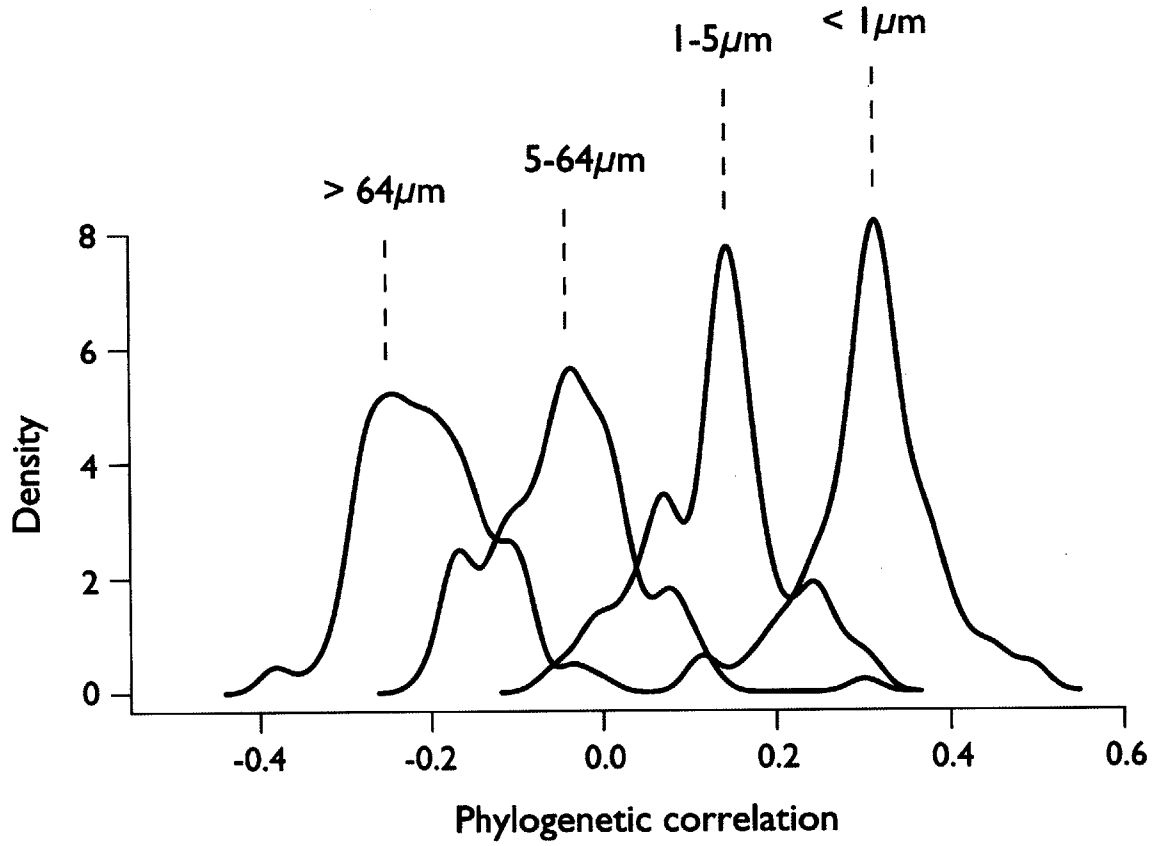


Figure 2.7. Siderophore production correlates positively with free-living lifestyle within the *V.splendidus*-like clade. Phylogenetic correlations as described in methods are calculated between the traits 'siderophore production' and 'sampling size-fraction'. The distribution of values corresponds to the different values calculated for 100 bootstraps of the *hsp60* tree for the 1,012 *Vibrio* isolates. We show a positive non-zero correlation for siderophore production and a sampling size-fraction <1 μm. At the other extreme, we show a negative non-zero correlation for siderophore production and a sampling size-fraction > 64 μm. The spearman correlation tests were performed using statistical software R (R Foundation for Statistical Computing 2012).

2.4.2 The Vibrionaceae produce a wide variety of siderophores

Realizing that siderophore production is a trait that can be observed across the *Vibrionaceae* phylogeny, we then asked whether all *Vibrionaceae* produced the same type of siderophore. We were curious to find out whether the distribution of different siderophores matched the phylogeny, whether some types of siderophores were ubiquitous across the phylogeny, and whether different types of siderophores and their known chemical properties could provide information on the potential environment of the producing strains. To obtain information on the different siderophore types being produced, the genomes of 69 of the strains screened for siderophore production were mined for siderophore biosynthetic clusters (Figure 2.2). The presence or absence of a siderophore biosynthesis cluster identified by this *in silico* approach overall matched very well with the results of the siderophore production CAS screen (Figure 2.8). Accordingly, strains that were predicted to produce a truncated pyoverdin did not produce a siderophore in the CAS screen (Figure 2.8). However, there were a few strains for which the bioinformatic analysis detected the presence of a siderophore biosynthetic cluster, but siderophore production was not detected on the functional screen (FF-85, 1F-211, 1F-230, FF-167, 9ZB36, 1S-124). This might be the result of regulatory mechanisms that repress siderophore production under the conditions of the assay.

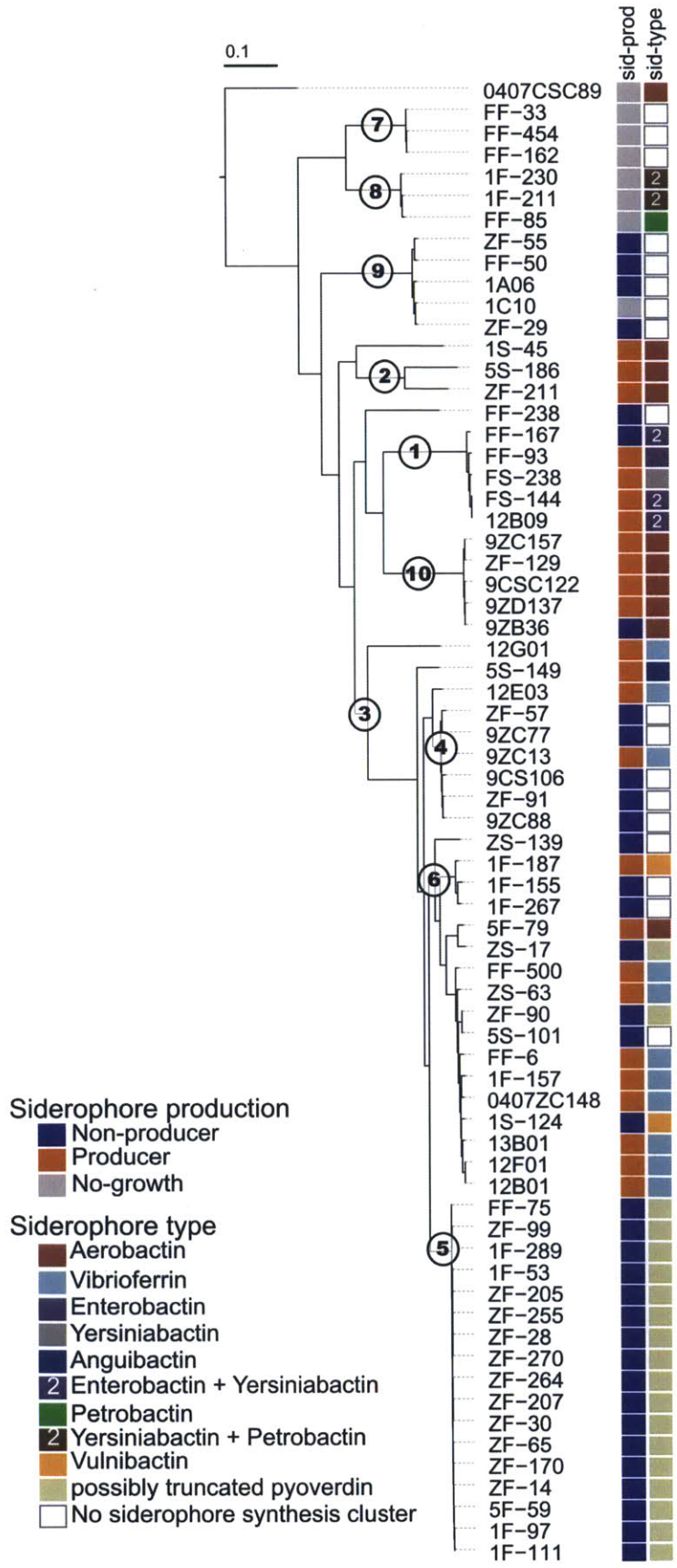

0.1
sid-prod
sid-type
0407CSC89
FF-33
FF-454
FF-162
1F-230
1F-211
FF-85
ZF-55
FF-50
1A06
1C10
ZF-29
1S-45
5S-186
ZF-211
FF-238
FF-167
FF-93
FS-238
FS-144
12B09
9ZC157
ZF-129
9CSC122
9ZD137
9ZB36
12G01
5S-149
12E03
ZF-57
9ZC77
9ZC13
9CS106
ZF-91
9ZC88
ZS-139
1F-187
1F-155
1F-267
5F-79
ZS-17
FF-500
ZS-63
ZF-90
5S-101
FF-6
1F-157
0407ZC148
1S-124
13B01
12F01
12B01
FF-75
ZF-99
1F-289
1F-53
ZF-205
ZF-255
ZF-28
ZF-270
ZF-264
ZF-207
ZF-30
ZF-65
ZF-170
ZF-14
5F-59
1F-97
1F-111
1
2
3
4
5
6
7
8
9
10
Siderophore production
Non-producer
Producer
No-growth
Siderophore type
Aerobactin
Vibrioferrin
Enterobactin
Yersiniabactin
Anguibactin
2 Enterobactin + Yersiniabactin
Petrobactin
2 Yersiniabactin + Petrobactin
Vulnibactin
possibly truncated pyoverdin
No siderophore synthesis cluster

Figure 2.8. The *Vibrionaceae* produce a variety of siderophores. Column 'sid-prod' refers to the outcome of the siderophore production assay. Column 'sid-type' refers to the type of siderophore biosynthetic cluster inferred by bioinformatics analysis as detailed in Figure 2.2. The tree is based on 66 single copy genes present in all sequenced strains. The number 2 in the siderophore-type column highlights the strains for which more than one siderophore synthesis cluster was detected. Ecologically-cohesive populations are marked. Populations 1 through 6 correspond to the 6 populations marked in Figures 2.4 and 2.5. 1: *Vibrio ordalii*, 2: *Allivibrio fischeri*, 3: *Vibrio splendidus*, 4: *Vibrio crassostreae*, 5: *Vibrio cyclotrophicus*, 6: *V. splendidus* 7: *Enterovibrio calviensis*, 8: *Enterovibrio norveginus*, 9: *Vibrio sp. F9 superstes/ breorganii*, 10: *Vibrio sp. F10*.

The types of siderophores produced by the 69 sequenced strains were inferred based on the *in silico* inspection of the biosynthetic pathway of each siderophore as described in the methods (Figure 2.2). Siderophores are either produced by non-ribosomal peptide synthetases (NRPS) or by a pathway independent of NRPSs, the NRPS independent synthesis (NIS) pathway (Crosa & Walsh 2002; Oves-Costales et al. 2009). AntiSMASH recognizes both types of pathways: the NRPS pathway is classified as 'nrps', while the NIS pathway is classified as 'siderophore' (Medema et al. 2011). The NIS pathway is characterized by the presence of distinct synthetases with high homology to the IucA/IucC synthetases involved in aerobactin biosynthesis (Oves-Costales et al. 2009). These synthetases may be flanked by genes coding for proteins responsible for the modification of specific substrates that will be incorporated in the siderophore or by specific transporters (Suzuki et al. 2006), allowing for more specific prediction of the siderophore product. While siderophores are readily recognized as products of the NIS pathway, NRPSs are responsible for the biosynthesis of a wide variety of small molecules, making the prediction that the NRPS codes for a siderophore, let a lone a specific type of siderophore more tenuous.

NRPSs are characterized by a set of distinct modular sections, each responsible for the incorporation of a different monomer (amino acids and other substrates): the number and order of each module determines the primary structure of the peptidic compound (Crosa & Walsh 2002). *In silico* inspection of these modules and more specifically of the adenylation domain of each module allows for the prediction of the NRPS product synthesized. Each NRPS module comprises a set of three catalytic domains necessary for the incorporation of each monomer: an adenylation domain (amp-binding: pfam00501) that recognizes and activates the monomeric substrate, a peptidyl-carrier-protein domain (pp-binding: pfam00550) that acts as a way station for the activated monomer as it is being chemically modified or incorporated into a growing peptide chain, and a condensation domain (condensation: pfam00668) that catalyzes the peptide bond formation (Crosa & Walsh 2002; Strieker et al. 2010). A sequence of 10 amino acid residues has been shown to confer specificity of the adenylation domain to a cognate monomeric substrate (Conti et al. 1997; Stachelhaus et al. 1999). NRPS Predictor takes advantage of this property to bioinformatically predict the affinity of the adenylation domain for a specific monomer and can hence help infer the type of NRPS product that may be synthesized (see Figure 2.2) (Röttig et al. 2011). Substrates that

are commonly used in siderophore biosynthesis include lysine, ornithine, histamine, citrate, serine, 2,3-dihydroxybenzoate and salicylate, to cite a few. NRPSs with adenylation domains that were specific only to hydrophobic substrates such as valine, alanine and leucine, which are unlikely to lead to the formation of a siderophore because they lack the appropriate side groups that would enable iron chelation were not taken into account in the analysis.

This approach revealed that the *Vibrionaceae* strains screened produce a wide variety of siderophore types (Figure 2.8). The inferred siderophore types encompass all three major types of iron-chelating groups commonly found in siderophores: catecholate (enterobactin) (Harris et al. 1979), hydroxamate (aerobactin, pyoverdin), carboxylate (vibrioferrin) (Yamamoto, et al. 1994), mixed, containing both hydroxamate and catecholate functional groups (anguibactin) or both catecholate and carboxylate funcional groups (petrobactin) (Homann et al. 2009). The siderophore biosynthetic gene clusters found belong to both the NRPS (enterobactin, anguibactin, yersiniabactin, pyoverdin) and the NIS (aerobactin, vibrioferrin, petrobactin) synthesis pathway types (Crosa & Walsh 2002; Challis 2005; Perry et al. 1999; López & Crosa 2007; Tanabe et al. 2003a; Lorenzo et al. 1986; Ravel & Cornelis 2003). While some siderophores, like enterobactin were found to be produced solely by one population (*Vibrio ordalii*), others, like aerobactin and vibrioferrin were found to be produced more widely across the tree (Figure 2.8). It is interesting to note that enterobactin, which was first characterized in *Escherichia coli* and *Salmonella typhimurium* (Earhart 1987) is produced by *V.ordalii,* thought to be animal-associated, more specifically a fish pathogen (Chart & Trust 1984), a likely environment where it could have encountered *E.coli* and acquired enterobactin synthesis genes from it.

An interesting observation is the consistent mismatch between the *in silico* analysis and the CAS assay results for the strains with the 'possibly truncated pyoverdin' biosynthesis cluster. Detailed analysis of the NRPS gene cluster detected in those strains (see representative strain FF-75 in the supplementary material, Table S2.2, section 2.9.2) shows similarities to the biosynthesis pathway of pyoverdin (the L-ornithine monooxygenase next to large NRPS enzymes). However, in contrast to pyoverdin biosynthesis, the clusters found in this study only contain two NRPS enzymes as opposed to four in pyoverdin

biosynthesis (Ravel & Cornelis 2003). Futhermore, the adenylation domains of these NRPSs show specificity to cysteine, serine and leucine for one NRPS, and to leucine for the other, which comes in contrast to the substrates usually used for pyoverdin biosynthesis. Because this NRPS configuration does not seem to match with that of other known siderophore NRPSs, it is unclear whether the negative result in the CAS functional screen (Figure 2.8) is caused by a truncated NRPS biosynthetic cluster or by a regulatory repression under the conditions of the assay. This points to the difficulty of properly predicting fully functional siderophore products of NRPS assembly from *in silico* analyses alone. The alternative of chemically identifying all siderophores for these 69 strains is unwieldy and not appropriate for the high-throughput nature of this study. Because of the relative simplicity of the NIS pathway compared to the NRPS pathway, the prediction of siderophore type produced is more robust for the NIS-type gene clusters (here: aerobactin, vibrioferrin and petrobactin) than for the NRPS-type.

Another interesting observation is that in all but six (1F-230, 1F-211, 5S-186, FF-167, FS-144 and 12B09) of the genomes analyzed only one siderophore biosynthesis cluster is detected (Figure 2.8 and Table S2.2, Section 2.9.2). Here, it is important to point out that the genomes analyzed are draft genomes, and that for some of the strains, we observe that the siderophore biosynthesis clusters coincide with the end or start of the contigs (Table S2.2, Section 2.9.2). This can be explained by the fact at these loci the assembler would have encountered sequences that were identical in two different regions in the genome. This may suggest that for these strains two or more identical copies of the detected siderophore cluster are found in the genome. An alternative, more likely explanation is the possibility that the cluster is found next to a highly repetitive element such as a transposase. However, most of these *Vibrionaceae* seem to have evolved to keep the least costly configuration of siderophore biosynthesis, maintaining one copy of the synthesis genes required to produce one type of siderophore.

2.4.3. Iron Transport Genes Are Ubiquitous Across the Vibrionaceae Phylogeny

Analysis of iron-related transport systems in the 69 *Vibrionaceae* genomes (Figure 2.9) reveals that all strains, including those that do not produce siderophores have evolved a strategy for iron transport. While receptors specific to the siderophores biosynthesized by this group of *Vibrio* tend to map to the groups that produce them, general iron transport receptors are widely distributed across the *Vibrionaceae* phylogeny. We find that virtually all *Vibrionaceae* genomes analyzed have the genes required for the transport of a catechol-type of siderophore, as well as for ferrous iron and hemin-bound ferric iron. On the other hand, the transporters for ferrichrome and other types of hydroxamate siderophores (FhuA-B), seem to be more abundant in the *V. splendidus*-like clade. Surprisingly, only a handful of strains have the genes required for transport of ferric-dicitrate, a by-product of metabolism and a fortuitous iron-chelator, which has recently been suggested as a siderophore with likely importance in the ocean (Toulza et al. 2012).

These findings are significant as they can help us understand the iron availability of the environment in which these *Vibrionaceae* have evolved (Winkelmann 2004). The Feo-ferrous iron transport genes are likely turned on under anaerobic conditions when ferrous iron is abundant (Cartron et al. 2006). This is especially relevant for the opportunistic pathogens or host-associated members of the *Vibrionaceae* family. Alternatively, the widespread presence of these ferrous iron transporters suggests that in spite of our expectations, ferrous iron may be a transiently useful source of iron in the ocean. It is probable that these vibrios could take advantage of the small ferrous iron pool generated by dynamic equilibrium between ferric and ferrous species or by photo-reduction of ferric-ligands (Barbeau 2006). Another hypothesis could be that ferrous iron is available on marine particles, where microaerobic environments are thought to persist (Simon et al. 2002).

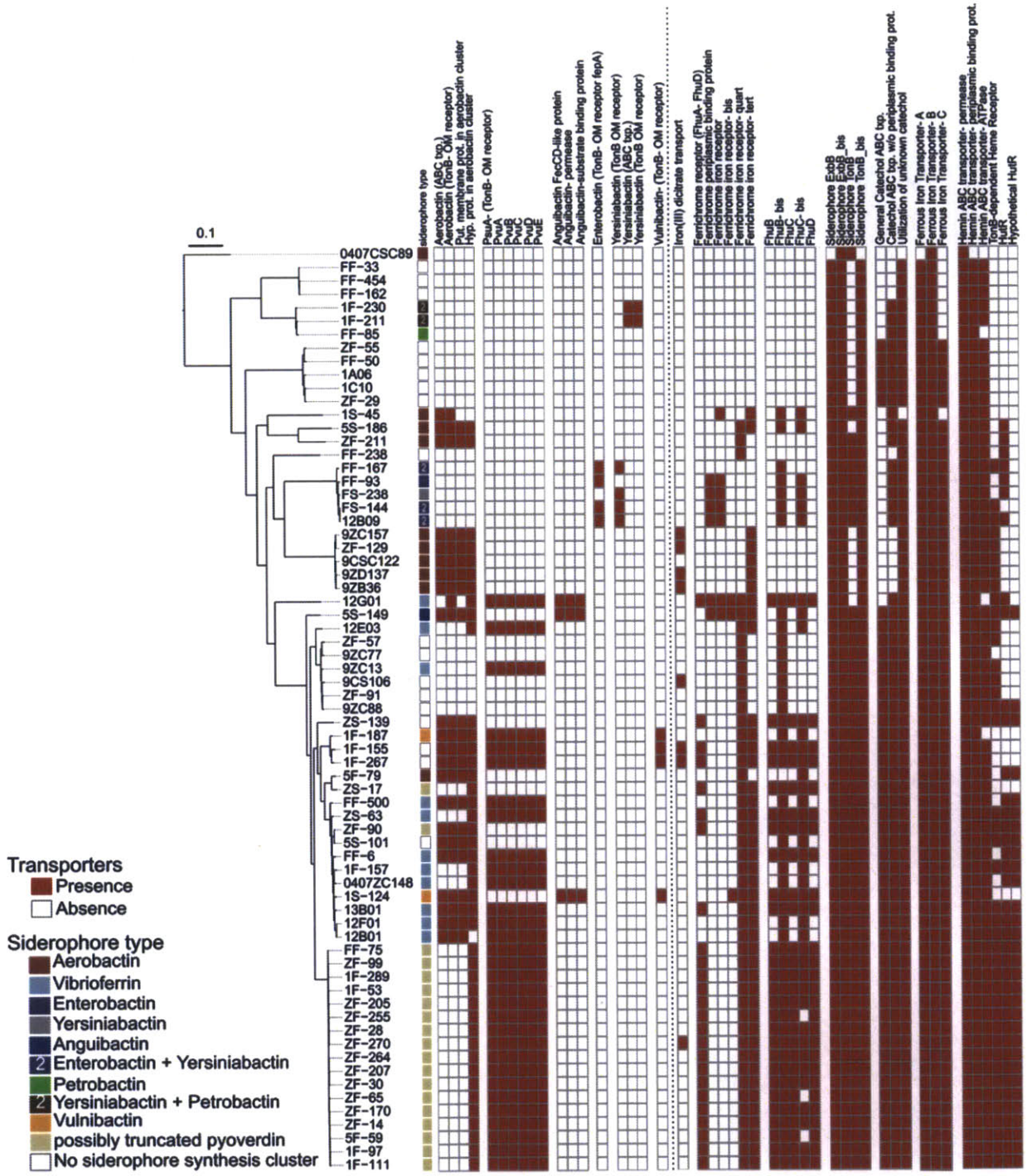


Figure 2.9. Distribution of iron transporters. Orthologous clusters of transporters are shown for each of the 69 strains with sequenced genomes. The first column, siderophore type is a reproduction from Figure 2.8. Subsequent columns show the presence (red) or absence (white) of the transporter in each genome. The dotted line separates the transporters associated with specific siderophores from the general iron transporters. Genes used to search for the orthologous clusters are listed in Table S2.3 in Section 2.9.3. The genes chosen to search for the specific siderophore transporters were those identified as either ABC transporters or TonB-dependent transporters in the vicinity of the siderophore synthesis clusters.

2.4.4. Genotypic and Phenotypic Signatures of 'Cheating'

Despite our observation that most siderophore-specific transporters are limited to the clades that produce them, aerobactin and vibrioferrin transporters are widely distributed across the phylogeny and are not limited to the strains for which aerobactin or vibrioferrin synthesis clusters are detected (Figure 2.9). Furthermore, we find that these transporters are present in a number of strains for which no siderophore production is detected in the functional CAS-screen (Figure 2.9). Detailed analysis of the genomic context surrounding these transporters in the siderophore non-producing *V. splendidus*-like strains shows that the vibrioferrin and aerobactin biosynthetic genes have in fact been excised or replaced, while the vibrioferrin and aerobactin-specific transporters have been retained (Figure 2.10). This suggests that these non-producing strains, by retaining specific siderophore transport genes are able to benefit from the siderophores produced by some of their close relatives.

These findings were confirmed by performing physiological assays on two siderophore producers, 5F79 (aerobactin producer) and 12B01 (vibrioferrin producer), and four non-producer representatives (ZS139, ZF90, 1F-53, ZF264). As expected, we found that in iron-poor media, the growth of the non-producer strains, ZS139, ZF90, 1F-53 and ZF264 was impaired, while the growth of the siderophore producer strains 5F79 (aerobactin producer) and 12B01 (vibrioferrin producer) was unaffected (Figure 2.11). Siderophore growth enhancements assays were performed to further demonstrate that the non-producers have lost the ability to produce a siderophore but maintained its specific transport genes and are in fact able to utilize that siderophore. Figure 2.12 shows that addition of pure aerobactin to ZF90 and pure vibrioferrin to ZF264, harboring the specific transporters for aerobactin and vibrioferrin respectively stimulated growth under iron-poor conditions. We further show that under iron-poor conditions, the growth of these non-producers is stimulated by the addition of cell-free supernatant obtained from stationary phase cultures of closely related strains that produce the siderophore for which ZF90 and ZF264 carry the transporters. This is further demonstrated in Figure 2.13, where micro-colonies of the non-producer ZS139 grow only in the vicinity of the spotted producer 5F79. Together, these findings suggest

that these non-producers can utilize the siderophore produced by their siderophore-producing close relatives. This confirms the hypothesis generated by *in silico* observations.

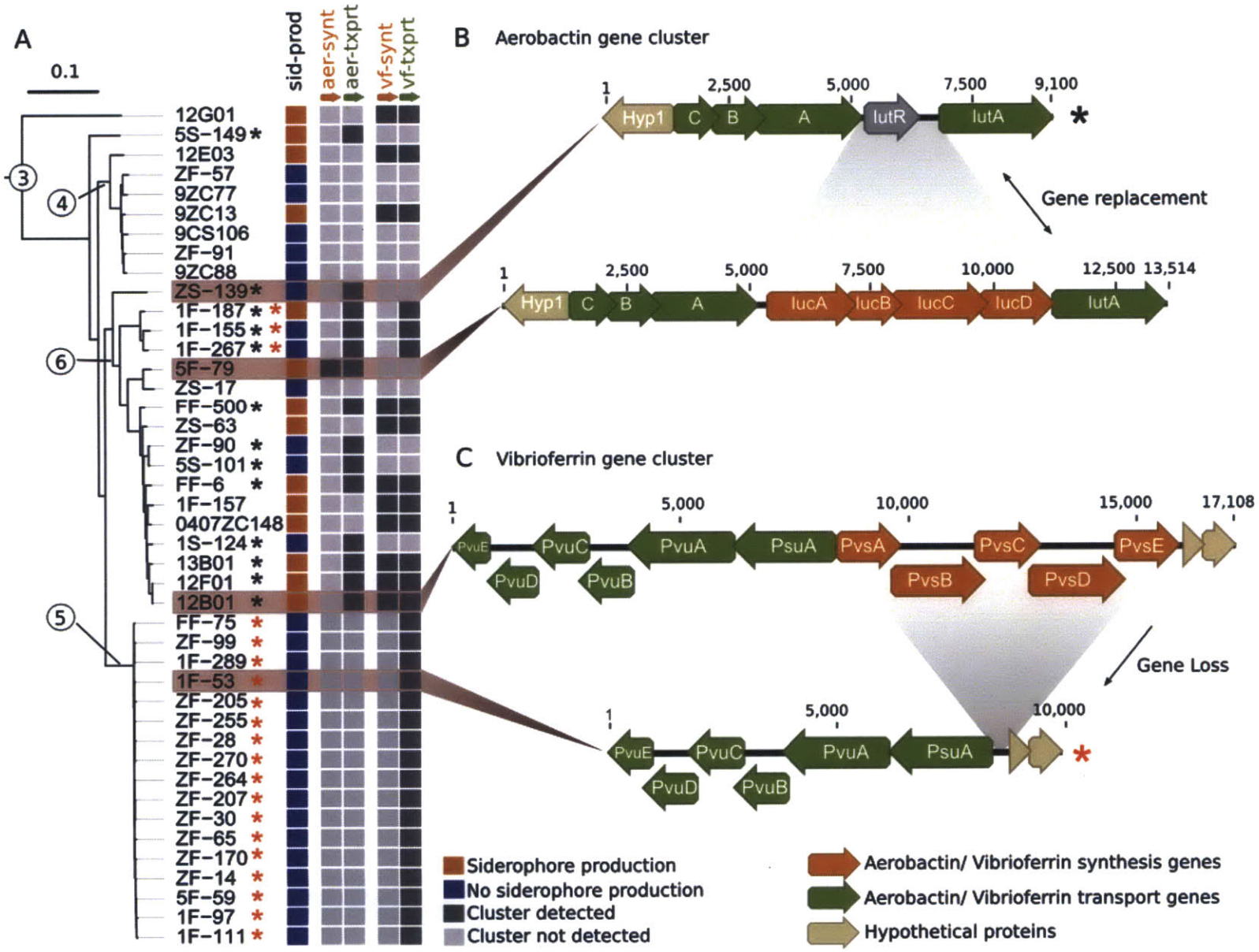

A
0.1
sid-prod
aer-synt
aer-txprt
vf-synt
vf-txprt
12G01
5S-149*
12E03
ZF-57
9ZC77
9ZC13
9CS106
ZF-91
9ZC88
ZS-139*
1F-187* *
1F-155* *
1F-267* *
5F-79
ZS-17
FF-500*
ZS-63
ZF-90 *
5S-101*
FF-6 *
1F-157
0407ZC148
1S-124*
13B01 *
12F01 *
12B01 *
FF-75 *
ZF-99 *
1F-289 *
1F-53 *
ZF-205*
ZF-255*
ZF-28 *
ZF-270*
ZF-264*
ZF-207*
ZF-30 *
ZF-65 *
ZF-170*
ZF-14 *
5F-59 *
1F-97 *
1F-111 *
3
4
6
5
B Aerobactin gene cluster
1
2,500
5,000
7,500
9,100
Hyp1
C
B
A
IutR
IutA
Gene replacement
1
2,500
5,000
7,500
10,000
12,500
13,514
Hyp1
C
B
A
IucA
IucB
IucC
IucD
IutA
C Vibrioferrin gene cluster
1
5,000
10,000
15,000
17,108
PvuE
PvuC
PvuA
PsuA
PvsA
PvsC
PvsE
PvuD
PvuB
PvsB
PvsD
Gene Loss
1
5,000
10,000
PvuE
PvuC
PvuA
PsuA
PvuD
PvuB
Siderophore production
No siderophore production
Cluster detected
Cluster not detected
Aerobactin/ Vibrioferrin synthesis genes
Aerobactin/ Vibrioferrin transport genes
Hypothetical proteins

Figure 2.10. Loss of siderophore synthesis genes explains patchiness of siderophore production trait in the *V. splendidus*-like clade. A) Phylogenetic relationship, siderophore production and siderophore synthesis and transport genes for sequenced strains in the *V. splendidus*-like group. The tree is based on 66 single copy genes present in all of the sequenced strains. The 'sid-prod' column refers to the outcome of the siderophore production assay; 'aer-syn' stands for aerobactin biosynthesis genes; aer-txpt for aerobactin-specific transport genes; vf-syn for vibrioferrin biosynthesis genes and vf-txpt for vibrioferrin-specific transport genes. B-C) Examples of configurations of siderophore synthesis and transporter gene clusters in producer and non-producer strains. The figures show that non-producer phenotype evolved from excision or replacement of the biosynthesis genes from the complete synthesis-transport cluster. The black and red stars are used to indicate strains that have a similar loss of synthesis gene cluster configuration in panel A. The dark shaded region shows regions of homology between the cooperator and cheater configurations.

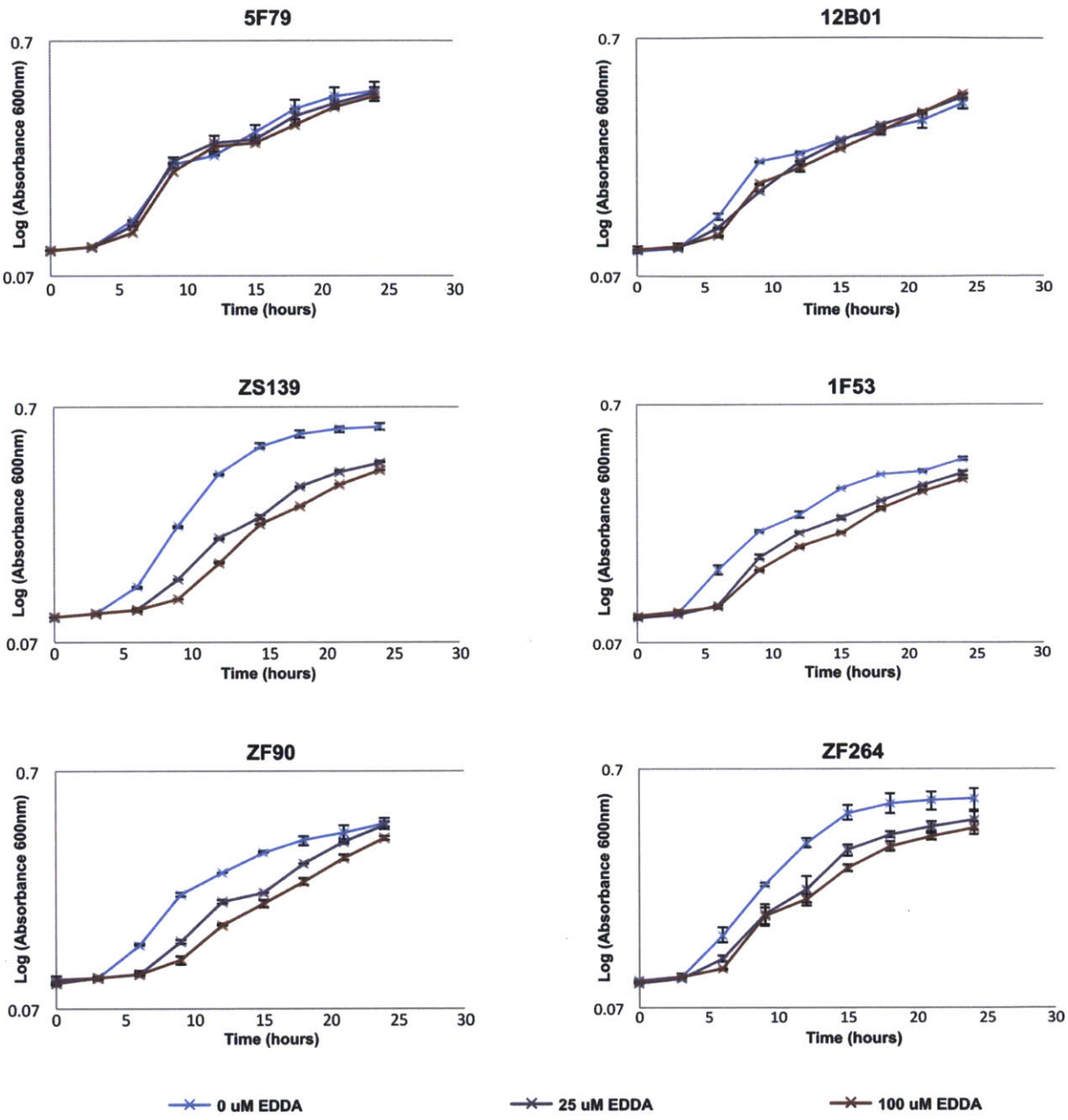


Figure 2.11. Growth of producers and non-producers in iron-poor conditions. Growth curves were obtained in iron-rich (growth media without EDDA) and iron-poor (growth media with EDDA) media for 5F79 (aerobactin producer), ZS139 and ZF90 (non-producers but with aerobactin specific transporters), 12B01 (vibrioferrin producer), 1F53 and ZF264 (non-producers but with vibrioferrin specific transporters). Growth was determined by measuring absorbance at 600 nm every 3 hours over a 24-hour period. Various concentrations of the iron-specific chelator EDDA were tested (2 μM, 10 μM, 20 μM, 25 μM, 50 μM, and 100 μM). For clarity's sake, only the results for three concentrations are displayed. Growth media was composed of 42.2 mM Na_2HPO_4, 22 mM KH_2PO_4, 18.7 mM NH_4Cl, 2% *w*/*v* NaCl, 0.3% casamino acids, 200 mM $MgSO_4$, 20 μM $CaCl_2$, and vitamins, as described in the Methods. For all four non-producer strains, ZS139, 1F53, ZF90 and ZF264 we observe a lag in the onset of exponential phase in the iron-poor conditions. The two siderophore producing strains, 5F79 and 12B01 cope well with the iron-limitation imposed by EDDA as their growth curves in the iron-poor media virtually overlap those in the iron-rich media.

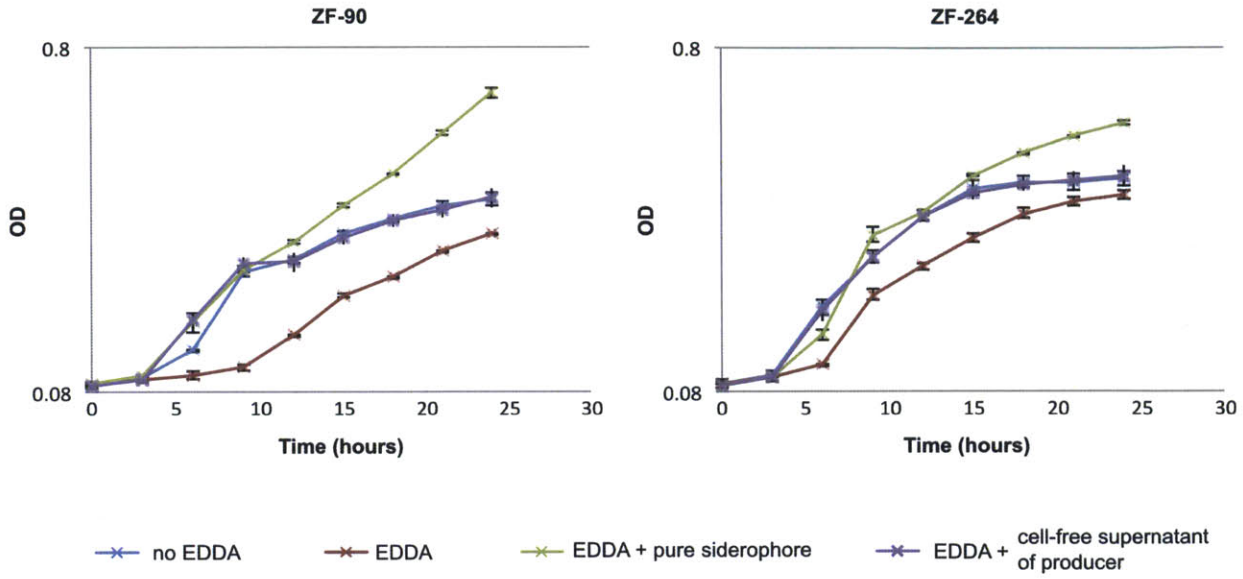


Figure 2.12. Siderophore utilization by non-producers. Growth curves of non-producer strains: blue, normal growth conditions (as described in methods); red, under iron-poor conditions (ZF-90: 100 μM EDDA, ZF-264: 50 μM EDDA); green, supplemented with pure ferric-siderophore (ZF-90: 100 μM EDDA and 80μM ferric-aerobactin; ZF-264: 50 μM EDDA and 70 μM ferric-vibrioferrin); and purple, supplemented with cell-free supernatant of siderophore producer culture (ZF-90: 100 μM EDDA and cell-free extract from 5F79 aerobactin producer, ZF-264: 50 μM EDDA and cell-free extract from 12B01 vibrioferrin producer). Growth was monitored spectrophotometrically, measuring absorbance at 600 nm every 3 hours. Data is plotted for three biological replicates. Error bars represent one standard deviation from the mean. For both non-producers, growth is impaired under iron-poor conditions but restored with the exogenously added pure siderophore as well as with the cell-free supernatant of the corresponding producer.

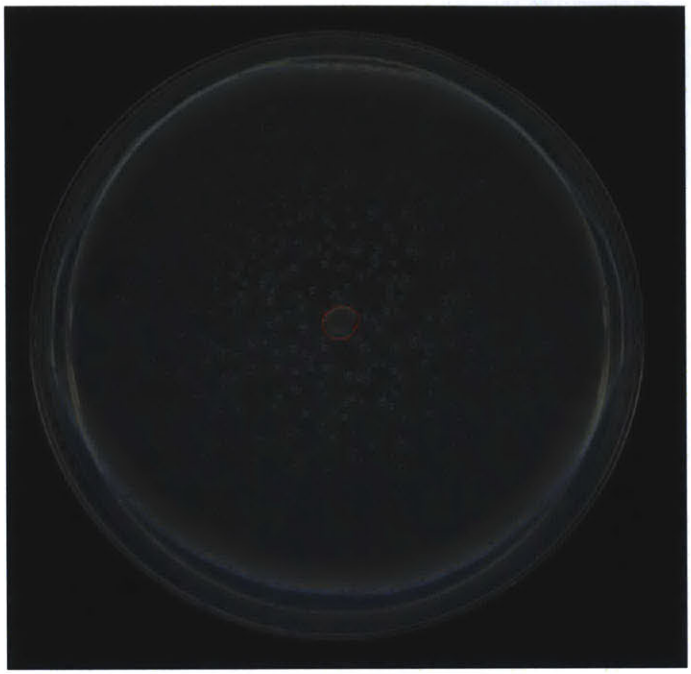


Figure 2.13. Growth of non-producer strains is facilitated by producer strains. Picture representation of spotting bioassay. The colony circled in red is the spotted aerobactin producing strain 5F-79, small colonies surrounding it are colonies from the streaked non-producing strain ZS-139. The growth media is high salt M9 media supplemented with 0.3% casamino acids, vitamins, calcium chloride, magnesium sulfate and 100 μM iron chelator EDDA. The fact that ZS-139 colonies do not grow beyond a certain distance from 5F-79 suggests that under these conditions, the non-producing strain is only able to grow due to the spotted producer.

2.5. Discussion

2.5.1. Patterns of siderophore production- information on iron in the marine micro-environment

Because of the dilute nature of the marine environment, the role of siderophores as a successful iron acquisition strategy in the ocean is often questioned (Volker & Wolf-Gladrow 1999; Hutchins & Rueter 1991). It has been suggested that siderophore production may only be beneficial in specific environments, where the collective behavior of siderophore producers, by increasing siderophore concentration in the local environment can counteract the effect of diffusion kinetics making siderophore re-uptake less likely (Volker & Wolf-Gladrow 1999). We show that for the group of *Vibrionaceae* studied, frequency of siderophore production centers around 40% (Figure 2.4). Interestingly, apart from a few populations where the frequency of siderophore production is close to 100%, siderophore production frequency also centers around middle values (30-60%) within populations (Figure 2.5). This suggests that, while siderophore production may be a superfluous and costly investment for specific organisms, it is a stable trait at the population level and therefore represents a worthwhile iron acquisition mechanism to maintain for this group of coastal vibrios.

Taking this study a step further, we were interested in determining whether siderophore production could be linked to micro-habitat preference. Indeed, marine particles constitute environments where marine microorganisms aggregate in large numbers, and have been hypothesized to be likely niches of siderophore producing marine bacteria (Hutchins & Rueter 1991). Contrary to our expectations, we found that siderophore production was positively correlated with *Vibrionaceae* isolates obtained from the <1μm sampling size-fraction, likely including strains that display a free-living lifestyle (Figure 2.7).

We propose a number of reasons that could help explain this surprising result. Iron is readily scavenged by particulate matter (Honeyman & Santschi 1989; Bruland & Lohan 2003), and is likely to be present at higher concentrations on particles than in the water column. This may preclude the need for siderophores, as microorganisms could utilize transiently available Fe^{2+}. Indeed, almost all of the 69 isolates for which a sequenced genome was available possess ferrous iron specific ABC transporters (Figure 2.9). Ferrous iron may also be more available within the context of biofilms, where inner layers may become

anaerobic (Stewart & Franklin 2008). In part II of this discussion, we propose that the negative correlation between siderophore production and large particle size (Figure 2.7) is driven by increased chances of encounters with closely related strains, facilitating the rise of *'cheaters'*. Finally, because our analysis does not include information on habitat-dependent siderophore production by microorganisms belonging to other families, we cannot rule out the possibility that overall siderophore production is indeed higher on particles, but performed by other microorganisms in the community. A large number of the isolates investigated possess a general catechol ABC transporter as well as (Figure 2.9), which could be utilized for exogenously produced siderophores.

Given the lack of siderophore biosynthetic genes in the genomes of the two most abundant free-living marine bacteria, *Prochlorococcus marinus* and *Candidatus Pelagibacter ubique*, the finding that siderophore production positively correlates with the <1μm fraction is all the more surprising. However, despite the fact that *P. marinus, Ca. P. ubique* and the *Vibrionaceae* isolates are obtained from the <1μm filter size-fraction, their lifestyles are extremely different. Indeed, both *P. marinus* and *Ca. P. ubique* are oligotrophs, adapted to consistently low nutrient environments, whereas vibrios are considered to be copiotrophs, poised to take advantage of transient nutrient patches (Lauro et al. 2009). Within this context, the evolutionary pressure to lose siderophore biosynthetic gene clusters may have been greater on the oligotrophs than the copiotrophs, which regularly find themselves in a high-cell environment during bursts of nutrients.

Knowledge of the various siderophore types produced by the *Vibrionaceae* (Figure 2.8) and their iron transporters (Figure 2.9) can provide information on the iron-environment experienced by these isolates. Comparing the complex formation constants of different siderophores with ferric iron can convey information about the relative availability of iron in various environments (Winkelmann 2004). For example, it is known that certain *Enterobacteriaceae* like *Escherichia coli* produce two types of siderophores: enterobactin with a strong affinity constant to iron, and aerobactin with a rather weak affinity constant to iron. In these cases, it has been suggested that aerobactin is excreted only under mild iron stress, whereas enterobactin is produced under severe iron limitation (Harris et al. 1979). Here, we find that the *Vibrionaceae*

isolates studied produce a variety of siderophores (Figure 2.8) with different ligand types and complex formation constants to iron (Table 2.2).

Table 2.2: Types and Strengths of Siderophores produced by the *Vibrionaceae* isolates investigated

Siderophore	Type of Chelating Functional Group	Proton-independent formation constant of ferric-siderophore complex
Enterobactin	catecholate	10^{-52} (Harris et al. 1979)
Yersiniabactin	5-member heterocycle	10^{-36} (Perry et al. 1999)
Anguibactin	mixed- catecholate, hydroxamate	unknown
Petrobactin	Citrate based catecholate	10^{-43} (Zhang et al. 2009)
Vulnibactin	catecholate	unknown
Vibrioferrin	carboxylate	10^{-24}(Amin et al. 2009a)
Aerobactin	mixed- carboxylate, hydroxamate	10^{-27}(Küpper et al. 2006)

The values for the ferric iron-siderophore complex formation constants vary by several orders of magnitude, from 10^{-24} for vibrioferrin to 10^{-52} for enterobactin. Interestingly, the siderophores with strong affinity to iron (enterobactin, petrobactin) tend to be produced by *Vibrionaceae* populations that are thought to be host-associated (Table 2.3). Indeed, the enterobactin synthesis genes are found in the genomes of isolates belonging to the *Vibrio ordalii* population cluster, while those for petrobactin synthesis are found in the *Enterovibrio* population cluster. Both populations are thought to be host associated (Chart & Trust 1984; Preheim et al. 2011). It is well known that within eukaryotic host environments, iron-binding proteins are part of the host-defense system against bacterial infections (Bullen et al., 2006). Our findings suggest that iron may be more tightly bound by such iron-binding proteins inside host environments than by iron-binding ligands present in ambient seawater. This confirms the well-accepted notion that pressure to evolve siderophore iron-transport systems is especially high for pathogens.

Table 2.3: Summary of vibrio populations and the corresponding siderophore synthesis genes identified in this study

Population	Population characteristic[1]	Siderophore type
Enterovibrio calviensis	Saprophytic life-style on live and dead zooplankton. Type I: on 64 μM particles, lower abundance in crabs and mussels.	None
Enterovibrio norveginus	Type III: Lower abundance on crab and mussels, higher on smaller particles and free-living. (pubmed search yields nothing?)	Yersiniabactin/ petrobactin
Vibrio sp. F9 – *V. breorganii*-like	Excluded from zooplankton. Found in mussels and crabs. Plant specialist. Type II	None
Allivibrio fischeri	Type II	Aerobactin
Vibrio ordalii	Type III: free living cells and smaller particles. Can grow at higher temperatures. Closely related to *V. cholerae*. Fish pathogen.	Enterobactin
Vibrio sp. F10	Type I Zooplankton specialist	Aerobactin
Vibrio crassostreae	Type II Plant specialist	No siderophore (1 vibrioferrin).
Vibrio cyclitrophicus		None
Vibrio splendidus	Variable	Vibrioferrin

[1.] Population characteristic were determined by Preheim 2010.

We find that within the group of *Vibrionaceae* isolates tested, the hydroxamate aerobactin and carboxylate vibrioferrin are the most widely represented types of siderophores (Figure 2.8). While many pathogens produce aerobactin, we show here that aerobactin is an important siderophore amongst marine *Vibrionaceae*, and especially amongst host-associated marine *Vibrionaceae*. Indeed, aerobactin synthesis clusters are found in *Vibrio* populations F10, thought to be zooplankton associated (either surface-attached or gut commensals) and *Allivibrio fischeri*, known to be associated with the squid *Euprymna scolopes* (Figure 2.8, Table 2.3). Vibrioferrin is found scattered across the *V. splendidus*-like clade, which consists of both free-living and particle associated population. Reduction of the ferric-siderophore complex by UV light has been demonstrated for both siderophores (Küpper et al. 2006; Amin et al. 2009b). The fact that these are common across the *Vibrionaceae* phylogeny contributes to support the new paradigm that photoreduction of

marine siderophores plays an important role in the cycling of iron in the ocean (Barbeau 2006). The recent finding that photoreactive vibrioferrin may be at the center of a mutualistic relationship between a dinoflagellate and a dinoflagellate-associated *Marinobacter* (Amin et al. 2009), suggests that aerobactin and vibrioferrin host-associated *Vibrio* producers may also provide bio-available iron to their algal hosts.

2.5.2. Patterns of siderophore production- imprints of microbe-microbe interactions.

In addition to contributing to our understanding of microbial iron acquisition in the marine environment, the data presented in this chapter provide a unique opportunity to explore how social evolutionary dynamics affect natural microbial communities. Siderophores mediate simple microbe-microbe interactions that may be thought of as illustrations of microbial cooperative behavior (Griffin et al. 2004). Because any cell that carries specific outer-membrane transporters can take up siderophores, these are considered a prime example of a *public good.* They can be exploited by *cheaters*, which do not bear the cost of production but nonetheless reap the benefits brought by the iron-siderophore complex. The *Vibrionaceae* isolates studied here are known to form ecologically cohesive populations (Hunt et al. 2008; Preheim, Timberlake, et al. 2011) within which we hypothesize that such cooperative interactions are likely to occur. Hence, the present study, which consists of an in-depth analysis of siderophore production in these closely-related *Vibrionaceae* isolates is well-suited to understanding the role of public good mediated interactions in natural systems.

Siderophores were amongst the first systems to be studied as public goods in *in vitro* experiments (Griffin et al. 2004; West & Buckling 2003). Most of these experiments were aimed at testing key social evolutionary biology concepts such as 'kin selection' theory (Hamilton 1964), as well as describing the conditions necessary for the maintenance of producers (*cooperators*) within a population (Chuang et al. 2009; Gore et al. 2009). For example, Chuang et al, 2009 shows that 'population structure' favors the rise of siderophore producers. In this experiment, 'population structure' is generated by the physical separation of different colonies in different test tubes (Chuang et al., 2009). It is unclear whether such artificial conditions used to model interactions between *cheaters* and producers are relevant within the natural environment and whether the constraints of the natural environment would be sufficient to favor cooperative behavior. Here,

we not only show that cooperative behavior mediated by public good interactions is observed within natural microbial populations, but also that it contributes to the microbial genetic diversity found in the environment.

One of the concepts that social evolutionary biologists have tried to tackle using synthetic microbial systems is that of 'collective action.' This can be summarized as follows: given the 'survival of the fittest' dogma, central tenet of the Darwinian theory of evolution, what is the incentive for microorganisms to produce costly siderophores if others can be the ones to incur the cost of production? Theory predicts that at high frequency of public good producers, there is a clear advantage in *cheating* and the frequency of *cheaters* in a population can be expected to rise (Nowak 2006). However, at high enough frequencies of *cheaters*, the public good would disappear, eventually leading to the collapse of the population. Hence, social evolutionary theory predicts a fine balance between frequencies of producers and *cheaters* (Nowak 2006). These predictions have been verified in synthetic microbial systems; most studies find that the public good trait stabilizes at intermediate frequencies within the population, with roughly equal representations of both producers and *cheaters* (Archetti et al. 2011). We show here that the average frequency of siderophore production across the ~1,000 *Vibrionaceae* isolates tested centers around 40% (Figures 2.4 and 2.5), suggesting that in nature as well, public good traits tend to stabilize at intermediate frequencies. Because the presence of a siderophore synthesis gene cluster corresponds well to the production of a siderophore as detected by the CAS assay (Figure 2.8), we assume that non-producers may potentially be *cheaters*. With this consideration in mind, our data suggests that for the *Vibrionaceae* isolates studied, 40% are siderophore producers while 60% are potential *cheaters*. This data provides the first confirmation that the theoretically predicted dynamics between producers and *cheaters* are validated in a natural system.

By showing that a number of *Vibrionaceae* non-producers actually have specific siderophore transporters (Figure 2.9), our analysis of iron-related transporters further suggests that public good dynamics can be observed in natural microbial populations. A number of isolates, while not producing any siderophore, have the outer-membrane transporters specific for aerobactin, vibrioferrin or vulnibactin transport (Figure 2.9). Upon closer inspection of these genomes, we found that the genes responsible for

siderophore synthesis have been excised, while the adjacent transporter genes have been retained (Figure 2.10). This could be interpreted as the rise of *cheaters* via the selective removal of costly siderophore synthesis genes. Confirming this hypothesis, putative *cheaters* that have lost the ability to produce siderophores showed a clear decrease in growth efficiency when grown in iron-poor media (Figure 2.11). By contrast, the growth of siderophore producers was similar in both iron-rich and iron-poor media (Figure 2.11). Moreover, when supplementing the iron-poor media with pure siderophores, putative *cheaters* were able to enhance their growth, showing that they could effectively take up the specific siderophore under those conditions (Figure 2.12). The growth enhancement effect was also observed when the cell-free supernatant of the siderophore producer was added to the putative *cheater* under iron-poor conditions (Figures 2.12 and 2.13), showing that the cross-feeding interaction is possible in nature. *In silico* predictions and physiological analysis confirm that non-producer strains, lacking siderophore biosynthesis genes, can act as *cheaters* that are dependent on exogenous siderophores, in contrast to producers, which are self-sufficient.

The finding that evolutionary transitions toward siderophore production in the *Vibrio splendidus*-like clade are negatively correlated with particle size (Figure 2.7) can also be interpreted in light of the dynamics between putative *cheaters* and producers. From a social evolutionary standpoint, the negative correlation between particle size and the relative frequency of siderophores suggests that the opportunity for *cheating* may be maximal in micro-environments where a large density of microorganisms with high encounter rates can be found, such as on large organic particles or organisms. One hypothesis for the observed trend is that higher local iron concentrations on large particles preclude the need for siderophore production. An alternative hypothesis is that the observed correlation is explained by differences in the structure of social interactions between particle-attached versus non-attached bacteria. The population dynamics of vibrios in aquatic environments is known to be punctuated by frequent blooms (Gilbert et al. 2012). Between blooms, some vibrios have a tendency to attach to particles while others remain planktonic (Hunt et al. 2008). Marine particles can harbor dense communities of attached bacteria with high metabolic activity (Simon et al. 2002), and therefore those vibrios attached to particles experience relatively constant high cell densities in their local environment. On the other hand those that remain free-living or attached to smaller

aggregates fluctuate from high cell densities during blooms, to low densities during their planktonic phase. This type of bottleneck population dynamics experienced by free-living vibrios has been shown to favor the abundance of cooperators in public good games (Chuang et al. 2009; Cremer et al. 2012). On the other hand, the constant high cell densities in the local environment of particle-attached bacteria could favor the establishment of cheaters that 'scrounge' public goods from producers (Nadell et al. 2009). This proposed model, based on the specific population dynamics associated to different micro-environments, provides an explanation for the variation in frequency of producers observed between particle attached and free-living vibrio.

Although our study is unable to detect other mechanisms (ie. regulatory repression of siderophore biosynthesis) potentially leading to the *cheating* phenotype, we find that the genomes analyzed act as records of the dynamics between producers and *cheaters*. These dynamics seem to favor the loss of genes responsible for the costly synthesis of siderophores, thereby introducing genetic diversity in groups of closely related *Vibrionaceae* isolates. Hence, in addition to suggesting that social interactions are common in natural populations, our results imply that population level processes such as cooperation can represent a major evolutionary force in natural populations, leaving its footprint on the gene content variability of microbial populations –and especially of those that are particle associated.

2.6. Conclusion

This study provides a unique insight in the role of siderophores in the marine environment at the micro-scale and population level. A high-throughput phenotypic and genomic approach allowed to establish that about half of the wild isolates investigated have the ability to produce siderophores. This suggests that other iron acquisition mechanisms, such as the utilization of exo-siderophores amongst others are prevalent, underlying the potential importance of interactions. Because the isolates investigated carry the genomic footprint of cooperative behavior occurring in the wild, we further suggest that microbe-microbe interactions are key drivers of iron acquisition in the marine environment. Finally, in addition to stressing the importance of microbial interactions in determining ecosystem function, this high throughput investigation of siderophore production by wild isolates provides unique insight into the bio-availability of iron in the micro-environments sampled.

2.7. Acknowledgements

I thank Otto X. Cordero for collaborating closely on this project. I also thank Martin Polz for valuable feedback and advice and for giving me complete access to his lab's collection of vibrios. Finally, thanks to Dr. Carl J. Carrano from San Diego State University for providing us with purified vibrioferrin.

2.8. References

Actis, L. a et al., 1986. Characterization of anguibactin, a novel siderophore from Vibrio anguillarum 775(pJM1). *Journal of Bacteriology*, 167(1), pp.57-65.

Altschul, S.F. et al., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 25, pp.3389-3402.

Amin, S. a, Green, D.H., Küpper, F.C. & Carrano, C.J., 2009a. Vibrioferrin, an unusual marine siderophore: iron binding, photochemistry, and biological implications. *Inorganic chemistry*, 48(23), pp.11451-8.

Amin, S. a, Green, D.H., Küpper, F.C. & Carrano, C.J., 2009b. Vibrioferrin, an unusual marine siderophore: iron binding, photochemistry, and biological implications. *Inorganic chemistry*, 48(23), pp.11451-8.

Amin, S.A., Green, D.H., Hart, M.C., Kupper, F.C., et al., 2009. Photolysis of iron–siderophore chelates promotes bacterial–algal mutualism. *Proceedings of the National Academy of Sciences*, 106(40), pp.901-910.

Archetti, M. et al., 2011. Economic game theory for mutualism and cooperation. *Ecology letters*, 14(12), pp. 1300-12.

Austin, B., 2011. Taxonomy of bacterial fish pathogens. *Veterinary research*, 42(1), p.20.

Azam, F. & Malfatti, F., 2007. Microbial structuring of marine ecosystems. *Nature reviews. Microbiology*, 5(10), pp.782-91.

Aziz, R.K. et al., 2008. The RAST Server: rapid annotations using subsystems technology. *BMC genomics*, 9, p.75.

Balado, M., Osorio, C.R. & Lemos, M.L., 2006. A gene cluster involved in the biosynthesis of vanchrobactin , a chromosome-encoded siderophore produced by Vibrio anguillarum. *Microbiology*, 152, pp.3517-3528.

Barbeau, K., 2006. Photochemistry of organic iron(III) complexing ligands in oceanic systems. *Photochemistry and photobiology*, 82(6), pp.1505-16.

Biosca, E.G. et al., 1996. Siderophore-mediated iron acquisition mechanisms in Vibrio vulnificus biotype 2. *Applied and environmental microbiology*, 62(3), pp.928-35.

Bruland, K.W. & Lohan, M.C., 2003. Controls of Trace Metals in Seawater. In H. Holland & K. Turekian, eds. *Treatise on Geochemistry*. Elsevier, pp. 23-47.

Bullen, J.J., Rogers, H.J., Spalding, P.B., Ward, C.G, 2006. Natural resistance, iron and infection: a challenge for clinical medicine. *Journal of Medical Microbiology*, 55, pp.251-58.

Cartron, M.L. et al., 2006. Feo--transport of ferrous iron into bacteria. *Biometals : an international journal on the role of metal ions in biology, biochemistry, and medicine*, 19(2), pp.143-57.

Challis, G.L., 2005. A widely distributed bacterial pathway for siderophore biosynthesis independent of nonribosomal peptide synthetases. *Chembiochem : a European journal of chemical biology*, 6(4), pp.601-11.

Chart, H. & Trust, T., 1984. Characterization of the surface antigens of the marine fish pathogens Vibrio anguillarum and Vibrio ordalii. *Canadian Journal of Microbiology*, 30, pp.703-710.

Chen, F. et al., 2006. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic acids research*, 34(Database issue), pp.D363-8.

Chuang, J.S., Rivoire, O. & Leibler, S., 2009. Simpson's Paradox in a Synthetic Microbial System. *Science*, 323(January), pp.272-275.

Computing, R.F. for S., 2012. R: A Language and Environment for Statistical Computing.

Conti, E. et al., 1997. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO Journal*, 16, pp.4174-83.

Cremer, J., Melbinger, A. & Frey, E., 2012. Growth dynamics and the evolution of cooperation in microbial populations. *Scientific reports*, 2, p.281.

Crosa, J.H., 1980. A plasmid associated with virulence in the marine fish pathogen Vibrio anguillarum specifies an iron sequestering system. *Nature*, pp.566-568.

Crosa, J.H. & Walsh, C.T., 2002. Genetics and Assembly Line Enzymology of Siderophore Biosynthesis in Bacteria. *Microbiology and Molecular Biology Reviews*, 66(2), pp.223-249.

DeLong, E.F. et al., 2006. Community genomics among stratified microbial assemblages in the ocean's interior. *Science (New York, N.Y.)*, 311(5760), pp.496-503.

Earhart, C.F., 1987. Ferrienterobactin transport in Escherichia coli. In Günther Winkelmann, D. van der Helm, & J. Neilands, eds. *Iron transport in microbes, plants and animals.* Weinheim, Germany: VCH Verlagsgesellschaft, pp. 67-84.

Eilers, H, Pernthaler, J & Amann, R., 2000. Succession of pelagic marine bacteria during enrichment: a close look at cultivation-induced shifts. *Applied and environmental microbiology*, 66(11), pp.4634-40.

Eilers, Heike, Pernthaler, Jakob, et al., 2000. Culturability and In Situ Abundance of Pelagic Bacteria from the North Sea. *Applied and Environmental Microbiology*, 66(7), pp.3044-3051.

Feingersch, R. et al., 2010. Microbial community genomics in eastern Mediterranean Sea surface waters. *The ISME journal*, 4(1), pp.78-87.

Felsenstein, J., 1985. Phylogenies and the Comparative Method. *The American Naturalist*, 125(1), pp.1-15.

Gauglitz, J.M., Zhou, H. & Butler, A., 2012. A suite of citrate-derived siderophores from a marine Vibrio species isolated following the Deepwater Horizon oil spill. *Journal of inorganic biochemistry*, 107(1), pp. 90-5.

Gilbert, J. a et al., 2012. Defining seasonal marine microbial community dynamics. *The ISME journal*, 6(2), pp.298-308.

Gilbert, J. a et al., 2010. The taxonomic and functional diversity of microbes at a temperate coastal site: a "multi-omic" study of seasonal and diel temporal variation. *PloS one*, 5(11), p.e15545.

Giovannoni, S.J. et al., 2005. Genome streamlining in a cosmopolitan oceanic bacterium. *Science (New York, N.Y.)*, 309(5738), pp.1242-5.

Gore, J., Youk, H. & van Oudenaarden, A., 2009. Snowdrift game dynamics and facultative cheating in yeast. *Nature*, 459(7244), pp.253-6.

Griffin, A.S., West, S.A. & Buckling, A., 2004. Cooperation and competition in pathogenic bacteria. *Nature*, 430, pp.1024-1027.

Griffiths, G.L. et al., 1984. Vibriobactin , a Siderophore from Vibrio cholerae. *J. Biol. Chem.*, 259(1), pp. 383-385.

Hamilton, W.D., 1964. The genetical evolution of social behaviour. I. *Journal of theoretical biology*, 7(1), pp. 1-16.

Harris, W.R. et al., 1979. Coordination chemistry of microbial iron transport compounds. 19. Stability constants and electrochemical behavior of ferric enterobactin and model complexes. *Journal of the American Chemical Society*, 101(20), pp.6097-6104.

Haygood, M.G., Holt, P.D. & Butler, A., 1993. Aerobactin production by a planktonic marine Vibrio sp. *Limnology and Oceanography*, 38(5), pp.1091-1097.

Homann, V.V. et al., 2009. Siderophores of Marinobacter aquaeolei: petrobactin and its sulfonated derivatives. *Biometals*, 22(4), pp.565-571.

Honeyman, B.D. & Santschi, P.., 1989. A Brownian pumping model for trace metal scavenging: evidence from Th isotopes. *Journal of Marine Research*, 47, pp.951-992.

Hunt, D.E. et al., 2008. Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science (New York, N.Y.)*, 320(5879), pp.1081-5.

Hutchins, D.A. & Rueter, J.G., 1991. Siderophore production and nitrogen fixation are mutually exclusive strategies in Anabaena 7120. *Limnology and Oceanography*, 36, pp.1-12.

Küpper, F.C. et al., 2006. Photoreactivity of iron(III)-aerobactin: photoproduct structure and iron(III) coordination. *Inorganic chemistry*, 45(15), pp.6028-33.

Lauro, F.M. et al., 2009. The genomic basis of trophic strategy in marine bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 106(37), pp.15527-33.

Letunic, I. & Bork, P., 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Nucleic acids research*, 23(1), pp.127-128.

Lorenzo, V.D.E. et al., 1986. Aerobactin Biosynthesis and Transport Genes of Plasmid Co1V-K30 in Escherichia coli K-12. *Journal of Bacteriology*, 165(2), pp.570-578.

López, C.S. & Crosa, J.H., 2007. Characterization of ferric-anguibactin transport in Vibrio anguillarum. *Biometals : an international journal on the role of metal ions in biology, biochemistry, and medicine*, 20(3-4), pp. 393-403.

Martin, J.D. et al., 2006. Structure and membrane affinity of new amphiphilic siderophores produced by Ochrobactrum sp. SP18. *Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry*, 11(5), pp.633-41.

Martinez, J.S. et al., 2003. Structure and membrane affinity of a suite of amphiphilic siderophores produced by a marine bacterium. *Proceedings of the National Academy of Sciences of the United States of America*, 100(7), pp.3754-9.

Medema, M.H. et al., 2011. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Research*, 39(Web Server), p.W339-W346.

Murakami, K. et al., 1998. Phylogenetic analysis of marine environmental strains of Vibrio that produce aerobactin. *Journal of Marine Biotechnology*, 6, pp.76-79.

Nadell, C.D., Xavier, J.B. & Foster, K.R., 2009. The sociobiology of biofilms. *FEMS microbiology reviews*, 33(1), pp.206-24.

Nowak, M. a, 2006. *Evolutionary Dynamics*, Cambridge, MA, USA: Belknap Press of Harvard University Press.

Nyholm, S.V. et al., 2009. Recognition between symbiotic Vibrio fischeri and the hemocytes of Euprymna scolopes. *Environmental microbiology*, 11(2), pp.483-493.

Nyholm, S.V. & McFall-Ngai, M.J., 2004. The winnowing: establishing the squid-vibrio symbiosis. *Nature reviews. Microbiology*, 2(8), pp.632-42.

Okujo, N & Yamamoto, S, 1994. Identification of the siderophores from Vibrio hollisae and Vibrio mimicus as aerobactin. *FEMS microbiology letters*, 118(1-2), pp.187-92.

Okujo, Noriyuki et al., 1996. Involvement of Vulnibactin in Utilization of Transferrin- and Lactoferrin-Bound Iron by Vibrio vulnificus. *Microbiol. Immunol.*, 40(8), pp.595-598.

Overbeek, R. et al., 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic acids research*, 33(17), pp.5691-702.

Oves-Costales, D., Kadi, N. & Challis, G.L., 2009. The long-overlooked enzymology of a nonribosomal peptide synthetase-independent pathway for virulence-conferring siderophore biosynthesis. *Chemical communications (Cambridge, England)*, (43), pp.6530-41.

Perry, R.D. et al., 1999. Yersiniabactin from Yersinia pestis: biochemical characterization of the siderophore and its role in iron transport and regulation. *Microbiology*, 145, pp.1181-1190.

Preheim, S.P., 2010. *Ecology and population structure of Vibrionaceae in the coastal ocean.* Woods Hole, MA: Massachusetts Institute of Technology and Woods Hole Oceanographic Institution.

Preheim, S.P., Boucher, Y., et al., 2011. Metapopulation structure of Vibrionaceae among coastal marine invertebrates. *Environmental microbiology*, 13(1), pp.265-75.

Preheim, S.P., Timberlake, S. & Polz, M.F., 2011. Merging taxonomy with ecological population prediction in a case study of Vibrionaceae. *Applied and environmental microbiology*, 77(20), pp.7195-206.

Punta, M. et al., 2012. The Pfam protein families database. *Nucleic acids research*, 40(Database issue), pp.D290-301.

Ravel, J. & Cornelis, P., 2003. Genomics of pyoverdine-mediated iron uptake in pseudomonads. *Trends in Microbiology*, 11(5), pp.192-195.

Reen, F.J. et al., 2006. The genomic code: inferring Vibrionaceae niche specialization. *Nature reviews. Microbiology*, 4(9), pp.697-704.

Rocap, G. et al., 2003. Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation. *Nature*, 424(6952), pp.1042-7.

Ruby, E.G., 1998. MINIREVIEW The Vibrio fischeri-Euprymna scolopes Light Organ Association : Current Ecological Paradigms. , 64(3), pp.805-812.

Rusch, D.B. et al., 2007. The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS biology*, 5(3), p.e77.

Röttig, M. et al., 2011. NRPSpredictor2--a web server for predicting NRPS adenylation domain specificity. *Nucleic acids research*, 39(May), pp.362-367.

Sandy, M. et al., 2010. Vanchrobactin and anguibactin siderophores produced by Vibrio sp. DS40M4. *Journal of natural products*, 73(6), pp.1038-43.

Sandy, M. & Butler, A., 2009. Microbial Iron Acquisition: Marine and Terrestrial Siderophores. *Chem. Rev.*, 109(10), pp.4580-4595.
Schwyn, B. & Neilands, J., 1987. Universal Chemical Assay for the Detection and Determination of Siderophores. *Analytical Biochemistry*, 160, pp.47-56.
Sigel, S.P., Stoebner, J. a & Payne, S.M., 1985. Iron-vibriobactin transport system is not required for virulence of Vibrio cholerae. *Infection and immunity*, 47(2), pp.360-2.
Simon, M. et al., 2002. Microbial ecology of organic aggregates in aquatic ecosystems. *Aquatic Microbial Ecology*, 28, pp.175-211.
Soengas, R.G. et al., 2006. Structural characterization of vanchrobactin, a new catechol siderophore produced by the fish pathogen Vibrio anguillarum serotype O2. *Tetrahedron Letters*, 47(39), pp. 7113-7116.
Stachelhaus, T., Mootz, H. & Marahiel, MA, 1999. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chemistry & biology*, 6, pp.493-505.
Stewart, P.S. & Franklin, M.J., 2008. Physiological heterogeneity in biofilms. *Nature reviews. Microbiology*, 6(3), pp.199-210.
Strieker, M., Tanović, A. & Marahiel, Mohamed a, 2010. Nonribosomal peptide synthetases: structures and dynamics. *Current opinion in structural biology*, 20(2), pp.234-40.
Suzuki, K. et al., 2006. Identification and transcriptional organization of aerobactin transport and biosynthesis cluster genes of Vibrio hollisae. *Research in microbiology*, 157(8), pp.730-40.
Suzuki, M.T. et al., 2004. Phylogenetic screening of ribosomal RNA gene-containing clones in Bacterial Artificial Chromosome (BAC) libraries from different depths in Monterey Bay. *Microbial ecology*, 48(4), pp.473-88.
Tanabe, T., Funahashi, T., Nakao, H., Miyoshi, S.-ichi, Shinoda, S. & Yamamoto, Shigeo, 2003a. Identification and Characterization of Genes Required for Biosynthesis and Transport of the Siderophore Vibrioferrin in Vibrio parahaemolyticus. *Journal of Bacteriology*, 185, pp.6938-6949.
Tanabe, T., Funahashi, T., Nakao, H., Miyoshi, S.-ichi, Shinoda, S. & Yamamoto, Shigeo, 2003b. Identification and Characterization of Genes Required for Biosynthesis and Transport of the Siderophore Vibrioferrin in Vibrio parahaemolyticus Identification and Characterization of Genes Required for Biosynthesis and Transport of the Siderophore Vibriof. *Society.*
Thompson, F.L., Iida, T. & Swings, J., 2004. Biodiversity of Vibrios. *Microbiology and Molecular Biology Reviews*, 68(3), pp.403-431.
Toulza, E. et al., 2012. Analysis of the Global Ocean Sampling (GOS) Project for Trends in Iron Uptake by Surface Ocean Microbes F. Rodriguez-Valera, ed. *PLoS ONE*, 7(2), p.e30931.
Volker, C. & Wolf-gladrow, D.A., 1999. Physical limits on iron uptake mediated by siderophores or surface reductases. *Marine Chemistry*, 65, pp.227-244.
Vraspir, J.M. & Butler, A., 2009. Chemistry of Marine Ligands and Siderophores. *Annual Review of Marine Science*, 1, pp.43-63.
Vraspir, J.M., Holt, P.D. & Butler, A., 2011. Identification of new members within suites of amphiphilic marine siderophores. *Biometals*, 24(1), pp.85-92.
Wandersman, C. & Delepelaire, P., 2004. Bacterial iron sources: from siderophores to hemophores. *Annual review of microbiology*, 58, pp.611-47.
Wang, Q. et al., 2008. Characterization of two TonB systems in marine fish pathogen Vibrio alginolyticus: their roles in iron utilization and virulence. *Archives of microbiology*, 190(5), pp.595-603.
Wang, Q. et al., 2007. Isolation, sequencing and characterization of cluster genes involved in the biosynthesis and utilization of the siderophore of marine fish pathogen Vibrio alginolyticus. *Archives of microbiology*, 188(4), pp.433-9.
West, S. a & Buckling, A., 2003. Cooperation, virulence and siderophore production in bacterial parasites. *Proceedings. Biological sciences / The Royal Society*, 270(1510), pp.37-44.
Winkelmann, Gunther, 2004. Ecology of Siderophores. In J. H. Crosa, A. R. Mey, & Shelley M Payne, eds. *Iron Transport in Bacteria.* Washington, DC: ASM Press, pp. 437-450.
Wright, A.C., Simpson, L.M. & Oliver, J.D., 1981. Role of Iron in the Pathogenesis of Vibrio vulnificus Infections. *Infection and immunity*, 34(2), pp.503-507.

Yamamoto, S et al., 1992. Isolation and partial characterization of a compound with siderophore activity from Vibrio parahaemolyticus. *FEMS microbiology letters*, 73(1-2), pp.181-6.
Yamamoto, Shigeo, Okujo, Z.N., et al., 1994. Structure and Iron Transport Activity of Vibrioferrin, a New Siderophore of Vibrio parahaemolyticus. *Journal of Biochemistry*, 115, pp.868-874.
Yamamoto, Shigeo, Okujo, Noriyuki, et al., 1994. Siderophore-Mediated Utilization of Iron Bound to Transferrin by Vibrio parahaemolyticus. *Microbiology and Immunology*, 38(9), pp.687-693.
Yamamoto, Shigeo et al., 1993. Structures of 2 Polyamine Containing Catecholate Siderophores From Vibrio Fluvialis. *Journal of Biochemistry*, 113(5), pp.538-544.
Zhang, G. et al., 2009. Ferric Stability Constants of Representative Marine Siderophores: Marinobactins, Aquachelins, and Petrobactin. *Inorg. Chem.*, 48(23), pp.11466-11473.

2.9. Supplementary Materials

2.9.1. Table S2.1: Results of the Siderophore Screen

Table S1: Results of the Siderophore Screen

We only show the results for the 1,012 strains (out of 1,711 screened) that grew either in liquid or on the CAS agar plates. When "Sampled Environment" is indicated by numbers, read as: 64, obtained from the >64 μm fraction; 5, >5 μm but < 64 μm; 1, >1 μm but < 5 μm; and 0.2, > 0.2 μm but < 1 μm (equated to "free-living"). "Growth in M9" shows the absorbance taken at 600 nm after 48 hours growth in iron poor media (modified M9 with 10 nM $FeCl_3$ and 100 μM 2,2'-bipyridyl) for two biological replicates. "Liquid CAS" records the outcome of the liquid CAS assay, and represents the absorbance at 630 nm after mixing the supernatant of overnight cultures with the CAS dye and shuttle solution. "CAS growth" represents whether cells grew on the CAS agar. "CAS halo" records whether colonies had an orange halo on the CAS agar, "++" means that the halo was very large, "+" means that the halo was clearly distinguishable, "+-" means that the halo was not clearly distinguishable, "-" means that a halo was clearly absent. "Siderophore Production" is the production data resolved from the outcomes of both the liquid and solid CAS assays. In order to read "yes", at least two siderophore positives were needed, which meant an absorbance < 0.3 for the liquid CAS or a "++", "+" or "+-" for the solid CAS assay.

Strain Name	Sampled Environment	Sampled Season	Population	Siderophore Production	Growth in M9, 1	Growth in M9, 2	liquid CAS,1	liquid CAS, 2	CAS growth, 1	CAS growth, 2	CAS halo, 1	CAS halo, 2
0407CG108	crab gill		NA	no	0.245	0.143	0.458	0.532	(+)	no	no	no
0407CG113	mussel hind gut		16	no	0.37	0.217	0.414	0.534	(+)	(+)	no	no
0407CG114	mussel hind gut		16	no	0.153	0.28	0.441	0.529	(+)	(+)	no	no
0407CG115	mussel hind gut		8	no	0.395	0.245	0.166	0.313	(+)	(+)	no	no
0407CG117	mussel hind gut		16	no	0.179	0.232	0.427	0.531	(+)	(+)	no	no
0407CG118	mussel hind gut		16	no	0.398	0.138	0.446	0.526	(+)	(+-)	no	no
0407CG120	mussel hind gut		16	no	0.154	0.197	0.39	0.514	(+)	(+)	no	no
0407CG121	crab gill		16	yes	1.111	0.762	0.159	0.348	(+)	(+)	(+-)	no
0407CG123	crab gill		NA	yes	0.121	0.801	0.169	0.323	(+)	(+)	(+-)	no
0407CG125	crab gill		NA	no	0.164	0.125	0.144	0.386	(+)	(+)	no	no
0407CG132	crab gill		16	no	0.289	0.137	0.474	0.523	(+)	no	no	no
0407CG133	crab gill		16	no	0.566	0.228	0.434	0.552	(+)	(+-)	no	no
0407CG136	crab gill		16	no	0.135	0.136	0.456	0.413	(+)	(+)	(+-)	no
0407CG145	crab gill		16	yes	0.152	0.159	0.159	0.167	(+)	(+)	(+)	(+-)
0407CG154	crab gill		2	yes	0.104	0.105	0.153	0.17	(+)	(+)	(+-)	(++)
0407CG81	crab gill		16	no	0.191	0.193	0.232	0.335	no	no	no	no
0407CS130	crab stomach		2	yes	0.1	0.11	0.17	0.181	(+)	(+)	(+)	(++)
0407CS145	crab stomach		16	yes	0.44	0.218	0.142	0.146	(+)	(+)	(++)	(++)
0407CS146	crab stomach		16	yes	0.183	0.177	0.152	0.18	(+)	(+)	(++)	(++)
0407CS152	crab stomach		2	yes	0.107	0.112	0.139	0.211	(+)	(+)	(+)	(+)
0407CS153	crab stomach		16	yes	0.456	0.217	0.144	0.166	no	no	no	no
0407CS154	crab stomach		16	no	0.159	0.157	0.453	0.447	(+)	(+)	no	no
0407CSC156	crab stomach		16	no	0.169	0.226	0.428	0.519	(+)	(+)	no	no
0407CSC159	crab stomach		NA	no	0.621	0.174	0.493	0.531	(+-)	no	no	no
0407CSC160	crab stomach		16	no	0.166	0.216	0.402	0.53	(+)	(+)	no	no
0407MG1	mussel gill		16	no	0.455	0.374	0.481	0.575	(+)	(+)	no	no
0407MG10	mussel gill		NA	yes	0.425	0.375	0.135	0.316	(+)	(+)	(+-)	(+)
0407MG11	mussel gill		16	yes	0.165	0.207	0.134	0.332	(+)	(+)	(++)	(+)
0407MG12	mussel gill		NA	no	0.175	0.267	0.146	0.33	no	no	no	no
0407MG13	mussel gill		16	no	0.41	0.2	0.136	0.332	no	no	no	no
0407MG15	mussel gill		NA	yes	0.505	0.318	0.145	0.341	(+)	(+)	(+-)	(+)
0407MG17	mussel gill		16	yes	0.367	0.208	0.141	0.322	(+)	(+)	(+)	(+)
0407MG18	mussel gill		16	yes	0.8	0.204	0.135	0.326	(+)	no	(+)	no
0407MG2	mussel gill		16	yes	0.388	0.246	0.185	0.35	(+)	(+)	(+-)	(+)
0407MG20	mussel gill		16	yes	0.156	0.214	0.136	0.336	(+)	(+)	(+)	(+)
0407MG21	mussel gill		NA	no	0.183	0.215	0.142	0.331	no	no	no	no
0407MG22	mussel gill		16	no	0.305	0.267	0.443	0.552	(+)	(+)	no	no
0407MG23	mussel gill		16	no	0.352	0.408	0.139	0.328	no	no	no	no
0407MG24	mussel gill		NA	no	0.388	0.208	0.146	0.341	no	no	no	no
0407MG25	mussel gill		16	yes	0.149	0.257	0.14	0.323	(+)	no	(+)	no
0407MG28	mussel gill		NA	no	0.159	0.252	0.167	0.337	(+)	(+)	no	no
0407MG29	mussel gill		16	no	0.157	0.496	0.134	0.324	no	no	no	no
0407MG30	mussel gill		NA	yes	0.439	0.222	0.139	0.355	no	(+)	no	(+-)
0407MG31	mussel gill		NA	no	0.63	0.36	0.142	0.338	no	(+)	no	no
0407MG33	mussel gill		16	no	0.376	0.188	0.141	0.344	no	no	no	no
0407MG34	mussel gill		NA	yes	0.426	0.513	0.209	0.323	(+)	(+)	no	(+)
0407MG35	mussel gill		16	yes	0.388	0.252	0.138	0.322	(+)	(+)	(+)	(+)

0407MG37	mussel gill		NA	yes	0.182	0.231	0.138	0.34	(+)	(+)	(++)	(+)
0407MG38	mussel gill		16	no	0.156	0.201	0.134	0.317	no	no	no	no
0407MG39	mussel gill		16	no	0.587	0.178	0.486	0.574	(+)	(+)	no	no
0407MG4	mussel gill		16	no	0.163	0.242	0.166	0.315	(+)	(+)	no	no
0407MG40	mussel gill		16	no	0.319	0.201	0.141	0.496	no	no	no	no
0407MG5	mussel gill		16	no	0.415	0.206	0.139	0.335	(+)	no	no	no
0407MG6	mussel gill		16	no	0.384	0.528	0.139	0.32	no	no	no	no
0407MG9	mussel gill		NA	no	0.594	0.34	0.493	0.583	(+)	(+)	no	no
0407ZA10	zooplankton		16	no	0.165	0.148	0.388	0.431	(+)	(+)	no	no
0407ZA12	zooplankton		16	no	0.234	0.164	0.439	0.392	(+)	(+)	no	no
0407ZA13	zooplankton		16	no	1.289	0.182	0.451	0.392	(+)	(+)	no	no
0407ZA14	zooplankton		16	no	0.159	0.162	0.451	0.454	(+)	(+)	no	no
0407ZA15	zooplankton		16	yes	0.191	0.195	0.23	0.178	(+)	(+)	no	no
0407ZA16	zooplankton		16	yes	0.176	0.18	0.247	0.195	(+)	(+)	no	no
0407ZA18	zooplankton		16	no	0.134	0.134	0.412	0.447	(+)	(+)	no	no
0407ZA2	zooplankton		8	yes	0.223	0.219	0.165	0.151	(+)	(+)	no	no
0407ZA20	zooplankton		16	no	0.309	0.159	0.47	0.407	(+)	no	no	no
0407ZA21	zooplankton		16	no	0.144	0.161	0.451	0.384	(+)	(+)	no	no
0407ZA22	zooplankton		16	no	0.167	0.178	0.437	0.417	(+)	(+)	no	no
0407ZA23	zooplankton		16	yes	0.164	0.164	0.181	0.161	(+)	(+)	no	no
0407ZA24	zooplankton		16	yes	0.165	0.169	0.284	0.23	(+)	(+)	no	no
0407ZA26	zooplankton		16	yes	0.169	0.155	0.172	0.161	(+)	(+)	no	no
0407ZA27	zooplankton		16	no	0.357	0.144	0.431	0.435	(+)	(+)	no	no
0407ZA28	zooplankton		16	yes	0.891	0.164	0.172	0.176	(+)	(+)	no	no
0407ZA29	zooplankton		16	no	0.389	0.142	0.45	0.394	(+)	(+-)	no	no
0407ZA30	zooplankton		16	no	0.144	0.147	0.422	0.405	(+)	(+)	no	no
0407ZA31	zooplankton		16	yes	0.16	0.163	0.188	0.161	(+)	(+)	no	no
0407ZA32	zooplankton		16	yes	0.171	0.17	0.294	0.247	(+)	(+)	no	no
0407ZA34	zooplankton		16	no	0.16	0.152	0.426	0.418	(+)	(+)	no	no
0407ZA36	zooplankton		NA	yes	0.275	0.189	0.188	0.189	(+)	(+)	no	no
0407ZA37	zooplankton		NA	no	0.137	0.177	0.406	0.358	(+)	(+)	no	no
0407ZA38	zooplankton		16	yes	0.169	0.173	0.257	0.224	(+)	(+)	no	no
0407ZA39	zooplankton		16	yes	0.168	0.164	0.253	0.21	(+)	(+)	no	no
0407ZA4	zooplankton		NA	no	0.208	0.152	0.444	0.5	(+)	(+-)	no	no
0407ZA5	zooplankton		16	no	0.523	0.172	0.372	0.293	(+)	(+)	no	no
0407ZA6	zooplankton		16	no	0.143	0.151	0.48	0.467	(+)	(+)	no	no
0407ZA7	zooplankton		16	yes	0.186	0.196	0.23	0.178	(+)	(+)	no	no
0407ZA8	zooplankton		16	yes	0.18	0.185	0.215	0.177	(+)	(+)	no	no
0407ZB16	zooplankton		16	yes	0.19	0.194	0.216	0.166	(+)	(+)	no	no
0407ZB18	zooplankton		16	yes	0.184	0.179	0.162	0.158	(+)	(+)	(+-)	(+-)
0407ZB19	zooplankton		16	no	0.316	0.14	0.426	0.433	(+-)	(+)	no	no
0407ZB33	zooplankton		NA	yes	0.156	0.152	0.273	0.186	(+)	(+)	no	(+-)
0407ZB36	zooplankton		NA	yes	0.208	0.181	0.19	0.162	(+)	(+)	no	no
0407ZB38	zooplankton		16	yes	0.18	0.182	0.206	0.161	(+)	(+)	no	no
0407ZC14	zooplankton		NA	yes	0.176	0.164	0.235	0.186	(+)	(+)	no	no
0407ZC3	zooplankton		16	yes	0.221	0.195	0.155	0.135	(+)	(+)	no	no
0407ZC6	zooplankton		NA	yes	0.18	0.185	0.196	0.156	(+)	(+)	no	no
0407ZC9	zooplankton		NA	no	0.122	0.131	0.38	0.251	(+)	(+)	no	no
1F_101	1	Fall	NA	yes	0.193	0.19	0.145	0.146	(+)	(+)	(+)	(+)
1F_104	1	Fall	2	no	0.24	0.204	0.455	0.426	no	no	no	no
1F_105	1	Fall	NA	no	0.234	0.207	0.466	0.43	(+-)	no	no	no

1F_110	1	Fall	16	no	0.208	0.206	0.482	0.43	(+)	(+)	no	no
1F_111	1	Fall	15B	no	0.225	0.218	0.347	0.33	no	no	no	no
1F_115	1	Fall	2	no	0.249	0.211	0.434	0.453	no	no	no	no
1F_128	1	Fall	3	yes	0.11	0.107	0.509	0.507	(+)	(+)	(++)	(++)
1F_157	1	Fall	18	yes	0.136	0.117	0.231	0.213	(+)	(+)	(++)	(++)
1F_165	1	Fall	13	no	0.089	0.085	0.524	0.528	(+)	(+)	no	no
1F_169	1	Fall	16	yes	0.094	0.092	0.521	0.523	(+)	(+)	(+-)	(+-)
1F_17	1	Fall	2	no	0.206	0.206	0.432	0.427	no	no	no	no
1F_185	1	Fall	16	no	0.072	0.072	0.53	0.538	(+)	(+)	(+-)	no
1F_20	1	Fall	2	no	0.267	0.241	0.448	0.439	no	no	no	no
1F_203	1	Fall	13	yes	0.095	0.094	0.527	0.524	(+)	(+)	(+-)	(+-)
1F_215	1	Fall	1	no	0.076	0.077	0.538	0.516	(+)	(+)	(+-)	no
1F_221	1	Fall	2	yes	0.235	0.232	0.222	0.213	(+)	(+)	(+)	(++)
1F_234	1	Fall	1	yes	0.116	0.112	0.478	0.499	(+)	(+)	(+)	(+-)
1F_236	1	Fall	3	yes	0.138	0.129	0.51	0.527	(+)	(+)	(+)	(++)
1F_24	1	Fall	2	no	0.191	0.172	0.486	0.469	no	no	no	no
1F_243	1	Fall	13	yes	0.103	0.101	0.459	0.456	(+)	(+)	(+-)	(+-)
1F_247	1	Fall	3	yes	0.156	0.243	0.454	0.354	(+)	(+)	(+)	(++)
1F_25	1	Fall	NA	no	0.13	0.116	0.553	0.533	(+)	(+)	no	no
1F_255	1	Fall	13	no	0.097	0.091	0.53	0.53	(+)	(+)	no	no
1F_259	1	Fall	3	yes	0.165	0.165	0.399	0.448	(+)	(+)	(+-)	(++)
1F_260	1	Fall	2	no	0.107	0.105	0.513	0.492	(+)	(+)	(+-)	no
1F_261	1	Fall	2	yes	0.117	0.116	0.451	0.461	(+)	(+)	(+-)	(++)
1F_269	1	Fall	16	yes	0.098	0.099	0.489	0.5	(+)	(+)	(+-)	(+-)
1F_273	1	Fall	NA	yes	0.109	0.102	0.484	0.494	(+)	(+)	(+-)	(++)
1F_279	1	Fall	16	yes	0.105	0.101	0.462	0.496	(+)	(+)	(+)	(+-)
1F_28	1	Fall	8	yes	0.238	0.218	0.163	0.149	(+)	(+)	(+-)	(+-)
1F_280	1	Fall	2	no	0.122	0.117	0.549	0.444	(+)	(+)	no	(+-)
1F_287	1	Fall	2	yes	0.102	0.104	0.514	0.471	(+)	(+)	(+-)	(+-)
1F_29	1	Fall	22	yes	0.183	0.174	0.136	0.14	(+)	(+)	(++)	(++)
1F_292	1	Fall	16	no	0.267	0.206	0.384	0.378	(+)	(+)	no	no
1F_293	1	Fall	16	no	0.163	0.155	0.444	0.428	(+)	(+)	no	no
1F_295	1	Fall	5	no	0.187	0.204	0.294	0.44	(+)	(+)	no	no
1F_297	1	Fall	NA	no	0.158	0.161	0.428	0.548	no	(+)	no	no
1F_303	1	Fall	16	no	0.189	0.194	0.418	0.451	no	no	no	no
1F_308	1	Fall	16	no	0.192	0.186	0.4	0.382	no	no	no	no
1F_31	1	Fall	16	no	0.121	0.118	0.519	0.502	(+)	(+)	no	no
1F_310	1	Fall	NA	no	0.165	0.165	0.144	0.447	no	no	no	no
1F_315	1	Fall	10	yes	0.132	0.142	0.176	0.432	(+)	(+)	(+-)	(+-)
1F_317	1	Fall	NA	no	0.118	0.12	0.637	0.435	(+)	(+)	no	no
1F_33	1	Fall	13	no	0.228	0.248	0.475	0.457	(+)	no	no	no
1F_36	1	Fall	2	no	0.256	0.224	0.456	0.46	no	no	no	no
1F_38	1	Fall	NA	yes	0.546	0.543	0.184	0.167	(+)	(+)	(+)	(+)
1F_39	1	Fall	2	no	0.264	0.2	0.442	0.431	no	no	no	no
1F_4	1	Fall	2	no	0.227	0.213	0.479	0.449	(+-)	no	no	no
1F_42	1	Fall	16	no	0.221	0.216	0.483	0.475	(+)	no	no	no
1F_53	1	Fall	15B	no	0.209	0.196	0.452	0.41	(+)	(+-)	(+-)	no
1F_55	1	Fall	18	yes	0.201	0.499	0.153	0.15	(+)	(+)	(+)	(+)
1F_57	1	Fall	NA	no	0.202	0.354	0.428	0.411	(+)	(+)	no	no
1F_63	1	Fall	16	no	0.24	0.239	0.406	0.398	(+)	(+-)	no	no
1F_66	1	Fall	2	no	0.197	0.173	0.487	0.47	no	no	no	no
1F_69	1	Fall	3	yes	0.311	0.312	0.265	0.295	(+)	(+)	(++)	(++)
1F_71	1	Fall	2	no	0.165	0.156	0.493	0.469	(+-)	(+-)	no	no

1F_77	1	Fall	10	no	0.153	0.163	0.545	0.511	(+)	no	no	no
1F_79	1	Fall	2	no	0.197	0.174	0.506	0.451	no	no	no	no
1F_81	1	Fall	2	no	0.214	0.195	0.471	0.463	no	no	no	no
1F_9	1	Fall	16	no	0.203	0.197	0.44	0.407	no	no	no	no
1F_90	1	Fall	2	no	0.198	0.192	0.447	0.42	no	no	no	no
1F_92	1	Fall	2	no	0.248	0.228	0.448	0.428	(+)	(+-)	no	no
1F_94	1	Fall	2	no	0.187	0.177	0.465	0.446	no	no	no	no
1F_97	1	Fall	15B	no	0.22	0.232	0.439	0.405	(+)	no	no	no
1F_99	1	Fall	2	no	0.204	0.183	0.44	0.472	(+-)	no	no	no
1S_105	1	Spring	11	no	0.114	0.113	0.493	0.476	(+)	(+)	no	no
1S_113	1	Spring	23	no	0.114	0.153	0.471	0.457	(+)	(+)	no	no
1S_120	1	Spring	14	no	0.113	0.123	0.437	0.423	(+)	(+)	no	(++)
1S_123	1	Spring	3	yes	0.226	0.171	0.376	0.39	(+)	(+)	(+-)	(+)
1S_128	1	Spring	7	yes	0.089	0.091	0.32	0.313	(+)	(+)	(+)	(+-)
1S_129	1	Spring	24A	no	0.1	0.1	0.507	0.495	(+)	(+)	no	(+-)
1S_134	1	Spring	23	no	0.123	0.119	0.487	0.459	(+)	(+)	no	no
1S_139	1	Spring	13	no	0.083	0.085	0.509	0.476	(+)	(+)	no	no
1S_14	1	Spring	24B	no	0.094	0.676	0.511	0.45	(+)	(+)	no	(+-)
1S_141	1	Spring	21	no	0.104	0.103	0.493	0.492	(+)	(+)	no	(+-)
1S_146	1	Spring	23	yes	0.143	0.33	0.17	0.171	(+)	(+)	(+)	(++)
1S_159	1	Spring	7	yes	0.078	0.081	0.455	0.474	(+)	(+)	(+)	(+)
1S_169	1	Spring	NA	no	0.091	0.094	0.513	0.5	(+)	(+)	no	no
1S_175	1	Spring	7	yes	0.088	0.09	0.339	0.333	(+)	(+)	(+-)	(+-)
1S_260	1	Spring	23	no	0.121	0.121	0.477	0.455	(+)	(+)	no	no
1S_278	1	Spring	NA	no	0.22	0.212	0.465	0.466	(+)	(+)	no	no
1S_296	1	Spring	19	yes	0.376	0.52	0.157	0.151	(+)	(+)	(++)	(+)
1S_297	1	Spring	13	no	0.159	0.176	0.448	0.423	(+)	(+)	no	no
1S_311	1	Spring	24B	no	0.227	0.371	0.404	0.446	(+)	(+)	no	no
1S_315	1	Spring	11	no	0.139	0.148	0.441	0.435	(+)	(+)	no	no
1S_317	1	Spring	NA	no	0.161	0.486	0.432	0.418	(+)	(+)	no	no
1S_319	1	Spring	NA	yes	0.387	0.374	0.182	0.166	(+)	(+)	(+)	(++)
1S_35	1	Spring	NA	no	0.376	0.364	0.428	0.431	(+)	(+)	no	no
1S_63	1	Spring	20	yes	0.194	0.217	0.302	0.338	(+)	(+)	(+-)	(+)
1S_65	1	Spring	20	yes	0.507	0.548	0.178	0.176	(+)	(+)	no	(+-)
1S_70	1	Spring	24A	yes	0.471	0.628	0.196	0.2	(+)	(+)	(+)	(+)
1S_77	1	Spring	14	no	0.119	0.106	0.445	0.43	(+)	(+)	no	(++)
1S_84	1	Spring	20	yes	0.477	0.434	0.202	0.168	(+)	(+)	(+-)	(+)
5F_102	5	fall	NA	yes	0.299	0.194	0.145	0.154	(+)	(+)	(++)	(++)
5F_104	5	fall	6	no	0.075	0.077	0.504	0.488	(+)	(+)	no	no
5F_106	5	fall	NA	yes	0.229	0.227	0.145	0.15	(+)	(+)	no	no
5F_112	5	fall	NA	no	0.306	0.312	0.435	0.424	(+)	(+)	no	no
5F_113	5	fall	NA	no	0.154	0.137	0.478	0.449	(+)	(+)	no	no
5F_117	5	fall	13	no	0.214	0.162	0.485	0.479	(+)	(+)	no	no
5F_119	5	fall	2	no	0.235	0.226	0.444	0.495	(+-)	(+-)	no	(+-)
5F_122	5	fall	13	no	0.108	0.106	0.498	0.501	(+)	(+)	no	no
5F_123	5	fall	13	no	0.156	0.148	0.49	0.482	(+)	(+)	no	no
5F_126	5	fall	NA	no	0.162	0.162	0.486	0.492	(+)	(+)	no	no
5F_129	5	fall	NA	yes	0.214	0.205	0.161	0.151	(+)	no	(+)	no
5F_132	5	fall	2	no	0.214	0.213	0.439	0.425	(+-)	no	no	no
5F_133	5	fall	15A	no	0.271	0.178	0.419	0.411	no	no	no	no
5F_135	5	fall	16	yes	0.22	0.23	0.146	0.184	(+)	(+)	(+-)	(+-)
5F_143	5	fall	2	no	0.201	0.226	0.448	0.425	no	no	no	no
5F_146	5	fall	8	yes	0.204	0.183	0.199	0.154	(+)	(+)	no	(+-)
5F_148	5	fall	2	no	0.229	0.226	0.456	0.482	no	no	no	no

5F_155	5	fall	2	no	0.2	0.234	0.44	0.457	no	no	no	no
5F_157	5	fall	10	yes	0.149	0.175	0.127	0.121	(+)	(+)	(+-)	(+-)
5F_161	5	fall	13	no	0.167	0.164	0.474	0.469	(+)	(+)	no	no
5F_165	5	fall	12	yes	0.293	0.293	0.162	0.155	(+)	(+)	no	(+)
5F_17	5	fall	8	yes	0.223	0.19	0.209	0.176	(+)	(+)	no	(+-)
5F_18	5	fall	NA	yes	0.282	0.224	0.214	0.124	(+)	no	no	no
5F_19	5	fall	NA	no	0.177	0.173	0.482	0.473	(+)	(+)	no	no
5F_192	5	fall	13	no	0.074	0.083	0.557	0.517	(+)	(+)	no	(+)
5F_205	5	fall	13	yes	0.153	0.151	0.201	0.223	no	no	no	no
5F_207	5	fall	13	no	0.145	0.15	0.504	0.513	(+)	(+)	no	no
5F_213	5	fall	13	no	0.076	0.074	0.513	0.526	(+)	(+)	no	no
5F_25	5	fall	NA	yes	0.216	0.211	0.163	0.176	(+)	(+)	no	(+)
5F_253	5	fall	NA	yes	0.161	0.165	0.275	0.28	no	no	no	no
5F_273	5	fall	16	no	0.088	0.152	0.529	0.505	(+)	(+-)	no	(+-)
5F_29	5	fall	16	yes	0.22	0.207	0.134	0.148	(+)	no	(+)	no
5F_31	5	fall	2	no	0.231	0.219	0.453	0.453	no	no	no	no
5F_322	5	fall	NA	no	0.091	0.09	0.369	0.364	(+)	(+)	(+-)	no
5F_324	5	fall	16	no	0.094	0.094	0.535	0.529	(+)	(+)	no	no
5F_33	5	fall	8	yes	0.233	0.192	0.203	0.153	(+)	(+)	(+-)	(+-)
5F_40	5	fall	13	no	0.173	0.166	0.488	0.489	(+)	(+)	no	no
5F_45	5	fall	NA	yes	0.186	0.169	0.258	0.312	(+)	(+)	no	(+-)
5F_47	5	fall	16	yes	0.219	0.208	0.159	0.146	(+)	no	(+)	no
5F_49	5	fall	2	no	0.205	0.194	0.441	0.429	no	no	no	no
5F_51	5	fall	NA	yes	0.153	0.563	0.161	0.156	(+)	(+)	(++)	(+)
5F_53	5	fall	8	yes	0.218	0.201	0.159	0.166	(+)	(+)	(+-)	(+)
5F_57	5	fall	NA	yes	0.471	0.173	0.144	0.152	(+)	(+-)	(++)	no
5F_59	5	fall	15A	no	0.205	0.173	0.438	0.453	(+-)	(+)	no	(+-)
5F_6	5	fall	NA	no	0.207	0.169	0.501	0.502	(+)	(+)	no	no
5F_60	5	fall	16	no	0.185	0.169	0.428	0.426	no	no	no	no
5F_61	5	fall	NA	no	0.267	0.257	0.419	0.377	(+)	no	no	no
5F_63	5	fall	13	no	0.121	0.124	0.476	0.472	(+)	(+)	no	no
5F_65	5	fall	2	no	0.213	0.218	0.48	0.473	no	no	no	no
5F_69	5	fall	NA	yes	0.358	0.18	0.139	0.138	(+)	(+)	(+-)	(+-)
5F_7	5	fall	8	yes	0.2	0.656	0.182	0.17	(+)	(+)	no	(+)
5F_73	5	fall	2	no	0.218	0.2	0.48	0.463	no	no	no	no
5F_76	5	fall	5	no	0.27	0.259	0.379	0.367	(+)	(+)	no	no
5F_79	5	fall	16	yes	0.21	0.208	0.153	0.218	(+)	(+)	(+)	(+)
5F_80	5	fall	NA	no	0.669	0.607	0.4	0.383	(+)	(+)	no	no
5F_85	5	fall	16	yes	0.192	0.159	0.133	0.175	(+)	(+)	(+)	(+)
5F_90	5	fall	NA	no	0.276	0.272	0.421	0.408	(+)	(+)	no	no
5F_92	5	fall	13	no	0.117	0.114	0.515	0.508	(+)	(+)	no	no
5F_94	5	fall	8	yes	0.185	0.475	0.17	0.371	(+)	(+)	no	(+-)
5F_97	5	fall	8	yes	0.22	0.205	0.171	0.161	(+)	(+)	(+-)	(+-)
5F_98	5	fall	2	no	0.208	0.201	0.44	0.45	no	no	no	no
5S_100	5	Spring	11	no	0.087	0.09	0.481	0.481	(+)	(+)	no	no
5S_101	5	Spring	19	no	0.11	0.114	0.495	0.469	(+)	(+)	no	no
5S_102	5	Spring	19	no	0.196	0.546	0.453	0.479	(+)	(+)	no	no
5S_104	5	Spring	NA	no	0.099	0.097	0.505	0.518	(+)	(+)	no	no
5S_106	5	Spring	NA	no	0.137	0.265	0.431	0.43	(+)	(+)	no	no
5S_114	5	Spring	NA	yes	0.525	0.181	0.181	0.163	(+)	(+)	(++)	(++)
5S_115	5	Spring	11	no	0.677	0.173	0.482	0.465	(+)	(+)	no	no
5S_118	5	Spring	NA	no	0.086	0.088	0.501	0.498	(+)	(+)	no	no
5S_122	5	Spring	20	no	0.095	0.093	0.494	0.499	(+)	(+)	no	no
5S_139	5	Spring	NA	yes	0.157	0.473	0.153	0.127	(+)	(+)	(+)	(+)

5S_141	5	Spring	NA	no	0.139	0.138	0.467	0.477	(+)	(+)	no	no
5S_143	5	Spring	25	no	0.111	0.102	0.5	0.483	(+)	(+)	no	no
5S_149	5	Spring	14	yes	0.127	0.13	0.418	0.436	(+)	(+)	(+-)	(++)
5S_151	5	Spring	NA	no	0.128	0.129	0.473	0.472	(+)	(+)	no	no
5S_156	5	Spring	21	no	0.135	0.13	0.481	0.496	(+)	(+)	no	no
5S_161	5	Spring	23	no	0.672	0.16	0.421	0.427	(+)	(+)	no	no
5S_162	5	Spring	19	no	0.732	0.153	0.415	0.469	(+)	(+)	no	no
5S_163	5	Spring	NA	yes	0.075	0.076	0.54	0.597	(+)	(+)	(+)	(+)
5S_18	5	Spring	NA	yes	0.308	0.187	0.164	0.177	(+)	(+)	(+)	(+)
5S_185	5	Spring	25	no	0.984	0.238	0.374	0.359	(+)	(+)	no	no
5S_186	5	Spring	7	yes	0.173	0.133	0.274	0.298	(+)	(+)	(+)	(+)
5S_187	5	Spring	4	yes	0.181	0.178	0.41	0.347	(+)	(+)	(+)	(+)
5S_200	5	Spring	11	no	0.149	0.158	0.438	0.453	(+)	(+)	no	no
5S_207	5	Spring	13	no	0.095	0.099	0.531	0.511	(+)	(+)	no	no
5S_208	5	Spring	24B	no	0.151	0.123	0.459	0.45	(+)	(+)	no	(+)
5S_209	5	Spring	23	no	0.983	0.35	0.44	0.429	(+)	(+)	no	no
5S_210	5	Spring	22	no	0.13	0.13	0.486	0.467	(+)	(+)	no	no
5S_213	5	Spring	NA	no	0.252	0.178	0.392	0.381	(+)	(+)	no	no
5S_217	5	Spring	NA	no	0.117	0.114	0.473	0.433	(+)	(+)	no	no
5S_225	5	Spring	21	no	0.213	0.175	0.471	0.467	(+)	(+)	no	no
5S_226	5	Spring	21	no	0.131	0.265	0.441	0.418	(+)	(+)	no	no
5S_228	5	Spring	23	no	0.254	0.373	0.432	0.428	(+)	(+)	no	no
5S_23	5	Spring	24A	no	0.104	0.109	0.47	0.464	(+)	(+)	no	no
5S_230	5	Spring	NA	no	0.099	0.105	0.455	0.455	(+)	(+)	no	no
5S_235	5	Spring	24A	no	0.111	0.117	0.477	0.472	(+)	(+)	no	no
5S_238	5	Spring	24B	no	0.111	0.11	0.486	0.505	(+)	(+)	no	(+-)
5S_239	5	Spring	14	no	0.13	0.197	0.434	0.359	(+)	(+)	no	(+)
5S_240	5	Spring	14	no	0.133	0.133	0.403	0.401	(+)	(+)	no	(++)
5S_242	5	Spring	24A	no	0.364	0.286	0.439	0.528	(+)	(+)	no	no
5S_245	5	Spring	21	no	0.112	0.114	0.496	0.484	(+)	(+)	no	(+-)
5S_246	5	Spring	NA	yes	0.538	0.4	0.136	0.131	(+)	(+)	(++)	(++)
5S_252	5	Spring	21	yes	0.154	0.147	0.17	0.159	(+)	(+)	no	(+)
5S_257	5	Spring	NA	no	0.116	0.097	0.501	0.461	(+)	(+)	no	no
5S_264	5	Spring	NA	no	0.093	0.092	0.499	0.463	(+)	(+)	no	no
5S_268	5	Spring	24B	yes	0.585	0.432	0.164	0.173	(+)	(+)	(+)	(+)
5S_269	5	Spring	11	no	0.096	0.096	0.501	0.514	(+)	(+)	no	no
5S_272	5	Spring	22	no	0.181	0.163	0.412	0.429	(+)	(+)	no	no
5S_279	5	Spring	19	no	0.714	0.418	0.394	0.454	(+)	(+)	no	no
5S_283	5	Spring	24B	no	0.099	0.103	0.497	0.482	(+)	(+)	no	no
5S_285	5	Spring	24A	no	0.439	0.359	0.406	0.407	(+)	no	no	no
5S_37	5	Spring	NA	yes	0.159	0.166	0.176	0.158	(+)	(+)	(++)	(++)
5S_4	5	Spring	NA	no	0.162	0.152	0.463	0.437	(+)	(+)	no	no
5S_41	5	Spring	14	yes	0.074	0.075	0.528	0.602	(+)	(+)	(+)	(+)
5S_46	5	Spring	11	no	0.095	0.626	0.504	0.507	(+)	(+)	no	no
5S_5	5	Spring	24A	yes	0.331	0.492	0.139	0.128	(+)	(+)	(++)	(++)
5S_51	5	Spring	NA	no	0.555	0.2	0.379	0.375	(+)	(+)	no	no
5S_57	5	Spring	21	yes	0.183	0.175	0.207	0.171	(+)	(+)	(+)	(+)
5S_76	5	Spring	13	yes	0.222	0.229	0.183	0.217	(+)	(+)	(+)	(+)
5S_78	5	Spring	13	yes	0.221	0.221	0.155	0.159	(+)	(+)	(+)	(+)
5S_81	5	Spring	21	no	0.18	0.202	0.43	0.421	(+)	(+)	no	no
5S_83	5	Spring	NA	yes	0.852	0.174	0.147	0.131	(+)	(+)	(++)	(++)
5S_86	5	Spring	NA	no	0.122	0.118	0.466	0.455	(+)	(+)	no	no
5S_9	5	Spring	NA	no	0.162	0.155	0.432	0.449	no	(+)	no	no

9CG153	crab gill		NA	no	0.157	0.163	0.4	0.519	(+)	(+)	(+--)	no
9CG154	crab gill		NA	no	0.161	0.161	0.41	0.537	(+)	(+)	(+--)	no
9CG155	crab gill		16	yes	0.197	0.206	0.163	0.312	(+)	(+)	(+--)	no
9CS100	crab stomach		8	no	0.142	0.229	0.491	0.541	(+)	(+)	no	no
9CS101	crab stomach		11	no	0.163	0.161	0.401	0.536	(+)	(+)	no	no
9CS102	crab stomach		11	no	0.197	0.17	0.371	0.52	(+)	(+)	no	no
9CS103	crab stomach		11	no	0.159	0.18	0.403	0.527	(+)	(+)	no	no
9CS104	crab stomach		11	no	0.164	0.172	0.439	0.562	(+)	(+)	no	no
9CS105	crab stomach		NA	no	0.18	0.179	0.339	0.502	(+)	(+)	no	no
9CS106	crab stomach		8	no	0.105	0.108	0.468	0.547	(+)	(+)	no	no
9CS107	crab stomach		11	no	0.161	0.165	0.37	0.522	(+)	(+)	no	no
9CS108	crab stomach		11	no	0.164	0.259	0.386	0.522	(+)	(+)	no	no
9CS110	crab stomach		11	no	0.244	0.173	0.379	0.538	(+)	(+)	no	no
9CS111	crab stomach		NA	no	0.161	0.16	0.4	0.522	(+)	(+)	no	no
9CS113	crab stomach		11	no	0.177	0.181	0.348	0.503	(+)	(+)	no	no
9CS114	crab stomach		NA	no	0.167	0.165	0.343	0.502	(+)	(+)	no	no
9CS116	crab stomach		7	no	0.167	0.299	0.399	0.519	(+)	(+)	no	no
9CS117	crab stomach		NA	no	0.163	0.155	0.373	0.514	(+)	no	no	no
9CS118	crab stomach		7	no	0.174	0.165	0.366	0.5	(+-)	no	no	no
9CS119	crab stomach		NA	no	0.158	0.162	0.392	0.522	(+)	(+)	no	no
9CS120	crab stomach		1	yes	0.091	0.098	0.252	0.42	(+)	(+)	no	(+-)
9CS121	crab stomach		15	no	0.159	0.153	0.404	0.548	(+)	(+)	no	no
9CS122	crab stomach		16	no	0.132	0.135	0.444	0.559	(+)	(+)	no	no
9CS123	crab stomach		NA	yes	0.172	0.174	0.147	0.281	(+)	(+)	(+-)	(+-)
9CS124	crab stomach		NA	no	0.183	0.234	0.367	0.509	(+)	no	no	no
9CS125	crab stomach		16	no	0.175	0.181	0.149	0.312	(+)	no	no	no
9CS126	crab stomach		8	no	0.251	0.173	0.441	0.573	(+)	no	no	no
9CS127	crab stomach		11	no	0.139	0.142	0.51	0.533	(+)	(+)	no	no
9CS128	crab stomach		15	no	0.166	0.167	0.429	0.552	(+)	(+)	no	no
9CS129	crab stomach		NA	yes	0.186	0.178	0.176	0.358	(+)	(+)	no	(+-)
9CS131	crab stomach		8	no	0.165	0.174	0.451	0.555	(+)	(+)	no	no
9CS132	crab stomach		8	yes	0.166	0.234	0.181	0.348	(+)	(+)	(+-)	no
9CS134	crab stomach		8	no	0.435	0.131	0.442	0.526	(+)	(+)	no	no
9CS135	crab stomach		16	yes	0.165	0.175	0.155	0.314	(+)	(+)	no	(+-)
9CS136	crab stomach		16	no	0.165	0.163	0.429	0.563	(+)	(+)	no	no
9CS137	crab stomach		16	no	0.165	0.168	0.401	0.546	(+)	(+)	no	no
9CS138	crab stomach		NA	yes	0.202	0.192	0.169	0.332	(+)	(+)	(+-)	no
9CS139	crab stomach		8	no	0.161	0.142	0.387	0.546	(+)	(+)	no	no
9CS142	crab stomach		8	no	0.159	0.151	0.458	0.571	(+)	(+)	no	no
9CS143	crab stomach		16	no	0.154	0.155	0.485	0.54	(+)	(+)	no	no
9CS144	crab stomach		NA	no	0.194	0.186	0.387	0.508	(+)	(+)	no	no
9CS145	crab stomach		16	no	0.159	0.165	0.443	0.535	(+)	(+)	no	no
9CS146	crab stomach		8	no	0.141	0.139	0.454	0.56	(+)	(+)	no	no
9CS147	crab stomach		15	no	0.164	0.152	0.382	0.525	(+)	(+)	no	no
9CS148	crab stomach		13	yes	0.203	0.249	0.206	0.348	(+)	(+)	no	(+-)
9CS149	crab stomach		15	no	0.171	0.157	0.412	0.514	(+)	(+-)	no	no
9CS151	crab stomach		8	no	0.152	0.157	0.473	0.537	(+)	(+)	no	no
9CS152	crab stomach		16	no	0.169	0.155	0.471	0.536	(+)	(+)	no	no
9CS155	crab stomach		8	no	0.124	0.119	0.442	0.537	(+)	(+)	no	no
9CS156	crab stomach		16	no	0.145	0.282	0.421	0.52	(+)	no	no	no
9CS157	crab stomach		NA	no	0.192	0.185	0.392	0.496	(+)	(+)	no	no
9CS158	crab stomach		NA	no	0.22	0.193	0.38	0.514	(+)	(+)	no	no
9CS159	crab stomach		NA	no	0.188	0.186	0.416	0.518	(+)	(+)	no	no

9CS82	crab stomach		16	yes	0.17	0.171	0.154	0.329	(+)	(+)	(+)	no
9CS83	crab stomach		8	no	0.139	0.131	0.452	0.547	(+)	(+)	no	no
9CS84	crab stomach		15	yes	0.214	0.263	0.141	0.279	(+)	(+)	(+--)	no
9CS85	crab stomach		15	yes	0.219	0.214	0.147	0.295	(+)	(+)	(+--)	no
9CS86	crab stomach		15	yes	0.237	0.213	0.145	0.295	(+)	(+-)	(+--)	no
9CS87	crab stomach		8	no	0.135	0.139	0.537	0.554	(+)	(+)	no	no
9CS89	crab stomach		8	no	0.13	0.128	0.456	0.535	(+)	(+)	no	no
9CS90	crab stomach		8	no	0.131	0.143	0.447	0.557	(+)	(+)	no	no
9CS92	crab stomach		3	yes	0.089	0.172	0.257	0.372	(+)	(+)	(++)	(++)
9CS95	crab stomach		15	yes	0.219	0.209	0.145	0.294	(+)	(+)	no	no
9CS96	crab stomach		16	no	0.185	0.174	0.25	0.311	(+)	(+)	no	no
9CS97	crab stomach		6	yes	0.219	0.218	0.143	0.276	no	no	no	no
9CS98	crab stomach		NA	no	0.112	0.118	0.474	0.561	(+)	(+)	no	no
9CS99	crab stomach		1	yes	0.159	0.165	0.244	0.389	(+)	(+)	(++)	(++)
9CSC81	crab stomach		8	no	0.162	0.162	0.446	0.537	(+)	(+)	no	no
9CSC82	crab stomach		8	no	0.135	0.138	0.44	0.517	(+)	(+)	no	no
9CSC86	crab stomach		15	yes	0.252	0.22	0.143	0.279	(+)	(+)	no	no
9CSC87	crab stomach		15	yes	0.217	0.224	0.14	0.277	(+)	(+)	no	no
9RW138	mussel		NA	yes	0.224	0.277	0.224	0.353	(+)	(+)	(+)	no
9RW139	mussel		NA	no	0.166	0.226	0.438	0.529	(+)	(+)	no	no
9RW140	mussel		8	no	0.155	0.214	0.503	0.538	(+)	(+)	no	no
9RW141	mussel		8	no	0.2	0.283	0.5	0.546	(+)	(+)	no	no
9RW142	mussel		16	yes	0.17	0.186	0.178	0.334	(+)	(+)	(++)	(+)
9RW144	seawater		NA	yes	0.088	0.098	0.41	0.455	(+)	(+)	(++)	(++)
9RW145	seawater		15	no	0.187	0.218	0.442	0.503	(+)	(+)	no	(+-)
9RW146	seawater		13	yes	0.214	0.264	0.226	0.372	(+)	(+)	(+)	no
9RW147	seawater		10	no	0.187	0.215	0.512	0.535	(+)	(+)	no	no
9RW148	seawater		NA	no	0.153	0.226	0.255	0.362	(+)	(+)	no	no
9RW149	seawater		11	no	0.156	0.224	0.392	0.52	(+)	(+)	no	no
9RW150	seawater		6	yes	0.16	0.182	0.214	0.367	(+)	(+)	(++)	(+)
9RW151	seawater		16	no	0.143	0.292	0.488	0.532	(+)	(+)	no	no
9RW153	seawater		10	no	0.172	0.191	0.498	0.553	(+)	(+)	no	no
9RW154	seawater		11	yes	0.19	0.259	0.209	0.352	(+)	(+)	(+)	(+)
9RW155	seawater		5	no	0.183	0.205	0.498	0.537	(+)	(+)	no	no
9RW156	seawater		NA	no	0.175	0.252	0.276	0.395	no	no	no	no
9RW159	seawater		NA	yes	0.171	0.353	0.188	0.35	(+)	(+)	(++)	(++)
9SW100	seawater		3	yes	0.143	0.149	0.288	0.419	(+)	(+)	(++)	(++)
9SW102	seawater		NA	no	0.152	0.28	0.497	0.556	no	no	no	no
9SW105	seawater		11	no	0.129	0.172	0.473	0.524	(+)	(+)	no	no
9SW106	seawater		NA	no	0.153	0.239	0.487	0.567	(+)	no	no	no
9SW107	seawater		NA	no	0.173	0.216	0.421	0.5	(+)	(+)	no	no
9SW108	seawater		11	no	0.366	0.214	0.396	0.498	(+)	(+)	no	no
9SW110	seawater		16	yes	0.213	0.508	0.238	0.377	(+)	(+)	(+)	(+)
9SW111	seawater		10	no	0.169	0.214	0.419	0.54	(+)	(+)	no	no
9SW113	seawater		NA	no	0.167	0.2	0.457	0.509	(+)	(+)	no	no
9SW114	seawater		1	yes	0.213	0.186	0.402	0.403	(+)	(+)	(+-)	(++)
9SW115	seawater		10	no	0.157	0.191	0.472	0.537	(+)	no	no	no
9SW116	seawater		10	no	0.27	0.231	0.452	0.53	(+)	(+)	no	no
9SW117	seawater		NA	yes	0.19	0.258	0.165	0.331	(+)	(+)	(+)	(+-)
9SW119	seawater		NA	yes	0.2	0.334	0.349	0.461	(+)	(+)	(+)	(+)
9SW121	seawater		11	no	0.157	0.202	0.418	0.518	(+)	no	no	no
9SW123	seawater		NA	yes	0.219	0.255	0.235	0.35	(+)	no	(+)	no
9SW125	seawater		16	no	0.128	0.201	0.482	0.521	(+)	no	no	no

9SW128	seawater		NA	yes	0.167	0.229	0.251	0.376	(+)	(+)	(+)	(+-)
9SW129	seawater		16	yes	0.166	0.275	0.193	0.371	(+)	no	(+)	no
9SW132	seawater		10	no	0.296	0.252	0.447	0.511	(+)	(+)	no	no
9SW136	seawater		NA	no	0.114	0.145	0.445	0.511	(+)	(+)	no	(+-)
9SW139	seawater		NA	no	0.208	0.346	0.41	0.478	(+)	(+-)	no	no
9SW140	seawater		10	no	0.198	0.253	0.39	0.527	no	no	no	no
9SW142	seawater		NA	yes	0.214	0.278	0.179	0.351	(+)	no	(+)	no
9SW144	seawater		16	no	0.132	0.155	0.48	0.541	(+)	(+)	no	no
9SW146	seawater		NA	no	0.185	0.226	0.431	0.527	(+)	(+)	(+-)	no
9SW149	seawater		NA	no	0.183	0.409	0.454	0.518	no	no	no	no
9SW150	seawater		NA	no	0.313	0.596	0.426	0.472	no	no	no	no
9SW151	seawater		NA	yes	0.159	0.191	0.361	0.43	(+)	(+)	(+-)	(+)
9SW81	seawater		8	no	0.112	0.125	0.532	0.557	(+)	(+)	no	no
9SW82	seawater		8	no	0.163	0.22	0.453	0.568	(+)	(+)	no	no
9SW83	seawater		5	no	0.193	0.236	0.401	0.529	(+)	(+)	no	no
9SW86	seawater		NA	yes	0.183	0.188	0.243	0.402	(+)	(+)	(+)	(+)
9SW87	seawater		NA	yes	0.194	0.363	0.243	0.381	(+)	(+)	(+)	(+)
9SW89	seawater		16	yes	0.167	0.204	0.186	0.356	(+)	(+)	(++)	(++)
9SW90	seawater		NA	no	0.174	0.257	0.27	0.377	(+)	(+)	no	no
9SW91	seawater		10	no	0.192	0.245	0.404	0.527	(+)	(+)	no	no
9SW92	seawater		5	no	0.216	0.241	0.383	0.534	(+)	(+)	no	no
9SW94	seawater		NA	no	0.092	0.09	0.502	0.528	(+)	(+)	no	no
9SW97	seawater		16	no	0.135	0.18	0.463	0.542	(+)	no	no	no
9SW98	seawater		10	no	0.164	0.184	0.487	0.552	(+)	(+)	no	no
9SW99	seawater		NA	yes	0.216	0.263	0.251	0.377	(+)	(+)	(+)	(+-)
9ZA1	zooplankton		NA	no	0.19	0.17	0.342	0.488	(+)	no	no	no
9ZA10	zooplankton		NA	no	0.221	0.184	0.486	0.578	no	no	no	no
9ZA100	zooplankton		NA	no	0.265	0.184	0.479	0.455	(+)	(+)	no	no
9ZA101	zooplankton		8	yes	0.249	0.178	0.232	0.28	(+)	(+)	(+-)	(+-)
9ZA102	zooplankton		10	no	0.273	0.22	0.408	0.4	(+-)	no	no	no
9ZA103	zooplankton		NA	yes	0.263	0.426	0.135	0.227	(+)	no	(+)	no
9ZA104	zooplankton		10	no	0.27	0.224	0.4	0.418	(+)	(+)	no	(+-)
9ZA105	zooplankton		7	no	0.184	0.151	0.376	0.441	(+)	(+)	no	no
9ZA107	zooplankton		8	yes	0.261	0.179	0.212	0.285	(+)	(+)	(+-)	(+-)
9ZA108	zooplankton		NA	no	0.192	0.155	0.425	0.436	no	(+)	no	no
9ZA109	zooplankton		NA	no	0.14	0.132	0.385	0.377	(+)	(+)	no	no
9ZA11	zooplankton		7	yes	0.187	0.184	0.233	0.481	(+)	(+)	(+)	(+)
9ZA111	zooplankton		8	yes	0.243	0.573	0.239	0.275	(+)	(+)	(+-)	(+-)
9ZA112	zooplankton		NA	no	0.228	0.159	0.473	0.468	(+)	(+)	no	no
9ZA113	zooplankton		7	no	0.369	0.135	0.146	0.31	(+)	(+)	no	no
9ZA114	zooplankton		7	yes	0.162	0.205	0.126	0.268	(+-)	(+-)	no	no
9ZA116	zooplankton		7	no	0.156	0.122	0.296	0.397	(+)	no	no	no
9ZA119	zooplankton		16	no	0.195	0.262	0.441	0.439	(+)	(+)	no	no
9ZA120	zooplankton		NA	no	0.161	0.154	0.434	0.466	(+)	(+)	no	no
9ZA121	zooplankton		NA	no	0.109	0.106	0.475	0.468	(+)	(+)	no	no
9ZA122	zooplankton		7	no	0.174	0.205	0.377	0.426	(+)	(+)	no	no
9ZA125	zooplankton		7	no	0.17	0.139	0.402	0.44	(+-)	(+)	no	no
9ZA128	zooplankton		NA	yes	0.168	0.153	0.266	0.378	(+)	(+)	(+-)	(+-)
9ZA129	zooplankton		7	no	0.17	0.159	0.399	0.425	(+)	(+)	no	no
9ZA13	zooplankton		NA	no	0.195	0.173	0.342	0.489	(+)	no	no	no
9ZA132	zooplankton		NA	no	0.238	0.163	0.468	0.457	(+)	(+)	no	no
9ZA133	zooplankton		NA	no	0.23	0.382	0.34	0.354	(+)	(+)	no	no
9ZA135	zooplankton		NA	yes	0.275	0.212	0.137	0.267	no	no	no	no
9ZA136	zooplankton		NA	yes	0.259	0.182	0.237	0.289	(+)	(+)	(+-)	(+-)

9ZA140	zooplankton		NA	no	0.243	0.131	0.451	0.486	(+)	(+)	no	no
9ZA142	zooplankton		NA	yes	0.266	0.186	0.146	0.218	(+)	(+)	(+)	(+-)
9ZA143	zooplankton		NA	yes	0.344	0.265	0.215	0.323	(+)	(+)	(+-)	(+-)
9ZA144	zooplankton		NA	yes	0.321	0.273	0.226	0.291	(+)	(+)	(+-)	(+-)
9ZA145	zooplankton		1	no	0.207	0.168	0.23	0.348	no	no	no	no
9ZA146	zooplankton		1	yes	0.203	0.353	0.14	0.248	(+)	(+)	(+++	(+++
9ZA147	zooplankton		12	no	0.216	0.182	0.422	0.431	(+)	(+)	no	no
9ZA15	zooplankton		7	no	0.174	0.138	0.359	0.502	(+)	no	no	no
9ZA151	zooplankton		10	no	0.261	0.196	0.42	0.434	(+)	(+)	no	no
9ZA152	zooplankton		16	no	0.188	0.141	0.419	0.433	(+)	(+)	no	no
9ZA155	zooplankton		NA	no	0.256	0.211	0.476	0.471	(+)	(+)	no	no
9ZA156	zooplankton		7	no	0.196	0.153	0.369	0.391	(+)	(+)	no	no
9ZA157	zooplankton		3	no	0.08	0.08	0.518	0.507	(+)	(+)	no	no
9ZA159	zooplankton		NA	yes	0.229	0.19	0.419	0.412	(+)	(+)	(++)	(++)
9ZA160	zooplankton		16	no	0.366	0.12	0.428	0.442	(+)	(+)	no	no
9ZA2	zooplankton		NA	no	0.182	0.159	0.433	0.57	(+)	no	no	no
9ZA21	zooplankton		NA	no	0.429	0.147	0.436	0.534	(+)	(+)	no	no
9ZA23	zooplankton		3	no	0.082	0.097	0.528	0.557	(+)	(+)	no	no
9ZA26	zooplankton		3	no	0.109	0.101	0.514	0.584	(+)	(+)	no	no
9ZA27	zooplankton		3	no	0.092	0.096	0.508	0.582	(+)	(+)	no	no
9ZA28	zooplankton		3	no	0.097	0.093	0.52	0.59	(+)	(+)	no	no
9ZA29	zooplankton		7	yes	0.253	0.159	0.268	0.481	(+)	(+)	(+)	(+)
9ZA3	zooplankton		11	yes	0.211	0.18	0.236	0.343	(+)	(+)	no	(+-)
9ZA30	zooplankton		7	yes	0.167	0.148	0.173	0.443	(+)	no	(+)	no
9ZA31	zooplankton		NA	no	0.147	0.194	0.325	0.408	(+)	(+)	no	no
9ZA32	zooplankton		11	no	0.211	0.164	0.357	0.539	(+)	(+)	no	no
9ZA34	zooplankton		NA	no	0.117	0.091	0.501	0.573	(+)	(+)	no	no
9ZA37	zooplankton		NA	no	0.337	0.227	0.418	0.534	(+)	(+)	no	no
9ZA39	zooplankton		NA	yes	0.177	0.192	0.227	0.405	(+)	no	(+)	no
9ZA4	zooplankton		16	no	0.274	0.136	0.446	0.522	(+)	(+)	no	no
9ZA40	zooplankton		7	no	0.168	0.141	0.332	0.481	(+)	(+)	no	no
9ZA41	zooplankton		8	no	0.299	0.112	0.479	0.555	(+)	(+-)	no	no
9ZA42	zooplankton		NA	no	0.339	0.182	0.423	0.564	(+)	(+)	no	no
9ZA43	zooplankton		NA	no	0.288	0.238	0.417	0.559	(+)	(+)	(+-)	no
9ZA46	zooplankton		16	no	0.181	0.142	0.408	0.52	(+)	(+)	no	no
9ZA48	zooplankton		16	no	0.448	0.143	0.157	0.401	(+)	(+)	no	no
9ZA5	zooplankton		NA	no	0.198	0.169	0.341	0.477	(+)	(+)	no	no
9ZA53	zooplankton		5	no	0.329	0.265	0.442	0.561	(+)	(+)	no	no
9ZA55	zooplankton		7	yes	0.183	0.138	0.166	0.449	(+)	(+)	(+)	(++)
9ZA6	zooplankton		NA	no	0.192	0.169	0.37	0.493	(+)	(+)	no	no
9ZA60	zooplankton		5	yes	0.349	0.168	0.289	0.441	(+)	(+)	(+-)	no
9ZA61	zooplankton		NA	no	0.209	0.187	0.375	0.552	no	no	no	no
9ZA63	zooplankton		NA	no	0.28	0.157	0.475	0.541	(+)	(+)	no	no
9ZA64	zooplankton		NA	no	0.337	0.16	0.126	0.394	no	no	no	no
9ZA65	zooplankton		NA	no	0.191	0.179	0.382	0.518	(+)	(+)	no	no
9ZA66	zooplankton		NA	no	0.303	0.226	0.427	0.537	(+)	(+)	no	no
9ZA67	zooplankton		10	no	0.213	0.18	0.427	0.546	(+)	(+)	no	no
9ZA68	zooplankton		NA	yes	0.284	0.185	0.137	0.423	(+)	(+)	(+-)	no
9ZA69	zooplankton		NA	no	0.278	0.281	0.305	0.49	(+)	(+)	no	(+-)
9ZA7	zooplankton		NA	no	0.113	0.117	0.446	0.549	(+)	(+)	no	no
9ZA72	zooplankton		7	yes	0.164	0.137	0.125	0.382	(+)	(+)	(+-)	no
9ZA74	zooplankton		NA	no	0.273	0.162	0.455	0.517	(+)	(+)	no	no
9ZA75	zooplankton		16	no	0.265	0.205	0.354	0.516	(+)	(+)	no	no

9ZA8	zooplankton		NA	no	0.189	0.177	0.386	0.475	(+)	(+)	no	no
9ZA80	zooplankton		NA	no	0.311	0.149	0.347	0.476	(+)	(+)	no	no
9ZA81	zooplankton		NA	no	0.188	0.153	0.455	0.436	no	no	no	no
9ZA82	zooplankton		NA	no	0.245	0.555	0.446	0.419	(+-)	(+-)	no	no
9ZA85	zooplankton		NA	no	0.24	0.179	0.443	0.425	(+-)	(+-)	no	no
9ZA87	zooplankton		NA	no	0.378	0.604	0.444	0.447	no	no	no	no
9ZA88	zooplankton		7	no	0.169	0.153	0.414	0.459	no	no	no	no
9ZA89	zooplankton		NA	no	0.186	0.151	0.463	0.43	no	no	no	no
9ZA90	zooplankton		NA	no	0.098	0.216	0.511	0.243	(+)	(+-)	no	no
9ZA92	zooplankton		NA	no	0.414	0.172	0.431	0.43	no	no	no	no
9ZA95	zooplankton		8	no	0.255	0.4	0.424	0.434	(+)	(+)	no	no
9ZA96	zooplankton		NA	yes	0.373	0.302	0.158	0.225	(+)	(+)	(+-)	(+-)
9ZA97	zooplankton		NA	no	0.208	0.173	0.377	0.361	(+)	(+)	no	no
9ZB1	zooplankton		NA	no	0.228	0.085	0.426	0.528	(+)	(+)	no	no
9ZB100	zooplankton		8	no	0.157	0.202	0.482	0.448	(+)	(+)	no	no
9ZB103	zooplankton		7	yes	0.199	0.234	0.257	0.27	no	no	no	no
9ZB105	zooplankton		NA	yes	0.241	0.249	0.162	0.204	(+)	(+)	no	no
9ZB106	zooplankton		1	yes	0.187	0.195	0.229	0.258	no	no	no	no
9ZB107	zooplankton		NA	no	0.142	0.205	0.512	0.505	(+)	(+)	no	no
9ZB108	zooplankton		7	no	0.16	0.18	0.463	0.43	(+)	no	no	no
9ZB110	zooplankton		16	no	0.176	0.23	0.462	0.412	(+)	(+)	no	no
9ZB113	zooplankton		3	no	0.123	0.113	0.535	0.493	(+)	(+)	no	no
9ZB114	zooplankton		10	no	0.154	0.272	0.51	0.484	(+-)	(+)	no	no
9ZB116	zooplankton		8	yes	0.188	0.217	0.208	0.203	(+)	(+)	(+)	(+)
9ZB118	zooplankton		NA	yes	0.2	0.285	0.147	0.163	(+)	(+)	(+-)	(++)
9ZB119	zooplankton		NA	yes	0.161	0.186	0.168	0.18	(+)	(+)	(+-)	(+)
9ZB121	zooplankton		NA	no	0.34	0.226	0.486	0.457	(+-)	(+-)	(+-)	no
9ZB122	zooplankton		15	no	0.158	0.187	0.469	0.425	(+)	(+)	no	(+-)
9ZB123	zooplankton		NA	no	0.164	0.227	0.456	0.435	(+)	(+)	no	(+-)
9ZB124	zooplankton		10	no	0.181	0.24	0.464	0.434	(+)	(+)	no	(+-)
9ZB131	zooplankton		16	yes	0.166	0.454	0.147	0.163	(+)	(+)	no	no
9ZB132	zooplankton		NA	yes	0.173	0.22	0.199	0.216	(+)	(+)	(+-)	(+-)
9ZB138	zooplankton		3	no	0.082	0.084	0.506	0.509	(+)	(+)	no	no
9ZB141	zooplankton		5	yes	0.167	0.506	0.148	0.182	(+)	(+)	no	no
9ZB143	zooplankton		16	no	0.219	0.292	0.303	0.298	(+)	(+)	no	no
9ZB146	zooplankton		NA	yes	0.121	0.097	0.248	0.476	(+)	(+)	(+-)	(++)
9ZB148	zooplankton		3	yes	0.243	0.14	0.198	0.221	(+)	(+)	(+)	(+)
9ZB149	zooplankton		16	no	0.136	0.298	0.48	0.444	(+)	(+)	no	no
9ZB150	zooplankton		3	no	0.099	0.249	0.538	0.522	(+-)	(+)	no	no
9ZB152	zooplankton		NA	yes	0.268	0.318	0.255	0.26	(+)	(+)	no	no
9ZB153	zooplankton		1	yes	0.177	0.22	0.218	0.238	(+-)	(+)	(+)	(++)
9ZB154	zooplankton		3	no	0.099	0.131	0.523	0.51	(+)	(+)	no	no
9ZB155	zooplankton		NA	yes	0.232	0.285	0.133	0.142	(+)	(+)	(+)	(+)
9ZB156	zooplankton		1	yes	0.285	0.248	0.142	0.179	(+)	(+)	(+)	(+)
9ZB159	zooplankton		1	yes	0.157	0.184	0.315	0.295	(+)	(+)	no	(+-)
9ZB16	zooplankton		11	no	0.218	0.159	0.361	0.504	(+)	(+)	no	no
9ZB160	zooplankton		1	yes	0.147	0.198	0.154	0.17	(+)	(+)	(++)	(++)
9ZB2	zooplankton		NA	no	0.328	0.141	0.423	0.563	(+)	(+)	no	no
9ZB20	zooplankton		NA	no	0.253	0.149	0.417	0.55	(+)	no	no	no
9ZB22	zooplankton		3	yes	0.374	0.17	0.149	0.141	(+)	(+)	(++)	(++)
9ZB23	zooplankton		NA	yes	0.527	0.158	0.16	0.159	(+)	(+)	(++)	(++)
9ZB24	zooplankton		NA	yes	0.35	0.165	0.153	0.155	(+)	(+)	(+-)	(+)
9ZB25	zooplankton		3	yes	0.31	0.165	0.15	0.154	(+)	(+)	(++)	(++)

9ZB27	zooplankton		NA	yes	0.184	0.28	0.151	0.15	(+)	(+)	(++)	(++)
9ZB28	zooplankton		NA	yes	0.189	0.17	0.154	0.153	(+)	(+)	(++)	(++)
9ZB29	zooplankton		3	yes	0.169	0.152	0.121	0.122	(+)	(+)	no	no
9ZB30	zooplankton		3	yes	0.109	0.097	0.184	0.177	(+)	(+)	no	no
9ZB31	zooplankton		5	yes	0.333	0.158	0.157	0.156	(+)	(+)	(+)	(+)
9ZB32	zooplankton		5	yes	0.742	0.158	0.154	0.154	(+)	(+)	(+)	(+)
9ZB33	zooplankton		5	yes	0.58	0.171	0.149	0.153	(+)	(+)	(+)	(+)
9ZB34	zooplankton		7	yes	0.3	0.169	0.153	0.159	(+)	(+)	(+)	(+)
9ZB35	zooplankton		7	yes	0.142	0.171	0.124	0.123	(+)	(+)	no	no
9ZB36	zooplankton		7	no	0.135	0.157	0.292	0.32	(+)	(+)	no	no
9ZB38	zooplankton		5	no	0.441	0.129	0.393	0.415	(+)	(+)	no	no
9ZB39	zooplankton		5	yes	0.157	0.119	0.284	0.285	(+)	(+)	no	no
9ZB40	zooplankton		7	yes	0.401	0.164	0.154	0.149	(+)	(+)	(+)	(+)
9ZB42	zooplankton		5	yes	0.287	0.172	0.156	0.159	(+)	(+)	(+)	(+)
9ZB43	zooplankton		5	yes	0.186	0.347	0.148	0.156	(+)	(+)	(++)	(++)
9ZB44	zooplankton		NA	no	0.162	0.158	0.309	0.342	(+)	(+)	no	no
9ZB46	zooplankton		16	no	0.258	0.133	0.396	0.406	(+)	(+)	no	no
9ZB47	zooplankton		5	no	0.382	0.125	0.368	0.388	(+)	(+)	(+-)	no
9ZB53	zooplankton		16	yes	0.142	0.143	0.161	0.152	(+)	(+)	(+-)	(+)
9ZB54	zooplankton		11	yes	0.358	0.131	0.146	0.146	(+)	(+)	no	no
9ZB56	zooplankton		NA	yes	0.382	0.135	0.195	0.18	(+)	(+)	(+-)	(+-)
9ZB59	zooplankton		5	no	0.132	0.33	0.437	0.44	(+)	(+)	no	no
9ZB60	zooplankton		5	no	0.139	0.137	0.433	0.445	(+)	(+)	no	no
9ZB62	zooplankton		NA	yes	0.248	0.135	0.142	0.143	(+)	(+)	(+-)	(+-)
9ZB63	zooplankton		NA	yes	0.334	0.225	0.245	0.256	(+)	(+)	no	no
9ZB64	zooplankton		1	yes	0.308	0.161	0.242	0.217	(+)	(+)	(+-)	no
9ZB65	zooplankton		16	no	0.574	0.13	0.43	0.444	(+)	(+)	no	no
9ZB66	zooplankton		NA	no	0.133	0.244	0.435	0.444	(+)	(+)	no	no
9ZB68	zooplankton		1	yes	0.198	0.185	0.129	0.128	(+)	(+)	(+-)	no
9ZB71	zooplankton		NA	no	0.319	0.216	0.361	0.368	(+)	(+)	no	no
9ZB72	zooplankton		NA	no	0.49	0.137	0.414	0.401	(+)	no	no	no
9ZB74	zooplankton		3	no	0.138	0.295	0.385	0.401	(+)	(+)	no	no
9ZB75	zooplankton		NA	no	0.206	0.197	0.353	0.353	(+)	(+)	(+-)	no
9ZB77	zooplankton		1	yes	0.187	0.152	0.154	0.154	(+)	(+)	(+-)	(+-)
9ZB78	zooplankton		NA	no	0.094	0.084	0.411	0.408	(+)	(+)	(+-)	no
9ZB79	zooplankton		7	no	0.139	0.116	0.354	0.379	(+)	(+)	no	(+-)
9ZB80	zooplankton		10	no	0.331	0.186	0.342	0.351	(+)	(+)	no	no
9ZB81	zooplankton		7	no	0.183	0.153	0.275	0.385	(+)	(+)	no	no
9ZB84	zooplankton		16	no	0.262	0.402	0.39	0.405	(+)	(+)	no	no
9ZB88	zooplankton		NA	no	0.18	0.144	0.508	0.497	(+)	(+)	no	no
9ZB90	zooplankton		7	no	0.142	0.161	0.415	0.47	(+)	(+)	no	no
9ZB93	zooplankton		NA	yes	0.258	0.334	0.125	0.168	(+)	(+)	(+-)	no
9ZB95	zooplankton		NA	no	0.165	0.146	0.603	0.4	no	(+)	no	no
9ZB96	zooplankton		NA	no	0.206	0.161	0.431	0.435	(+)	(+)	no	no
9ZB98	zooplankton		8	no	0.178	0.416	0.463	0.488	(+)	(+)	no	no
9ZB99	zooplankton		1	yes	0.219	0.429	0.173	0.326	(+)	(+)	(++)	(+-)
9ZC1	zooplankton		3	yes	0.33	0.135	0.395	0.397	(+)	(+)	(+-)	(+-)
9ZC10	zooplankton		NA	no	0.22	0.597	0.39	0.395	(+)	(+)	no	no
9ZC101	zooplankton		NA	yes	0.351	0.224	0.156	0.194	no	(+-)	no	no
9ZC104	zooplankton		8	no	0.153	0.187	0.482	0.472	(+)	(+)	no	no
9ZC106	zooplankton		1	yes	0.125	0.153	0.207	0.242	(+)	(+)	(++)	(++)
9ZC107	zooplankton		8	no	0.183	0.253	0.446	0.436	(+)	(+)	no	no
9ZC108	zooplankton		NA	no	0.175	0.166	0.442	0.465	no	no	no	no
9ZC11	zooplankton		6	no	0.243	0.244	0.375	0.386	(+)	(+)	no	no

9ZC110	zooplankton		NA	no	0.172	0.216	0.374	0.36	(+)	(+)	no	no
9ZC112	zooplankton		8	no	0.129	0.181	0.642	0.492	(+)	(+)	no	no
9ZC113	zooplankton		NA	yes	0.287	0.367	0.164	0.174	(+)	(+)	no	(+-)
9ZC114	zooplankton		3	yes	0.261	0.325	0.26	0.237	(+)	(+)	no	no
9ZC115	zooplankton		1	yes	0.17	0.2	0.254	0.258	(+)	(+)	(++)	(++)
9ZC116	zooplankton		8	no	0.156	0.191	0.454	0.438	no	no	no	no
9ZC117	zooplankton		8	no	0.172	0.16	0.458	0.453	no	no	no	no
9ZC118	zooplankton		NA	yes	0.242	0.244	0.154	0.192	(+)	(+)	no	no
9ZC119	zooplankton		5	no	0.168	0.21	0.386	0.397	(+)	(+)	no	no
9ZC12	zooplankton		6	yes	0.19	0.187	0.146	0.146	(+)	(+)	(+-)	(+-)
9ZC13	zooplankton		8	yes	0.318	0.164	0.143	0.147	(+)	(+)	(+-)	(+-)
9ZC14	zooplankton		8	yes	0.28	0.154	0.15	0.154	(+)	(+)	no	(+-)
9ZC15	zooplankton		8	yes	0.239	0.15	0.158	0.154	(+)	(+)	no	(+-)
9ZC16	zooplankton		8	yes	0.338	0.16	0.139	0.144	(+)	(+)	(+-)	(+-)
9ZC17	zooplankton		6	no	0.922	0.156	0.395	0.406	(+)	(+)	no	no
9ZC2	zooplankton		6	no	0.141	0.354	0.365	0.378	(+)	(+)	no	no
9ZC21	zooplankton		10	no	0.271	0.194	0.368	0.361	(+)	(+)	no	no
9ZC22	zooplankton		8	yes	0.395	0.162	0.17	0.161	(+)	(+)	no	no
9ZC23	zooplankton		11	yes	0.565	0.162	0.142	0.142	(+)	(+)	no	(+-)
9ZC24	zooplankton		NA	yes	0.363	0.164	0.145	0.136	(+)	(+)	no	(+-)
9ZC25	zooplankton		NA	yes	0.737	0.174	0.138	0.139	(+)	(+)	(+-)	(+)
9ZC26	zooplankton		6	yes	0.206	0.257	0.133	0.185	(+)	(+)	no	(+)
9ZC27	zooplankton		NA	yes	0.198	0.196	0.235	0.276	(+)	(+)	no	no
9ZC28	zooplankton		11	no	0.162	0.152	0.34	0.35	(+)	(+)	no	no
9ZC29	zooplankton		NA	no	0.233	0.211	0.387	0.361	(+)	(+)	no	no
9ZC3	zooplankton		6	no	0.257	0.255	0.377	0.387	(+)	(+)	no	no
9ZC30	zooplankton		10	yes	0.419	0.168	0.153	0.172	(+)	(+)	(+-)	no
9ZC31	zooplankton		16	yes	0.752	0.183	0.137	0.136	(+)	(+)	(+)	(+)
9ZC32	zooplankton		8	yes	0.864	0.184	0.138	0.131	(+)	(+)	no	(+)
9ZC33	zooplankton		5	yes	0.587	0.205	0.132	0.13	(+)	(+)	(+-)	(+)
9ZC35	zooplankton		16	yes	0.199	0.196	0.128	0.135	(+)	(+)	(+-)	(+)
9ZC36	zooplankton		8	no	0.199	0.312	0.511	0.483	(+)	(+)	no	no
9ZC37	zooplankton		16	no	0.554	0.514	0.475	0.471	(+)	(+)	no	no
9ZC38	zooplankton		8	no	0.128	0.194	0.508	0.489	(+)	(+)	no	no
9ZC39	zooplankton		8	no	0.181	0.216	0.439	0.422	(+)	(+)	no	no
9ZC4	zooplankton		NA	yes	0.163	0.164	0.165	0.154	(+)	(+)	no	(+-)
9ZC40	zooplankton		8	no	0.149	0.237	0.468	0.458	(+)	(+)	no	no
9ZC42	zooplankton		16	yes	0.192	0.232	0.133	0.142	(+)	(+)	(+-)	(+-)
9ZC46	zooplankton		10	no	0.164	0.192	0.436	0.422	(+-)	(+-)	no	no
9ZC47	zooplankton		NA	no	0.18	0.188	0.378	0.387	(+)	no	no	no
9ZC48	zooplankton		16	no	0.188	0.416	0.452	0.454	(+)	(+)	no	no
9ZC5	zooplankton		NA	yes	0.351	0.17	0.165	0.163	(+)	(+)	no	no
9ZC50	zooplankton		8	no	0.154	0.19	0.478	0.506	(+)	(+)	no	no
9ZC51	zooplankton		NA	no	0.182	0.221	0.373	0.32	(+)	(+)	no	no
9ZC52	zooplankton		NA	no	0.214	0.497	0.51	0.477	(+-)	no	no	no
9ZC53	zooplankton		16	no	0.144	0.766	0.422	0.403	(+)	(+)	no	no
9ZC56	zooplankton		11	no	0.17	0.228	0.398	0.398	(+)	(+)	no	no
9ZC57	zooplankton		15	yes	0.231	0.28	0.141	0.142	(+)	(+)	no	(+)
9ZC6	zooplankton		6	yes	0.315	0.157	0.153	0.152	(+)	(+)	no	(+-)
9ZC62	zooplankton		16	yes	0.193	0.512	0.137	0.139	(+)	no	(+)	no
9ZC63	zooplankton		1	yes	0.167	0.189	0.159	0.191	(+)	(+)	(++)	(++)
9ZC64	zooplankton		8	no	0.139	0.144	0.436	0.437	(+)	(+)	no	no
9ZC65	zooplankton		8	no	0.13	0.5	0.514	0.489	(+)	(+)	no	no

9ZC69	zooplankton		7	yes	0.16	0.208	0.339	0.324	(+)	(+)	(+-)	(+-)
9ZC70	zooplankton		8	yes	0.202	0.227	0.148	0.148	(+)	(+)	no	no
9ZC73	zooplankton		8	yes	0.202	0.242	0.201	0.205	(+)	(+)	(+)	(+)
9ZC74	zooplankton		NA	no	0.135	0.156	0.397	0.387	(+)	(+)	no	no
9ZC76	zooplankton		8	no	0.194	0.196	0.476	0.449	(+)	(+)	no	no
9ZC77	zooplankton		8	no	0.083	0.084	0.477	0.492	(+)	(+)	no	no
9ZC78	zooplankton		8	no	0.153	0.171	0.489	0.501	(+)	(+)	no	no
9ZC79	zooplankton		8	no	0.164	0.197	0.457	0.453	(+)	(+-)	no	no
9ZC8	zooplankton		NA	no	0.458	0.24	0.359	0.371	(+)	(+)	no	no
9ZC80	zooplankton		8	no	0.088	0.116	0.511	0.511	(+)	(+)	no	no
9ZC81	zooplankton		NA	no	0.173	0.411	0.459	0.442	(+-)	(+-)	no	(+-)
9ZC83	zooplankton		NA	yes	0.246	0.241	0.173	0.194	(+)	(+)	no	(+)
9ZC85	zooplankton		8	no	0.171	0.208	0.42	0.404	(+)	(+)	no	no
9ZC87	zooplankton		3	yes	0.272	0.322	0.267	0.527	(+)	(+)	no	(+-)
9ZC88	zooplankton		8	no	0.154	0.212	0.45	0.46	(+)	(+)	no	no
9ZC9	zooplankton		6	no	0.276	0.235	0.385	0.397	(+)	(+)	no	no
9ZC90	zooplankton		8	no	0.1	0.111	0.49	0.49	(+)	(+)	no	no
9ZC92	zooplankton		16	yes	0.181	0.25	0.162	0.178	(+)	(+)	no	no
9ZC93	zooplankton		NA	no	0.15	0.194	0.488	0.475	(+-)	(+)	no	no
9ZC95	zooplankton		8	yes	0.204	0.227	0.241	0.284	(+)	(+)	no	(+-)
9ZC96	zooplankton		3	no	0.096	0.109	0.52	0.504	(+)	(+)	no	(+-)
9ZC98	zooplankton		NA	no	0.08	0.082	0.479	0.46	(+)	(+)	no	no
9ZD1	zooplankton		8	yes	0.202	0.285	0.159	0.175	no	no	no	no
9ZD14	zooplankton		NA	no	0.166	0.209	0.437	0.51	(+)	(+)	no	no
9ZD15	zooplankton		NA	no	0.155	0.193	0.406	0.422	(+)	no	no	no
9ZD17	zooplankton		NA	no	0.185	0.208	0.42	0.422	(+)	no	no	no
9ZD19	zooplankton		8	no	0.088	0.105	0.517	0.442	(+)	(+)	no	no
9ZD2	zooplankton		NA	no	0.151	0.443	0.451	0.47	(+)	(+)	no	no
9ZD21	zooplankton		8	no	0.131	0.215	0.477	0.462	(+)	(+)	no	no
9ZD23	zooplankton		8	no	0.164	0.22	0.429	0.421	(+)	(+)	no	no
9ZD25	zooplankton		NA	no	0.278	0.432	0.386	0.388	(+)	(+)	no	no
9ZD26	zooplankton		NA	no	0.175	0.249	0.454	0.454	(+)	(+)	no	no
9ZD27	zooplankton		16	yes	0.164	0.166	0.183	0.18	(+)	(+)	(+)	(+-)
9ZD29	zooplankton		NA	no	0.085	0.108	0.521	0.47	(+)	(+)	no	no
9ZD30	zooplankton		16	no	0.158	0.377	0.469	0.486	(+)	(+)	no	no
9ZD32	zooplankton		NA	no	0.156	0.225	0.465	0.45	no	no	no	no
9ZD34	zooplankton		16	no	0.165	0.353	0.47	0.48	(+)	(+)	no	no
9ZD4	zooplankton		NA	no	0.081	0.086	0.539	0.523	(+)	(+)	no	no
9ZD40	zooplankton		16	no	0.162	0.443	0.433	0.418	(+-)	(+-)	no	no
9ZD44	zooplankton		11	yes	0.196	0.235	0.264	0.238	(+)	(+)	no	no
9ZD47	zooplankton		16	no	0.171	0.102	0.46	0.507	(+)	(+)	no	no
9ZD48	zooplankton		NA	no	0.162	0.178	0.428	0.418	(+)	(+)	no	(+-)
9ZD50	zooplankton		NA	yes	0.19	0.231	0.299	0.29	(+)	(+)	no	(+-)
9ZD51	zooplankton		NA	no	0.169	0.425	0.411	0.419	(+)	(+-)	no	no
9ZD7	zooplankton		8	yes	0.195	0.258	0.144	0.151	(+)	(+)	no	no
9ZD9	zooplankton		NA	no	0.164	0.426	0.454	0.471	(+-)	no	no	no
FF_1	0.2	fall	NA	yes	0.182	0.201	0.151	0.413	(+)	(+)	(+)	(+)
FF_101	0.2	fall	NA	yes	0.155	0.16	0.185	0.21	no	no	no	no
FF_104	0.2	fall	1	yes	0.16	0.162	0.277	0.175	(+)	(+)	no	no
FF_107	0.2	fall	16	no	0.096	0.105	0.518	0.498	(+)	(+)	no	no
FF_110	0.2	fall	6	yes	0.155	0.144	0.278	0.292	(+)	(+-)	(++)	no
FF_112	0.2	fall	NA	no	0.096	0.1	0.465	0.465	(+)	(+)	no	no
FF_115	0.2	fall	NA	yes	0.153	0.175	0.276	0.293	(+)	(+)	(+-)	(+--)

FF_124	0.2	fall	NA	yes	0.097	0.099	0.498	0.509	(+)	(+)	(+)	(+-)
FF_126	0.2	fall	NA	yes	0.122	0.122	0.469	0.491	(+)	(+)	(++)	(++)
FF_13	0.2	fall	16	yes	0.171	0.181	0.254	0.4	(+)	(+)	(+)	no
FF_132	0.2	fall	16	yes	0.11	0.112	0.491	0.491	(+)	(+)	(+-)	(+-)
FF_134	0.2	fall	3	no	0.128	0.131	0.452	0.453	(+)	(+)	(+-)	no
FF_136	0.2	fall	6	yes	0.17	0.187	0.21	0.231	(+)	(+)	(+-)	(+--)
FF_14	0.2	fall	16	no	0.362	0.186	0.378	0.54	(+)	(+)	no	no
FF_140	0.2	fall	16	yes	0.172	0.186	0.177	0.198	(+)	(+)	(++)	(+-)
FF_142	0.2	fall	1	no	0.099	0.102	0.488	0.495	(+)	(+)	no	no
FF_144	0.2	fall	18	no	0.094	0.094	0.523	0.517	(+)	(+)	no	no
FF_145	0.2	fall	NA	no	0.123	0.122	0.484	0.475	(+)	(+)	no	no
FF_146	0.2	fall	6	no	0.143	0.141	0.49	0.503	(+)	(+)	no	(+--)
FF_149	0.2	fall	16	yes	0.189	0.212	0.179	0.182	no	no	no	no
FF_15	0.2	fall	2	no	0.162	0.18	0.401	0.389	no	no	no	no
FF_152	0.2	fall	NA	yes	0.159	0.165	0.205	0.237	(+)	(+)	(++)	(+)
FF_170	0.2	fall	NA	yes	0.097	0.096	0.505	0.496	(+)	(+)	(++)	(++)
FF_172	0.2	fall	NA	yes	0.133	0.145	0.312	0.298	(+)	(+)	(++)	(++)
FF_178	0.2	fall	5	no	0.075	0.075	0.537	0.516	(+)	(+)	no	no
FF_179	0.2	fall	NA	yes	0.09	0.096	0.552	0.494	(+)	(+)	(+)	(+-)
FF_182	0.2	fall	2	no	0.075	0.077	0.529	0.53	(+)	(+)	no	(+-)
FF_187	0.2	fall	3	yes	0.075	0.079	0.547	0.493	(+)	(+)	(+)	(+-)
FF_188	0.2	fall	3	yes	0.075	0.075	0.532	0.51	(+)	(+)	(+--)	(+--)
FF_190	0.2	fall	1	no	0.08	0.084	0.513	0.524	(+)	(+)	no	no
FF_191	0.2	fall	NA	yes	0.076	0.079	0.568	0.515	(+)	(+)	(+)	(+-)
FF_197	0.2	fall	3	yes	0.158	0.158	0.192	0.2	(+)	(+)	(++)	(++)
FF_2	0.2	fall	3	yes	0.276	0.263	0.286	0.139	(+)	(+)	(+-)	(+)
FF_217	0.2	fall	NA	yes	0.158	0.157	0.434	0.466	(+)	(+)	(++)	(++)
FF_22	0.2	fall	16	yes	0.319	0.173	0.166	0.401	(+)	(+)	(++)	(++)
FF_222	0.2	fall	3	yes	0.271	0.279	0.341	0.351	(+)	(+)	(+--)	(+-)
FF_224	0.2	fall	6	yes	0.138	0.077	0.504	0.52	(+)	(+)	(+--)	(+--)
FF_225	0.2	fall	16	yes	0.191	0.153	0.251	0.259	(+)	no	(+)	no
FF_233	0.2	fall	16	yes	0.09	0.094	0.47	0.424	(+)	(+)	(++)	(++)
FF_236	0.2	fall	3	yes	0.071	0.077	0.555	0.553	(+)	(+)	(+-)	(+-)
FF_238	0.2	fall	6	no	0.075	0.077	0.548	0.522	(+)	(+)	no	no
FF_239	0.2	fall	NA	yes	0.185	0.325	0.223	0.178	no	no	no	no
FF_24	0.2	fall	NA	yes	0.228	0.214	0.134	0.258	(+)	(+)	(++)	(++)
FF_263	0.2	fall	3	yes	0.144	0.162	0.548	0.331	(+)	(+)	(+)	(+)
FF_265	0.2	fall	16	yes	0.122	0.1	0.505	0.496	(+)	(+)	(+--)	(+--)
FF_268	0.2	fall	16	yes	0.146	0.153	0.228	0.236	no	(+)	no	(+-)
FF_274	0.2	fall	15B	yes	0.109	0.106	0.426	0.41	(+)	(+)	(++)	(+)
FF_279	0.2	fall	3	yes	0.143	0.154	0.229	0.242	no	(+)	no	(++)
FF_292	0.2	fall	16	no	0.132	0.146	0.296	0.315	(+)	(+)	no	no
FF_296	0.2	fall	NA	no	0.085	0.089	0.503	0.489	(+)	(+)	(+-)	no
FF_298	0.2	fall	1	no	0.117	0.118	0.474	0.478	(+)	(+)	(+--)	no
FF_299	0.2	fall	16	no	0.08	0.081	0.489	0.505	(+)	(+)	no	no
FF_3	0.2	fall	16	no	0.221	0.207	0.383	0.142	(+)	(+)	no	no
FF_30	0.2	fall	NA	no	0.11	0.118	0.422	0.471	(+)	(+)	(+-)	no
FF_301	0.2	fall	NA	yes	0.155	0.166	0.22	0.221	(+)	(+)	(++)	(+-)
FF_302	0.2	fall	2	yes	0.152	0.163	0.202	0.223	(+)	(+)	(+--)	(+--)
FF_304	0.2	fall	NA	no	0.079	0.081	0.517	0.537	(+)	(+)	no	no
FF_305	0.2	fall	16	yes	0.102	0.105	0.487	0.487	(+)	(+)	(+)	(+-)
FF_308	0.2	fall	NA	no	0.115	0.108	0.481	0.481	(+)	(+)	(+--)	no
FF_309	0.2	fall	NA	yes	0.174	0.316	0.186	0.225	(+)	(+)	(+)	no

FF_310	0.2	fall	5	no	0.117	0.131	0.45	0.45	(+)	(+)	no	no
FF_317	0.2	fall	3	no	0.217	0.185	0.452	0.429	(+)	no	(+--)	no
FF_319	0.2	fall	18	yes	0.126	0.128	0.268	0.271	(+)	(+)	no	(+--)
FF_320	0.2	fall	NA	no	0.078	0.162	0.507	0.463	no	(+)	no	no
FF_344	0.2	fall	3	yes	0.132	0.131	0.443	0.443	(+)	(+)	(+--)	(+--)
FF_348	0.2	fall	NA	no	0.101	0.1	0.539	0.513	(+)	(+)	no	no
FF_349	0.2	fall	NA	no	0.075	0.076	0.564	0.574	(+)	(+)	no	no
FF_352	0.2	fall	NA	yes	0.152	0.147	0.239	0.25	(+)	no	(+-)	no
FF_36	0.2	fall	16	yes	0.229	0.198	0.128	0.684	no	(+)	no	(+)
FF_371	0.2	fall	NA	yes	0.154	0.152	0.269	0.295	no	no	no	no
FF_372	0.2	fall	13	no	0.092	0.098	0.548	0.531	(+)	(+)	no	no
FF_373	0.2	fall	1	yes	0.126	0.107	0.498	0.522	(+)	(+)	(+-)	(+)
FF_379	0.2	fall	NA	yes	0.178	0.215	0.284	0.298	(+)	(+)	(+)	(+)
FF_380	0.2	fall	3	yes	0.129	0.118	0.48	0.498	(+)	(+)	(+--)	(+)
FF_385	0.2	fall	NA	yes	0.087	0.086	0.569	0.552	(+)	(+)	(+--)	(+-)
FF_386	0.2	fall	16	no	0.079	0.071	0.516	0.53	(+)	(+)	no	no
FF_443	0.2	fall	3	yes	0.106	0.377	0.47	0.327	(+)	(+)	(+--)	(+--)
FF_451	0.2	fall	3	yes	0.085	0.084	0.483	0.52	(+)	(+)	(+--)	(+-)
FF_452	0.2	fall	2	yes	0.117	0.106	0.477	0.48	(+)	(+)	(+-)	(+-)
FF_457	0.2	fall	5	no	0.11	0.112	0.5	0.51	(+)	(+)	(+--)	no
FF_458	0.2	fall	1	yes	0.185	0.152	0.347	0.377	(+)	(+)	(++)	(+)
FF_459	0.2	fall	16	yes	0.152	0.152	0.241	0.318	no	(+)	no	(++)
FF_461	0.2	fall	NA	no	0.174	0.134	0.428	0.428	(+)	(+)	no	no
FF_464	0.2	fall	NA	yes	0.075	0.08	0.218	0.245	(+)	(+)	(+-)	(+-)
FF_465	0.2	fall	NA	yes	1.265	0.175	0.526	0.475	(+)	(+)	(+--)	(+--)
FF_466	0.2	fall	16	no	0.156	0.157	0.34	0.357	(+)	no	(++)	no
FF_47	0.2	fall	16	no	0.258	0.189	0.136	0.481	no	no	no	no
FF_472	0.2	fall	2	no	0.1	0.094	0.482	0.518	(+)	(+)	no	no
FF_475	0.2	fall	16	no	0.096	0.218	0.491	0.413	(+)	(+)	no	no
FF_481	0.2	fall	16	yes	0.134	0.14	0.245	0.264	(+)	(+)	(++)	(++)
FF_482	0.2	fall	NA	no	0.098	0.089	0.481	0.501	(+)	(+)	(+--)	no
FF_483	0.2	fall	3	no	0.111	0.109	0.457	0.463	(+)	(+)	no	(+-)
FF_484	0.2	fall	16	yes	0.175	0.277	0.2	0.224	(+)	(+)	(++)	(++)
FF_488	0.2	fall	NA	yes	0.15	0.147	0.46	0.465	(+)	(+)	(+--)	(+--)
FF_489	0.2	fall	3	yes	0.275	0.114	0.353	0.508	(+)	(+)	(+--)	(+)
FF_491	0.2	fall	5	no	0.141	0.114	0.443	0.432	(+)	(+)	(+--)	no
FF_495	0.2	fall	3	yes	0.103	0.105	0.515	0.515	(+)	(+)	(+-)	(+-)
FF_498	0.2	fall	3	yes	0.125	0.113	0.485	0.476	(+)	(+)	(+)	(++)
FF_500	0.2	fall	16	yes	0.236	0.148	0.274	0.264	(+)	(+)	(++)	(+)
FF_51	0.2	fall	2	no	0.156	0.166	0.452	0.435	(+-)	no	no	no
FF_52	0.2	fall	10	no	0.157	0.163	0.21	0.379	no	(+)	no	no
FF_53	0.2	fall	16	yes	0.224	0.208	0.126	0.501	no	(+)	no	(+)
FF_54	0.2	fall	NA	yes	0.206	0.194	0.141	0.226	(+)	(+)	(+)	(++)
FF_57	0.2	fall	9	no	0.102	0.108	0.47	0.161	(+)	(+)	no	no
FF_58	0.2	fall	13	no	0.106	0.106	0.49	0.4	(+)	(+)	no	no
FF_59	0.2	fall	NA	no	0.163	0.173	0.433	0.513	(+)	(+)	no	no
FF_6	0.2	fall	18	yes	0.389	0.553	0.146	0.171	no	no	no	no
FF_61	0.2	fall	NA	no	0.146	0.16	0.478	0.349	(+)	(+)	no	no
FF_62	0.2	fall	2	no	0.224	0.233	0.427	0.142	no	(+)	no	no
FF_68	0.2	fall	3	yes	0.387	0.388	0.208	0.338	(+)	(+)	(+)	(+)
FF_72	0.2	fall	2	no	0.155	0.174	0.435	0.445	no	no	no	no
FF_75	0.2	fall	15A	no	0.16	0.172	0.558	0.389	(+)	(+)	no	no
FF_77	0.2	fall	2	yes	0.358	0.361	0.232	0.39	(+)	(+)	(+)	(++)

FF_8	0.2	fall	2	no	0.161	0.2	0.523	0.393	no	no	no	no
FF_85	0.2	fall	2	no	0.197	0.175	0.443	0.415	(+)	(+)	no	no
FF_9	0.2	fall	2	no	0.154	0.178	0.446	0.473	no	no	no	no
FF_91	0.2	fall	NA	no	0.107	0.112	0.689	0.411	(+)	(+)	no	no
FF_93	0.2	fall	3	yes	0.135	0.159	0.275	0.292	no	(+)	no	no
FF_99	0.2	fall	3	yes	0.088	0.089	0.489	0.501	(+)	(+)	(+-)	(+-)
ZF_10	64	fall	13	no	0.1	0.125	0.512	0.487	(+)	(+)	no	no
ZF_100	64	fall	13	no	0.124	0.131	0.488	0.46	(+)	(+)	no	no
ZF_103	64	fall	NA	yes	0.171	0.167	0.483	0.152	(+)	(+)	(+-)	(+-)
ZF_105	64	fall	NA	no	0.114	0.116	0.497	0.481	(+)	(+)	no	no
ZF_107	64	fall	10	yes	0.078	0.077	0.531	0.534	(+)	(+)	(+--)	(+--)
ZF_109	64	fall	10	yes	0.082	0.081	0.517	0.5	(+)	(+)	(+--)	(+--)
ZF_11	64	fall	8	no	0.094	0.091	0.448	0.477	(+)	(+)	no	no
ZF_113	64	fall	13	no	0.088	0.102	0.53	0.536	(+)	(+)	no	no
ZF_115	64	fall	NA	no	0.097	0.11	0.522	0.497	(+)	(+)	no	no
ZF_129	64	fall	10	yes	0.081	0.082	0.381	0.391	(+)	(+)	(+--)	(+--)
ZF_135	64	fall	13	no	0.079	0.08	0.532	0.53	(+)	(+)	no	no
ZF_14	64	fall	15A	no	0.117	0.118	0.504	0.481	(+)	(+)	(+--)	no
ZF_15	64	fall	13	no	0.079	0.079	0.529	0.522	(+)	(+)	no	no
ZF_155	64	fall	NA	no	0.126	0.126	0.394	0.387	(+)	(+)	no	no
ZF_157	64	fall	NA	no	0.141	0.13	0.356	0.404	(+)	(+)	no	no
ZF_159	64	fall	13	no	0.089	0.087	0.513	0.509	(+)	(+)	no	no
ZF_17	64	fall	13	no	0.091	0.087	0.531	0.539	(+)	(+)	no	no
ZF_170	64	fall	15A	yes	0.207	0.211	0.403	0.394	(+)	(+)	(+-)	(+-)
ZF_173	64	fall	10	no	0.156	0.156	0.429	0.387	(+)	(+)	no	(+-)
ZF_175	64	fall	12	yes	0.277	0.293	0.144	0.135	(+)	(+)	(+-)	(+)
ZF_18	64	fall	5	no	0.117	0.13	0.427	0.437	(+)	(+)	no	no
ZF_181	64	fall	NA	yes	0.225	0.209	0.148	0.133	(+)	(+)	(+)	(+)
ZF_189	64	fall	NA	yes	0.144	0.148	0.169	0.181	(+)	(+)	no	(+-)
ZF_19	64	fall	NA	no	0.081	0.077	0.516	0.522	(+)	(+)	no	no
ZF_192	64	fall	12	yes	0.245	0.223	0.162	0.143	(+)	(+)	(+-)	(+-)
ZF_193	64	fall	16	no	0.224	0.258	0.429	0.457	(+)	(+)	no	no
ZF_195	64	fall	13	no	0.098	0.097	0.556	0.546	(+)	(+)	no	no
ZF_197	64	fall	25	no	0.179	0.189	0.385	0.382	(+)	(+)	no	no
ZF_2	64	fall	NA	no	0.113	0.128	0.473	0.523	(+)	(+)	no	no
ZF_20	64	fall	16	yes	0.157	0.156	0.18	0.184	(+)	(+)	(++)	(++)
ZF_201	64	fall	10	yes	0.161	0.158	0.193	0.206	(+)	(+)	no	no
ZF_203	64	fall	16	no	0.257	0.257	0.408	0.4	(+)	(+)	no	no
ZF_205	64	fall	15A	no	0.167	0.189	1.068	0.425	(+)	(+)	no	(+-)
ZF_207	64	fall	15A	no	0.231	0.231	0.409	0.368	(+)	no	no	no
ZF_209	64	fall	NA	yes	0.175	0.671	0.287	0.267	(+)	(+)	(+)	(+)
ZF_21	64	fall	13	no	0.083	0.083	0.524	0.522	(+)	(+)	no	no
ZF_211	64	fall	8	yes	0.204	0.218	0.203	0.157	(+)	(+)	(+-)	(+-)
ZF_215	64	fall	13	no	0.095	0.096	0.554	0.53	(+)	(+)	no	no
ZF_219	64	fall	13	no	0.246	0.244	0.426	0.422	(+)	(+)	no	no
ZF_22	64	fall	5	no	0.112	0.108	0.476	0.478	(+)	(+)	no	no
ZF_221	64	fall	8	yes	0.189	0.185	0.176	0.146	(+)	(+)	(+-)	(+-)
ZF_223	64	fall	12	yes	0.25	0.257	0.173	0.153	(+)	(+)	(+-)	(+-)
ZF_225	64	fall	5	no	0.261	0.255	0.358	0.404	(+)	(+)	no	no
ZF_25	64	fall	8	no	0.099	0.093	0.354	0.368	(+)	(+)	(+-)	no
ZF_255	64	fall	15A	no	0.138	0.353	0.444	0.458	(+)	(+)	no	(+-)
ZF_257	64	fall	10	yes	0.14	0.144	0.189	0.445	(+)	(+)	(+-)	(+-)
ZF_261	64	fall	NA	no	0.145	0.148	0.419	0.416	(+)	(+)	no	no

ZF_265	64	fall	NA	no	0.155	0.347	0.416	0.462	(+)	(+)	no	no
ZF_269	64	fall	13	no	0.132	0.134	0.49	0.127	(+)	(+)	no	no
ZF_27	64	fall	10	yes	0.095	0.095	0.453	0.478	(+)	(+)	(+-)	(+-)
ZF_270	64	fall	15A	no	0.195	0.185	0.392	0.138	(+)	(+)	no	no
ZF_31	64	fall	NA	yes	0.151	0.151	0.156	0.164	no	no	no	no
ZF_32	64	fall	13	no	0.084	0.084	0.51	0.536	(+)	(+)	no	no
ZF_33	64	fall	13	no	0.085	0.083	0.503	0.54	(+)	(+)	no	no
ZF_34	64	fall	5	no	0.116	0.134	0.429	0.437	(+)	(+)	no	no
ZF_35	64	fall	5	no	0.123	0.117	0.422	0.444	(+)	(+)	no	no
ZF_36	64	fall	8	no	0.079	0.083	0.505	0.499	(+)	(+)	no	no
ZF_41	64	fall	16	yes	0.158	0.156	0.199	0.188	(+)	(+)	(+)	(++)
ZF_45	64	fall	NA	no	0.137	0.134	0.469	0.454	(+)	(+)	no	no
ZF_47	64	fall	8	yes	0.105	0.104	0.369	0.365	(+)	(+)	(+-)	(+-)
ZF_49	64	fall	13	no	0.09	0.087	0.542	0.525	(+)	(+)	no	no
ZF_5	64	fall	NA	no	0.086	0.084	0.508	0.518	(+)	(+)	(+--)	no
ZF_50	64	fall	10	yes	0.124	0.14	0.161	0.242	(+)	(+)	(+-)	(+-)
ZF_53	64	fall	NA	no	0.136	0.129	0.453	0.454	(+)	(+)	no	no
ZF_59	64	fall	NA	yes	0.152	0.152	0.164	0.159	no	no	no	no
ZF_61	64	fall	NA	yes	0.222	0.216	0.146	0.161	(+)	(+)	(+-)	(+-)
ZF_63	64	fall	13	no	0.087	0.084	0.524	0.54	(+)	(+)	no	no
ZF_67	64	fall	10	yes	0.088	0.088	0.454	0.441	(+)	(+)	(+-)	(+-)
ZF_7	64	fall	NA	yes	0.159	0.156	0.166	0.169	(+)	(+)	(+)	no
ZF_71	64	fall	NA	yes	0.179	0.18	0.227	0.239	(+)	(+)	(+--)	no
ZF_72	64	fall	12	yes	0.508	0.196	0.134	0.476	(+)	(+)	(+)	(+)
ZF_73	64	fall	8	yes	0.103	0.103	0.319	0.315	(+)	(+)	(+--)	(+--)
ZF_76	64	fall	16	no	0.158	0.166	0.436	0.43	no	no	no	no
ZF_77	64	fall	NA	yes	0.164	0.169	0.193	0.195	(+)	(+)	(+)	(+)
ZF_79	64	fall	NA	yes	0.155	0.157	0.195	0.203	(+)	(+)	(+-)	(+)
ZF_8	64	fall	13	no	0.088	0.087	0.524	0.523	(+)	(+)	no	no
ZF_81	64	fall	10	yes	0.082	0.081	0.497	0.484	(+)	(+)	(+-)	(+-)
ZF_83	64	fall	8	yes	0.111	0.298	0.336	0.402	(+)	(+)	(+-)	(+)
ZF_89	64	fall	12	yes	0.154	0.155	0.188	0.19	(+)	(+)	(+-)	(+-)
ZF_9	64	fall	13	no	0.088	0.083	0.524	0.529	(+)	(+)	no	no
ZF_90	64	fall	19	no	0.139	0.167	0.421	0.131	(+)	(+)	no	no
ZF_91	64	fall	13	no	0.086	0.085	0.526	0.544	(+)	(+)	no	no
ZF_92	64	fall	13	no	0.148	0.15	0.452	0.509	(+)	(+)	no	no
ZF_97	64	fall	NA	yes	0.287	0.138	0.218	0.226	(+)	(+)	(+-)	(+-)
ZS_10	64	spring	21	yes	0.623	0.451	0.162	0.186	(+)	(+)	(+)	(++)
ZS_103	64	spring	NA	no	0.162	0.137	0.613	0.5	(+)	(+)	no	no
ZS_107	64	spring	22	yes	0.176	0.51	0.15	0.144	(+)	(+)	(+)	(++)
ZS_109	64	spring	NA	no	0.23	0.236	0.469	0.473	(+)	(+-)	no	no
ZS_11	64	spring	11	no	0.555	0.307	0.352	0.374	(+)	(+)	no	(+-)
ZS_111	64	spring	NA	no	0.43	0.157	0.414	0.401	(+)	(+)	no	no
ZS_113	64	spring	NA	no	0.179	0.076	0.492	0.487	(+)	(+)	no	no
ZS_117	64	spring	22	no	0.406	0.187	0.413	0.433	no	(+)	no	no
ZS_119	64	spring	19	no	0.262	0.077	0.406	0.435	(+)	(+)	no	no
ZS_121	64	spring	NA	no	0.15	0.155	0.554	0.439	(+)	(+)	no	no
ZS_125	64	spring	22	yes	0.407	0.238	0.155	0.124	(+)	(+)	(++)	(++)
ZS_136	64	spring	19	yes	0.146	0.17	0.176	0.158	(+)	(+)	(+)	(+)
ZS_138	64	spring	20	no	0.15	0.124	0.442	0.472	(+)	(+)	no	no
ZS_139	64	spring	11	no	0.228	0.282	0.361	0.419	(+)	(+)	no	(+-)
ZS_15	64	spring	11	no	0.261	0.298	0.355	0.368	(+)	(+)	no	(+-)
ZS_151	64	spring	NA	no	0.978	0.224	0.436	0.37	(+)	(+)	no	no
ZS_161	64	spring	NA	no	0.26	0.262	0.447	0.449	(+)	(+)	no	no

ZS_163	64	spring	NA	no	0.15	0.144	0.471	0.487	(+)	(+)	no	no
ZS_166	64	spring	NA	no	0.185	0.17	0.454	0.401	(+)	(+)	no	no
ZS_168	64	spring	21	no	0.309	0.251	0.406	0.375	(+)	(+)	no	no
ZS_178	64	spring	20	no	0.16	0.146	0.479	0.47	(+)	no	no	no
ZS_18	64	spring	24B	yes	0.16	0.679	0.154	0.159	no	no	no	no
ZS_180	64	spring	21	no	0.131	0.133	0.496	0.513	(+)	(+)	no	no
ZS_181	64	spring	21	no	0.131	0.133	0.502	0.507	(+)	(+)	no	no
ZS_183	64	spring	19	no	0.147	0.149	0.469	0.476	(+)	(+)	no	no
ZS_185	64	spring	21	no	0.655	0.208	0.397	0.447	(+)	(+)	no	no
ZS_186	64	spring	17	no	0.084	0.157	0.491	0.526	(+)	(+-)	no	no
ZS_187	64	spring	21	no	0.109	0.188	0.497	0.502	(+)	(+)	no	no
ZS_189	64	spring	11	yes	0.218	0.143	0.183	0.187	(+)	(+)	(+)	(+)
ZS_19	64	spring	20	yes	0.162	0.367	0.162	0.143	(+)	(+)	(++)	(++)
ZS_190	64	spring	11	no	0.219	0.094	0.423	0.428	(+)	(+)	no	no
ZS_192	64	spring	NA	no	0.218	0.181	0.451	0.444	(+)	(+)	no	no
ZS_193	64	spring	21	no	0.172	0.074	0.46	0.473	(+)	(+)	no	no
ZS_195	64	spring	21	no	0.11	0.11	0.479	0.509	(+)	(+)	no	no
ZS_196	64	spring	NA	no	0.477	0.399	0.368	0.396	(+)	(+)	no	no
ZS_198	64	spring	21	no	0.11	0.123	0.497	0.501	(+)	(+)	no	no
ZS_2	64	spring	22	no	0.114	0.176	0.471	0.482	(+)	(+)	no	no
ZS_201	64	spring	19	yes	0.186	0.387	0.175	0.172	(+)	(+)	(+-)	(+-)
ZS_202	64	spring	NA	no	0.358	0.077	0.433	0.434	(+)	(+)	no	(+-)
ZS_205	64	spring	11	no	0.277	0.075	0.429	0.431	(+)	(+)	no	(+-)
ZS_208	64	spring	21	no	0.373	0.189	0.442	0.453	(+)	no	no	no
ZS_210	64	spring	21	no	0.109	0.144	0.483	0.485	(+)	(+)	no	no
ZS_211	64	spring	20	yes	0.196	0.074	0.2	0.209	(+)	(+)	(+)	(+)
ZS_213	64	spring	21	no	0.117	0.076	0.492	0.515	(+)	(+)	no	no
ZS_22	64	spring	11	no	0.18	0.243	0.444	0.478	(+)	(+)	no	no
ZS_23	64	spring	NA	no	0.226	0.072	0.484	0.503	(+)	(+)	no	no
ZS_25	64	spring	NA	yes	0.354	0.264	0.149	0.139	(+)	(+)	(++)	(+-)
ZS_29	64	spring	21	yes	0.434	0.439	0.196	0.193	(+)	(+)	(++)	(+)
ZS_33	64	spring	11	no	0.14	0.242	0.452	0.447	(+)	(+)	no	no
ZS_36	64	spring	24A	yes	0.698	0.394	0.157	0.158	(+)	(+)	(++)	(+)
ZS_41	64	spring	20	no	0.237	0.208	0.453	0.481	(+)	(+)	no	no
ZS_46	64	spring	20	no	0.406	0.075	0.405	0.458	(+)	(+)	no	no
ZS_5	64	spring	19	no	0.386	0.257	0.417	0.421	(+)	(+)	no	no
ZS_53	64	spring	25	yes	0.132	0.134	0.163	0.175	(+)	(+)	(+)	(++)
ZS_55	64	spring	NA	no	0.209	0.074	0.443	0.467	(+)	(+)	no	no
ZS_58	64	spring	20	no	0.116	0.074	0.481	0.498	(+)	(+)	no	no
ZS_64	64	spring	NA	yes	0.156	0.136	0.139	0.138	(+)	(+)	(++)	(++)
ZS_66	64	spring	21	no	0.241	0.37	0.464	0.483	(+-)	no	no	no
ZS_74	64	spring	24B	yes	0.146	0.128	0.145	0.136	(+)	(+)	(++)	(++)
ZS_76	64	spring	24B	yes	0.172	0.082	0.138	0.141	(+)	(+)	(++)	(++)
ZS_8	64	spring	21	yes	0.285	0.335	0.148	0.123	no	no	no	no
ZS_84	64	spring	24B	yes	0.184	0.127	0.143	0.163	(+)	no	(++)	no
ZS_88	64	spring	24B	yes	0.143	0.077	0.209	0.155	(+)	(+)	(++)	(++)
ZS_90	64	spring	24B	yes	0.145	0.076	0.141	0.135	(+)	(+)	(++)	(++)
ZS_92	64	spring	NA	no	0.22	0.079	0.378	0.375	(+)	(+)	no	no
ZS_99	64	spring	11	no	0.194	0.141	0.444	0.464	no	(+)	no	no

2.9.2. Table S2.2. Result of bioinformatic searches for siderophore biosynthetic genes in sequenced vibrio strains.

Strain	# clusters detected	Inferred siderophore type	Gene #	RAST Annotation	Domain name (pfam number)	Adenylation domain specificity
0407CSC89	2 (1 siderophore in 2 contigs)	aerobactin - non functional (missing IucB)	prot_04034	-none	- none	
			prot_04035	citrate 6-N-acetyl-6-Nhydroxy-L-lysine ligase-aerobactin biosynthesis protein IucA, siderophore synthetase superfamily group C	-IucA/IucC family (pfam04183)	
			prot_04036	-none	MSF_1- major facilitator family (pfam07690)	
			prot_04037	-none	-protein of unknown function (DUF2834) (pfam11196)	
			prot_04038	-none	-none	
			prot_04039	-none	-none	
			prot_04040	-none	-none	
			prot_04041	-none		
			prot_04042	-none		
			prot_04043	-none		
			prot_04044	-none		
			prot_04045	-none		
			prot_04046 (end of contig)	-none		
		aerobactin - non-functional missing IucB	prot_01492	FIG01065164- hypothetical	- type1 phosphodiesterase (pfam01663)	
			prot_01493	none	- spondin N-terminal (pfam06468)	
			prot_01494	none		
			prot_01495	FIG01067830- hypothetical	- MFS_1 major facilitator superfamily (pfam07690)	
			prot_01496	none	-IucA/IucC family (pfam04183) -Ferric iron reductase FhuF like transporter (pfam06276) - Acetyltransferase 8 (GNAT) domain (pfam13523)	
			prot_01497	L-lysine 6-monooxygenase, aerobactin biosynthesis protein IucD	- L-lysine 6-monooxygenase (pfam13434)	
			prot_01498	Aerobactin siderophore receptor IutA	-TonB-dependent receptor plug domain (pfam07715) -TonB-dpendent receptor (pfam00593)	
			prot_01499	dTDP-4-dehydrorhamnose reductase		
FF-33	1 cluster detected	non-siderophore nrps cluster				

			prot_01499	dTDP-4-dehydrorhamnose reductase		
FF-33	1 cluster detected	non-siderophore nrps cluster				
FF-162	1 cluster detected	non-siderophore nrps cluster				
1F-230	2	nrps cluster: yersiniabactin (missing the salicylate ligase?)	prot_02593	none	no hit	
			prot_02594	none	AraC regulator HTH_AraC: pfam00165: helix-turn-helix regulatory protein.	
			prot_02595	none	ACPS superfamily: pfam01648 (4'-phosphopantetheinyl transferase) entD multidomain	
			prot_02596	none	NADB Rossmann superfamily	
			prot_02597	none	Lys9 superfamily, Saccharopine dehydrogenase : pfam 03435 NADB Rossman superfamily (NAD dehydrogenase)	
			prot_02598	TonB dependent receptor		
			prot_02599	none	FepC ABC Iron Siderophore B12 Hemin: COG1120 ATPase component	
			prot_02600	none	TM ABC iron siderophore like FecCD: pfam01032- transport family-permease subunity. FecD	
			prot_02601	none	-TonB siderophore OM receptor -TM ABC iron siderophore: FeCD (pfam 01032)	
			prot_02602	none	-TroA like superfamily (cd 01142) -FepB, FecB periplasmic binding protein (pfam 01497)	
			prot_02603	none	GrsT Thioesterase (pfam 00975)	
			prot_02604	Irp2 (yersiniabactin synthesis enzyme)	-condensation (pfam00668) -nrps (pfam08415) -amp-binding (pfam00501) -pp-binding (pfam00550) -NADB Rossmann superfamily (pfam07993)	cysteine
			prot_02605 (end of contig)	none	-AMP-binding (pfam00501) -pp-binding (pfam 00550)	No A-domain detected in NRPS-predictor2
		siderophore	prot_02665	AsbA	-IucA/IucC family (pfam04183)	

		cluster: anthrachelin			-FhuF (ferric iron reductase family) (pfam06276)	
			prot_02666	AsbB	-IucA/IucC family (pfam04183) -FhuF (ferric iron reductase family) (pfam06276)	
			prot_02667	long chain fatty acid coA ligase	- amp-binding (pfam00501) - amp-binding (pfam00501)	No A-domain detected in NRPS predictor 2
			prot_02668	acyl carrier protein	- pp-binding (pfam00550)	
			prot_02669	hypothetical GBAA1986		
1F-211	2	siderophore cluster: anthrachelin	IucA-like, prot_01246	anthrachelin biosysnthesis AsbA, siderophore synthetase superfamily, groupA. large component acetyl transferase	-IucA/IucC family (pfam04183) -FhuF (ferric iron reductase family) (pfam06276)	N/A
			IucC-like, prot_01247	anthrachelin biosythesis protein AsbB, siderophore synthetase super family group C, ligase.	-IucA/IucC family (pfam04183) -FhuF (ferric iron reductase family) (pfam06276)	N/A
			prot_01248	CoA ligase associated with anthrachelin synthesis	- phosphomethyl pyrimidine kinase (pfam08543)	N/A
			prot_01249	acyl carrier protein associated with anthrachelin biosynthesis		N/A
			prot_01250	hypothetical protein GBAA1986		
		nrps cluster: yersiniabactin (pks domain missing)	prot_01506	none	- ABC transporter- transmembrane region (pfam00664) - ABC transporter (pfam00005)	
			prot_01507	none	- ABC transporter- transmembrane region (pfam00664) - ABC transporter (pfam00005)	
			prot_01508	none	-AMP-binding (pfam00501)	2,3-dihydroxybenzoic acid salicylate
			prot_01509	none	- chorismate binding (pfam00425)	
			prot_01510	dihydroaeruginoate synthetase PchE, nrps module	-pp-binding (pfam00550) -condensation (pfam00668) -nrps (pfam08415)	
			prot_01511	none	none	
			prot_01512	long chain fatty acid coA ligase	-pp-binding (pfam00550) -condensation: (pfam00668) -nrps (pfam08415) -condensation (pfam00668) -nrps (pfam08415) -amp-binding (pfam00501) -pp-binding (pfam00550)	cysteine
			prot_01513	iron acquisition yersiniabactin synthesis	-condensation (pfam00668)	cysteine

				enzyme	-nrps (pfam08415) -amp-binding (pfam00501) -pp-binding (pfam00550) -nad-binding:	
			prot_01514	none	Thioesterase domain (pfam00975)	
			prot_01515	none	no hit	
			prot_01516		periplasming binding protein: pfam01497 TroA like superfamily (FhuD like, FeuA, FatB, BtuF, HemV-2, TroA-a, TroA like etc) FepB like, FecB	
			prot_01517		FecCD: pfam01032 FepG _ TM_ABC_iron siderophores like superfamily (FecD, FepD, BtuC, FecD)	
			prot_01518		FecCD transport family: pfam01032 TM_ABC_iron siderophore like super family, FepD, FepG , fecD fecC BtuC etc PFAM01032: FecCD ferric siderophore transpor system permease	
			prot_01519	none	ABC transporter: pfam00005 ABC-Iron sderophores B12_Hemin (FepC family)	
			prot_01520	TonB receptor		
			prot_01521		saccharopine dehydrogenase: pfam03435	
			prot_01522	none	NAD-binding 10: pfam13460	
			prot_01523		ACPS: 4'-phosphopantetheinyl transferase pfam01648	
FF-85	1	anthrachelin	prot_02226	AsbA		
			prot_02227	AsbB		
			prot_02228	long chain fatty acid coA ligase		
			prot_02229	acyl carrier protein		
			prot_02230	hypothetical GBAA1986		
ZF-55	none					
FF-50	none					
1A06	none					
1C10	none					
ZF-29	none					
1S-45	1	aerobactin	prot_02472	Leucine responsive regulatory protein (lrp regulon)		
			prot_02473	Xaa-Pro aminopeptidase		
			prot_02474	Aerobactin siderophore receptor IutA		
			prot_02475	Ferric aerobactin ABC transporter, permease component		

			prot_02476	none BLASTp against NR: ferric aerobactin ABC transporter periplasmic substrate binding protein (V. vulnificus MO6-24/O), 92% cov., 51% id.	-periplasmic binding protein (pfam01497)	
			prot_02477	Ferric aerobactin ABC transporter, ATPase component		
			prot_02478	IucC - citrate: 6-N-acetyl-6-N-hydroxy-L-lysine ligase, aerobactin biosynthesis protein IucA, siderophore synthetase superfamily group C		
			prot_02479	- none - BLASTp vs. NR:	no match	
			prot_02480	IucB- N6-hydroxylysine O-acetyl transferase aerobactin biosynthesis protein		
			prot_02481	IucD- L-lysine 6-monooxygenase aerobactin biosynthesis protein		
			prot_02482	membrane protein, suppressor for copper sensitivity ScsB		
			prot_02483	membrane protein, suppressor for copper sensitiity ScsD		
5S-186	2	unknown siderophore?	prot_02011	FIG00921876-hypothetical	- N-formylglutamate amidohydrolase (pfam05013) - Domain of unknown function (DUF1704) (pfam08014)	
			prot_02012	none	- ABC transporter (pfam00005)	
			prot_02013	none	-none	
			prot_02014	ABC-type Fe(III) siderophore transport system permease 2 component	- FecCD transport family (pfam01032)	
			prot_02015	none	- FecCD transport family (pfam01032)	
			prot_02016	none	- Periplasmic binding protein (pfam01497)	
			prot_02017	putative OMR family iron-sederophore receptor precursor	- TonB-dependent receptor plug domain (pfam07715) -TonB dependent receptor (pfam00593)	
			prot_02018	siderophore synthetase superfamily group C (blastp: alcaligin biosynthesis complex protein, 94% cov. 45% id)		
			prot_02019	siderophore biosynthesis protein monooxygenase		
			prot_02020	siderophore biosynthesis L-2,4-diaminobutyrate decarboxylase		
			prot_02021	FIG01202209: hypothetical	-none	
			prot_02022	none	-none	
			prot_02023	FIG01202373: hypothetical	- domain of unknown function DUF3087	

					(pfam11286)	
			prot_02024	none	- fusaric acid resistance family like protein (pfam12805) -fusaric acid resistance family protein (pfam13515)	
			prot_02025	FIG01200908: hypothetical	- DUF318- predicted permease (pfam03773)	
		aerobactin	prot_03982	ATP-dependent RNA helicase DbpA		
			prot_03983	none	- GGDEF domain (pfam00990), oxygen sensing.	
			prot_03984	hypothetical portein in aerobactin uptake cluster		
			prot_03985	putative membrane protein		
			prot_03986	ferric aerobact ABC transporter, ATPase component		
			prot_03987	ferric aerobactin ABC transporter, periplasmic substrate binding protein		
			prot_03988	ferric aerobactin ABC transporter, permease component		
			prot_03989	IucA: citrate 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group A		
			prot_03990	IucB: N6-hydroxylysine O-acetyl transferase aerobactin biosynthesis protein		
			prot_03991	IucC: citrate 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group C		
			prot_03992	IucD: L-lysine 6-monooxygenase		
			prot_03993	aerobactin siderophore receptor IutA- TonB-dependent siderophore receptor		
			prot_03994	putative stomatin/prohibitin family membrane protease subunit YbbK		
			prot_03995	putative activity regulator of membrane protein YbbK		
ZF-211	1	aerobactin	prot_02614	ATP-dependent RNA helicase DbpA		
			prot_02615	none	- GGDEF domain (pfam00990), oxygen sensing.	
			prot_02616	hypothetical portein in aerobactin uptake cluster		
			prot_02617	putative membrane protein		
			prot_02618	ferric aerobact ABC transporter, ATPase component		
			prot_02619	ferric aerobactin ABC transporter, periplasmic substrate binding protein		

			prot_02620	ferric aerobactin ABC transporter, permease component		
			prot_02621	IucA: citrate 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group A		
			prot_02622	IucB: N6-hydroxylysine O-acetyl transferase aerobactin biosynthesis protein		
			prot_02623	IucC: citrate 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group C		
			prot_02624	IucD: L-lysine 6-monooxygenase		
			prot_02625	aerobactin siderophore receptor IutA- TonB-dependent siderophore receptor		
			prot_02626	putative stomatin/prohibitin family membrane protease subunit YbbK		
			prot_02627	putative activity regulator of membrane protein YbbK		
FF-238	none					
FF-167	2	cluster 1: enterobactin (nrps)	prot_03560	2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I alpha (involved in chorismate synthesis)		
			prot_03561	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (leads to 2,3 DHB - precursor) (entA)		
			prot_03562	isochorismate synthase (chorismate to isochorismate) (*entC*)		
			prot_03563	2,3-dihydroxybenzoate AMP ligase (entE)	AMP-binding (pfam 00501)	dhb
			prot_03564	isochorismatase (isochorismate to 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase or salicylate) (entB)		
			prot_03565	enterobactin exporter EntS	MFS_3: transmembrane secretion effector (pfam05977)	
			prot_03566	none	no hit	
			prot_03567	enterobactin synthetase component F (*entF*)	condensation (pfam00668) amp-binding (pfam 00501) pp-binding (pfam00550) condensation (pfam00668) condensation (pfam00668) amp-binding (pfam 00501) pp-binding (pfam00550) thioesterase (pfam00975)	domain 1: ornithine/arginine/lysine domain 2: serine
			prot_03568			
			prot_03569	enterobactin esterase (*fes*) (hydrolyzes ferric-enterobactin complex to relase Fe(III)).	putative esterase (pfam00756)	

	prot_03570	TonB dependent receptor (*fepA*)	TonB dependent receptor (pfam 00593)	
	prot_03571	none	ACPS- 4'-phosphopantetheinyl transferase (entD)	
cluster 2: yersiniabactin (nrps)	prot_01860	TonB dependent receptor		
	prot_01861	AmpG permease		
	prot_01862	iron acquisition yersiniabactin synthesis enzyme Irp2	pp-binding (pfam00550) condensation (pfam00668) nrps (pfam08415) amp-binding (pfam00501) methyltransferase_12 (pfam08242) DUF2009: unknown function (pfam13193) pp-binding (pfam00550) condensation (pfam00668) nrps (pfam08415) pp-binding (pfam00550)	cysteine
	prot_01863	iron acquisition yersiniabactin synthesis enzyme Irp2	beta-ketoacyl synthase, N terminal domain (pfam00109) beta-ketoacyl synthase, C-terminal domain (pfam02801) acyl transferase domain (pfam00698) KR domain (pfam08659) pp-binding (pfam00550) condensation (pfam00668)	N/A
	prot_01864	iron acquisition yersiniabactin synthesis enzyme Irp2	nrps (pfam08415) methyltransferase_12 pp-binding (pfam00550) condensation (pfam00668) nrps (pfam08415) amp-binding (pfam00501) pp-binding (pfam00550) thioesterase (pfam00975)	cysteine
	prot_01865	iron acquisition yersiniabactin synthesis enzyme Irp3	Oxidoreductase, NAD binding Rossmann fold (pfam01408)	
	prot_01866	iron acquisition yersiniabactin synthesis enzyme Irp3 (Ybt like)	thioesterase (pfam00975)	
	prot_01867	2,3-dihydroxybenzoate AMP ligase	chorismate binding (pfam00425)	No A-domain detected?
	prot_01868	2,3-dihydroxybenzoate-AMP ligase	amp-binding (pfam00501)	dhb or salicylate
	prot_01869	high affinity choline uptake protein BetT	BCCT transporter - betaine, carnitine, choline (pfam02028)	
	prot_01870	HTH-type transcriptional regulator BetI		
	prot_01871	betaine aldehyde dehydrogenase		
	prot_01872	choline dehydrogenase		
	prot_01873	Transport ATP binding protein CydC	ABC transporter transmembrane region (pfam00664)	

					ABC transporter (pfam00005)	
			prot_01874	FIG01204006 hypothetical	ABC transmembrane region (pfam0664)	
FF-93	1	enterobactin	prot_02049 (start of contig)	Chitinase		
			prot_02050	none		
			prot_02051	2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I alpha		
			prot_02052	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase		
			prot_02053	isochorismate synthase		
			prot_02054 (end of contig)	2,3-dihydroxybenzoate-AMP ligase	amp-binding (pfam00501)	
			prot_02685	none	ACPS- 4' phosphopantheteinyl transferase (pfam01648) (entD like)	
			prot_02686	TonB-dependent receptor		
			prot_02687	enterobactin esterase		
			prot_02688	none	MbtH-like protein (fragment) (pfam03621)	
			prot_02689	enterobactin synthetase component F	condensation (pfam00668) amp-binding (pfam00501) pp-binding (pfam00550) condensation (pfam00668) condensation (pfam00668) amp-binding (pfam00501) pp-binding (pfam00550) thioesterase (pfam00975)	domain 1: lysine, ornithine, arginine domain 2: serine
			prot_02690	none	none	
			prot_02691	enterobactin exporter EntS	MFS_3: transmembrane secretion effector (pfam05977)	
			prot_02692 (end of contig)	isochorismatase	isochorismatase (pfam00857) pp-binding	
FS-238	2 (but one siderophore)	yersiniabactin, 1?	prot_02035 (beginning of new contig)	Irp2	-pp binding domain: pfam 00550 -NRP: pfam08415 -AMP binding: pfam00501 -methyltransferase pfam08242 -ppbinding : pfam 00550 -NRP: pfam 08415 -pp binding: pfam 00550 entF (condensation)	cysteine
			prot_02036	AmpG permease		
			prot_02037	TonB dependent receptor		
		yersiniabactin, 2?	prot_02390 (start of contig)	none	no significant similarity	

			prot_02391	none	no significant similarity	
			prot_02392	YbtT, resembles thioesterase	GrsT : COG3208 thioesterase: pfam00975	
			prot_02393	Irp3	-oxidoreductase family, NAD binding Rossman fold (pfam 01408)	
			prot_02394 (end of contig)	Irp2	-acyltransferase (pfam00698) -beta-ketoacyl reductase domain (cd 08955) -nrp: (pfam 008415) -methyltransferase (pfam 08242) -pp-binding (pfam00550) -nrps (pfam008415) -amp binding (pfam00501) -pp-binding (pfam00550)	cysteine
FS-144	2	cluster 1: enterobactin (nrps)	prot_01965 (start of contig)	none		
			prot_01966	chitinase		
			prot_01967	2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I alpha		
			prot_01968	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase		
			prot_01969	isochorismate synthase		
			prot_01970	2,3-dihydroxybenozate -AMP ligase	amp-binding	dhb, salicylate
			prot_01971	isochorismatase		
			prot_01972	enterobactin exporter EntS		
			prot_01973	none	none	
			prot_01974	Enterobactin synthetase component F	condensation amp-binding pp-binding condensation condensation amp-binding pp-binding thioesterase	domain 1: ornithine, lysine, arginine domain 2: serine
			prot_01975	none		
			prot_01976	enterobactin esterase		
			prot_01977	TonB dependent receptor		
			prot_01978	none	ACPS: 4'-phosphopantetheinyl transferase	
		cluster 2: yersiniabactin (nrps)	prot_01608	transposase		
			prot_01609	none		
			prot_01610	none		
			prot_01611	FIG01200201: hypothetical		

			prot_01612	4-hydroxybenzoate transporter		
			prot_01613	pantothenate Na+ symporter		
			prot_01614	L-serine dehydratase		
			prot_01615	FIG01204006: hypothetical	ABC transporter transmembrane region ABC transporter	
			prot_01616	Transport ATP-binding protein CydC	ABC transporter transmembrane region ABC transporter	
			prot_01617	2,3-dihydroxybenzoate-AMP ligase	amp-binding	dhb, salicylate
			prot_01618	2,3-dihydroxybenzoate AMP ligase (Irp5)	chorismate binding	no A domain
			prot_01619	YbT- resembles thioesterase	thioesterase	
			prot_01620	Yersiniabactin synthesis enzyme Irp3	oxidoreductase family, NAD binding Rossmann fold (pfam01408)	
			prot_01621	Yersiniabactin synthesis enzyme Irp2	beta-ketoacyl synthase, N-terminal beta-ketoacyl synthase, C-terminal domain Acyl transferase domain KR domain pp-binding condensation nrps methyltransferase_12 pp-binding condensation nrps amp-binding pp-binding thioesterase	cysteine
			prot_01622	Yersiniabactin synthesis enzyme Irp2	pp-binding condensation nrps amp-binding methyltransferase 12 DUF4009 pp-binding condensation nrps pp-binding	cysteine
			prot_01623	AmpG permease		
			prot_01624	TonB dependent receptor		
			prot_01625	none		
			prot_01626		HTH_18: helix-turn-helix transcriptional regulator	
12B09	2	cluster 1: enterobactin (nrps)	prot_00994	electron transport RnfB		
			prot_00995	RnfA		
			prot_00996	none	ACPS- 4'-phosphopantetheinyl transferase family	

	prot_00997	TonB dependent receptor		
	prot_00998	enterobactin esterase		
	prot_00999	none		
	prot_01000	enterobactin synthetase component F	condensation amp-binding pp-binding condensation condensation amp-binding pp-binding thioesterase	- domain 1: ornithine/ lysine/arginine - domain 2: serine
	prot_01001	none	none	
	prot_01002	enterobactin exporter EntS		
	prot_01003	isochorismatase		
	prot_01004	2,3-dihydroxybenzoate AMP ligase	amp-binding enzyme	dhb, salicylate
	prot_01005	isochorismate synthase		
	prot_01006	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase		
	prot_01007	2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I alpha		
	prot_01008	chitinase		
cluster 2: yersiniabactin	prot_03210	none		
	prot_03211	TonB-dependent receptor		
	prot_03212	AmpG permease		
	prot_03213	Irp2 (comment: hmwp2 in yersiniabactin synthesis)	pp-binding condensation nrps amp-binding methyltransferase 12 DUF4009 pp-binding condensation	cysteine
	prot_03214	Irp2	nrps pp-binding	
	prot_03215	Irp2 (comment: hmwp1 in yersiniabactin synthesis)	beta-ketoacyl synthase, N terminal domain beta-ketoacyl synthase, C-terminal domain acyl transferase domain KR domain pp-binding condensation nrps methyltransferase 12 pp-binding condensation nrps	cysteine

					amp-binding pp-binding thioesterase	
			prot_03216	Irp3	Oxidoreductase family, NAD binding Rossmann fold	
			prot_03217	YbT, resembles thioesterase		
			prot_03218	2,3-dihydroxybenzoate AMP ligase (Irp5)	chorismate binding enzyme	No A-domain detected
			prot_03219	2,3-dihydroxybenzoate AMP-ligase	amp-binding DUF4009	dhb, salicylate
			prot_03220	Transport ATP binding protein CydC		
			prot_03221	FIG01204006	ABC transmembrane region ABC transporter	
			prot_03222	L-serine dehydratase		
			prot_03223	pantothenate Na+ symporter		
			prot_03224	4-hydroxybenzoate transporter		
			prot_03225	FIG01200201 hypothetical		
			prot_03226	none		
			prot_03227	none		
			prot_03228	transposase		
			prot_03229	transposase		
9ZC157	1 cluster	aerobactin	prot_02879 (start of contig)	IutA- Aerobactin sideorphore receptor-TonB-dependent receptor		
			prot_02880	IucD: L-lysine 6-monooxygenase aerobactin synthesis protein		
			prot_02881	IucC: citrate- 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group C		
			prot_02882	IucB: N6-hydroxylysine O-acetyltransferase		
			prot_02883	IucA: citrate- 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group A		
			prot_02884	ferric aerobactin ABC transporter permease component		
			prot_02885	ferric aerobactin ABC transporter periplasmic substrate binding protein		
			prot_02886	ferric aerobactin ABC transporter ATPase component		
			prot_02887	none	-none	
			prot_02888	putative membrane portein		
			prot_02889	hypothetical protein in aerobactin uptake cluster		
			prot_02890	glutathione S-transferase		

			prot_02891	none	- bacterial regulatory helix turn helix protein, LysR family (pfam00126) - LysR substrate (pfam03466)	
			prot_02892	methylenetetrahydrofolate dehydrogenase		
			prot_02893	membrane-bound lytic murein transglycosylate B		
ZF-129	1 cluster	aerobactin	prot_03010	membrane-bound lytic murein transglycosylate B		
			prot_03011	methylenetetrahydrofolate dehydrogenase		
			prot_03012	none		
			prot_03013	glutathione S-transferase		
			prot_03014	hypothetical protein in aerobactin uptake cluster		
			prot_03015	putative membrane protein		
			prot_03016	none		
			prot_03017	ferric aerobactin ABC transporter ATPase component		
			prot_03018	ferric aerobactin ABC transporter periplasmic substrate binding protein		
			prot_03019	ferric aerobactin ABC transporter permease component		
			prot_03020	IucA: citrate- 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group A		
			prot_03021	IucB: N6-hydroxylysine O-acetyltransferase		
			prot_03022	IucC: citrate- 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group C		
			prot_03023	IucD: L-lysine 6-monooxygenase aerobactin synthesis protein		
			prot_03024 (end of contig)	IutA- Aerobactin sideorphore receptor-TonB-dependent receptor		
9CSC122	1 cluster	aerobactin	prot_02345 (start of contig)	IutA- Aerobactin sideorphore receptor-TonB-dependent receptor		
			prot_02346	IucD: L-lysine 6-monooxygenase aerobactin synthesis protein		
			prot_02347	IucC: citrate- 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group C		
			prot_02348	IucB: N6-hydroxylysine O-acetyltransferase		
			prot_02349	IucA: citrate- 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group A		

			prot_02350	ferric aerobactin ABC transporter permease component		
			prot_02351	ferric aerobactin ABC transporter periplasmic substrate binding protein		
			prot_02352	ferric aerobactin ABC transporter ATPase component		
			prot_02353	none		
			prot_02354	putative membrane portein		
			prot_02355	hypothetical protein in aerobactin uptake cluster		
			prot_02356	glutathione S-transferase		
			prot_02357	none		
			prot_02358	methylenetetrahydrofolate dehydrogenase		
			prot_02359 (end of contig)	membrane-bound lytic murein transglycosylate B		
9ZD137	1 cluster	aerobactin	prot_02903 (start of contig)	IutA- Aerobactin sideorphore receptor-TonB-dependent receptor		
			prot_02904	IucD: L-lysine 6-monooxygenase aerobactin synthesis protein		
			prot_02905	IucC: citrate- 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group C		
			prot_02906	IucB: N6-hydroxylysine O-acetyltransferase		
			prot_02907	IucA: citrate- 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group A		
			prot_02908	ferric aerobactin ABC transporter permease component		
			prot_02909	ferric aerobactin ABC transporter periplasmic substrate binding protein		
			prot_02910	ferric aerobactin ABC transporter ATPase component		
			prot_02911	none	none	
			prot_02912	none	none	
			prot_02913 (end of contig)	none	none	
9ZB36	1 cluster	aerobactin	prot_02594	none	- autoinducer binding domain (pfam03472) -bacterial regulatory proteins luxR (pfam00196)	
			prot_02595	none	- predicted membrane protein (DUF2306) (pfam10067)	
			prot_02596	none	none	
			prot_02597	IutA- Aerobactin sideorphore receptor-		

			prot_02350	ferric aerobactin ABC transporter permease component		
			prot_02351	ferric aerobactin ABC transporter periplasmic substrate binding protein		
			prot_02352	ferric aerobactin ABC transporter ATPase component		
			prot_02353	none		
			prot_02354	putative membrane portein		
			prot_02355	hypothetical protein in aerobactin uptake cluster		
			prot_02356	glutathione S-transferase		
			prot_02357	none		
			prot_02358	methylenetetrahydrofolate dehydrogenase		
			prot_02359 (end of contig)	membrane-bound lytic murein transglycosylate B		
9ZD137	1 cluster	aerobactin	prot_02903 (start of contig)	IutA- Aerobactin sideorphore receptor-TonB-dependent receptor		
			prot_02904	IucD: L-lysine 6-monooxygenase aerobactin synthesis protein		
			prot_02905	IucC: citrate- 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group C		
			prot_02906	IucB: N6-hydroxylysine O-acetyltransferase		
			prot_02907	IucA: citrate- 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group A		
			prot_02908	ferric aerobactin ABC transporter permease component		
			prot_02909	ferric aerobactin ABC transporter periplasmic substrate binding protein		
			prot_02910	ferric aerobactin ABC transporter ATPase component		
			prot_02911	none	none	
			prot_02912	none	none	
			prot_02913 (end of contig)	none	none	
9ZB36	1 cluster	aerobactin	prot_02594	none	- autoinducer binding domain (pfam03472) -bacterial regulatory proteins luxR (pfam00196)	
			prot_02595	none	- predicted membrane protein (DUF2306) (pfam10067)	
			prot_02596	none	none	
			prot_02597	IutA- Aerobactin sideorphore receptor-		

				TonB-dependent receptor		
			prot_02598	IucD: L-lysine 6-monooxygenase aerobactin synthesis protein		
			prot_02599	IucC: citrate- 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group C		
			prot_02600	IucB: N6-hydroxylysine O-acetyltransferase		
			prot_02601	IucA: citrate- 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group A		
			prot_02602	ferric aerobactin ABC transporter permease component		
			prot_02603	ferric aerobactin ABC transporter periplasmic substrate binding protein		
			prot_02604	ferric aerobactin ABC transporter ATPase component		
			prot_02605	none		
			prot_02606	putative membrane protein		
			prot_02607	hypothetical protein in aerobactin uptake cluster		
			prot_02608	glutathione S-transferase		
			prot_02609			
			prot_02610	methylenetetrahydrofolate dehydrogenase		
			prot_02611	membrane bound lytic murein transglycosylase		
12G01	1 cluster	Vibrioferrin	prot_00102	AttH component of AttEFGH ABC transport		
			prot_00103	AttF component of AttEFGH ABC transport		
			prot_00104	AttE component of AttEFGH ABC transport		
			prot_00105	none	none	
			prot_00106	FIG01203004- hypothetical	-protein of unknown function (DUF3297) (pfam11730)	
			prot_00107	putative protease		
			prot_00108	none	- domain of unknown function (DUF302) (pfam03625)	
			prot_00109	PvsE- vibrioferrin decarboxylase		
			prot_00110	PvsD- vibrioferrin amide bond forming protein, siderophore synthetase superfamily group A		
			prot_00111	PvsC- vibrioferrin membrane spanning transport protein		
			prot_00112	PvsB- vibrioferrin amide bond forming protein, siderophore synthetase superfamily		

				group B		
			prot_00113	PvsA- vibrioferrin ligase/carboxylase		
			prot_00114	PsuA- ferric siderophore receptor	-TonB-dependent receptor plug domain (pfam07715) -TonB dependent receptor (pfam00593)	
			prot_00115	PvuA- vibrioferrin receptor	-TonB-dependent receptor plug domain (pfam07715) -TonB dependent receptor (pfam00593)	
			prot_00116	PvuB- ferrichrome ABC transporter-periplasmic iron binding protein	- periplasmic binding protein (pfam01497)	
			prot_00117	none	none	
			prot_00118	PvuC- ferrichrome ABC transporter-permease	- FecCD family - (pfam01032)	
			prot_00119	PvuD-ferrichrome ABC transporter - permease	- FecCD family - (pfam01032)	
			prot_00120	PvuE- ferrichrome ABC transporter - ATP binding subunit	- ABC transporter (pfam00005)	
			prot_00121	N-acetylglucosamine regulated methyl-accepting chemotaxis protein		
5S-149	1	yersiniabactin or anguibactin?	prot_00134 (beginning of contig 2)	annotation by RAST Ferrichrome-iron receptor		
			prot_00135	Iron compound ABC uptake transporter substrate-binding protein		
			prot_00136	ferric siderophore ABC transporter, permease protein		
			prot_00137	Iron transport protein *(fatD like)*		
			prot_00138	Peptide synthetase	-pp-binding (pfam00550) -condensation (pfam00668)	
			prot_00139	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EC 1.3.1.28) - product is DHBA		
			prot_00140			
			prot_00141			
			prot_00142	none	none	
			prot_00143	none	none	
			prot_00144		none	
			prot_00145	none	phosphoribosyl transferase domain (pfam00156)	
			prot_00146	none	none	
			prot_00147	none	none	
			prot_00148	none	none	
			prot_00149	none	hth24 wing helix turn helix DNA binding (pfam13412)	

			prot_00150	none	histidine kinase, hsp90 like ATPase (pfam02518)	
			prot_00151	none	none	
			prot_00152	none	none	
			prot_00153	none	resolvase, N-term. domain (pfam00239) hth_7, helix-turn-helix domain (pfam02796)	
			prot_00154	none	none	
			prot_00155	none	resolvase (pfam00239)	
			prot_00156	none	AAA_25 (pfam13481)	
			prot_00157	none	none	
			prot_00158	none	none	
			prot_00159	FIG01206744	none	
			prot_00160	FIG01206255	none	
			prot_00161	none	DUF457- predicted membrane bound metal dependent hydrolase (pfam04307)	
			prot_00162	none	ACPS: 4'-phosphopantetheinyl transferase (pfam01648)	
			prot_00163	isochorismatase	isochorismatase (pfam00857) pp-binding (pfam00550)	
			prot_00164	2,3-dihydroxybenzoate-AMP-ligase	AMP-binding (pfam00501)	dhb or salicylate
			prot_00165	isochorismate synthase	chorismate-binding (pfam00425)	
			prot_00166	2-keto-3-deoxy-D-arabino-heptulosonate-7phosphate synthase I alpha	DAHP synthetase I family (EC 2.5.1.54) (pfam00793)	
			prot_00167	hypothetical	ABC transporter (pfam00005)	
			prot_00168	CydC (transport)	-ABC transporter transmembrane region (pfam 00664) -ABC transporter (pfam 00005)	
			prot_00169	histidine decarboxylase	-Pyridoxal-dependent decarboxylase conserved domain (pfam00282)	
			prot_00170	nonribosomal peptide synthetase (pcp and c domains- angM? --> more similar to nrp in V. anguillarum 775 than in Yersinia)	-condensation (pfam00668) -nrps (pfam08415) -condensation (pfam00668) -nrps (pfam08415)	
			prot_00171	siderophore protein monooxygenase	-L-lysine 6 monooxygenase (pfam13434)	
			prot_00172	Ybt, resembles thioesterase	-Thioesterase (pfam00975)	
			prot_00173 (end of contig 2)	Irp2	-AMP-binding enzyme (pfam00501) -PP-binding (pfam00550)	cysteine
12E03	1 cluster	vibrioferrin	prot_01905	permease of the drug/metabolite transporter (DMT)		
			prot_01906	none	- glyoxalase domain (pfam12681)	
			prot_01907	LysR-family tanscriptional reculator		

			prot_01908	FIG00949965- hypothetical	- carboxylesterase family (pfam00135)	
			prot_01909	PvuE- ferrichrome ABC transporter, ATP binding subunit		
			prot_01910	PvuD- ferrichrome ABC transporter, permease		
			prot_01911	PvuC- ferrichrome ABC transporter, permease		
			prot_01912	PvuB- ferrichrome ABC transporter, periplasmic iron binding protein		
			prot_01913	PvuA- ferrichrome ABC transporter		
			prot_01914	PsuA- siderophore receptor		
			prot_01915	PvsA- vibrioferrin ligase/carboxylase		
			prot_01916	none BLASTp: PvsB (V. splendidus LGP32)- 100% cov. 96% id.		
			prot_01917	PvsC- vibrioferrin membrane spanning transport protein		
			prot_01918	PvsD- vibrioferrin amide bond forming protein, siderophore synthetase superfamily group A		
			prot_01919	PvsE- vibrioferrin decarboxylase protein		
			prot_01920	FIG01203944- hypothetical	- protein of unknown function DUF3081 (pfam11280)	
			prot_01921	FIG01200795- hypothetical	- complex I intermediate associated protein (pfam08547)	
ZF-57	none					
9ZC77	none					
9ZC13	1 cluster	vibrioferrin	prot_00530	FIG012064343-hypothetical protein	- complexI intermediate associated protein (pfam08547)	
			prot_00531	FIG01203944- hypothetical protein	- protein of unknown function DUF3081 (pfam11280)	
			prot_00532	PvsE- vibrioferrin decarboxylase		
			prot_00533	PvsD- vibrioferrin amide bond forming protein, siderophore synthetase superfamily group A		
			prot_00534	PvsC- vibrioferrin membrane spanning transport protein		
			prot_00535	none BLASTp: PvsB in Vibrio splendidus LGP32 (cov. 100%, id. 97%)		
			prot_00536	PvsA- vibrioferrin ligase/carboxylase		
			prot_00537	PsuA- ferric siderophore receptor		
			prot_00538	PvuA- vibrioferrin receptor		

			prot_00539	PvuB- ferrichrome ABC transporter-periplasmic iron binding protein		
			prot_00540	PvuC- ferrichrome ABC transporter-permease		
			prot_00541	PvuD-ferrichrome ABC transporter - permease		
			prot_00542	PvuE- ferrichrome ABC transporter - ATP binding subunit		
			prot_00543	none	-glyoxalase domain (pfam12681)	
			prot_00544	FIG01202053	none	
			prot_00545	permease of drug/metabolite transporter DMT		
9CSC106	none					
ZF-91	none					
9ZC88	none					
ZS-139	1 cluster	pyoverdin-like	prot_03549	glycerophosphoryl diester phosphodiesteras		
			prot_03550	glycerol-3-phosphate transporter		
			prot_03551	none		
			prot_03552	FIG01201856: hypothetical protein		
			prot_03553	FIG01204141: hypothetical protein		
			prot_03554	FIG01205527: hypothetical protein		
			prot_03555	L-ornithine 5-monooxygenase, PvdA of pyoverdin synthesis		
			prot_03556	hypothetical protein PvdY		
			prot_03557	ferric hydroxamate outer membrane receptor FhuA		
			prot_03558	long-chain fatty acid CoA ligase		
			prot_03559	Long-chain fatty acid CoA ligase		
			prot_03560	Hypothetical MbtH like protein PA2412 homolog		
			prot_03561	none		
			prot_03562	ferric hydroxamate ABC transporter permease component FhuB		
			prot_03563	ferric hydroxamate ABC transporter periplasmic substrate binding protein FhuD		
			prot_03564	ferric hydroxamate ABC transporter, ATP binding protein FhuC		
			prot_03565	none		
			prot_03566	FIG01203301- hypothetical		
			prot_03567	FIG01203225: hypothetical		
1F-187	1	Vulnibactin	prot_02639	none	MFS_1: major facilitator superfamily	
			prot_02640	oxidoreductase	alpha-beta hydrolase family	

			prot_02641	amide synthase component of siderophore synthesis	condensation	
			prot_02642	ferric vulnibactin VuuA	TonB dependent receptor	
			prot_02643	none	pp-binding	
			prot_02644	2,3-dihydroxybenzoate AMP ligase	amp-binding DUF4009	dhb, salicylate
			prot_02645	isochorismate pyruvate lyase	chorismate mutase type II (pfam01817)	
			prot_02646	isochorismatase		
			prot_02647	Vulnibactin utlization protein VuuB	siderophore interacting FAD-binding domain (pfam08021) siderophore interacting protein (pfam04954)	
			prot_02648	2,3-dihydroxybenzoate AMP ligase	amp-binding DUF4009	dhb, salicylate
			prot_02649	isochorismate synthase		
			prot_02650	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase		
			prot_02651	2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I alpha	DAHP synthetase I family	
			prot_02652 (end of contig)	nrps module	condensation nrps condensation nrps AMP-binding	cysteine
1F-155	none					
1F-267	none					
5F-79	1 cluster	aerobactin	prot_03971	Oligopeptide ABC transporter, periplasmic oligopeptide binding protein OppA		
			prot_03972	FIG01200669- hypothetical		
			prot_03973	Oligopeptide ABC transporter periplasmic oligopeptide binding protein OppA		
			prot_03974	hypothetical protein in aerobactin uptake cluster		
			prot_03975	putative membrane protein		
			prot_03976	ferric aerobactin ABC transporter ATPase component		
			prot_03977	ferric aerobactin ABC transporter periplasmic substrate binding protein		
			prot_03978	ferric aerobactin ABC transporter permease component		
			prot_03979	IucA: citrate- 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group A		
			prot_03980	IucB: N6-hydroxylysine O-acetyltransferase		

			prot_03981	IucC: citrate- 6-N-acetyl-6-N-hydroxy-L-lysine ligase. siderophore synthetase superfamily group C		
			prot_03982	IucD: L-lysine 6-monooxygenase aerobactin synthesis protein		
			prot_03983	IutA- Aerobactin siderophore receptor-TonB-dependent receptor		
			prot_03984 (end of contig)	none		
ZS-17	1	pyoverdin like	prot_03431	L-ornithine 5-monooxygenase, PvdA of pyoverdin biosynthesis	- L-lysine 6-monooxygenase (pfam13434)	
			prot_03432	Hypothetical protein PvdY	-acetyltransferase domain (pfam13523)	
			prot_03433	ferric hydroxamate outer membrane receptor FhuA	- TonB dependent receptor plug domain (pfam07715) - TonB dependent receptor (pfam00593)	
			prot_03434	Long chain fatty acid coA ligase	- condensation (pfam00668) -HxxPF rpt (pfam13745) - amp-binding (pfam00501) - pp-binding (pfam00550) -condensation (pfam00668) -thioesterase (pfam00975)	Leucine
			prot_03435	Long-chain fatty acid coA ligase (pyoverdin has 4 nrps)	- condensation (pfam00668) - amp-binding (pfam00501) - pp-binding (pfam00550) - condensation (pfam00668) - condensation (pfam00668) - amp-binding (pfam00501) - pp-binding (pfam00550) - condensation (pfam00668) - HxxPF rpt (pfam13745) - amp-binding (pfam00501) - pp-binding (pfam00550)	domain 1: cysteine domain 2: leucine domain 3: serine
			prot_03436	hypothetical MbtH like protein	- MbtH-like (pfam03621)	
			prot_03437	none	- ABC transporter (pfam00005)	
			prot_03438	ferric hydroxamate ABC transporter, permease component FhuB	- FecCD transport family (pfam01032) - FecCD transport family (pfam01032)	
			prot_03439	ferric hydroxamate ABC transporter periplasmic substrate binding protein FhuD	- periplasmic binding protein (pfam01497)	
			prot_03440	ferric hydroxamate ABC transporter ATP binding protein FhuC	- ABC transporter (pfam00005)	
			prot_03441	none	- ACPS 4'-phosphopantetheinyl transferase family (pfam01648)	
			prot_03442	FIG01203301	- acetyltransferase (GNAT) family (pfam00583)	

			prot_03443	FIG01202600	- none	
			prot_03444	FIG01203225	- none	
FF-500	1	vibrioferrin	prot_01753	FIG01206343		
			prot_01754	FIG01203944		
			prot_01755	PvsE- vibrioferrin decarboxylase		
			prot_01756	PvsD- vibrioferrin amide bond forming protein, siderophore synthetase superfamily group A		
			prot_01757	PvsC- vibrioferrin membrane spanning transport protein		
			prot_01758	PvsB- Vibriferrin amid bon foring protein, siderophore synthetase superfamily group B		
			prot_01759	PvsA- vibrioferrin ligase/carboxylase		
			prot_01760	PsuA- ferric siderophore receptor		
			prot_01761	PvuA- vibrioferrin receptor		
			prot_01762	none		
			prot_01763	PvuB- ferrichrome ABC transporter-periplasmic iron binding protein		
			prot_01764	PvuC- ferrichrome ABC transporter-permease		
			prot_01765	PvuD-ferrichrome ABC transporter - permease		
			prot_01766	PvuE- ferrichrome ABC transporter - ATP binding subunit		
			prot_01767	none		
ZS-63	1	vibrioferrin	prot_04151	FIG01206343 hypothetical protein		
			prot_04152	FIG01203944 hypothetical protein		
			prot_04153	PvsE- vibrioferrin decarboxylase		
			prot_04154	PvsD- vibrioferrin amide bond forming protein, siderophore synthetase superfamily group A		
			prot_04155	PvsC- vibrioferrin membrane spanning transport protein		
			prot_04156	PvsB- Vibriferrin amid bon foring protein, siderophore synthetase superfamily group B		
			prot_04157	PvsA- vibrioferrin ligase/carboxylase		
			prot_04158	PsuA- ferric siderophore receptor		
			prot_04159	PvuA- vibrioferrin receptor		
			prot_04160	PvuB- ferrichrome ABC transporter-periplasmic iron binding protein		
			prot_04161	PvuC- ferrichrome ABC transporter-permease		
			prot_04162	PvuD-ferrichrome ABC transporter - permease		

			prot_04163	PvuE- ferrichrome ABC transporter - ATP binding subunit		
			prot_04164	none		
ZF-90	1	pyoverdin like	prot_02252	glycerophosphoryl diester phosphodiesteras	- glycerophoryl diester phosphodiesterase (pfam03009)	
			prot_02253	glycerol-3-phosphate transporter		
			prot_02254	FIG01205527: hypothetical protein	- DUF3297: protein of unknown function (pfam11730)	
			prot_02255	L-ornithine 5-monooxygenase, PvdA of pyoverdin synthesis	- L-lysine 6-monooxygenase (pfam13434)	
			prot_02256	hypothetical protein PvdY	- acetyl transferase 8 (GNAT domain) (pfam13523)	
			prot_02257	ferric hydroxamate outer membrane receptor FhuA	- TonB dependent receptor plug domain (pfam07715) - TonB dependent receptor (pfam00593)	
			prot_02258	long-chain fatty acid CoA ligase	- condensation (pfam00668) -HxxPF rpt (pfam13745) - amp-binding (pfam00501) - pp-binding (pfam00550) -condensation (pfam00668) -thioesterase (pfam00975)	leucine
			prot_02259	Long-chain fatty acid CoA ligase	- condensation (pfam00668) - amp-binding (pfam00501) - pp-binding (pfam00550) - condensation (pfam00668) - condensation (pfam00668) - amp-binding (pfam00501) - pp-binding (pfam00550) - condensation (pfam00668) - HxxPF rpt (pfam13745) - amp-binding (pfam00501) - pp-binding (pfam00550)	domain 1: cysteine domain 2: leucine domain 3: serine
			prot_02260	Hypothetical MbtH like protein PA2412 homolog		
			prot_02261	none	- ABC transporter (pfam00005)	
			prot_02262	ferric hydroxamate ABC transporter permease component FhuB		
			prot_02263	ferric hydroxamate ABC transporter periplasmic substrte binding protein FhuD		
			prot_02264	ferric hydroxamate ABC transporter, ATP binding protein FhuC		
			prot_02265	none	- ACPS 4'-phosphopantetheinyl (pfam01648)	
			prot_02266	FIG01203301- hypothetical	- acetyltransferase (GNAT) family (pfam00583)	

			prot_02267	FIG01202600 - hypothetical	- none	
			prot_02268	FIG01203225- hypothetical	- none	
5S-101	none	none				
FF-6		vibrioferrin	prot_04630	FIG01206343 hypothetical protein		
			prot_04631	FIG01203944 hypothetical protein		
			prot_04632	PvsE- vibrioferrin decarboxylase		
			prot_04633	PvsD- vibrioferrin amide bond forming protein, siderophore synthetase superfamily group A		
			prot_04634	PvsC- vibrioferrin membrane spanning transport protein		
			prot_04635	PvsB- Vibriferrin amid bon foring protein, siderophore synthetase superfamily group B		
			prot_04636	PvsA- vibrioferrin ligase/carboxylase		
			prot_04637	PsuA- ferric siderophore receptor		
			prot_04638	PvuA- vibrioferrin receptor		
			prot_04639	PvuB- ferrichrome ABC transporter-periplasmic iron binding protein		
			prot_04640	PvuC- ferrichrome ABC transporter-permease		
			prot_04641	PvuD-ferrichrome ABC transporter - permease		
			prot_04642	PvuE- ferrichrome ABC transporter - ATP binding subunit		
			prot_04643	none		
1F-157	1 cluster	vibrioferrin (transporters on a different contig)	prot_03091 (start of contig)	PvsA- vibrioferrin ligase/carboxylase		
			prot_03092	PvsB- vibrioferrin amide bond forming-siderophore synthetase superfamily group B		
			prot_03093	PvsC- vibrioferrin membrane spanning transport protein		
			prot_03094	PvsD- vibrioferrin amide bond forming protein, siderophore synthetase superfamily group A		
			prot_03095	PvsE- vibrioferrin decarboxylase protein		
			prot_03096	FIG0123944- hypothetical		
			prot_03097	FIG01206343- hypothetical		
0407ZC148	1 cluster	vibrioferrin (note: no PvsE)	prot_03312	FIG01206343-hypothetical protein		
			prot_03313	FIG01203944-hypothetical		
			prot_03314	PvsD- vibrioferrin amide bond forming protein, siderophore synthetase superfamily group A		
			prot_03315	PvsC- vibrioferrin membrane spanning		

				transport protein		
			prot_03316	PvsB- Vibriferrin amid bon foring protein, siderophore synthetase superfamily group B		
			prot_03317	PvsA- vibrioferrin ligase/carboxylase		
			prot_03318	PsuA- ferric siderophore receptor		
			prot_03319	PvuA- vibrioferrin receptor		
			prot_03320	PvuB- ferrichrome ABC transporter-periplasmic iron binding protein		
			prot_03321	PvuC- ferrichrome ABC transporter-permease		
			prot_03322	PvuD-ferrichrome ABC transporter - permease		
			prot_03323	PvuE- ferrichrome ABC transporter - ATP binding subunit		
1S-124	1	Vulnibactin	prot_02570	nrps module	-condensation (pfam00668) -nrps (pfam08415) -condensation (pfam00668) -nrps (pfam08415) - amp-binding (pfam00501)	cysteine
			prot_02571	2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I alpha	- DAHP synthase I family (pfam00793)	
			prot_02572	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase	- enoyl acyl carrier protein reductase (pfam13561)	
			prot_02573	isochorismate synthase	- chorismate binding enzyme (pfam00425)	
			prot_02574	2,3-dihydroxybenzoate AMP ligase	- amp-binding (pfam00501) - DUF4009 (pfam13193)	dhb/salicylate
			prot_02575	Vulnibactin utilization VuuB	- Siderophore interacting FAD-binding domain (pfam08021) - siderophore interacting protein (pfam04954)	
			prot_02576	isochorismatase		
			prot_02577	isochorismate lyase		
			prot_02578	2,3-dihydroxybenzoate amp-ligase	- amp-binding (pfam00501) - DUF4009 (pfam13193)	dhb/salicylate
			prot_02579	none	- pp-binding (pfam00550)	
			prot_02580	ferric vulnibactin receptor VuuA	- TonB dependent receptor plug domain (pfam07715) -TonB dependent receptor (pfam00593)	
			prot_02581	amide synthase of siderophore synthetase	-condensation (pfam00668)	
			prot_02582	oxidoreductase	-alpha/beta hydrolase family (pfam12697)	
			prot_02583	none	- MFS_1- major facilitator family (pfam07690)	
			prot_02584	quinolate synthetase	- quinolinate synthetase A protein (pfam02445)	

			prot_02585 (end of contig)	TPR repeat		
13B01	1 cluster	vibrioferrin	prot_01060 (start of contig)	bifunctional protein - zinc containing alcohol dehydrogenase		
			prot_01061	FIG01206343- hypothetical protein		
			prot_01062	FIG01203944- hypothetical protein		
			prot_01063	PvsE- vibrioferrin decarboxylase protein		
			prot_01064	PvsD- vibrioferrin amide bond forming protein, siderophore synthetase superfamily group A		
			prot_01065	PvsC- vibrioferrin membrane spanning transport protein		
			prot_01066	PvsB- Vibriferrin amid bon foring protein, siderophore synthetase superfamily group B		
			prot_01067	PvsA- vibrioferrin ligase/carboxylase		
			prot_01068	PsuA- ferric siderophore receptor		
			prot_01069	PvuA- vibrioferrin receptor		
			prot_01070	PvuB- ferrichrome ABC transporter-periplasmic iron binding protein		
			prot_01071	PvuC- ferrichrome ABC transporter-permease		
			prot_01072	PvuD-ferrichrome ABC transporter - permease		
			prot_01073	PvuE- ferrichrome ABC transporter - ATP binding subunit		
			prot_01074	none		
			prot_01075	FIG01202053- hypothetical		
			prot_01076	permease of the drug/metabolite transporter DMT superfamily		
12F01	1 cluster	vibrioferrin	prot_01798	FIG01202053-hypothetical		
			prot_01799	permease of the drug/metabolite transporter DMT superfamily		
			prot_01800	none		
			prot_01801	PvuE- ferrichrome ABC transporter, ATP binding subunit		
			prot_01802	PvuD- ferrichrome ABC transporter, permease		
			prot_01803	PvuC- ferrichrome ABC transporter, permease		
			prot_01804	PvuB- ferrichrome ABC transporter, periplasmic iron binding protein		
			prot_01805	PvuA- ferrichrome ABC transporter		
			prot_01806	PsuA- siderophore receptor		

			prot_01807	PvsA- vibrioferrin ligase/carboxylase		
			prot_01808	PvsB- vibrioferrin amide bond forming protein- siderophore synthetase superfamily group B		
			prot_01809	PvsC- vibrioferrin membrane spanning transport protein		
			prot_01810	PvsD- vibrioferrin amide bond forming protein, siderophore synthetase superfamily group A		
			prot_01811	PvsE- vibrioferrin decarboxylase protein		
			prot_01812	FIG01203944- hypothetical		
			prot_01813	FIG01206343		
12B01	cluster 1	vibrioferrin	prot_00086	none		
			prot_00087	PvuE- ferrichrome ABC transporter, ATP binding subunit		
			prot_00088	PvuD- ferrichrome ABC transporter, permease		
			prot_00089	PvuC- ferrichrome ABC transporter, permease		
			prot_00090	PvuB- ferrichrome ABC transporter, periplasmic iron binding protein		
			prot_00091	PvuA- ferrichrome ABC transporter		
			prot_00092	PsuA- siderophore receptor		
			prot_00093	PvsA- vibrioferrin ligase/carboxylase		
			prot_00094	PvsB- vibrioferrin amide bond forming protein- siderophore synthetase superfamily group B		
			prot_00095	PvsC- vibrioferrin membrane spanning transport protein		
			prot_00096	PvsD- vibrioferrin amide bond forming protein, siderophore synthetase superfamily group A		
			prot_00097	PvsE- vibrioferrin decarboxylase protein		
			prot_00098	FIG01203944- hypothetical protein		
			prot_00099	FIG01206343- hypothetical protein		
			prot_00100	bifunctional protein- zinc containing alcohol dehydrogenase.		
FF-75	1	pyoverdin-like	prot_00318	FIG01203225	- none	
			prot_00319	FIG01202600	- none	
			prot_00320	FIG01203301	- acetyltransferase 1 (pfam00583)	
			prot_00321	none	- ACPS 4' phosphopantetheinyl transferase (pfam01648)	
			prot_00322	ferric hydroxamte ABC transporter ATP-binding protein FhuC		

			prot_00323	ferric hydroxamte ABC transporter periplasmic substrate binding protein FhuD		
			prot_00324	ferric hydroxamte ABC transporter, permease component FhuB		
			prot_00325	none	- ABC transporter (pfam00005)	
			prot_00326	hypothetical MbtH-like		
			prot_00327	long-chain fatty acid CoA ligase	- condensation (pfam00668) - amp-binding (pfam00501) - pp-binding (pfam00550) - condensation (pfam00668) - condensation (pfam00668) - amp-binding (pfam00501) - pp-binding (pfam00550) - condensation (pfam00668) - HxxPF rpt (pfam13745) - amp-binding (pfam00501) - pp-binding (pfam00550)	domain 1: cysteine domain 2: serine domain 3: leucine
			prot_00328	long-chain fatty acid CoA ligase	- condensation (pfam00668) -HxxPF rpt (pfam13745) - amp-binding (pfam00501) - pp-binding (pfam00550) -condensation (pfam00668) -thioesterase (pfam00975)	leucine
			prot_00329	ferric hydroxamate outer membrane receptor FhuA		
			prot_00330	hypothetical protein PvdY	- acetyltransferase 8 (GNAT) (pfam13523)	
			prot_00331	L-ornithine 5-monooxygenase PvdA of pyoverdin biosynthesis	- L-lysine 6-monooxygenase (pfam13434)	
			prot_00332	FIG01205527	- DUF3297 (pfam11730)	
			prot_00333	FIG-1204141	- none	
			prot_00334	FIG01201856	- none	
			prot_00335	glycerol-3-phosphate transporter		
			prot_00336	glycerophosphoryl diester phosphodiesterase		
1F-111	1	pyoverdin-like	prot_02261	FIG01203225- hypothetical		
			prot_02262	FIG01202600- hypothetical		
			prot_02263	FIG01203301- hypothetical		
			prot_02264	none	ACPS- 4'phosphopantetheinyl transferase (pfam01648)	
			prot_02265	ferric hydroxamate ABC transporter, ATP binding protein FhuC		
			prot_02266	ferric hydroxamate ABC transporter, periplasmic substrate binding protein FhuD		
			prot_02267	ferric hydroxamate ABC transporter, permease component FhuB		

		prot_02268	none	-ABC transporter (pfam00005)	
		prot_02269	hypothetical MbtH like		
		prot_02270	long-chain fatty acid CoA ligase	- condensation (pfam00668) - amp-binding (pfam00501) - pp-binding (pfam00550) - condensation (pfam00668) - condensation (pfam00668) - amp-binding (pfam00501) - pp-binding (pfam00550) - condensation (pfam00668) - HxxPF rpt (pfam13745) - amp-binding (pfam00501) - pp-binding (pfam00550)	
		prot_02271	long-chain fatty acid CoA ligase	- condensation (pfam00668) -HxxPF rpt (pfam13745) - amp-binding (pfam00501) - pp-binding (pfam00550) -condensation (pfam00668) -thioesterase (pfam00975)	
		prot_02272	ferric hydroxamate outer membrane receptor FhuA		
		prot_02273	hypothetical protein PvdY		
		prot_02274	L-ornithine 5-monooxygenase, PvdA of pyoverdin biosynthesis		
		prot_02275	FIG01205527: hypothetical		
		prot_02276	FIG01204141: hypothetical		
		prot_02277	FIG01201856: hypothetical		
		prot_02278	glycerol-3-phosphate transporter		
		prot_02279	glycerophosphoryl diester phosphodiesterase		

2.9.3. Table S2.3: List of Iron Transport Genes Used to look for Orthologs in Sequenced Strains.

Gene Name	Gene id in specified strain	Strain
PvuE	prot_01909	12E03
PvuD	prot_01910	12E03
PvuC	prot_01911	12E03
PvuB	prot_01912	12E03
PvuA	prot_01913	12E03
PsuA	prot_01914	12E03
PvsA	prot_01915	12E03
PvsB	prot_01916	12E03
PvsC	prot_01917	12E03
PvsD	prot_01918	12E03
PvsE	prot_01919	12E03
IucA	prot_03979	5F-79
IucB	prot_03980	5F-79
IucC	prot_03981	5F-79
IucD	prot_03982	5F-79
IutA	prot_03983	5F-79
Aerobactin_permease	prot_03978	5F-79
Aerobactin_PBP	prot_03977	5F-79
Aerobactin_ATPase	prot_03976	5F-79
Hyp_aerobactin_uptake	prot_03974	5F-79
Aerobactin_Put_membrane	prot_03975	5F-79
Yersinibactin_TonB_OM_1F-230	prot_02601	1F-230
Yersiniabactin_ATPase_FepC	prot_02599	1F-230
Yersiniabactin_Permease_FecCD	prot_02600	1F-230
Yersiniabactin_PBP_FecB	prot_02602	1F-230
Yersiniabactin_TonB_FF-167	prot_01861	FF-167
Anguibactin_fatD_5S-149	prot_00137	5S-149
Pyoverdin_FhuB_permease_ZS-139	prot_03562	ZS-139
Pyoverdin_FhuA_OM_ZS-139	prot_03557	ZS-139
Pyoverdin_FhuD_PBP_ZS139	prot_03563	ZS-139
Pyoverdin_FhuC_ATPbinding	prot_03564	ZS-139
Vulnibactin_TonB_receptor	prot_02642	1F-187
Enterobactin_TonB_fepA	prot_03569	FF-167
Catechol_PBP	prot_03697	5F-59
Catechol_permease1	prot_03698	5F-59
Catechol_permease2	prot_03699	5F-59
Catechol_ATP	prot_03700	5F-59
Util_unknown_catechol	prot_00054	5F-59
FecD	prot_03423	1F-155
FhuA	prot_03947	1F-155

FhuB	prot_04384	1F-155
FhuD	prot_04383	1F-155
FhuC	prot_04382	1F-155
FhuB_bis	prot_00352	12G01
Ferrichrome_PBP	prot_00353	12G01
Ferrichrome_iron_receptor	prot_00355	12G01
FhuC_bis	prot_00354	12G01
Ferrous_Iron_Transport_A	prot_03887	12B01
Ferrous_Iron_Transport_B	prot_03886	12B01
Ferrous_Iron_Transport_C	prot_03885	12B01
Siderophore_TonB	prot_00179	12B01
Siderophore_ExbB	prot_00180	12B01
Siderophore_TonB_1	prot_03694	12B01
Siderophore_ExbB_1	prot_03696	12B01
Ferrichrome_iron_receptor_bis	prot_00404	12G01
iron_compound_substrate_binding	prot_00405	12G01
ferric_siderophore_permease	prot_00406	12G01
iron_transport_protein	prot_00407	12G01
Ferrichrome_iron_receptor_tert	prot_01473	12G01
Ferrichrome_iron_receptor_quart	prot_03949	12G01
TonB_HemeReceptor_Ferrichrome_HutA	prot_01497	12F01
HutA_bis	prot_05131	12F01
TonB_HutR	prot_01994	12F01
Hypotetical_HutR	prot_01992	12F01
Hemin_ABC_Transporter_permease	prot_00183	12B01
Hemin_ABC_transporter_pbinding	prot_00182	12B01
Hemin_ATPase	prot_00184	12B01

Public good dynamics drive evolution of iron acquisition strategies in natural bacterioplankton populations

Otto X. Cordero[a,1], Laure-Anne Ventouras[b,1], Edward F. DeLong[a,b], and Martin F. Polz[a,2]

Departments of [a]Civil and Environmental Engineering and [b]Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

Edited by W. Ford Doolittle, Dalhousie University, Halifax, NS, Canada, and approved October 24, 2012 (received for review August 1, 2012)

A common strategy among microbes living in iron-limited environments is the secretion of siderophores, which can bind poorly soluble iron and make it available to cells via active transport mechanisms. Such siderophore–iron complexes can be thought of as public goods that can be exploited by local communities and drive diversification, for example by the evolution of "cheating." However, it is unclear whether bacterial populations in the environment form stable enough communities such that social interactions significantly impact evolutionary dynamics. Here we show that public good games drive the evolution of iron acquisition strategies in wild populations of marine bacteria. We found that within nonclonal but ecologically cohesive genotypic clusters of closely related Vibrionaceae, only an intermediate percentage of genotypes are able to produce siderophores. Nonproducers within these clusters exhibited selective loss of siderophore biosynthetic pathways, whereas siderophore transport mechanisms were retained, suggesting that these nonproducers can act as cheaters that benefit from siderophore producers in their local environment. In support of this hypothesis, these nonproducers in iron-limited media suffer a significant decrease in growth, which can be alleviated by siderophores, presumably owing to the retention of transport mechanisms. Moreover, using ecological data of resource partitioning, we found that cheating coevolves with the ecological specialization toward association with larger particles in the water column, suggesting that these can harbor stable enough communities for dependencies among organisms to evolve.

microbial diversity | ocean bacteria | population structure | genome dynamics

Ecological populations of bacteria, comprising clusters of coexisting close relatives with similar resource preference in the environment, are often composed of organisms that can have hundreds of unique genes per genome (1). This enormous diversity among organisms with similar function and niche association remains one of the most puzzling phenomena in microbiology. Identifying the selective pressures that drive the evolution of variable gene content—the so-called flexible genome—is essential to understand the functional impact of genomic diversity on the ecological function of microbial populations in the environment. Functional gene annotations across close relatives in different bacterial species show that variable gene content is relatively enriched in genes involved in synthesis of secondary metabolites, proteins exposed on the cell surface, and defense mechanisms, as well as transport of exogenous compounds (2–4). This suggests that the high turnover of genes in bacteria could result from adaptations to rapidly changing biotic interactions. This idea supports the emerging view that microbial diversity is not solely explained by adaptation to abiotic conditions (5, 6) and indicates that interactions within populations and communities play a large role in creating genomic diversity.

Some of the better-understood types of interactions in bacteria are those mediated by siderophores, small molecules that strongly bind poorly soluble iron. Under iron-limiting conditions, siderophores are excreted outside the cell, where they form complexes with ferric-iron (7). The ferric-iron–siderophore complex is then recognized by specific outer-membrane receptors that transport it back inside the cell. Because any cell that carries specific outer-membrane transporters can take them up, siderophores can be considered a prime example of a public good. They can be exploited by "cheaters," which do not bear the cost of production but nonetheless reap the benefits brought by the iron–siderophore complex. These systems were among the first to be studied as public goods in in vitro experiments (8, 9). They have also been widely studied for their role in pathogenicity (7, 10) and in natural environments such as the ocean, where they may play an important role in iron acquisition under limiting conditions (11–13). Because of the dilute nature of the aquatic environment, it has been suggested that production of public goods such as siderophores is an efficient strategy only in environments where local cell densities are high (14). In line with this idea, siderophore production has been described in heterotrophic marine bacteria such as *Vibrio*, *Alteromonas*, and *Marinobacter* (13), which normally exploit nutrient patches with high local cell densities and potentially high siderophore reabsorption rates.

Although the dynamic emerging from the interaction between cheaters and producers has been extensively studied using mathematical models and synthetic microbial systems (8, 15), it has not been conclusively demonstrated that complex bacterial populations in the environment engage in the stable or recurrent interactions that are assumed in these models. For example, in environments such as the ocean, where sympatric microbial populations and species are diverse and dispersal rates are high, it is unclear whether closely related genotypes assemble reproducibly into nonclonal, locally interacting and coevolving populations, as opposed to forming random assemblages. Therefore, validating the role of social interactions such as public good games in natural populations has potential to provide support for the existence of stable social structure and for the hypothesis that pervasive genomic microdiversity in the environment in part reflects differences in social roles.

The present study seeks to unravel the evolutionary and ecological dynamics of siderophore production and cheating in

Author contributions: O.X.C., L.-A.V., and M.F.P. designed research; O.X.C. and L.-A.V. performed research; O.X.C. and L.-A.V. analyzed data; and O.X.C., L.-A.V., E.F.D., and M.F.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: Whole Genome Shotgun projects have been deposited in the DNA Data Bank of Japan (DDBJ)/European Molecular Biology Laboratory (EMBL)/GenBank databases (accession nos. AJWN00000000, AHTI00000000, AICZ00000000, AIDA00000000-AID500000000, AJYD00000000-AJYZ00000000, and AJZA00000000-AJZQ00000000).

[1]O.X.C. and L.-A.V. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: mpolz@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1213344109/-/DCSupplemental.

natural microbial populations, using a collection of >1,700 marine isolates of Vibrionaceae. These isolates are organized in phylogenetic clusters of highly related but nonclonal genotypes. Clusters of isolates are differentiated by their propensity to associate with different classes of resources, operationally defined by separation into four size classes in the planktonic environment. The smallest class encompasses dissolved organic and inorganic nutrients, primarily exploited by free-living cells, whereas all other size classes represent collections of different types of particles and organisms, onto which bacteria can attach and form biofilms. Because these genotypic clusters group samples of ecologically and genetically similar organisms, they are hypothesized to represent samples from wild ecological populations, and as such they serve as a platform for studies of evolutionary and ecological dynamics of bacteria in the wild.

To explore the evolutionary and ecological dynamics of siderophores, we performed a large-scale phenotypic screen aimed at discerning the variability of the siderophore production among ecologically distinct Vibrionaceae populations. Next, we leveraged genomic analysis of 61 sequenced isolates from these populations to identify the genomic basis of phenotypic variation. Finally, we integrated these results with the habitat partitioning of the Vibrionaceae in the ocean to provide an environmental perspective on how the interaction between ecological and evolutionary phenomena influences iron-acquisition strategies.

Results and Discussion

Our high-throughput phenotypic screening (*Materials and Methods*) indicated that siderophore production is a patchy trait within *Vibrio* populations: clusters of close relatives showed an average frequency of producers of approximately 40% even in the most shallow branches of the phylogeny where the most recently diverged isolates are found (Fig. 1). Three exceptions with high frequency of producers were observed in clusters of the *Vibrio* sp. F12 (16, 17), *Allivibrio fischerii*, and *Vibrio ordalii*. These latter two are known from the literature to be associated with animal hosts (18, 19), where iron acquisition is likely a critical factor in colonization (20). For all other sampled populations, the production trait is patchily distributed across the tree, implying that it is frequently gained and lost within lineages. This suggests variable selective pressure along each lineage to either produce siderophores or to cheat.

Using 61 sequenced genomes from the isolate collection (Table S1), we established the genetic basis of siderophore production variability in the natural *Vibrio* populations. We found a strong match between measured production and the presence of siderophore biosynthesis gene clusters in the genome annotation. Of the 61 isolates tested (25 siderophore positives, 36 negatives) there were only three discrepancies (false negatives) between the functional assay and the genome annotation data, indicating that the patchiness of the trait is primarily the result of gene loss and/or independent acquisitions, as opposed to differential gene regulation or errors in the phenotypic screening assay.

Because siderophore biosynthesis genes are often coded in large gene clusters linked with their specific transporters (21), we expected that nonproducers would lack both biosynthesis genes and the cognate high-affinity receptors. Further inspection of the genomes, however, showed that nonproducers have evolved by

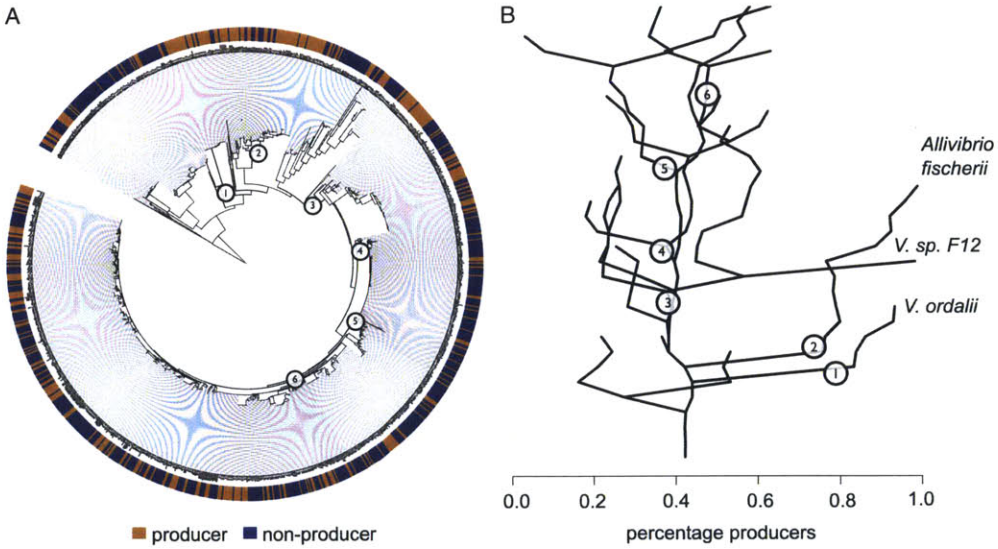


Fig. 1. Siderophore production is an intermediate frequency trait in natural *Vibrio* populations. (*A*) The distribution of producers and nonproducers mapped onto the isolate phylogeny (*Materials and Methods*). The tree is based on the genetic marker hsp60 and comprises different genotypic clusters previously found to have cohesive ecology and hypothesized to represent samples from natural ecological populations. (*B*) Representation of the data shown in *A* in terms of percentage of producers descending from each internal node of the phylogeny. Among the populations with a high incidence of producers are animal host-associated *V. ordalii* and *A. fischerii*. The clade descending from node 3 corresponds to close relatives of *V. splendidus* (referred to as the *V. splendidus*-like group). Nodes 4, 5, and 6 correspond to the most recent common ancestors of *V. crassostreae*, *V. cyclotrophicus*, and *V. splendidus*. Within this large group, the patchy distribution of producers suggests that siderophore biosynthesis genes are rapidly gained and lost, possibly by recombination.

selective loss of biosynthesis genes but retention of the specific receptors, consistent with the canonical model of a cheater. For this analysis, we focused on the *Vibrio splendidus*-like clade, consisting of distinct subpopulations of species *V. splendidus*, *Vibrio crassostreae*, *Vibrio tasmaniensis*, and *Vibrio cyclotrophicus*, thought to specialize in exploiting different marine particulate nutrient patches (16). In these populations, transitions between production and nonproduction seem to have occurred multiple times because of the extreme patchiness of the production trait. The most common siderophores in this clade, vibrioferrin and aerobactin, are encoded by superoperons in which biosynthesis genes are adjacent to the specific siderophore receptors (22, 23). Contrary to our expectation, whereas biosynthesis genes had been excised or replaced, the adjacent receptors in the nonproducers were kept intact (Figs. 2, S2, and S3). Although putative aerobactin cheaters were distributed on different branches within the *V. splendidus*-like clade, the genomic region harboring the siderophore receptors was identical in gene composition and order in these strains (Fig. S1). This observation indicates that nonbiosynthetic gene clusters had a common evolutionary origin and that these clusters are alternative operon variants that are maintained in the population.

Consistent with their putative cheater role, individuals that have lost the ability to produce siderophores suffered a clear decrease in growth efficiency when grown in iron-poor media (*Materials and Methods*). By contrast, the growth of producers was similar in both iron-rich and iron-poor media (Fig. 3 *A* and *C*). Moreover, when supplementing the iron-poor media with pure siderophores, putative cheaters were able to enhance their growth, showing that they could effectively take up the specific siderophores under those conditions (Fig. 3). Insensitivity to iron-limited conditions was also observed in nonproducers when the media was supplemented with the <3 kDa fraction (small molecules) of the cell-free supernatant from siderophore producers (*Materials and Methods*), showing that the cross-feeding interaction is possible in nature. Based on a combination of genomic comparison and biological assays, our analysis suggests that within the studied *Vibrio* populations there is fine-grained differentiation in social roles, whereby some individuals have

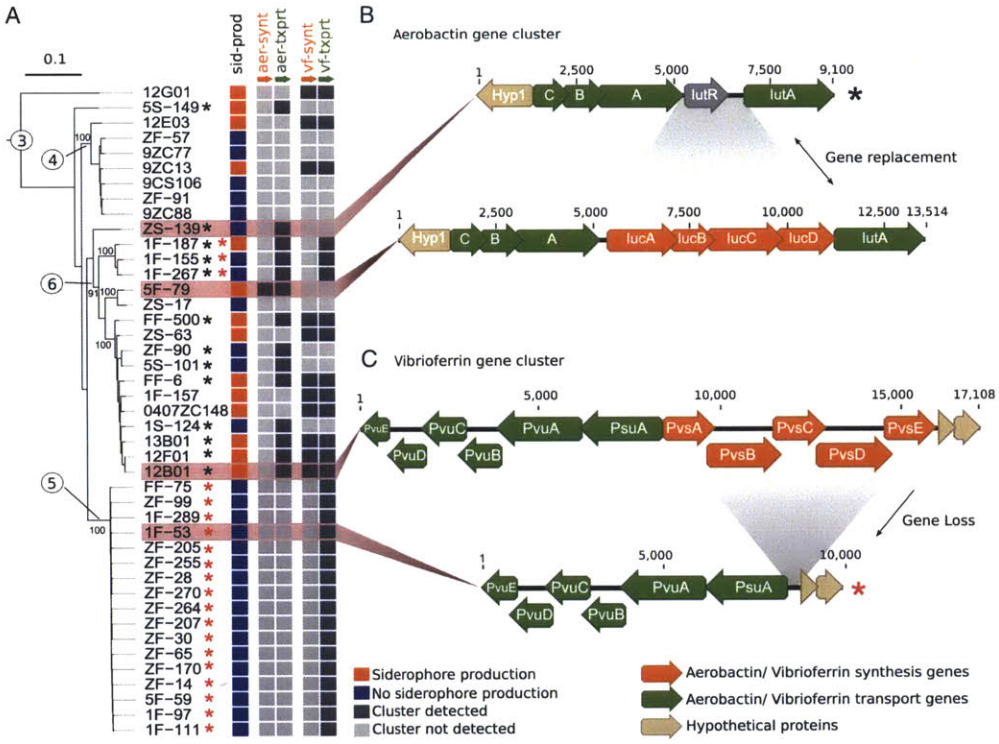


Fig. 2. Evolution of "cheating" explains patchiness of siderophore production trait in the *V. splendidus*-like group. (*A*) Phylogenetic relationship, siderophore production, and siderophore synthesis and transport genes for sequenced strains in the *V. splendidus*-like group. The tree is based on 66 single-copy genes present in all of the sequenced isolates. The "sid-prod" column refers to the outcome of the phenotypic screen; "aer-syn" stands for aerobactin biosynthesis genes, "aer-txpt" for aerobactin-specific transport genes, "vf-syn" for vibrioferrin biosynthesis genes, and "vf-txpt" for vibrioferrin-specific transport genes. (*B* and *C*) Examples of configurations of siderophore synthesis and transporter gene clusters in producer and nonproducer strains. The figures show that nonproducer phenotype evolved from excision of the biosynthesis genes from the complete synthesis-transport cluster. Black and red stars in *A* indicate the presence of similar "cheater" gene cluster configurations as those shown in *B* and *C* (Figs. S1 and S2).

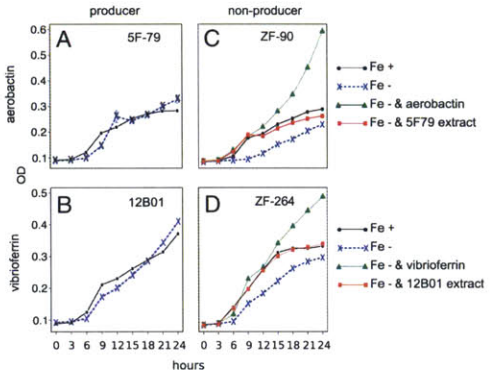


Fig. 3. Growth assays confirm bioinformatic predictions of siderophore biosynthesis and utilization capabilities. (*A* and *B*) Growth of producer strains 5F-79 and 12B01 in M9-based media (Fe^{+}) and in media supplemented with 100 μM EDDA (Fe^{-}), an iron-specific chelator. The results show that both siderophore producer strains 5F-79 (aerobactin) and 12B01 (vibrioferrin) are insensitive to EDDA addition. By contrast, nonproducers ZF-90 and ZF-264 (*C* and *D*), which have specific receptors for aerobactin (aer) and vibrioferrin (vf), respectively, suffered an appreciable effect when grown in Fe^{-} media (100 μM EDDA for ZF-90, 50 uM EDDA for ZF-264). This effect was counteracted by addition of the corresponding siderophore aerobactin (80 μM) or vibrioferrin (70 μM) and by cell-free extract from corresponding producers 5F-79 and 12B01.

evolved to exploit exogenous pools of siderophores, whereas others remain self-sufficient.

Having established the variability in siderophore production and its genetic basis, we asked what habitats and population dynamics favor cheaters or producers in the marine environment. Focusing on the *V. splendidus*-like group containing small subpopulations with recent and frequent transitions between particle attached and free-living life-styles (Fig. S3), we measured the correlation between the siderophore production trait and the size fraction associated to each isolate. To that end, we used an adaptation of Felsenstein's independent contrast method to control for the fact that traits measured for each strain (e.g., siderophore production) are not independent observations but related through the phylogeny (*Materials and Methods*) (24). Doing this we could control for spurious correlations that may appear as a result of the differential associations between clades and traits and not because of a real dependency between the traits. In addition, we accounted for the effect of phylogenetic uncertainty by calculating the distribution of correlation values on a set of 100 bootstrap trees.

The results of the correlation analysis (Fig. 4) show that association with large particle sizes in the ocean is correlated with a higher frequency of cheaters, whereas the free-living stage (<1 μm) is positively correlated with siderophore production. We find that these correlations are significant, with P values <0.01, and not driven by a few transitions on a few clades (Fig. S4). The correlation between siderophore production and the free-living lifestyle is the strongest among the four size fractions. For the other size fractions, the distribution of correlation values shifts progressively toward negative values, with the lowest correlation values observed for the largest size fraction. This result shows that the frequency of producers changes in a predictable manner, based on the size of the particle to which the population is specialized.

These results suggest that the opportunity for cheating is highest on large particles. One hypothesis for the observed trend is that higher local iron concentrations on large particles preclude the need for siderophore production. However, the observation that nonproducers maintain receptors while selectively losing biosynthesis genes (Fig. 2) indicates that social cheating—and not only environmental selection—drives changes in siderophores producer abundance. We hypothesize that the observed correlation is explained by differences in the structure of social interactions between particle-attached vs. nonattached bacteria. Vibrios are copiotrophs, which can rapidly expand their populations under favorable conditions, in the extreme leading to blooms that can dominate a bacterioplankton community (25). Between blooms, some vibrios have a tendency to attach to particles, whereas others remain planktonic (16). Marine particles can harbor dense communities of attached bacteria with high metabolic activity (26), and therefore vibrios attached to particles experience relatively constant high cell densities in their local environment. On the other hand, those that remain free-living or attached to smaller aggregates fluctuate from high cell densities during blooms to low densities during their planktonic

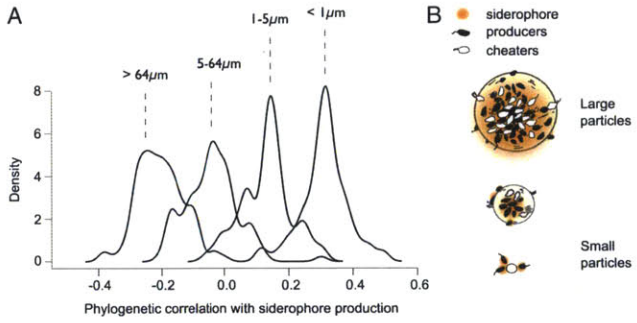


Fig. 4. Cheaters within the *V. splendidus*-like clade are more successful in larger particles than in smaller ones. (*A*) Spearman correlations calculated between the phylogenetic contrasts of the "siderophore production" trait and the different sampling size fractions (SOM). The distributions cover values calculated for 100 bootstraps of the hsp60 tree for the 1,013 vibrio isolates. (*B*) Cartoon representation of the impact of particle size on frequency of siderophore producers. Our results suggests that social interactions are more relevant on large particles, possibly owing to the higher cell densities and long periods of attachment facilitating the accumulation and exploitation of public goods.

phase. This type of bottleneck population dynamics experienced by free-living vibrios has been shown to favor the abundance of cooperators in public good games (27, 28). On the other hand, the constant high cell densities in the local environment of particle-attached bacteria could favor the establishment of cheaters that "scrounge" public goods from producers (29). This proposed model, based on the specific population dynamics associated to different microenvironments, provides an explanation for the variation in frequency of producers observed between particle attached and free-living *Vibrio*.

Conclusion

Our results suggest that public good interactions—in particular those mediated by siderophores—could play an important role in structuring natural bacterial populations in the ocean. By studying the patterns of social interactions imprinted on bacterial genomes, we have documented the genomic events that led to transitions between siderophore production and cheating. The evidence suggests that these trends are mediated by dynamic differential gain and loss of siderophores gene cluster variants. Phylogenetic analysis shows that although populations retain an intermediate frequency of producers, transitions occur frequently along individual lineages, so that producers and cheaters coexist in a dynamic equilibrium whereby public good output is stable at the population level but not at the individual level. These results contrast with the idea that functional characterizations of genotypes and populations are interchangeable, as is frequently assumed. Our results further provide evidence that the incentive to cheat in public good games can drive genomic diversity, as recently suggested by the Black Queen Hypothesis (30), and that this force is relevant at the scale of closely related but nonclonal genotypes in natural populations.

Finally, the data indicate that the marine microenvironment influences public good interactions in predictable ways, in this case with particle size influencing the frequency of cheaters via its effect on population dynamics. Marine particles are known to harbor dense microbial communities with high metabolic activity and potential for interactions (26, 31, 32), and they are important features of the marine ecosystem (26). Here, we propose that public good interactions impact the evolution and dynamics of the microbial communities that colonize and grow on marine particles and that public good game theory could be applied to understand microbial diversity in these important microenvironments.

Because of the dilute nature of the marine environment, the role of siderophores as a successful iron acquisition strategy in the ocean is often questioned (14). It has been suggested that siderophore production may only be beneficial in high cell density environments where high local siderophore concentrations can counteract the effect of diffusion kinetics (14). On the surface, this idea would seem to contradict our result showing a negative correlation between the frequency of producers and particle size. However, as mentioned above, the *Vibrio* strains isolated from the free-living fraction most likely experience episodic population dynamics reaching high cell densities during punctual bursts of growth. We hypothesize that the difference in population dynamics between small size fraction and large particle populations could explain the correlations reported in Fig. 4. Therefore, our study suggests that not only diffusion kinetics, but also the dynamics of social interactions need to be considered to explain the distribution of siderophore producers in the marine environment.

These hypotheses open avenues for further research, which should explore the role of public good games among other particle-attached species. In contrast to classic public good game models, the populations studied here are not clonal, and although we have focused our study on one particular type of public good, because of their genome diversity cheaters with respect to one public good could act as cooperators with respect to another. Further research is needed to investigate these types of synergisms. Overall, our study highlights the importance of describing population-based phenomena, for interpreting the mechanisms that generate and sustain the enormous genetic diversity observed in wild populations bacteria.

Materials and Methods

Siderophore Screen of Wild Isolates. We screened a total of 1,710 Vibrionaceae isolates from seawater collected at the Plum Island Estuary, MA using the well-established Chrome-Azurol S (CAS) assay (33). Biological replicate measurements for both the liquid and solid versions of the assay were performed in high-throughput using a 96-well format. For the solid and liquid versions of the CAS assay, strains were first grown overnight in duplicates in 2216 marine broth (Difco, Becton Dickinson) at room temperature. Overnight cultures were stamped directly onto CAS agar plates (solid assay) and transferred into iron-poor media to induce siderophore production in liquid (*SI Materials and Methods*). After 48 h cell-free supernatants were mixed in a 1:1 ratio with liquid CAS dye and 2 μL shuttle solution (33). The mixture was incubated in the dark for 15 min, after which absorbance at 630 nm was measured on a Synergy2 filter-based multimode plate reader (Biotek). Siderophore production was considered positive for all absorbances <0.3. The liquid CAS dye and shuttle solution were prepared as previously described (33).

Bioinformatic Analysis of Siderophore Synthesis and Transport Gene Clusters. Gene finding and annotations were performed using the SEED subsystems (34) and the RAST server (35) on the 61 genomes. Siderophore biosynthesis clusters were identified using a combination of annotation text searches and the software AntiSMASH developed to identify biosynthetic clusters of secondary metabolites (36). Details of the search criteria used for these programs are presented in *SI Materials and Methods*.

Phylogenetic Correlation Analysis. We calculated the Spearman rank correlation between the siderophore production and particle size traits associated with each isolate. To correct for the fact that values measured for each strain are not independent of each other (being related through the phylogeny), we used Felsenstein's independent contrast method (24). Our data are composed of binary values on the leaves of the tree (e.g., 1 production, 0 nonproduction; 1 large-particle associated, 0 otherwise), and we reconstructed states for internal nodes of the tree by calculating the fraction of nonzero values under each node. We then applied Felsenstein's independent contrast method (24) to obtain a vector of differences between the fractions of producers under sister nodes of the phylogeny. This vector corresponds to evolutionary transitions increasing or decreasing the frequency of producers or strains associated to a specific size fraction. Using the vectors of transitions for the siderophore production and size fraction variables, Spearman rank correlations were calculated (removing nodes where either transition was 0) between the vectors to measure the association between the traits. To reduce the noise in the estimation of frequencies, we did not consider evolutionary transitions between leaves or between leaves and internal nodes. This process was repeated for each of the 100 maximum likelihood bootstrap trees calculated with the FastTree program (37) using the evolutionarily conserved housekeeping gene *hsp60*. The distributions in Fig. 4 show the results of this analysis for each size fraction. Distributions were smoothed using kernel density estimation as implemented in the function "density" provided in the statistical computing language R.

Growth Enhancement Experiments. Isolates were grown overnight at room temperature in 2216 marine broth and pelleted (2 min at 9,300 × *g*) to wash cells using a modified M9 salts solution. Cells were then inoculated (1:1,000) into modified M9 salts growth media (*SI Materials and Methods*) with (iron-poor) or without (iron-replete) the iron-specific chelator EDDA (ethylenediamine-*N*,*N*′-diacetic acid) (Sigma-Aldrich). During incubation at room temperature absorbance at 600 nm was measured every 3 h for 24 h. Purified ferric-aerobactin (EMC Microcollections) and vibrioferrin (generously provided by Carl Carrano, San Diego State University, San Diego, CA) were added at specified concentrations. Extracts from the aerobactin producer 5F-79 and the vibrioferrin producer 12B01 grown under iron-poor conditions were obtained by filtering cell-free supernatant through a 3-kDa membrane using an Amicon Ultra centrifugal unit (Millipore).

ACKNOWLEDGMENTS. We thank Carl J. Carrano for generously providing purified vibrioferrin. Funding for this work was provided by National Science

Foundation (NSF) Grant DEB 0821391; the Gordon and Betty Moore Foundation; the Cooperative Agreement between the Masdar Institute of Science and Technology (Masdar Institute), Abu Dhabi, United Arab Emirates and the Massachusetts Institute of Technology (MIT), Cambridge, MA (Reference 02/MI/MI/CP/11/07633/GEN/G/00); Center for Microbial Oceanography: Research and Education NSF Science and Technology Center Award EF0424599; and the Broad Institute's Scientific Planning and Allocation of Resources Committee program.

1. Shapiro BJ, et al. (2012) Population genomics of early events in the ecological differentiation of bacteria. *Science* 336(6077):48–51.
2. Nakamura Y, Itoh T, Matsuda H, Gojobori T (2004) Biased biological functions of horizontally transferred genes in prokaryotic genomes. *Nat Genet* 36(7):760–766.
3. Cordero OX, Hogeweg P (2009) The impact of long-distance horizontal gene transfer on prokaryotic genome size. *Proc Natl Acad Sci USA* 106(51):21748–21753.
4. Boucher Y, et al. (2011) Local mobile gene pools rapidly cross species boundaries to create endemicity within global *Vibrio cholerae* populations. *mBio* 2(2):e00335-10.
5. Kolenbrander PE, Egland PG, Diaz PI, Palmer RJ, Jr. (2005) Genome-genome interactions: Bacterial communities in initial dental plaque. *Trends Microbiol* 13(1):11–15.
6. Lawrence D, et al. (2012) Species interactions alter evolutionary responses to a novel environment. *PLoS Biol* 10(5):e1001330.
7. Miethke M, Marahiel MA (2007) Siderophore-based iron acquisition and pathogen control. *Microbiol Mol Biol Rev* 71(3):413–451.
8. Griffin AS, West SA, Buckling A (2004) Cooperation and competition in pathogenic bacteria. *Nature* 430(7003):1024–1027.
9. West SA, Buckling A (2003) Cooperation, virulence and siderophore production in bacterial parasites. *Proc Biol Sci* 270(1510):37–44.
10. Dale SE, Doherty-Kirby A, Lajoie G, Heinrichs DE (2004) Role of siderophore biosynthesis in virulence of *Staphylococcus aureus*: Identification and characterization of genes involved in production of a siderophore. *Infect Immun* 72(1):29–37.
11. Mawji E, et al. (2008) Hydroxamate siderophores: Occurrence and importance in the Atlantic Ocean. *Environ Sci Technol* 42(23):8675–8680.
12. Amin SA, et al. (2009) Photolysis of iron-siderophore chelates promotes bacterial-algal mutualism. *Proc Natl Acad Sci USA* 106(40):17071–17076.
13. Sandy M, Butler A (2009) Microbial iron acquisition: Marine and terrestrial siderophores. *Chem Rev* 109(10):4580–4595.
14. Völker C, Wolf-Gladrow DA (1999) Physical limits on iron uptake mediated by siderophores or surface reductases. *Mar Chem* 65:227–244.
15. Damore JA, Gore J (2012) Understanding microbial cooperation. *J Theor Biol* 299: 31–41.
16. Hunt DE, et al. (2008) Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* 320(5879):1081–1085.
17. Preheim SP, et al. (2011) Metapopulation structure of Vibrionaceae among coastal marine invertebrates. *Environ Microbiol* 13(1):265–275.
18. Nyholm SV, Stewart JJ, Ruby EG, McFall-Ngai MJ (2009) Recognition between symbiotic *Vibrio fischeri* and the haemocytes of *Euprymna scolopes*. *Environ Microbiol* 11(2):483–493.
19. Austin B (2011) Taxonomy of bacterial fish pathogens. *Vet Res* 42(1):20.
20. Ratledge C, Dover LG (2000) Iron metabolism in pathogenic bacteria. *Ann Rev Microbiol* 54:881–941.
21. Crosa JH, Walsh CT (2002) Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. *Microbiol Mol Biol Rev* 66(2):223–249.
22. Suzuki K, et al. (2006) Identification and transcriptional organization of aerobactin transport and biosynthesis cluster genes of *Vibrio hollisae*. *Res Microbiol* 157(8): 730–740.
23. Tanabe T, et al. (2003) Identification and characterization of genes required for biosynthesis and transport of the siderophore vibrioferrin in *Vibrio parahaemolyticus*. *J Bacteriol* 185(23):6938–6949.
24. Felsenstein J (1985) Phylogenies and the comparative method. *Am Nat* 125:1–15.
25. Gilbert JA, et al. (2012) Defining seasonal marine microbial community dynamics. *ISME J* 6(2):298–308.
26. Simon M, Grossart HP, Schweitzer B, Ploug H (2002) Microbial ecology of organic aggregates in aquatic ecosystems. *Aquat Microb Ecol* 28:175–211.
27. Chuang JS, Rivoire O, Leibler S (2009) Simpson's paradox in a synthetic microbial system. *Science* 323(5911):272–275.
28. Cremer J, Melbinger A, Frey E (2012) Growth dynamics and the evolution of cooperation in microbial populations. *Sci Rep* 2:281.
29. Nadell CD, Xavier JB, Foster KR (2009) The sociobiology of biofilms. *FEMS Microbiol Rev* 33(1):206–224.
30. Morris JJ, Lenski RE, Zinser ER (2012) The Black Queen Hypothesis: Evolution of dependencies through adaptive gene loss. *mBio* 3(2):e00036-12.
31. Long RA, et al. (2005) Antagonistic interactions among marine bacteria impede the proliferation of *Vibrio cholerae*. *Appl Environ Microbiol* 71(12):8531–8536.
32. Grossart HP, Kiørboe T, Tang K, Ploug H (2003) Bacterial colonization of particles: Growth and interactions. *Appl Environ Microbiol* 69(6):3500–3509.
33. Schwyn B, Neilands JB (1987) Universal chemical assay for the detection and determination of siderophores. *Anal Biochem* 160(1):47–56.
34. Overbeek R, et al. (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* 33(17):5691–5702.
35. Aziz RK, et al. (2008) The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics* 9:75.
36. Medema MH, et al. (2011) antiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res* 39(Web Server issue):W339-46.
37. Price MN, Dehal PS, Arkin AP (2009) FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 26(7):1641–1650.

Supporting Information

Cordero et al. 10.1073/pnas.1213344109

SI Materials and Methods

Siderophore Screen of Wild Isolates. We screened a total of 1,710 Vibrionaceae isolates from seawater collected at the Plum Island Estuary, MA (1, 2). The isolates were chosen such that each ecologically cohesive population previously identified as part of these studies was represented. Siderophore production was assessed using the well-established Chrome-Azurol S (CAS) assay (3). Biological replicate measurements for both the liquid and solid versions of the assay were performed in high-throughput, using either 96-well 2-mL culture blocks or 96-well microtiter plates for absorbance measurements.

For the solid version of the CAS assay, strains were grown in duplicates overnight in 2216 marine broth (Difco, Becton Dickinson) at room temperature and stamped directly onto CAS agar plates. CAS agar plates were prepared by mixing a dye made of CAS, Fe, and hexadecyl-trimethyl-ammonium bromide (HDTMA) with M9-based growth media (see below) appropriate for Vibrionaceae. For 1 L of CAS-agar, 100 mL of CAS-Fe-HDTMA dye was mixed with 900 mL of freshly prepared growth media. The CAS-Fe-HDTMA dye was prepared in advance as follows, for 1L: 10 mL of a 10 mM ferric chloride ($FeCl_3$) in 100 mM hydrochloric acid (HCl) solution was mixed with 590 mL of a 1-mM aqueous solution of CAS. The Fe-CAS solution was then added to 400 mL of a 2-mM aqueous solution of HDTMA. The resulting CAS-Fe-HDTMA solution was autoclaved for 25 min in a polycarbonate bottle that had previously been soaked overnight in 10% (vol/vol) HCl then rinsed five times with MilliQ water. The CAS-Fe-HDTMA dye was stored at room temperature covered from light until use.

The growth media was prepared as follows, for 1L of CAS-agar: 30.24 g of 1,4-piperazine-diethanesulfonic acid (Pipes), together with 1 g of ammonium chloride (NH_4Cl), 3 g potassium phosphate (KH_2PO_4), and 20 g sodium chloride (NaCl) was dissolved into MilliQ water by adjusting the pH with 10 M NaOH to 6.8. Note that the commonly used phosphate buffer disodium phosphate (Na_2HPO_4-$_7H_2O$) was omitted because phosphate can chelate iron and lead to a discoloration of the CAS dye. As a solidifying agent, 9 g of agar noble (Difco) were added to the solution. We found that the more commonly used solidifying agents, agarose, and agar also led to a discoloration of the CAS dye, likely owing to higher phosphate content. The volume was adjusted to 860 mL, and the solution was autoclaved. After cooling, 30 mL of a sterile 10% (wt/vol) Casamino acids (Difco) aqueous solution and 10 mL of a sterile 20% (wt/vol) glucose aqueous solutions were added. Finally, the 100 mL of CAS-Fe-HDTMA were added to the growth media. The final concentrations of the CAS-agar components are as follows: 100 mM Pipes, 18 mM NH_4Cl, 22 mM KH_2PO_4, 2% (wt/vol) NaCl, 0.3% casamino acids, 0.2% (wt/vol) glucose, 10 μM $FeCl_3$, 58 μM CAS, and 80 μM HDTMA. Unless otherwise noted, all chemicals were obtained from Sigma-Aldrich.

For the liquid version of the assay, isolates were first grown overnight in duplicates in 2216 marine broth (Difco) at room temperature, then transferred into the Vibrionaceae growth media described above (without the solidifying agent or the CAS-Fe-HDTMA dye). After overnight growth at room temperature, isolates were transferred into iron-poor media, consisting of the Vibrionaceae growth media described above amended with 100 μM 2,2′-bipyridyl and 10 nM FeCl3 to induce siderophore production. After 48 h, cultures were centrifuged and 99 μL of supernatant were mixed in a 1:1 ratio with liquid CAS dye and 2 μL shuttle solution (3). The mixture was incubated in the dark for 15 min, after which absorbance at 630 nm was measured on a Synergy2 filter-based multimode plate reader (Biotek). In the presence of a siderophore, the absorbance of the dye at 630 nm is quenched. Siderophore production was considered positive for all absorbances <0.3. The liquid CAS dye and shuttle solution were prepared as previously described (3).

Bioinformatic Analysis of Siderophore Synthesis and Transport Gene Clusters. Gene finding and annotations were performed using the SEED subsystems (4) and the RAST server (5). Siderophore biosynthesis clusters were identified using a combination of annotation text searches and the software AntiSMASH developed to identify biosynthetic clusters of secondary metabolites (6). AntiSMASH was run on all of the contigs for all 61 genomes using default parameters. For the purpose of this study, we focused on the "nrps" and "siderophore" metabolite clusters identified by AntiSMASH. These matched well with the annotation text searches that were concurrently performed. Text searches were performed for the following expressions: "siderophore," "actin" (encompassing many of the names of known siderophores), "ferrin," "pyoverdin," "chelin," "ferrichrome," "heme," "hydroxamate," "catechol." In some cases, genes belonging to a siderophore synthesis cluster were missed by the RAST annotation software. The presence of those genes in the siderophore synthesis clusters was confirmed by performing a BLASTp (7) search against the National Center for Biotechnology Information nonredundant protein database. For the genomes in which a "truncated pyoverdin-like" cluster was found, only two nonribosomal peptide-synthetase (NRPS) enzymes, homologs of PvdI and PvdD, were found as opposed to three or more NRPS needed for the synthesis of pyoverdin (8).

Growth Enhancement Experiments. Growth enhancement experiments were performed for the following isolates: ZS139, ZF90, 12B01, and ZF264. Isolates were grown overnight at room temperature in 2216 marine broth (Difco). Cultures were pelleted (2 min at 10,000 rpm in a microcentrifuge), the supernatant was discarded, and pellets were washed once in modified M9 salts [42.2 mM dibasic sodium phosphate (Na_2HPO_4), 22 mM monobasic potassium phosphate (KH_2PO_4), 18.7 mM ammonium chloride (NH_4Cl), and 2% (wt/vol) sodium chloride (NaCl)] to remove any carry-over 2216 media. Washed cells were pelleted and resuspended in modified M9 salts. Cells were then inoculated (1:1,000) into growth media [modified M9 salts, 0.3% (wt/vol) casamino acids, 200 mM magnesium sulfate ($MgSO_4$), 20 μM calcium chloride ($CaCl_2$), and vitamins: 0.1 $\mu g \cdot L^{-1}$ vitamin B12, 2 $\mu g \cdot L^{-1}$ biotin, 5 $\mu g \cdot L^{-1}$ calcium pantothenate, 2 $\mu g \cdot L^{-1}$ folic acid, 5 $\mu g \cdot L^{-1}$ nicotinamide, 10 $\mu g \cdot L^{-1}$ pyridoxin hydrochloride, 5 $\mu g \cdot L^{-1}$ riboflavin, 5 $mg \cdot L^{-1}$ thiamin hydrochloride] with (iron-poor) or without (iron-replete) the iron-specific chelator EDDA (ethylenediamine-N,N′-diacetic acid). Cells were incubated at room temperature. Absorbance at 600 nm was measured every 3 h for 24 h. Purified ferric-aerobactin (EMC Microcollections) and vibrioferrin (generously provided by Carl Carrano, San Diego State University, San Diego, CA) were added at specified concentrations. Extracts from the aerobactin producer 5F-79 and the vibrioferrin producer 12B01 grown under iron-poor conditions were obtained by filtering cell-free supernatant through a 3-kDa membrane using an Amicon Ultra centrifugal unit (Millipore).

1. Hunt DE, et al. (2008) Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* 320(5879):1081–1085.
2. Preheim SP, et al. (2011) Metapopulation structure of Vibrionaceae among coastal marine invertebrates. *Environ Microbiol* 13(1):265–275.
3. Schwyn B, Neilands JB (1987) Universal chemical assay for the detection and determination of siderophores. *Anal Biochem* 160(1):47–56.
4. Overbeek R, et al. (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* 33(17): 5691–5702.
5. Aziz RK, et al. (2008) The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics* 9:75.
6. Medema MH, et al. (2011) antiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res* 39(Web Server issue):W339-46.
7. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410.
8. Ravel J, Cornelis P (2003) Genomics of pyoverdine-mediated iron uptake in pseudomonads. *Trends Microbiol* 11(5):195–200.

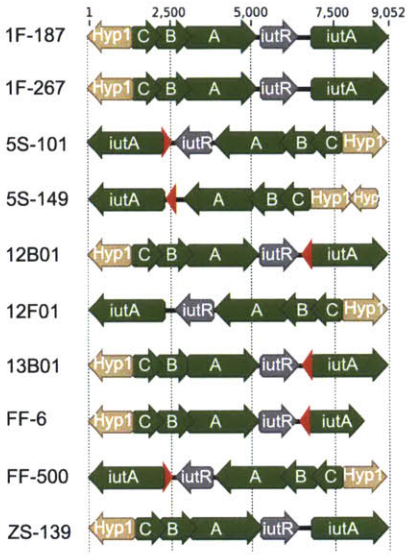


Fig. S1. List of aerobactin clusters in nonproducers. Genes in orange correspond to biosynthesis and genes in green to transport. Other colors correspond to genes annotated as hypothetical or unknown function. With the exception of 5S-149, there is basically only one cluster variant present across nonproducers, suggesting that the cluster was created once and shared by recombination.

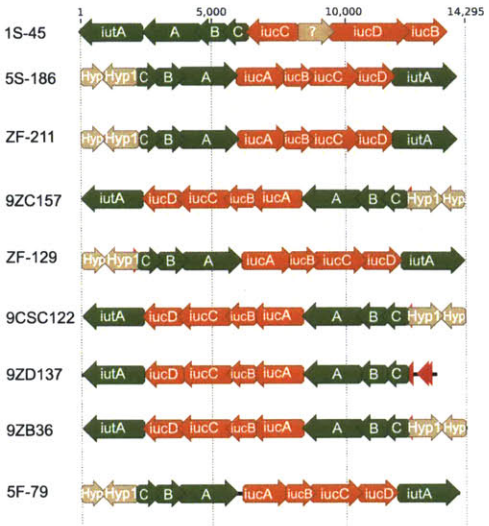


Fig. S2. Aerobactin biosynthesis and transport gene clusters present in producers. Genes in orange correspond to biosynthesis and genes in green to transport. Other colors correspond to genes annotated as hypothetical or unknown function. The figure shows that across the studied *Vibrio* populations there is essentially one version of the biosynthesis gene cluster.

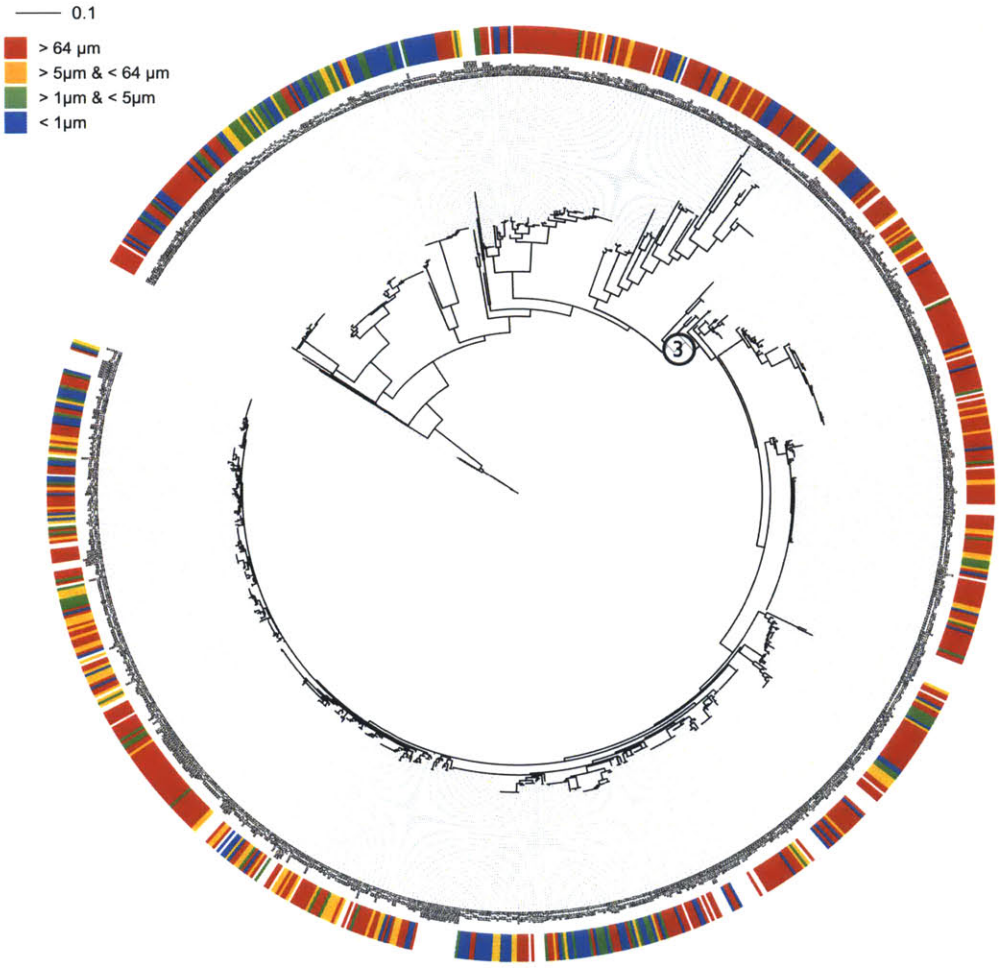


Fig. S3. Hsp60 phylogeny of vibrios indicating the corresponding size fractions for those strains isolated via the size fractionation scheme. The figure shows that in the *Vibrio splendidus*-like group (under node 3), the majority of strains are particle attached and that transitions to nonattached lifestyle have occurred multiple times on different parts of the phylogeny. Approximately 70% of the strains in the figure are found in particles larger than 5 µm, and the remainder in particles smaller than 5 µm.

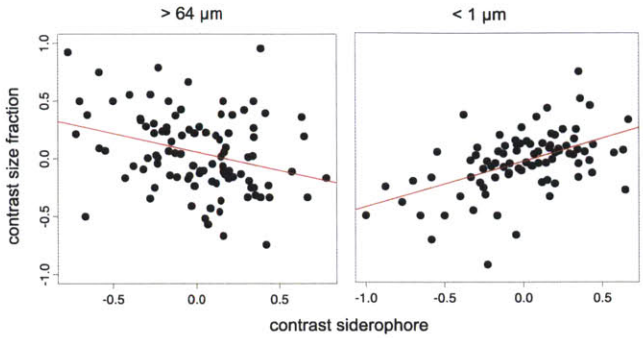


Fig. S4. Scatter plot of the independent contrasts between the siderophore trait and the two extreme size fractions in a randomly chosen bootstrap tree. The figure illustrates that the measured correlations are supported by multiple phylogenetic transitions. Accordingly, correlations calculated on these data are significant at *P* value 0.01.

Table S1. List of 61 isolates with full genome sequences

Strain	Production measured	Siderophore anNotated
ZF-55	No	No
FF-50	No	No
1A06	No	No
ZF-29	No	No
1S-45	Yes	Yes
5S-186	Yes	Yes
ZF-211	Yes	Yes
FF-238	No	No
FF-167	No	Yes
FF-93	Yes	Yes
FS-238	Yes	Yes
FS-144	Yes	Yes
12B09	Yes	Yes
9ZC157	Yes	Yes
ZF-129	Yes	Yes
9CSC122	Yes	Yes
9ZD137	Yes	Yes
9ZB36	No	Yes
12G01	Yes	Yes
5S-149	Yes*	Yes
12E03	Yes	Yes
ZF-57	No	No
9ZC77	No	No
9ZC13	Yes	Yes
9CS106	No	No
ZF-91	No	No
9ZC88	No	No
ZS-139	No	No
1F187	Yes	Yes
1F-155	No	No
1F-267	No	No
5F-79	Yes	Yes
ZS-17	No	No†
FF-500	Yes	Yes
ZS-63	Yes	Yes
ZF-90	No	No†
5S-101	No	No
FF-6	Yes	Yes
1F-157	Yes	Yes
0407ZC148	Yes	Yes
1S-124	No	Yes
13B01	Yes	Yes
12F01	Yes	Yes
12B01	Yes	Yes
FF-75	No	No†
ZF-99	No	No†
1F-289	No	No†
1F-53	No	No†
ZF-205	No	No†
ZF-255	No	No†
ZF-28	No	No†
ZF-270	No	No†
ZF-264	No	No†
ZF-207	No	No†
ZF-30	No	No†
ZF-65	No	No†
ZF-170	No	No†
ZF-14	No	No†
5F-59	No	No†
1F-97	No	No†
1F-111	No	No†

*Siderophore biosynthesis genes identified on plasmid sequence.
†Truncated pyoverdin-like cluster.

Chapter 3

Changes in Microbial Community Structure in Response to Relief of Iron Limitation in the Central Equatorial Pacific Ocean.

3.1. Abstract

Iron is a limiting micronutrient in vast regions of the world's ocean. As a necessary cofactor of photosynthetic reactions, iron availability has been shown to impact phytoplankton growth and activity, and hence modulate the global carbon cycle in the Equatorial Pacific Ocean. With naturally fluctuating iron inputs to the region, determining the response of the microbial community to iron addition is key in furthering our understanding of the carbon cycle in the region. Using next-generation high-throughput sequencing we characterized iron-induced changes in community structure of the whole microbial community in the Central Equatorial Pacific. We confirmed previous results that iron leads to a community shift, selecting for a phytoplankton community dominated by large cell diatom species, the majority of which are from the *Pseudo-nitzschia* spp. type. We further show that shifts in community structure are accompanied by changes in bacterial and viral assemblages, with a notable enrichment of Bacteroidetes. Functional gene analysis shows that iron selects for a functionally distinct Bacteroidetes population. We suggest that the Bacteroidetes enrichment is an indirect effect of iron addition, as the functional gene analysis corroborated with the literature seems to indicate that they live in close association with diatom species. Whether this enrichment of Bacteroidetes impacts bacterial production (and hence the carbon cycle) at the ecosystem level remains to be elucidated. This work represents the first metagenomic effort to characterize the impact of iron addition on microbial community structure in the Central Equatorial Pacific Ocean.

With René Boiteau[1,2], Dawn Moran[1], Dan Repeta[1], Mak Saito[1], Edward DeLong[3,4].

[1]Marine Chemistry and Geochemistry Department, WHOI, Woods Hole, MA, USA
[2]Earth Atmospheric and Planetary Sciences Department, MIT, Cambridge, MA, USA
[3]Civil and Environmental Engineering Department, MIT, Cambridge, MA, USA
[4]Biological Engineering Department, MIT, Cambridge, MA, USA

3.2. Background

In large regions of the world's oceans, iron limits phytoplankton growth. This is especially true in regions characterized by low biomass abundances despite relatively high nutrient concentrations such as in the Equatorial Pacific, the Southern Ocean and the Subarctic North Pacific Ocean. As these regions comprise ~20% of the oceans (Martin et al., 1994) and constitute a significant biome of our planet, it is important to unravel the factors that may govern microbial activity and hence impact global geochemical cycles in these regions.

In an effort to understand the factors governing phytoplankton growth and photosynthetic activity in these high-nitrate low chlorophyll (HNLC) regions, a number of large-scale (mesocosm) iron addition experiments were performed in the late 1990's and early 2000's (Boyd et al., 2007). Proving that iron is in fact a limiting nutrient in these regions, these mesocosm iron enrichment experiments all led to distinct phytoplankton blooms accompanied by increases in primary production rates and photochemical efficiency (Boyd et al., 2000; Boyd et al., 2005; Coale et al., 1996; Takeda et al. 2005). The addition of iron to HNLC waters generally resulted in notable shifts from photosynthetic assemblages initially dominated by pico-phytoplankton such as the cyanobacteria *Synechococcus* to communities dominated by large-cell diatom micro-phytoplankton. While these experiments all demonstrated that addition of iron to HNLC waters selects for diatom species, the specific types of blooming diatoms differed in the different experiments. The addition of iron in the IronEx II experiment (Equatorial Pacific) resulted in a bloom of pennate diatoms (Coale et al., 1996), as well as in SOIREE (Southern Ocean) where the dominant blooming phytoplankton was identified as the pennate diatom *Fragilariopsis kerguelensis* (Boyd et al., 2000) and in SOFEX (Southern Ocean) where the largest change in phytoplankton biomass abundance was observed for the pennate diatom *Pseudo-nitzschia* spp. (Coale et al., 2004). In contrast, the addition of iron to Subarctic Pacific waters led to the formation of a bloom dominated by the fast-growing centric diatom *Chaetoceros* in SEEDS (Takeda & Tsuda, 2005) and by *Chaetoceros, Thalassiosira* spp., and *Thalassiothrix longissima* in SERIES (Boyd et al., 2005). Given that diatoms constitute one of the most diverse phytoplankton groups (with an estimated 200,000 different species) (Armbrust, 2009) with likely diverse lifestyles and function, identifying the

specific types and functional gene content of diatoms likely to bloom during natural iron inputs (aerial dust deposition or upwelling) in specific HNLC regions is key.

In addition to selecting for large cell diatom species, iron enrichment of HNLC waters has been shown to affect the various groups of the microbial food web. Using flow-cytometric techniques researchers demonstrated that in the IronEx II experiment (Equatorial Pacific Ocean), while pennate diatom cell abundances increased 85-fold, cell abundances of picophytoplankton species such as *Synechococcus* increased modestly (1.5 fold), and in the case of *Prochlorococcus* even decreased (Cavender-Bares et al. 1999). Furthermore, despite the decrease in *Prochlorococcus* abundance, *Prochlorococcus* division rate increased (Mann & Chisholm, 2000). Together with the observed increase in microzooplankton abundance (ciliates and flagellates <200μm) these findings suggested that the pico-phytoplankton were also stimulated by iron addition but were more intensely grazed upon than the large-cell diatoms, which could evade grazing pressure thanks to their hard silicified frustules (Coale et al., 1996). Using a combination of flow-cytometry, microscopy and pigment analyses of different size fractions, similar changes in food-web dynamics were observed in subsequent mesocosm iron enrichment experiments where the addition of iron stimulated (to various extents) most photosynthetic groups accompanied by increases in abundance (and hence grazing pressure) of both mesozooplankton (< 200 μm) and microzooplankton (>200 μm) (Boyd et al., 2000; Gervais et al. 2002; Takeda & Tsuda, 2005).

The bacterioplankton community was also affected by the addition of iron to HNLC waters. Bacterial abundance (enumerated by epifluorescence microscopy), bacterial production (measured by ^{3}H-leucine incorporation) and bacterial growth rate all increased during IronEx II (Equatorial Pacific) (Cochlan, 2001). In the SOIREE (Southern Ocean) mesocosm iron enrichment experiment, while bacterial abundances did not increase over the course of the experiment, bacterial production rates tripled, suggesting that in this environment the grazing pressure is not only exerted on pico-phytoplankton but also on the bacterioplankton (Boyd et al., 2000). Increases in bacterial production, while revealing that enhanced iron availability may also impact some members of the bacterial community do not, however, provide information on the identity (and hence potential function) of the specific stimulated bacteria. A recent

investigation of iron-induced changes in bacterial community composition (via Denaturing Gradient Gel Electrophoresis profiling) in the Equatorial Pacific Ocean identified four distinct response groups, suggesting that a higher resolution of iron-induced changes within the bacterial community is necessary (Eldridge et al. 2007).

While these experiments have significantly furthered our understanding of the impact of iron availability on the ecology of microbial communities in HNLC regions, recent advances in molecular approaches could enable us to complement this understanding by more specifically characterizing the stimulated microorganisms both in terms of identity and metabolic potential. Indeed, studies employing metagenomic approaches have been successful at not only revealing the astounding diversity of marine microorganisms (Rusch et al., 2007), but also at describing functional gene signatures specific to various environments (Tringe et al., 2005) and across physico-chemical gradients in the ocean (DeLong et al., 2006; Hewson et al. 2009). Similarly, metagenomic approaches used to study eukaryotic phytoplankton, while still in their infancy, have revealed how as-of-yet uncultured phytoplankton species may adapt to low-nutrient conditions in the ocean (Cuvelier et al., 2010). A recent study using metagenomic and metatranscriptomic approaches characterized changes in diatom-specific gene expression in response to addition of iron in a microcosm experiment in the northeastern Pacific Ocean (Marchetti et al., 2012). Authors showed that the addition of iron stimulated a bloom dominated by *Pseudo-nitzschia* and *Fragilariopsis* and led to the notable over-representation of diatom transcripts involved in the synthesis of photosynthetic proteins and chlorophyll (Marchetti et al., 2012). Here, using a metagenomic approach, we explore the response of the whole microbial community (including pico-eukaryotes and bacteria) to iron addition in the Central Equatorial Pacific Ocean.

Along the Equator, the Equatorial Pacific Ocean is characterized by two distinct provinces with different hydrological features and biological characteristics. Extending from the American continent to the ~140°W meridian is the Eastern Equatorial Pacific, the Equatorial Upwelling zone, characterized by a shallow thermocline (~40 m) and surface waters rich in macronutrients (Pennington et al., 2006). To the west of the 140°W meridian is the Western Pacific Warm Pool, where warm surface waters brought from

the East via the action of the trade winds accumulate and contribute to the deepening of the thermocline (~150 m) (Pennington et al., 2006). The accumulated waters in the western Equatorial Pacific raise the sea level (~30-50 cm higher relative to the Eastern Equatorial Pacific), leading to the formation of the eastward-flowing Equatorial Undercurrent, which shoals from west to east and is one of the drivers of the Equatorial Upwelling in the east (Pennington et al., 2006). In the Western Pacific Warm Pool, due to low surface macronutrient concentrations, biomass is scarce. The high surface macronutrient concentrations in the Eastern Equatorial Pacific should in theory support the growth of large-cell diatom and dinoflagellate species. Instead biomass in this region is dominated by small-celled picophytoplankton (Landry & Kirchman, 2002). Because iron inputs via dust deposition to the region are minimal (Duce et al., 1991; Gordon et al. 1997) the Eastern Equatorial Pacific is a typical HNLC region, and phytoplankton growth here is iron limited. Despite the hypothesis that iron could also be supplied from the shelf off the coast of Papua New Guinea to the Eastern Equatorial Pacific via the Equatorial Undercurrent (Mackey et al. 2002; Slemons et al., 2010), iron concentrations in surface waters of the Eastern Equatorial Pacific have been recorded to be as low as <0.1 nmol.kg^{-1} (Gordon et al., 1997). In the present study, we performed a microcosm (ship-board bottle incubation) iron enrichment experiment to waters obtained from the Central Equatorial Pacific (157.5° W, 0.21° N) a transition zone between the Eastern Equatorial Pacific and the Western Pacific Warm Pool, where conditions are similar to the iron-limited equatorial upwelling site, albeit slightly more oligotrophic than the typical HNLC region farther east.

Using cultivation-independent molecular approaches, we set out to explore how the microbial community of the Central Equatorial Pacific responds to enhanced iron availability. We specifically ask: **How does addition of iron affect microbial community structure- both in terms of diversity and functional gene content?** We look at the microbial community as a whole, exploring not only how the phytoplankton community is affected by iron addition, but also how the bacterial community changes.

3.3. Materials and Methods

3.3.1 Experimental Set-up and Sample Collection

Water from 50 m depth was collected from the Equatorial Pacific (157.5°W, 0.21°N) aboard the R/V Kilo Moana during the 2011 MetZyme cruise (KM1128). Water was collected on October 15, 2011 (at 21:30 local time) from 10 trace-metal clean X-Niskin bottles on a trace metal rosette suspended on a non-metallic line. The Niskin bottles were drained within a trace-metal clean positive pressure chamber directly into four 20-liter acid-washed carboys. Two of the carboys served as a control (control samples), while 1 nM $FeCl_3$ was added to the other two carboys (+Fe samples). Carboys were then incubated for four days in an on-deck flow-through incubator simulating light level at ~25 m depth (~20% of incident light). At the end of the incubation period, all 20 liters were filtered onto 0.2 μm polyethersulfone (PES) filters (Pall, Port Washington, NY). Filters from one set of duplicate incubations were stored in 10 mL sterile storage buffer (40 mM EDTA, 0.75 M sucrose, 50 mM Tris pH 8.3) and stored at -80°C until further processing (DNA extraction). Filters from the other set of duplicates were directly stored at -80°C until further processing (imaging). Acid washing of the incubation carboys was performed in a trace-metal clean lab following protocols established by the Saito lab. Briefly, carboys were soaked for a few days in ~0.1% citranox (Alconox White Plains, NY, USA), followed by a 2-week soak in 1 N Baker Instra-analyzed HCl (Mallinckrodt Baker, Phillispburg, PA, USA) –flipping the carboys to the other side after 1 week, then rinsed first with pH 2 water and then 5 times with MilliQ H_2O. See Figure 3.1 for experimental set-up.

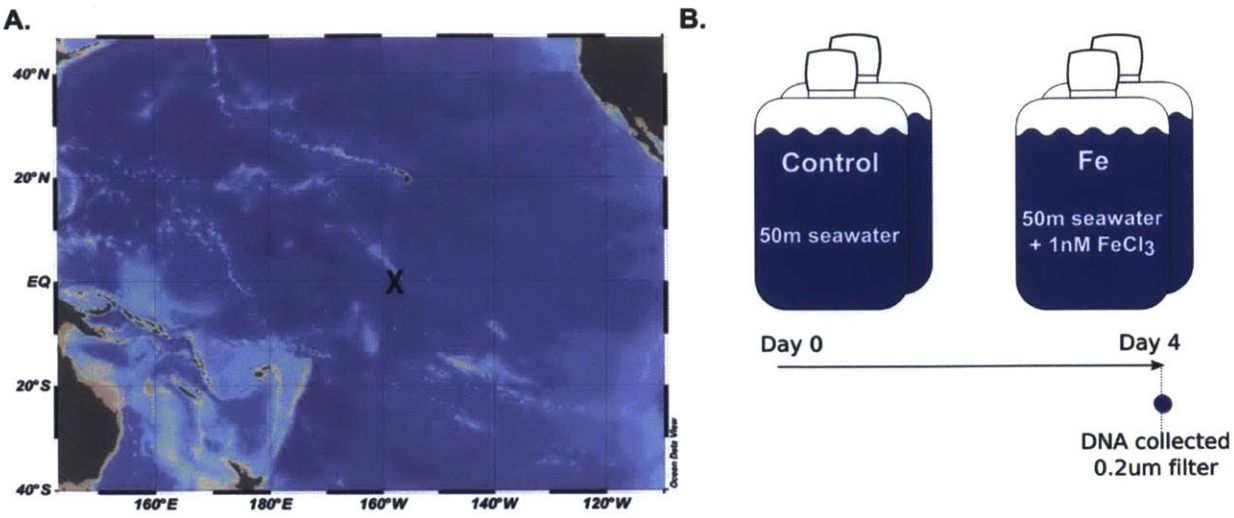


Figure 3.1. Experimental Set-up. A. Site in the Central Equatorial Pacific (indicated by an X) where 50 m seawater used for the microcosm experiment was collected. B. Seawater was incubated in four different incubation bottles, two unamended incubation bottles (Control), and two incubation bottles amended with 1 nM $FeCl_3$ (Fe). Incubation bottles were kept in an on-deck flow-through incubator simulating a light level corresponding to ~20% of incident light. Cells from two incubation bottles (1 control and 1 Fe) were collected onto a 0.2 μm filter for DNA extraction and sequencing after 4 days of incubation.

3.3.2 DNA Extraction

DNA extraction was performed as previously described (DeLong et al., 2006). The PES filters were thawed on ice, removed from the storage buffer and cut in half with a clean razor blade. Cells on the filter were lysed by incubating the half filters in the original 10 mL storage buffer (40 mM EDTA, 0.75 M sucrose, 50 mM Tris pH 8.3) along with an additional 3 mL storage buffer with an aqueous solution of 5 mg/mL lysozyme (Sigma-Aldrich, St. Louis, MO, USA) at 37°C for 45 min. Proteins were then degraded by adding 1 mL of 10 mg/mL proteinase K (EMD Chemicals, Millipore, Billerica, MA) and 1.5 mL of 10% SDS (Ambion, Applied Biosystems, Foster City, CA, USA) and incubating for 2 hours at 55°C. Proteins were separated from nucleic acids by extraction of the resulting lysate in 1 volume of phenol-chloroform-isoamyl alcohol, pH 8 (25:24:1) (Sigma-Aldrich). Samples were spun for 10 min at 1,000 rpm. The top aqueous phase was retained and extracted once more with phenol-chloroform-isoamyl alcohol. After centrifugation, the top aqueous phase was retained and gently mixed with 1 volume of chloroform: isoamyl alcohol (24:1) (Sigma-Aldrich). Samples were centrifuged for 5 min at 1,000 rpm and the top aqueous phase was transferred to an Amicon Ultra 10K tube (Millipore, Billerica, MA) to remove salts. Samples were spun for 30 min at 5,000 x g in a fixed rotor centrifuge. Samples were then washed twice with TE buffer (10 mM Tris pH 8.0, 1 mM EDTA) (Ambion) and spun twice at 5,000 x g for 20 min.

3.3.3. Clone Library Construction

The 18S ribosomal RNA gene clone libraries were generated as previously described (Moon-van der Staay et al. 2001; Not et al. 2007; Romari et al. 2004; Shi et al. 2009). Nearly full length 18S rRNA gene was amplified from the extracted DNA (from both the Control and the Fe samples) by PCR using the primers (5'-ACCTGGTTGATCCTGCCAG-3') and (5'-TGATCCTTCYGCAGGTTCAC-3') complimentary to regions near the termini of the 18S rRNA gene. The 25 µL PCR reactions each contained 10 ng of template DNA, 1.5 mM MgSO4, 500 nM of each primer, 800 µM of dNTPs (200 µM each), 2.5 units of Taq DNA polymerase (Invitrogen, Carlsbad, CA, USA), 1X of Invitrogen PCR buffer supplied with the Taq polymerase (producing products with an A overhang). The polymerase chain reaction was run using the following parameters: denaturation at 94°C for 1 min (initial denaturation at 94°C for 5

min), annealing at 55°C for 2 min, extension at 72°C for 3 min, followed by a final extension at 72°C for 10 min. The reaction was run over 35 thermal cycles. PCR products were visualized on a 0.8% agarose gel to confirm the presence of the 1.8kb DNA fragment, and purified following the manufacturer's instructions for the Qiaquick PCR purification Kit (Qiagen, Germantown, MD, USA). No amplified fragment of the expected size (1.8kb) was obtained from the Control DNA, suggesting that there were too few eukaryotes in the Control bottle.

The purified PCR product was cloned into TOPO® TA Cloning Kit (Invitrogen) per the manufacturer's instructions. Briefly, the cloning reaction (2 μL PCR purified product, 1 μL thawed salt solution from the kit, 2 μL water, 1 μL pCR™ 2.1-TOPO® vector from the kit) was incubated for 5 min at room temperature, then diluted 4-fold in water. Chemical transformation was performed by mixing 8 μL of the diluted cloning reaction with 50 μL of TOP10 chemically competent *E.coli* cells and incubating on ice for 5-30 min (here: 18min). After heat shocking the cells (42°C for 30 seconds), 250 μL of room temperature SOC media was added and cells were incubated for 1 hour at 37°C and 200 rpm to allow for expression of antibiotics resistance. About 250 μL of cells were spread on Luria-Bertani (LB) agar plates containing 50 μg/mL of kanamycin and 40 μg/mL of X-gal (5-bromo-4-chloro-3-indolyl- beta-D-galactopyranoside). After overnight incubation at 37°C, white colonies (with an insert) were picked and grown overnight at 37°C in liquid LB supplemented with 50 μg/mL kanamycin prior to DNA extraction and sequencing.

A total of 96 colonies were picked and sequenced using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) with the M13 forward (5'-GTA AAA CGA CGG CCA G-3') and M13 reverse (5'-CAG GAA ACA GCT ATG AC-3') primers. Sequencing was performed in-house on an ABI PRISM® 3700 DNA Analyzer (Applied Biosystems). Sequences were informatically and manually curated. Contaminating vector sequence as well as low quality sequence were trimmed (from both sides with the parameters: the first 25 bases contain less than 3 bases with confidences below 25%), using the Sequencher 4.10 software (Gene Codes, Ann Arbor, MI, USA). These are listed in Table S3.1, section 3.9.1.

3.3.4. Phylogenetic Assignment and Tree Construction

Because sequences of suitable quality did not exceed ~800bp in length, we chose to build a tree only for the sequences mapping to the 3'-end of the 18S rRNA gene sequence. Sequences were aligned to reference SSU (small subunit) sequences from the SILVA database of 16S and 18S rRNA gene sequences (release 111) using the online tool SINA aligner (http://www.arb-silva.de/aligner/) (Pruesse et al., 2007). Assignment to close relatives was performed using the ARB software and a set of already aligned reference 18S rRNA gene sequences from the SILVA Ref SSU (release 111) database (Ludwig et al., 2004). Phylogenetic trees of selected close relatives and cloned sequences were constructed in ARB using the neighbor-joining method and the Phylip distance matrix. We specified that the treeing algorithm only take into account alignment positions with a base call for all of the sequences considered. Hence, the phylogenetic tree constructed for the sequences mapping to the 3'-end sequences was based on a total of 644 aligned bases. The phylogenetic tree was visualized using the interactive Tree Of Life (iTOL) program (Letunic & Bork, 2007).

3.3.5. High-Throughput Sequencing

The extracted genomic DNA was sequenced on a GS FLX pyrosequencing platform (454 Life Sciences, Roche, Branford, CT, USA) using Titanium technology. Library preparation and sequencing was performed per the manufacturer's recommendation. Quality control was performed following the manufacturer's recommendation. The two DNA samples were each sequenced on a half plate yielding 244,697,589 (control sample) and 195,262,718 (+Fe sample) sequenced bases. Refer to Table 3.1 for a summary of read statistics.

3.3.6. Bioinformatic Analysis

Duplicate reads (defined as: 99% sequence identity over 100% length +/- 1 bp of the longest read) were identified by clustering reads using the CD-HIT software (Li & Godzik, 2006). Reads coding for ribosomal RNA (rRNA) were identified by mapping (BLASTn) non-duplicate reads to a database of reference 18S, 16S (Silva RefSSU database release 106) and 5S ribosomal RNA gene sequences (Pruesse et al., 2007). Reads for which the BLASTn alignment yielded hits with a bit-score greater than 50 were

considered significant. To determine the rough representation of the cloned 18S rRNA gene sequences in the metagenomic dataset, 454 reads that were identified as ribosomal RNA gene sequences were compared against a blastable database of the cloned 18S rRNA gene sequences using megablast (Altschul et al. 1990).

Protein-coding genes were identified comparing non-duplicate, non-rRNA sequence reads to the NCBI non-redundant (nr) protein database (downloaded on 04/27/2012) using BLASTx (Altschul et al. 1990). Hits for which the BLAST alignment was characterized by a bit-score of 50 or greater were considered significant. Sequencing and annotation statistics for each sample are summarized in Table 3.1. Community composition was determined by retrieving the species name associated to each significant (bit score > 50) BLASTx hit to the NCBI nr-protein database. The taxonomy of the retrieved species name was then established using the NCBI taxonomy names and hierarchy files (downloaded on 05/22/2012). When a read had two or more equal top hits (as determined by the highest bit score), counts were assigned to the lowest common ancestor. Based on the NCBI taxonomy, reads were then binned at the kingdom or genus-level.

Non-duplicate, non-rRNA reads were compared against the KEGG (Kyoto Encyclopedia of Genes and Genome) ontology of peptides (Kanehisa & Goto, 2000, downloaded on 10/09/2012) using BLASTx (Altschul et al., 1990). Hits with a bitscore > 50 were considered significant. Significant KEGG hits of reads belonging to the Bacteroidetes phylum (identified as reads with a significant BLASTx hit against the NCBI nr-protein database for which the associated species name belongs to the Bacteroidetes) were used for the functional profile comparison of the Bacteroidetes population in the control versus the iron sample.

3.3.7. Imaging

After thawing the PES filters that had been stored at -80°C without storage buffer, cells were gently resuspended in 7.5 mL (control) and 12.5 mL (+Fe) filtered seawater. Because of variations in cell densities, less volume was added to the control sample to optimize microscope view in one field. Both samples were observed in a Palmer Maloney counting chamber, which holds exactly 0.1 mL of volume

sample. Samples were viewed on a Zeiss Axioplan 2 Imaging microscope at 400X magnification and images were captured with Zeiss Axiovision 4.8 software.

3.4. Results

3.4.1. Read Statistics

Pyrosequencing of community DNA for the control and the +Fe samples yielded a total of 665,956 and 650,429 sequence reads respectively, with average read length of 367 bp and 300 bp (Table 3.1). Exact duplicate reads, a common artifact of 454 sequencing technology (Gomez-Alvarez et al. 2009) were removed from further analysis. Roughly 0.35% of the non-duplicate reads matched ribosomal RNA sequences. Of the non-rRNA encoding reads, 65% had a significant hit in the NCBI non-redundant (nr) protein database for the control sample and 51% for the +Fe sample. The large difference in the number of significant hits to the nr-protein database is likely due to the dearth of eukaryotic microbe sequences in the database, as these seem to be enriched in the +Fe sample (see Figure 3.4).

Table 3.1: Read statistics Parentheses indicate the percent of total sequence reads unless otherwise noted.

	Control	+Fe
Total number of reads	665,956	650,429
Mean length	367bp	300bp
Mean GC%	38	41
Duplicate reads[a]	2,012 (0.3)	1,329 (0.2)
rRNA reads	2,303 (0.35)	2,269 (0.35)
Mean length	415bp	343bp
Mean GC%	47	47
Reads assigned to nr-protein[b]	435,937 (65)[d]	330,147 (51)[d]
Reads assigned to KEGG ko[c]	322,315(49)[d]	214,994(33)[d]

[a] duplicates: 99% id percent over 100% percent length +/- 1bp of the longest read.

[b] number of unique reads with a significant BLASTx hit to the NCBI nr-protein database (bitscore > 50)- database downloaded on 04/27/2012

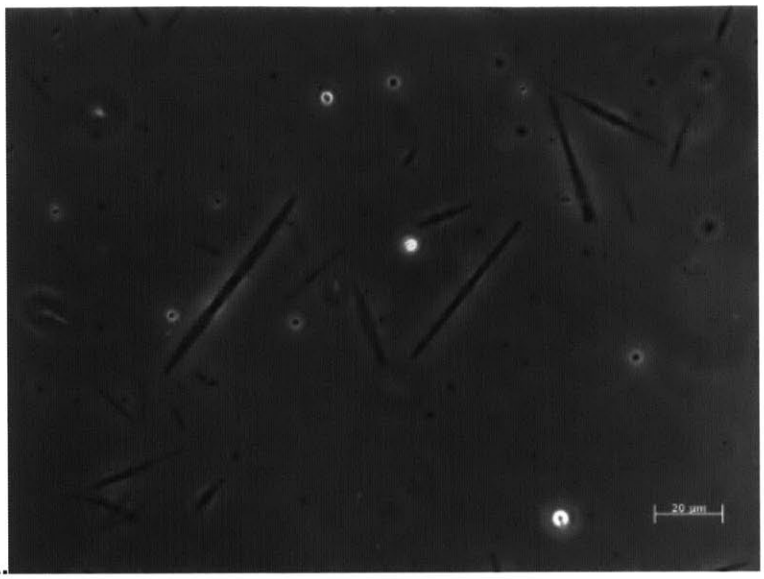
[c] number of unique reads with a significant BLASTx hit to the KEGG ontology (bitscore > 50)- database downloaded on 10/09/2012

[d] Indicated percentages are relative to the total number of non-duplicate, non-rRNA reads.

3.4.2. Diatoms Bloom In Response to Iron Addition

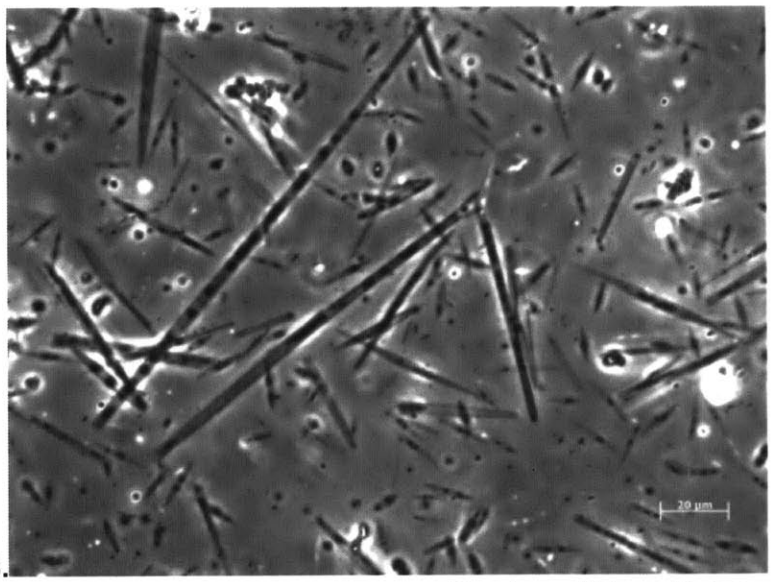
The addition of iron led to a distinct bloom of algal cells. This bloom was visible to the naked eye when comparing filters from the +Fe sample to those obtained from the control sample (filters from the Fe sample had a distinct brownish pigmentation). This observation was confirmed by light microscopy, enabling the visualization of cells resuspended from these filters (Figure 3.2). Cells with the typical pennate diatom morphology could be observed in both the control and the +Fe samples. Interestingly, these cells were a lot more abundant and larger in the +Fe sample. A rough quantification –only taking into account cells with the typical pennate diatom morphology that were greater than 20 μm in length– revealed that these pennate diatoms were ~6 fold more abundant in the +Fe sample than in the control sample (Figure 3.2). Together, these observations suggest that the iron amendment induced a notable shift in the microbial community structure. To further characterize this shift, genomic DNA extracted from cells (>0.2 μm) retrieved in the control and the Fe incubation bottles at the end four days of incubation was pyrosequenced. The average percent GC content of pyrosequenced reads shifted towards higher values in the +Fe sample, further suggesting that addition of iron led to a notable community shift (Figure 3.3).

To obtain information on the microbial community composition of each sample, metagenomic reads were assigned taxonomically based on the species name associated with each significant top BLASTx hit to the NCBI-nr protein database (see Methods, Section 3.3.6). The microbial community in both the control and the +Fe samples was dominated by bacterial reads that respectively accounted for 64% and 47% of all non-rRNA non duplicate reads (88% and 85% of reads with a significant NR hit) respectively (Figure 3.4). Reads assigned to the Eukaryota accounted for ~1% of non-rRNA non-duplicate reads in the control sample and almost ~4% in the iron sample. This ~3.5 fold increase in relative abundance of Eukaryota reads confirms the visual observations that addition of iron led to an enrichment of eukaryotic algae (Figure 3.4).



A.

Courtesy of Dawn Moran



B.

Courtesy of Dawn Moran

Figure 3.2: Visual assessment of phytoplankton bloom. Control (panel A) and +Fe (panel B) samples visualized on a Zeiss Axioplan 2 Imaging microscope at 400X magnification. Cells were rinsed off of the 0.2 µm PES filters with 7.5 mL (control) and 12.5 mL of filtered seawater. Images represent the field of view for 0.1 mL of sample volume. Pennate diatoms are the dominant type of cell visualized in both samples. They are a lot more abundant, larger and form longer chains in the +Fe (panel B) sample. Scale bars at the bottom right represent 20 µm.

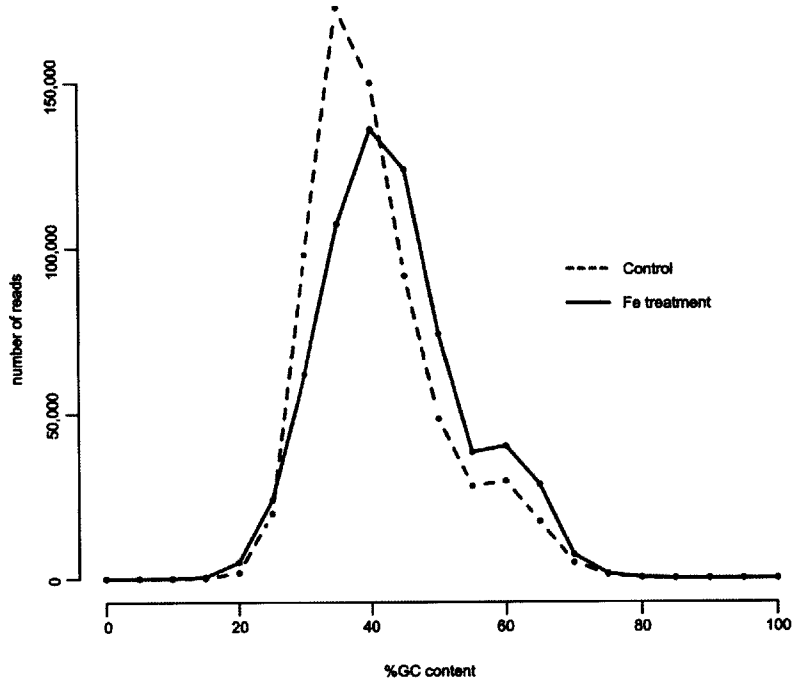


Figure 3.3. Shift in average percent GC content between the Fe and the control samples. Distribution of reads binned by their average %GC content. The frequency of reads per average %GC content of each read is plotted. Average percent GC was determined for every read post removal of duplicate reads. We observe a shift in the distribution towards higher average %GC content in the Fe sample.

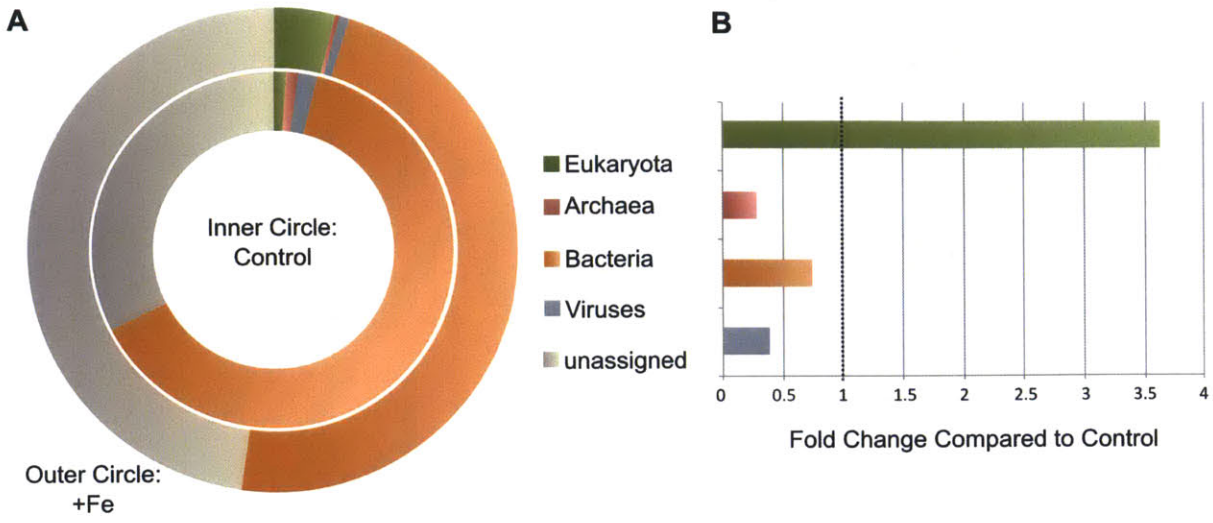


Figure 3.4. Iron addition leads to an enrichment in eukaryotes. A). Relative abundance of reads with significant hits (by BLASTx) to NCBI non-redundant protein database broken down by kingdom. Read counts are normalized to the total number of non-duplicate non-rRNA reads. Reads successfully assigned to one of the 4 superkingdoms account for 68% (control) and 52% (Fe) of non-rRNA, non-duplicate pyrosequence reads. Unassigned reads refer to those reads with no significant BLASTx hit in the NCBI-nr protein database. B). Fold change in relative abundance of each kingdom between the iron treatment and the control. Any value greater than 1 indicates enrichment in the iron treatment.

To gain a more detailed understanding of these community shifts, we compared the relative abundance of metagenomic sequence reads in the iron sample with the relative abundance of metagenomic sequence reads in the control sample, at the genus level (Figure 3.5). For better visualization, we removed any genus having 20 or less sequence reads for both the control and the iron samples and represented the log fold change in relative abundance of genus-specific sequence reads in the +Fe sample to the control sample (Figure 3.6). The genera with the largest log fold-change indeed belonged to the Eukaryota, some of which were enriched between 10 and 100 fold in the iron sample compared to the control sample. Most of these genera (Cylindrotheca, Fistulifera, Phaeodactylum, Thalassiosira, Odontella and Synedra) belonged to the Bacillariophyta (diatom) phylum (Figure 3.6). This enrichment in diatom specific reads is in agreement with the results of other iron amendment experiments performed in high-nitrate low chlorophyll regions (HNLC), as most of these have resulted in diatom blooms (Boyd et al., 2007). However, the determination of the specific identity of the blooming diatoms is difficult using metagenomic data alone, since this is database dependent and only a limited number of diatom genomes have been deposited in publicly available databases (as of 04/27/2012, *Phaeodactylum tricornutum* CCAP 1055/1 and *Thalassiosira pseudonana* CCMP 1335 were the only diatom whole genomes in the NCBI database). For the same reason, due to database limitations it is difficult to infer whether other types of microbial eukaryotes were also enriched in the Fe sample. For example, as of 04/27/2012, the only publicly available dinoflagellate genomes were plastid genomes of *Durinskia* and *Kryptoperidinium*. Because plastids in both *Durinskia* and *Kryptoperidinium* evolved from diatom endosymbionts (Imanian et al. 2010), it is likely that the reads with significant hits to these genomes could in fact come from natural diatom populations that may be related to those endosymbionts rather than from actual dinoflagellates.

To address this problem, we constructed a clone library of 18S ribosomal RNA gene sequences from both the control and the iron samples. We PCR amplified an 1.8kb conserved region of the 18S ribosomal RNA gene (Moon-van der Staay et al., 2001). PCR amplification of the 18S rRNA gene did not yield any product from the DNA obtained from the control sample. We picked and sequenced 96 clones, yielding 27 high quality sequences with an average length of ~800 bp. The large number of low-quality sequences was the result of both poor sequencing and the selection of clones with short inserts. Because

the cloning vector used allowed for non-directional cloning and because the sequences obtained had an average length of ~800bp (precluding the assembly of full-length sequences), we constructed a phylogenetic tree (Figure 3.7) based on the 3'-end of the gene. We found that all the cloned 18S rRNA sequences clustered with diatom-like 18S rRNA gene sequences (Figure 3.7). Based on our phylogenetic analysis, the majority (88.9% of high quality clone sequences) of the clone sequences were most closely related to *Pseudo-nitzschia* species (Figure 3.7). In addition to *Pseudo-nitzschia*, we found a few clone sequences closely related to two other diatom genera: *Thalassionema* (7.4% of high quality clone sequences) and *Cylindrotheca* (3.7% of high quality clone sequences). To better assess the abundance of these groups in the overall microbial community, we recruited pyrosequenced DNA reads identified as ribosomoal RNA genes to the cloned 18S rRNA gene sequences (Figure 3.7). We found that 3.4% of the total rRNA gene sequence reads in the +Fe metagenome recruited to a cloned 18S rRNA gene sequence, and 0.56% of the total number of rRNA gene sequence reads in the control sample metagenome recruited to the same cloned 18S rRNA gene sequences. These numbers agree well with the increase in the relative abundance of Eukaryota protein-coding sequence reads in the +Fe metagenome compared to that of the control sample (Figure 3.4). As *Thalassionema, Cylindrotheca* and *Pseudo-nitzschia* are all pennate diatoms, these molecular data reflect well the visual observation that pennate diatoms were a lot more abundant in the +Fe sample than in the control sample (Figure 3.2). Together, these findings suggest that the bloom in eukaryotic algae is largely driven by pennate diatoms, more specifically by *Pseudo-nitzschia*-like species, and to a lesser extent by *Thalassionema* and *Cylindrotheca*-like species.

Analysis of reads recruited to diatom genomes confirmed these observations. Indeed, the average percent identity of reads recruiting to the *Phaeodactylum tricornutum* CCAP 1055/1 and *Thalassiosira pseudonana* CCMP1335 genomes both in the control and in the Fe sample is rather low (Table 3.2), suggesting that the natural community was likely composed of diatom species related to but quite distinct from *T. pseudonana* and *P. tricornutum*. We reason that the natural diatom community harbors genes similar to the *T. pseudonana* genes, and other genes more similar to *P. tricornutum* genes, suggesting that the ambient diatom community may be composed of two distinct groups of diatoms. Interestingly, the ratio of the relative abundance of reads (as a percentage of total non-duplicate, non rRNA reads) recruiting to *P. tricornutum* CCAP 1055/1 to

the relative abundance of reads recruiting to *T. pseudonana* CCMP1355 is the same in both the control and the iron samples (Table 3.2). The fact that *T.pseudonana* to *P.tricornutum* ratios remained the same in both the control and the Fe sample implies that the diatom community may have been similar in terms of composition in both the control and the Fe samples, albeit present at very different abundances.

Roughly 60% (both in the control and the Fe samples) of reads recruiting to the Bacillariophyta phylum mapped to *P. tricornutum* CCAP1055/1 and 26% (both in the control and Fe samples) to *T. pseudonana* CCMP1355, suggesting that roughly ~14% of diatom genes in the natural population were more similar to as-of yet un-sequenced diatoms. This seems rather low considering the fact that genome coverage of both *T. pseudonana* CCMP1355 and *P. tricornutum* CCAP1055/1 genomes was low. Indeed, the 12,456 reads in the Fe sample recruiting to the *P. tricornutum* CCAP1055/1 genome mapped to a total 5,042 unique protein coding genes in the genome- that is roughly 50% of the protein coding genes (10,408) found in the *P. tricornutum* CCAP1055/1 genome. Similarly, the 5,329 reads in the Fe sample recruited to the *T. pseudonana* CCMP1335 genome mapped to 2,858 unique protein coding genes, which represents roughly 24% of the protein coding genes (11,849) found in the *T. pseudonana* genome (Figure 3.8). This low genomic coverage is in agreement with the low overall sequence similarity between metagenome sequence reads and each the two diatom genomes (Table 3.2). Together, these data underline how different the natural blooming diatom community is from these sequenced diatom species, and is likely to carry a number of genes specific for its open-ocean, low nutrient conditions lifestyle.

Table 3.2: Summary statistics of diatom-specific reads.

	Control	+Fe
Bacillariophyta - NR hits	4,072 (0.6)[a]	20,794 (3.2)[a]
Phaeodactylum tricornutum CCAP 1055/1	2,444 (0.36)[a]	12,456 (1.9)[a]
Thalassiosira pseudonana CCMP 1335	1,075 (0.16)[a]	5,329(0.82)[a]
Phaeodactylum : Thalassiosira ratio	2.25	2.3
Average % id of Thalassiosira reads	54.3	56.6
Average % id of Phaeodactylum reads[c]	55.7	58

a. Values in parentheses represent number of reads as a percentage of total non-duplicate non-rRNA reads (646,831 in Fe/ 661,641 in Control).
b. Average % identity of significant BLASTx hits recruiting to *Thalassiosira pseudonana* CCMP 1335 genome.
c. Average % identity of significant BLASTx hits recruiting to *Phaeodactylum tricornutum* CCAP 1055/1 genome.

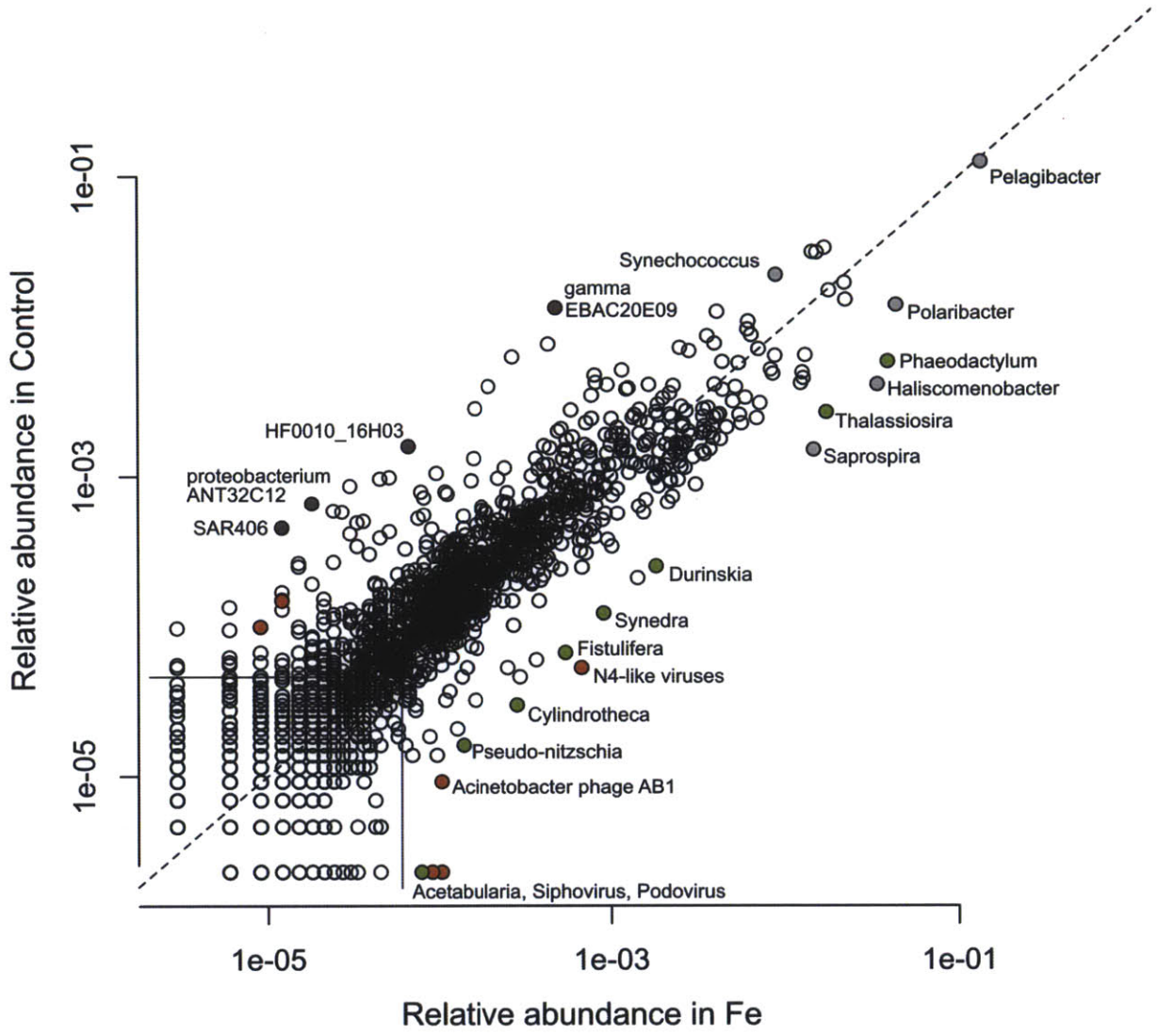


Figure 3.5: Relative abundance of genus-specific sequence reads. Reads with a significant hit (bit score > 50) to the NCBI-nr protein database were binned by genus. For a specific genus, the relative abundance of reads in the iron sample (x-axis) is plotted against the relative abundance of reads binned to the same genus in the control sample (y-axis). Selected genera are color coded to indicate which kingdom they belong to (refer to Figure 3.6 for legend). To compute relative abundance of each genus, genus-specific sequence read counts were normalized to the total number of metagenomic sequence reads assigned at the genus level (435,937 for the control sample, 330,147 for the +Fe sample). Both axes are in log-scale. The dotted line represents the x=y line. The continuous lines represent a read count of 20 normalized to the total genus-specific read counts for each treatment. This shows that by retaining genera with at least 20 sequence read counts in either treatment, noise stemming from low-count genera is removed.

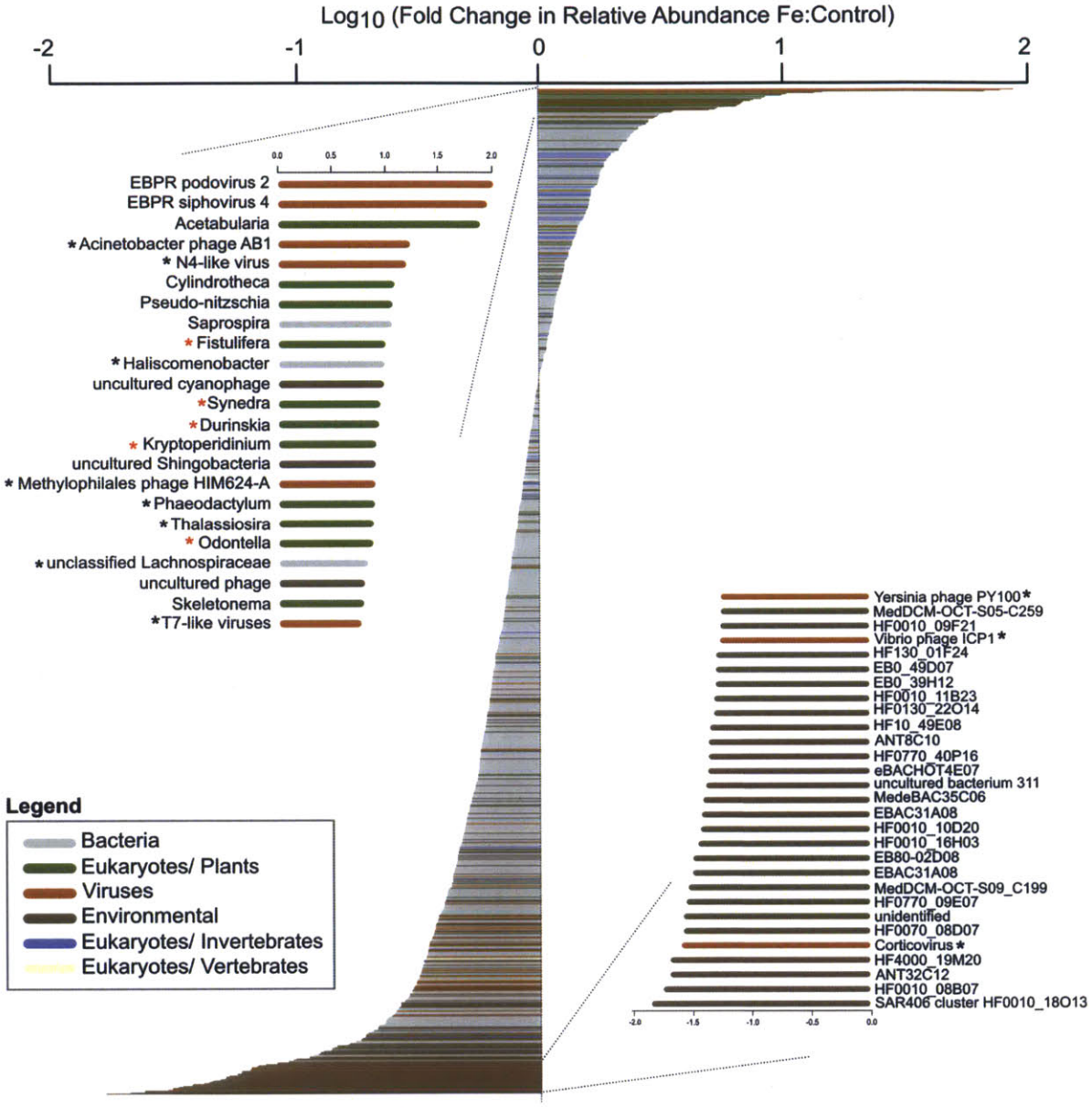


Figure 3.6. Genus-level community changes. Reads with a significant hit (bitscore >50) to the NCBI nr-protein database were binned by genus. We plot the log of the ratio of the relative abundance of reads in the iron treatment to the relative abundance of reads in the control sample. Only genera with at least 20 read counts in either the control or the iron treatment are represented (see Figure 3.5). The red asterisks indicate that the genome in the database is that of an organelle (mitochondria or chloroplast). The black asterisks indicate the genera for which a fully sequenced genome is available in the NCBI database. To compute the relative abundance, counts summed for a specific genus are normalized to the total number of reads assigned at the genus level (435,937 for the control sample, 330,147 for the +Fe sample).

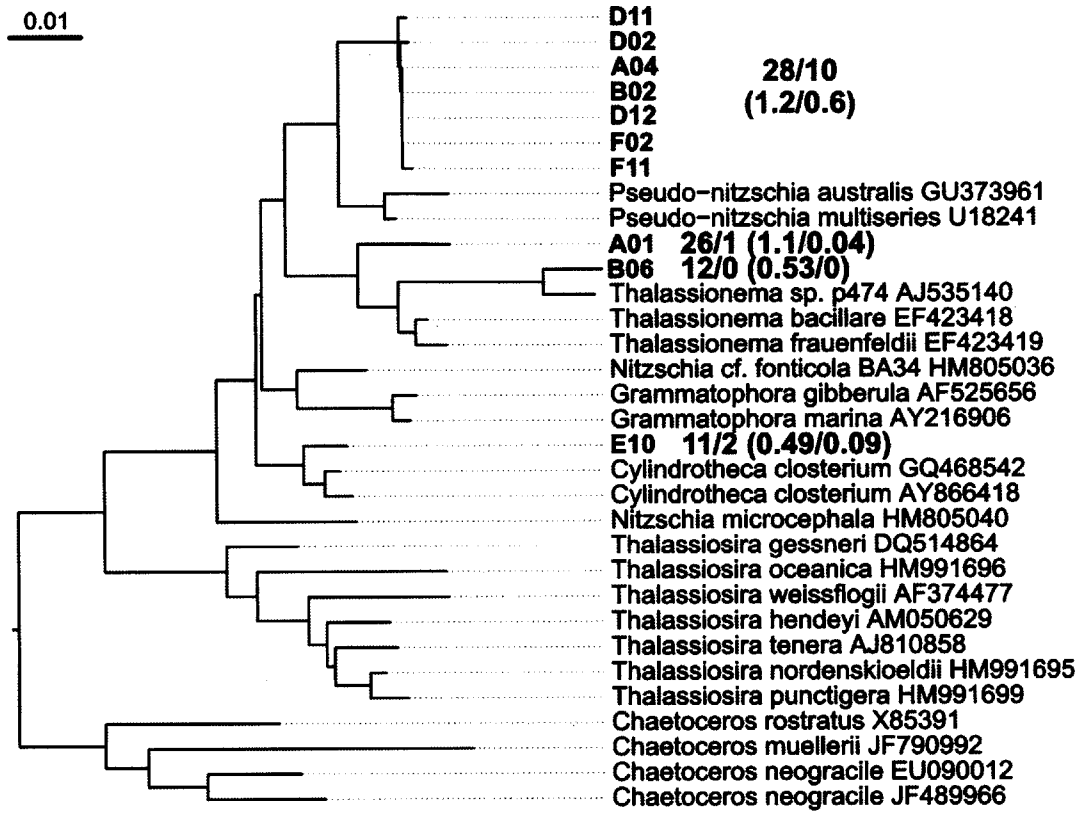


Figure 3.7. Phylogenetic representation of the cloned 18S rRNA gene sequences and selected close relatives. Neighbor-joining tree (Phylip distance matrix) constructed in ARB (Ludwig et al., 2004) for sequences mapping to the 3'-end of the 18S rRNA gene using a multi-sequence alignment of 644 bases. The names of clone sequences obtained from the +Fe treatment are bolded. The majority of clone sequences (89%) mapped closely to *Pseudo-nitzschia australis* and *Pseudo-nitzschia multiseries*. Next to the name of each clone, we indicate in bold the number of 454 reads identified as ribosomal RNAs in the Fe and the control samples that recruit to the cloned 18S rRNA sequence (Fe/Control). The value in parentheses is the same number as a percentage of the total number of 454 reads identified as ribosomal RNA. In the iron sample, a total of 77 reads (3.4% of ribosomal RNA reads) mapped to any of the 18S rRNA sequences obtained by cloning. In the control sample, 13 reads (0.56% or ribosomal RNA reads) mapped to the clone sequences.

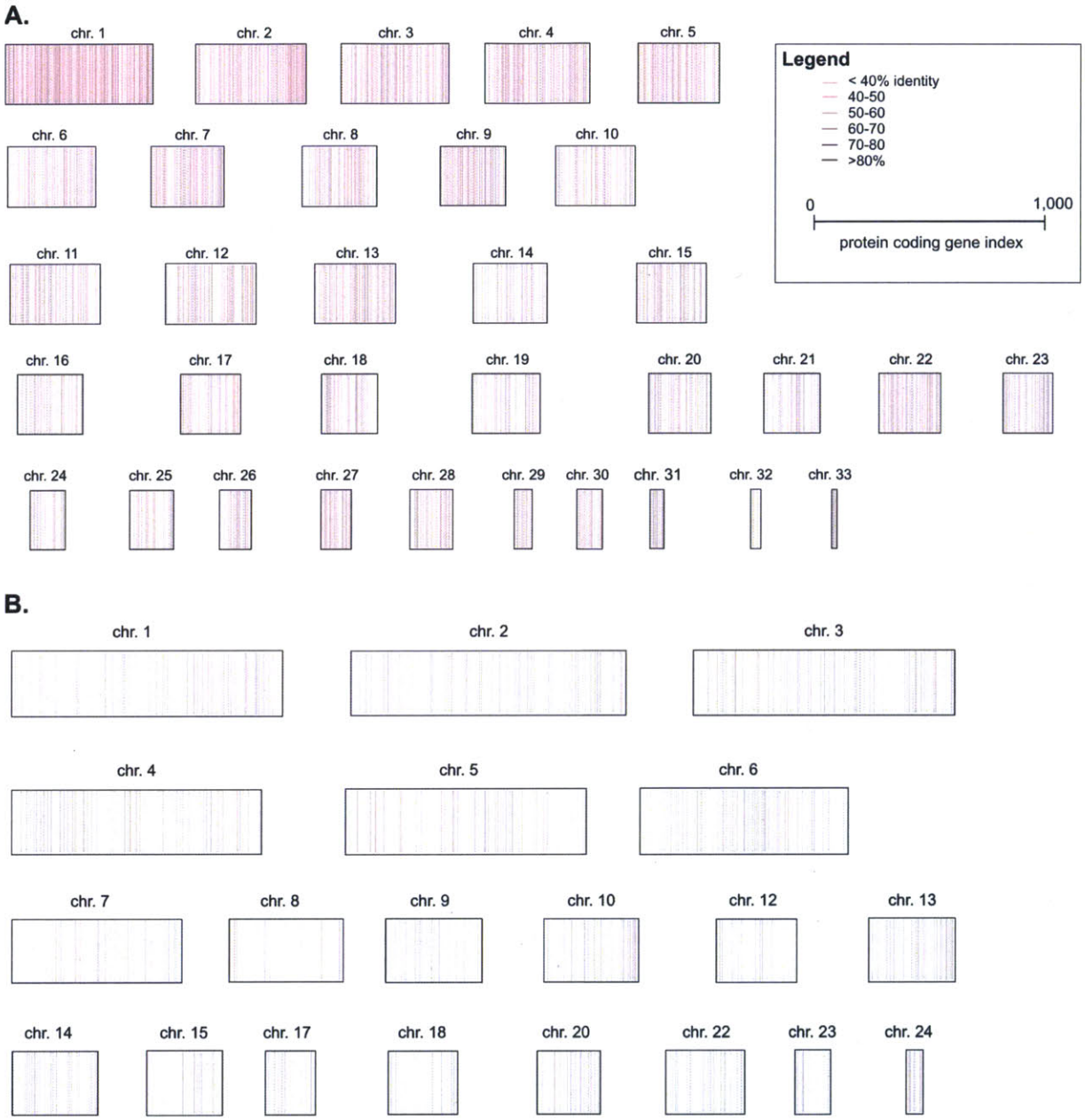


Figure 3.8. Coverage of metagenomic reads in the Fe sample of *Phaeodactylum tricornutum* and *Thalassiosira pseudonana* genomes. The average percent identity of pyrosequence reads in the Fe sample to hits in the two diatom genomes: *Phaeodactylum tricornutum* CCAP 1055/1 (panel A) and *Thalassiosira pseudonana* CCMP 1335 (panel B) is represented (color-coded) along each chromosome . We represent only protein-coding genes as they are ordered along the chromosome (protein coding gene index). The index position of proteing-coding genes that did not recruit any metagenomic read are left blank. The *P. tricornutum* CCAP 1055/1 genome contains 10,408 protein coding genes, of which the metagenomic reads in the Fe sample hit 5,042. The *Thalassiosira pseudonana* CCMP 1335 genome contains 11,849 protein coding genes of which the metagenomic reads in the Fe sample hit 2,858. The low genome coverage of *T. pseudonana* (~25%) is evident from the much sparser (a lot lighter pink) distribution of hits. For the *T. pseudonana* genome only 20 chromosomes are represented, comprising 10,432 proteins (the others were not available on Genbank).

3.4.3. Iron Addition leads to a Shift in the Bacterial Community

The addition of iron to 50 m-seawater from the Central Equatorial Pacific also resulted in interesting shifts within the bacterial community. The relative abundance of bacterial reads was lower in the +Fe sample compared to the control sample. This could be explained by a decrease in total bacterial cells, or by the observed increase in eukaryotic reads. Despite this we noticed that certain bacterial genera (*Haliscomenobacter, Saprospira,* - to a lesser extent *Polaribacter*) were enriched in the +Fe sample compared to the control sample (Figure 3.5 and Figure 3.6). To investigate whether other bacterial genera were enriched in the +Fe sample, we represented the log fold-change of the relative abundance of bacterial genera in the +Fe sample to the Control sample (Figure 3.9). Interestingly, we found that most bacterial genera enriched in the +Fe sample compared to the control sample belonged to the Bacteroidetes, including: *Haliscomenobacter, Saprospira, Polaribacter, Niastella, Chitinophaga, Kordia* and *Microscilla*, among others. In fact, while the relative abundance of bacterial reads decreased from 64% in the control to 47 % in the +Fe sample (Figure 3.4), the relative abundance of Bacteroidetes reads increased from 10% in the control to 16% in the +Fe sample (as a percentage of total non-duplicate, non rRNA reads) (Table 3.3). Certain Proteobacteria genera such as *Jannaschia, Dinoroseobacter* (both from the Rhodobacteraceae) and *Glaciecola* (Alteromonads) were also found to be enriched in the Fe sample. On the other hand, bacterial reads that were under-represented in the +Fe sample compared to the control sample belonged to the Proteobacteria phylum (Myxococales, SAR86 cluster, a number of unclassified environmental Proteobacteria) as well as to other unclassified environmental bacterial sequences.

The addition of iron to waters of the Central Equatorial Pacific also had some impact on the viral community. While a number of cyanophages (*Synechococcus* phages S-SM2, S-SSM7, S-CAM8, S-CRM01, *Prochlorococcus* phage Syn1, P-SSM7, P-HM2 to cite a few) were enriched in the control sample relative to the +Fe sample, a number of other phages N4-like viruses, T7-like viruses, Methylophilales HIM624-A were enriched in the +Fe sample relative to the Control (Figure 3.6). Most of the viral sequences enriched in the +Fe sample come from bacteriophages (not eukaryotic viruses). It is possible that the types of phages enriched in the +Fe sample may infect the blooming Bacteroidetes community.

Closer inspection indicates that iron addition did not only lead to a community shift amongst bacteria but also amongst the Bacteroidetes. Indeed, there were notable differences in the percent sequence similarity representation of different Bacteroidetes genera between the control and the +Fe samples (Figure 3.9). In some cases (*Haliscomenobacter* and *Saprospira*), the average percent similarity differed by 3 and 4% between the control and the Fe sample (respectively). To further characterize this apparent shift, we compared the functional gene profile of the Bacteroidetes population in the control and in the +Fe samples (Figure 3.10). One hypothesis is that the Bacteroidetes population present in the control sample was better adapted to low iron or the low abundance of eukaryotes. Significant hits to the KEGG ko database from reads that had been identified as Bacteroidetes-specific (see Methods) were retrieved and used to determine changes in the functional profile of the Bacteroidetes. While only 55% (control) and 49% (+Fe) of Bacteroidetes-specific reads were assigned a KEGG ko number (Table 3.3), we nonetheless observe that two functionally distinct Bacteroidetes communities were found in the control sample versus the +Fe sample (Figure 3.10).

Table 3.3. Sequence reads recruiting to Bacteroidetes genomes.

	Control	+Fe
Bacteroidetes- NR hits	66,771 (10)*	105,114(16)*
Significant KEGG hit	64,051	99,286
Assigned ko number	37,081(55)	51,576(49)
Assigned pathway	23,262	31,290

*Value in parenthesis is the percentage of total non-rRNA non-duplicate reads. (661,641 in control; 646,831 in Fe)

A number of KEGG ko were specifically enriched in the Fe sample. These included serine endopeptidases (K08676- EC), proline aminopeptidase (K01259) which may be involved in degrading various exogenous peptides, certain glycosidases specific to starch and various polysaccharides (K01176, EC: 3.2.1.1), beta-D-glucosides (K05349, EC: 3.2.1.21), or on N-acetylglucosides and N-acetylgalactosides (K12373, EC: 3.2.1.52), various poorly characterized esterases (K06889, K07017, K06978), as well as genes (long-chain acyl-coA synthetase) involved in fatty acid metabolism (K01897: EC.6.2.1.3). This indicates that the Bacteroidetes population stimulated by the addition of iron was potentially well poised to degrade large

polysaccharide molecules. Because diatoms in particular are known to produce exo-polysaccharides in large quantitites, at all stages of growth (Passow, 2002), it is possible that the Bacteroidetes community stimulated by the addition of iron could have been associated with the blooming diatom community. Interestingly, one of the KEGG ko functional categories enriched in the Fe sample is a cytochrome c peroxidase likely responsible for the detoxification of hydrogen peroxide. Hydrogen peroxide is produced intra-cellularly as a by-product of respiration. Recent studies have also showed that diatom exudates are enriched in hydrogen peroxide (Kustka et al. 2005; Steigenberger & Statham, 2010). Together, these findings suggest that the Bacteroidetes population stimulated after the addition of iron could have been stimulated in response to the blooming diatom community.

Finally, we found the enrichment of DNA-processing types of genes (K07497: integrase and K03427 and K01153 both genes from the type I hsd restriction enzyme system) in the Fe sample intriguing. While not sufficient to prove the existence of a prophage in a bacterial genome, most phages carry integrase genes and their presence in a bacterial genome may indicate the presence of a lysogen (Casjens, 2003). As activation of a phage lysogen in response to relief of nutrient limitation in marine bacterial communities has been reported in the past (Shi et al. 2011), we propose that the Bacteroidetes population enriched in the Fe sample could potentially be enriched in phage lysogens. This is interesting to ponder, especially in light of recent studies showing that the abundance of prophages in genomes of Bacteroidetes varies between Bacteroidetes populations specialized for growth in different environments (Krupovic & Forterre, 2011; Reyes et al. 2012). While not providing any formal proof, our findings could be suggestive of the possibility that prophage abundance is different between free-living populations of marine Bacteroidetes and particle-associated ones.

Some of the KEGG ko functional genes enriched in the control sample also included peptidases (general dipeptidases: K01273: EC.3.4.13.19, K06013: EC 3.4.24.84) as well as one glycosidase (K01206, EC 3.2.1.51) specifically involved in the cleavage of fucose, a monosaccharide that is often implicated as a terminal modification of glycan structures (Becker & Lowe, 2003). Fucose has been identified as one of the key components of saccharides produced by brown algae (Nishide et al., 1990). Other KEGG ko

functional genes that were enriched in the control sample relative to the Fe sample are arylsulfatases, which catalyze the hydrolysis of sulfate esters, which may be useful in the degradation of degradation of sulfated sugars. In fact, one important structural carbohydrate of certain brown algae is fucoidan (a sulfated fucose) (Chizhov et al., 1999). Interestingly, some KEGG ko functional genes that may be involved in nitrogen transport and metabolism were also found to be enriched in the control sample: glutamine synthetase (K01915) and a glycine cation symporter (K03310). Other interesting KEGG ko functional genes that were enriched in the control sample relative to the Fe sample is an amidohydrolase (K12941, *abgB*) which cleaves p-aminobenzo glutamate into p-aminobenzoate and glutamate as well as ferrochelatase (K01772, *hemH*, EC: 4.99.1.1), which catalyzes the last step in heme biosynthesis. Finally, we notice that the most abundant KEGG ko in both control and Fe Bacteroidetes communities was the a *fhuA*-like gene (ferrichrome outer-membrane receptor), suggesting that transport of a ferrichrome-like siderophore is important for both natural populations of Bacteroidetes.

The two Bacteroidetes populations observed in the control and Fe samples both displayed the metabolic potential to break down sugars and peptides. Nevertheless, the analysis described above suggests that the Bacteroidetes population in the Fe sample could have potentially been able to degrade a broader array of sugars. While no single pathway-level enrichment was obvious from this analysis, we can nonetheless conclude that addition of iron to Central Equatorial Pacific waters selected for a functionally distinct Bacteroidetes community.

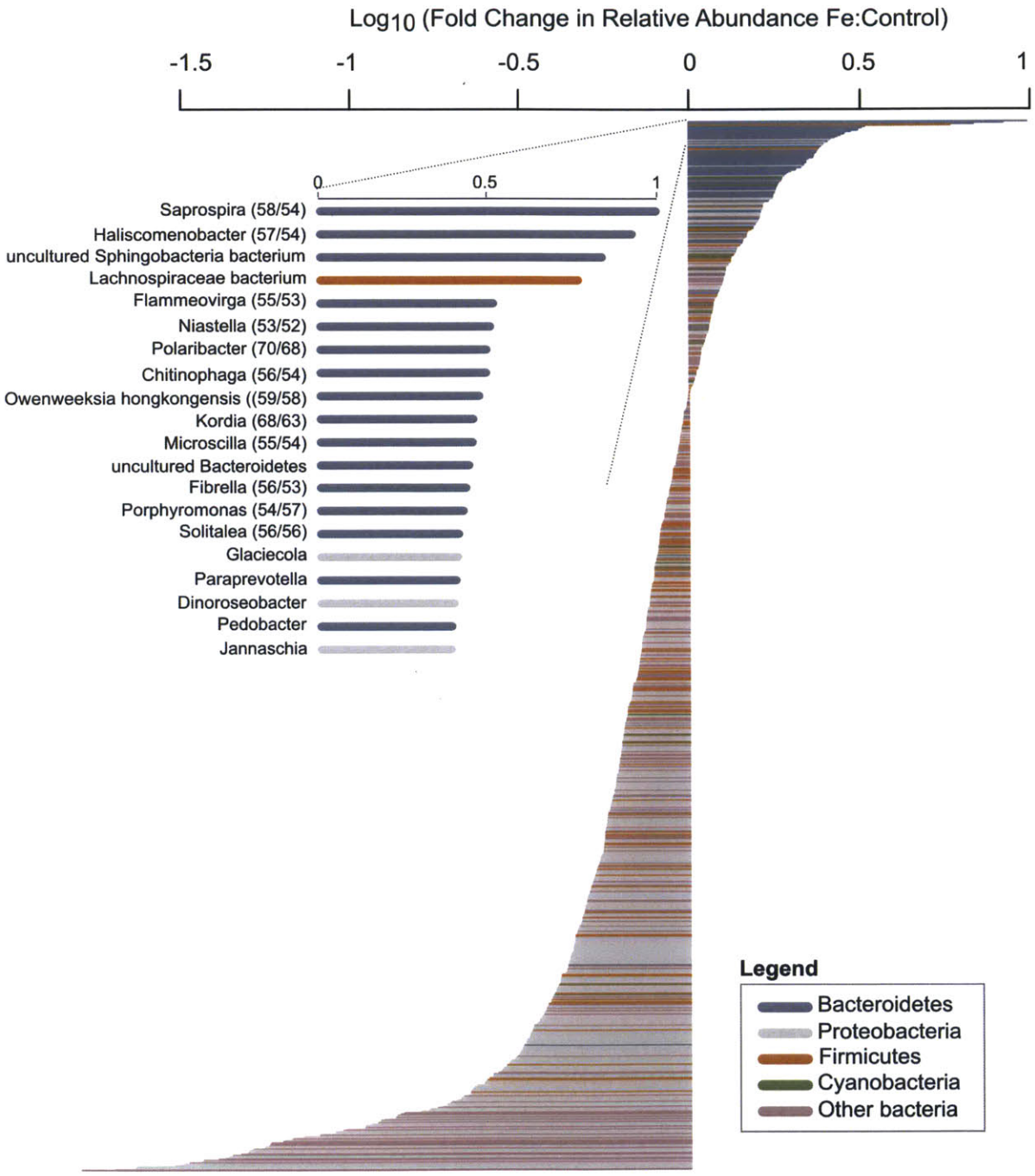


Figure 3.9. Genus-level Changes Specific to the Bacterial Community. This figure is essentially the same as Figure 4, with the one difference that only bacterial genera are represented. Reads with a significant hit (bitscore >50) to the NCBI nr-protein database were binned by genus. We plot the log of the ratio of the relative abundance of reads in the iron treatment to the relative abundance of reads in the control sample. Only genera with at least 20 read counts in either the control or the iron treatment are represented (see Figure 3.5). In parentheses we indicate the average percent identity of reads to each specified genus in the Fe sample and the Control sample (Fe/Control). To compute the relative abundance, counts summed for a specific genus are normalized to the total number of reads assigned at the genus level.

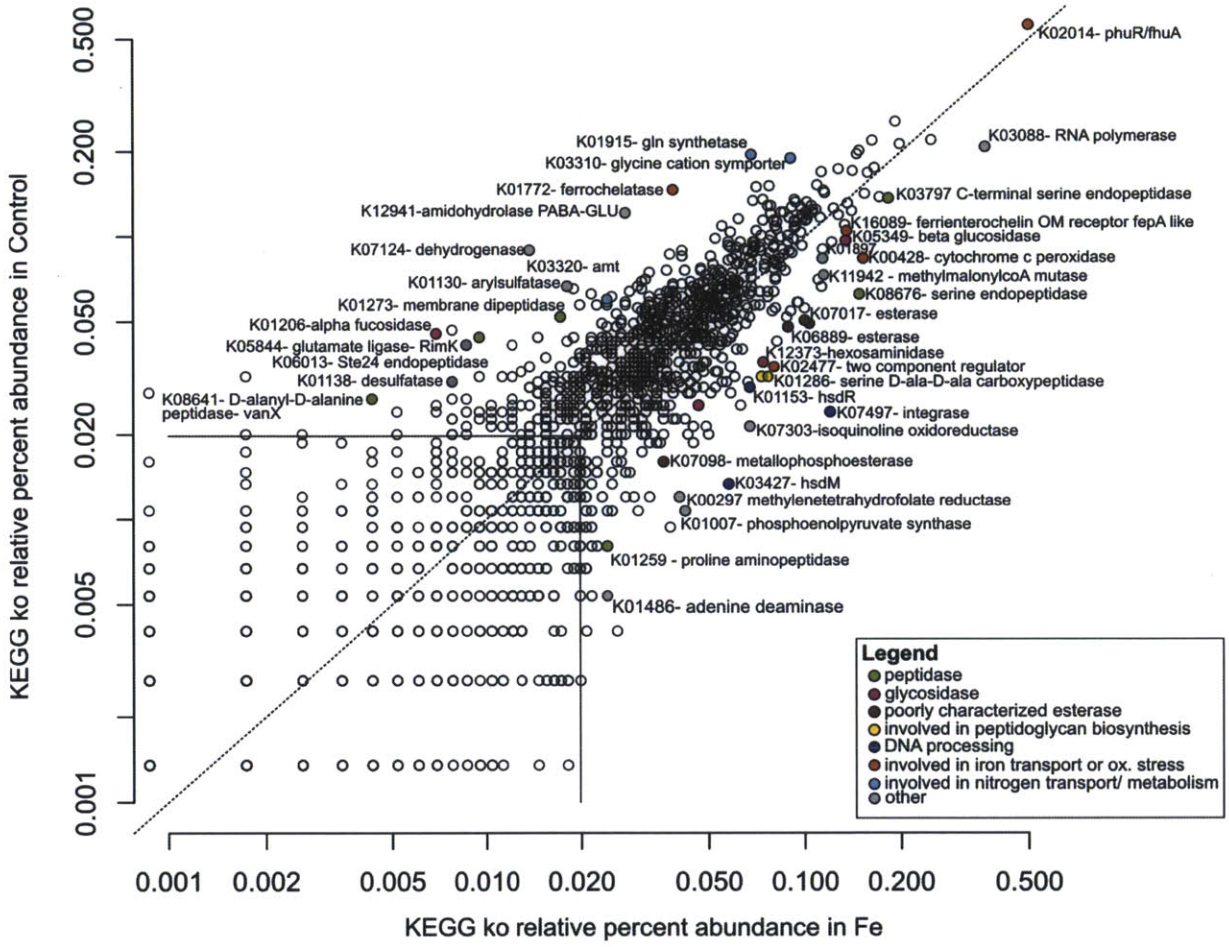


Figure 3.10. Shift in the functional profile of the Bacteroidetes population. We represent the percent relative abundance of KEGG ko numbers for Bacteroidetes specific reads in the Control versus the Fe sample. Ko counts are normalized to the total number of Bacteroidetes reads (74,766 in the Control, 115,885 in the Fe sample). The dotted line represents the x=y line. The vertical line represents x=22 and the horizontal line represents y=15, demonstrating that ko's with low counts are noise and are ignored in the analysis. Ko's are color-coded based on their function.

3.5 Discussion

Diatoms are known to have high requirements for iron. Iron addition to HNLC regions, including in the Equatorial Pacific Ocean invariably leads to diatom-dominated phytoplankton blooms (Coale et al., 1996), pointing to their success at rapidly responding to new inputs of iron as well as evading increased grazing pressure. This suggests that while diatoms have a high requirement for iron, they nonetheless persist under low iron conditions. The persistence of diatom species in open-ocean iron-limited waters is accompanied (or made possible) by a number of acclimation strategies, explaining their lower iron requirement when compared to coastal counterparts (Sunda & Huntsman, 1995; Sunda et al. 1991). Our findings suggest that *Pseudo-nitzschia* spp. may persist at low abundances under ambient (iron-limiting conditions) in the Central Equatorial Pacific Ocean. Metagenomic analysis of the diatom community in the control and in the +Fe samples in our experiment suggest that they were similar in composition. This could suggest that the ambient diatom community in the Central Equatorial Pacific is able to cope with ambient low-iron conditions but responds rapidly to iron addition, possibly experiencing a number of changes at the transcriptional and post-transcriptional level, as was recently suggested in a meta-transcriptomics analysis of an iron-induced diatom bloom in the North Pacific (Marchetti et al., 2012). However, because low-level iron contamination is possible with such bottle incubation experiments, the presence of *Pseudo-nitzschia* spp. observed in the control sample could also be explained by low-level iron stimulation. Until the microbial community composition from samples obtained directly from ambient seawater (prior to any bottle incubation) is determined, we cannot conclude with certainty that the diatom community observed in the control sample accurately reflects the iron-starved ambient diatom community. Despite this caveat, the clear enrichment of *Pseudo-nitzschia* spp. in the +Fe sample compared to the control sample confirms that these diatoms most definitely respond to the addition of iron in these waters, and could hence represent part of the ambient diatom community.

Our results further show that a number of bacterial genera were specifically enriched in the Fe sample compared to the control sample (Figure 3.9). These findings complement previous studies where increases in abundance and bacterial production of the whole bacterial fraction were observed in both

microcosm (Price & Ahner, 1994) and mesocosm (Cochlan, 2001) iron addition experiments to the Equatorial Pacific. As iron is not only necessary for photosynthetic reactions but also as a co-factor of the respiratory electron chain, this stimulation of bacteria could be a direct effect of iron, whereby iron-limited bacteria would be relieved of their micronutrient limitation. Both Price and Cochlan, however suggested that a more probable explanation for the stimulation of bacteria would be an indirect effect of iron, whereby the increased bacterial cell abundances and rates of bacterial production would instead be due to more abundant dissolved organic matter (DOM) released by the blooming phytoplankton (Cochlan, 2001; Price & Ahner, 1994). In fact, the specific enrichment of Bacteroidetes in the Fe sample relative to the control sample in our experiment (Figure 3.9) supports this hypothesis.

Indeed, a growing body of work points to the specialization of Bacteroidetes as consumers of phytoplankton-derived organic matter. Early molecular surveys of particle-attached (composed predominantly of material from a Rhizosolenia/Phaeocystis diatom bloom) versus free-living bacteria showed that the particle-attached communities were dominated by members of the Bacteroidetes phylum (Delong et al. 1993). Furthermore, more recent studies have showed that members of the Bacteroidetes phylum can better degrade polymeric compounds such as proteins and polysaccharides (notably chitin) than their free-living Proteobacteria counterparts, which instead can better assimilate monomeric substrates such as amino acids (Cottrell & Kirchman, 2000). These findings were corroborated by recent analyses of sequenced genomes from *Polaribacter* sp. MED152 (González et al., 2008) and *Gramella forsetii* (Bauer et al., 2006) revealing that these two representatives of marine Bacteroidetes possess a high proportion of hydrolytic enzymes, including peptidases and glycosyl hydrolases, further demonstrating the likelihood that Bacteroidetes could be directly stimulated by phytoplankton-derived high-molecular weight dissolved organic matter (HMWDOM). In fact, characterization of the bacterial community associated with diatom blooms shows that these are usually dominated by α- and γ-Proteobacteria as well as Bacteroidetes (Pinhassi et al., 2004; Riemann et al. 2000; Rinta-Kanto et al. 2012). Therefore it is highly likely that the specific enrichment of Bacteroidetes observed in the Fe sample could have been a direct result of the blooming diatom community. A recent study demonstrated that the bacterial community in the iron-limited

region of the Ross Sea was co-limited by iron and DOM (Bertrand et al., 2011). Hence, it is possible that the enrichment of Bacteroidetes in the Fe sample could be the result of both the blooming diatom community and the enhanced iron availability.

The stimulation (both in terms of growth and activity) of bacteria is common during phytoplankton blooms. In fact bacterial consumption of DOM released by phytoplankton is an integral step in the microbial loop and the remineralization of carbon in the ocean (Azam et al., 1983). The type of stimulated bacteria is likely to depend on the specific composition of the released DOM and hence on the specific composition of the blooming phytoplankton community (Kerkhof et al. 1999). Diatoms are known to harbor interactions with specific bacteria, which include (but are not limited to) members of the *Cytophaga* and *Flavobacteria* (Amin et al. 2012). Our results show that the iron-induced *Pseudo-nitzschia* spp.-dominated bloom specifically selects for a wide variety of Bacteroidetes (likely not limited to one genus), and specific members of the Proteobacteria (*Dinoroseobacter*, *Jannaschia* and *Glaciecola*) (Figure 3.9). Interestingly, the bacterial genus commonly associated with phytoplankton blooms, *Roseobacter* (Buchan & Moran, 2005) was not stimulated by the *Pseudo-nitzschia* spp.-dominated bloom of this experiment.

Our results also show that the bloom-associated Bacteroidetes community was functionally distinct from the non-bloom community (Figure 3.10), harboring specific genes for fatty acid, protein and polysaccharide consumption. Further suggesting a specific interaction with diatoms, is the enriched peroxidase gene, that could presumably enable the bloom-associated Bacteroidetes community to reduce hydrogen peroxide produced as a by-product of the reaction of light with DOM produced by the diatoms (Steigenberger & Statham, 2010). The resolution of this functional gene analysis does not clarify whether these Bacteroidetes were just feeding on senescing phytoplankton cells or if they may have been involved in a mutualistic interaction with the diatom cells. Representing ~10% of sequence reads in the control sample, the non-bloom-Bacteroidetes community was enriched in genes likely to be relevant under severe nitrogen limitation (Figure 3.10), which is highly probable under free-living, low DOM concentration conditions Interestingly, the non-bloom Bacteroidetes community also possessed glycosidases and peptidases. Distinct from those found in the bloom-associated Bacteroidetes community, these suggest that the non-bloom

Bacteroidetes population could still consume DOM derived from phytoplankton. Given that diatom species were present at low-levels in our control sample, we propose that the functionally distinct Bacteroidetes communities in the control and the Fe-samples could reflect a difference in the composition of diatom exudate under bloom or non-bloom conditions (Barofsky et al. 2009).

These results have implications in terms of understanding the fate of the newly fixed carbon as a result of the iron-induced diatom bloom. Whether this newly fixed carbon will be sequestered via the sinking of diatom cells (Smetacek, 1985) or will be quickly remineralized by heterotrophs depends on the rate of DOM release (dependent on the specific type of blooming diatom) but also on the rate of DOM consumption by the stimulated bacterial community (linked to the DOM preference of the specific type of the stimulated bacteria)(Obernosterer et al., 2008). With the knowledge that a specific group of Bacteroidetes is stimulated by an iron-induced *Pseudo-nitzschia* dominated bloom in the Central Equatorial Pacific, targeted experiments aimed at measuring rates of Bacteroidetes DOM consumption can be designed.

3.6 Conclusions and Future Directions

We show that the addition of iron to the iron-limited equatorial upwelling site of the Central Equatorial Pacific Ocean substantially impacts the microbial community structure both in terms of composition and functional gene content. In agreement with previous findings, iron addition resulted in a bloom of eukaryotic phytoplankton species, mostly the pennate diatoms *Pseudo-nitzschia* spp., *Thalassionema* and *Cylindrotheca*. Accompanying the diatom bloom was an enrichment of functionally distinct Bacteroidetes, which were likely associated with the blooming diatoms. These findings show how iron availability in this region of the ocean dramatically influences the ecology of the ambient microbial community, where Bacteroidetes are allowed to bloom in response to the enrichment of specific diatoms. Finally, our results point to the dearth of marine pico-eukaryotic sequences in publicly available databases as proper functional characterization of diatom-specific genes as well as detection of potential enrichment of micro- and meso-zooplankton grazers were hampered by this.

The present study represents the first application of metagenomic approaches to explore iron-induced changes in microbial community structure in the iron-limited Central Equatorial Pacific biome. We present our results as shifts in metabolic potential of the microbial community between a control and an iron-amended sample. To obtain a more accurate understanding of these shifts, sampling at different time points, especially at the start time of the incubation would be beneficial. This would enable for a proper characterization of the ambient microbial community. Sampling at multiple time-points would further enable for a dynamic characterization of the observed community structure changes. In particular, it would be interesting to observe the time lapse between the onset of iron amendment and the detection of the diatom bloom, as well as between the onset of the diatom bloom and the Bacteroidetes bloom. Furthermore, sampling at multiple time-points could reveal whether other groups bloom prior to the diatom and Bacteroidetes. One limitation of the present study is the use of a metagenomic approach alone. Such an approach allows for the characterization of the metabolic potential of a community. However, as iron-induced changes within a species group are likely to occur at the transcriptional (or even post-transcriptional) level, a useful complementary approach would be the use of metranscriptomics (and even

proteomics) to investigate how the input of iron to an iron-limited environment impacts the activity of specific functional genes. The hypothesized association between the enriched Bacteroidetes and the blooming diatom community could be confirmed by scanning electron microscopy, enabling the visualization of bacteria on the diatom frustules. Finally, as the links between molecular signatures and ecosystem-level processes are being currently unraveled, it would be beneficial to supplement such –omic approaches with rate measurements of not only photosynthesis but also species-specific bacterial production.

3.7 Acknowledgements

I thank René Boiteau for conducting the ship-board incubation experiment, collecting and kindly sharing the samples with me. I also thank Dawn Moran who visualized the samples by microscopy. Finally, I thank Tsultrim Palden for preparing and sequencing the 454 libraries and John Eppley and Elizabeth Ottesen for useful conversations and feedback on the bioinformatic analysis.

3.8 References

Altschul, S. F., Gish, W., Miller, W., Myers, E., & Lipman, D. (1990). Basic Local Alignment Search Tool. *Journal of Molecular Biology*, *215*, 403-410.

Amin, S. a, Parker, M. S., & Armbrust, E. V. (2012). Interactions between diatoms and bacteria. *Microbiology and molecular biology reviews : MMBR*, *76*(3), 667-84. doi:10.1128/MMBR.00007-12

Armbrust, E. V. (2009). The life of diatoms in the world ' s oceans. *Nature*, *459*(May), 185-192. doi: 10.1038/nature08057

Bauer, M., Kube, M., Teeling, H., Richter, M., Lombardot, T., Allers, E., Würdemann, C. a, et al. (2006). Whole genome analysis of the marine Bacteroidetes'Gramella forsetii' reveals adaptations to degradation of polymeric organic matter. *Environmental microbiology*, *8*(12), 2201-13. doi:10.1111/j. 1462-2920.2006.01152.x

Becker, D. J., & Lowe, J. B. (2003). Fucose: biosynthesis and biological function in mammals. *Glycobiology*, *13*(7), 41R-53R. doi:10.1093/glycob/cwg054

Bertrand, E.M., Saito, M.A., Lee, P.A., Dunbar, R.B., Sedwick, P.N., DiTullio, G.R. (2011). Iron limitation of a springtime bacterial and phytoplankton community in the Ross Sea: implications for vitamin B12 nutrition. *Frontiers in Microbiology*, *2*, 1-12. doi: 10.3389/fmicb.2011.00160

Boyd, P W, Watson, a J., Law, C. S., Abraham, E. R., Trull, T., Murdoch, R., Bakker, D. C., et al. (2000). A mesoscale phytoplankton bloom in the polar Southern Ocean stimulated by iron fertilization. *Nature*, *407*(6805), 695-702. doi:10.1038/35037500

Boyd, P W, Jickells, T., Law, C. S., Blain, S., Boyle, E. a, Buesseler, K. O., Coale, K. H., et al. (2007). Mesoscale iron enrichment experiments 1993-2005: synthesis and future directions. *Science (New York, N.Y.)*, *315*(5812), 612-7. doi:10.1126/science.1131669

Boyd, Philip W, Strzepek, R., Jackson, G., Wong, C. S., Mckay, R. M., Law, C., Sherry, N., et al. (2005). The evolution and termination of an iron-induced mesoscale bloom in the northeast subarctic Pacific. *Limnology and Oceanography*, *50*(6), 1872-1886.

Buchan, A., & Moran, M. A. (2005). Overview of the Marine Roseobacter Lineage. *Limnology and Oceanography*, *71*(10), 5665-5677. doi:10.1128/AEM.71.10.5665

Casjens, S. (2003). Prophages and bacterial genomics: what have we learned so far? *Molecular Microbiology*, *49*(2), 277-300. doi:10.1046/j.1365-2958.2003.03580.x

Cavender-Bares, K. K., Mann, E. L., Chisholm, S. W., Ondrusek, M. E., & Bidigare, R. R. (1999). Differential response of equatorial Pacific phytoplankton to iron fertilization. *Limnology and Oceanography*, *44*(2), 237-246. doi:10.4319/lo.1999.44.2.0237

Chizhov, A. O., Dell, A., Morris, H. R., Haslam, S. M., McDowell, R. a, Shashkov, a S., Nifant'ev, N. E., et al. (1999). A study of fucoidan from the brown seaweed Chorda filum. *Carbohydrate research*, *320*(1-2), 108-19.

Coale, K. H., Johnson, K. S., Chavez, F. P., Buesseler, K. O., Barber, R. T., Brzezinski, M. a, Cochlan, W. P., et al. (2004). Southern Ocean iron enrichment experiment: carbon cycling in high- and low-Si waters. *Science (New York, N.Y.)*, *304*(5669), 408-14. doi:10.1126/science.1089778

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

Cochlan, W. (2001). The heterotrophic bacterial response during a mesoscale iron enrichment experiment (IronEx II) in the eastern equatorial Pacific Ocean. *Limnology and Oceanography*, *46*(2), 428-435.

Cottrell, M. T., & 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. *Applied and environmental microbiology*, *66*(4), 1692. doi:10.1128/AEM. 66.4.1692-1697.2000.Updated

Cuvelier, M. L., Allen, A. E., Monier, A., McCrow, J. P., Messié, M., Tringe, S. G., Woyke, T., et al. (2010). Targeted metagenomics and ecology of globally important uncultured eukaryotic phytoplankton. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(33), 14679-84. doi: 10.1073/pnas.1001665107

DeLong, E. F., Preston, C. M., Mincer, T., Rich, V., Hallam, S. J., Frigaard, N.-U., Martinez, A., et al. (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science (New York, N.Y.), 311*(5760), 496-503. doi:10.1126/science.1120250

Delong, E. F., Franks, D. G., & Alldredge, A. L. (1993). Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. *Limnology and Oceanography, 38*(5), 924-934. doi:10.4319/lo.1993.38.5.0924

Duce, R., Liss, P. S., Merrill, J. T., Atlas, E. L., Hicks, B. B., Millertl, J. M., Prospero, J. M., et al. (1991). The atmospheric input of trace species to the world ocean. *Global Biogeochemical Cycles, 5*(3), 193-259.

Eldridge, Melanie, Cadotte, Marc, Rozmus, Alison, Wilhelm, S. (2007). The response of bacterial groups to changes in available iron in the Eastern subtropical Pacific Ocean. *Journal of Experimental Marine Biology and Ecology, 348*, 11-22.

Gervais, F., Riebesell, U., & Gorbunov, M. Y. (2002). Changes in primary productivity and chlorophyll a in response to iron fertilization in the Southern Polar Frontal Zone. *Limnology and Oceanography, 47*(5), 1324-1335. doi:10.4319/lo.2002.47.5.1324

Gomez-Alvarez, V., Teal, T. K., & Schmidt, T. M. (2009). Systematic artifacts in metagenomes from complex microbial communities. *The ISME Journal, 3*(11), 1314-1317. Nature Publishing Group. doi:10.1038/ismej.2009.72

González, J. M., Fernández-Gómez, B., Fernàndez-Guerra, A., Gómez-Consarnau, L., Sánchez, O., Coll-Lladó, M., Del Campo, J., et al. (2008). Genome analysis of the proteorhodopsin-containing marine bacterium Polaribacter sp. MED152 (Flavobacteria). *Proceedings of the National Academy of Sciences of the United States of America, 105*(25), 8724-9. doi:10.1073/pnas.0712027105

Gordon, R. M., Coale, K. H., & Johnson, K. S. (1997). Iron distributions in the equatorial Pacific: Implications for new production. *Limnology and Oceanography, 42*(3), 419-431. doi:10.4319/lo.1997.42.3.0419

Hewson, I., Paerl, R. W., Tripp, H. J., Zehr, J. P., & Karl, D. M. (2009). Metagenomic potential of microbial assemblages in the surface waters of the central Pacific Ocean tracks variability in oceanic habitat. *Limnology and Oceanography, 54*(6), 1981-1994.

Imanian, B., Pombert, J.-F., & Keeling, P. J. (2010). The complete plastid genomes of the two "dinotoms" Durinskia baltica and Kryptoperidinium foliaceum. *PloS one, 5*(5), e10711. doi:10.1371/journal.pone.0010711

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research, 28*(1), 27-30.

Kerkhof, L. J., Voytek, M. A., Sherrell, R. M., Millie, D., & Schofield, O. (1999). Variability in bacterial community structure during upwelling in the coastal ocean. *Hydrobiologia, 401*, 139-148.

Krupovic, M., & Forterre, P. (2011). Microviridae goes temperate: microvirus-related proviruses reside in the genomes of Bacteroidetes. *PloS one, 6*(5), e19893. doi:10.1371/journal.pone.0019893

Kustka, A. B., Shaked, Y., Milligan, A. J., King, D. W., & Morel, F. M. M. (2005). Extracellular production of superoxide by marine diatoms: Contrasting effects on iron redox chemistry and bioavailability. *Limnology and Oceanography, 50*(4), 1172-1180. doi:10.4319/lo.2005.50.4.1172

Landry, M. R., & Kirchman, D. L. (2002). Microbial community structure and variability in the tropical Pacific. *Deep Sea Research Part II: Topical Studies in Oceanography, 49*(13-14), 2669-2693. doi:10.1016/S0967-0645(02)00053-X

Letunic, I., & Bork, P. (2007). Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Nucleic acids research, 23*(1), 127-128.

Li, W., & Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics (Oxford, England), 22*(13), 1658-9. doi:10.1093/bioinformatics/btl158

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., et al. (2004). ARB: a software environment for sequence data. *Nucleic acids research, 32*(4), 1363-71. doi:10.1093/nar/gkh293

Mackey, D. J., O'Sullivan, J. E. Os., & Watson, R. J. (2002). Iron in the western Pacific: a riverine or hydrothermal source for iron in the Equatorial Undercurrent? *Deep Sea Research Part I: Oceanographic Research Papers, 49*(5), 877-893. doi:10.1016/S0967-0637(01)00075-9

Mann, E. L., & Chisholm, S. W. (2000). Iron limits the cell division rate of Prochlorococcus in the eastern equatorial Pacific. *Limnology and Oceanography, 45*(5), 1067-1076. doi:10.4319/lo.2000.45.5.1067

Marchetti, A., Schruth, D. M., Durkin, C. a., Parker, M. S., Kodner, R. B., Berthiaume, C. T., Morales, R., et al. (2012). Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1118408109

Martin, J., Coale, K., Johnson, K., Fitzwater, S., Gordon, R., Tanner, S., Hunter, C., et al. (1994). Testing the iron hypothesis in ecosystems of the equatorial Pacific Ocean.

Moon-van der Staay, S. Y., De Wachter, R., & Vaulot, D. (2001). Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature*, *409*(6820), 607-10. doi:10.1038/35054541

Nishide, E., Anzai, H., Uchida, N., & Nisizawa, K. (1990). Sugar constituents of fucose-containing polysaccharides from various Japanese brown algae. *Hydrobiologia*, *204/205*, 573-576.

Not, F., Gausling, R., Azam, F., Heidelberg, J. F., & Worden, A. Z. (2007). Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environmental microbiology*, *9*(5), 1233-52. doi:10.1111/j.1462-2920.2007.01247.x

Obernosterer, I., Christaki, U., Lefèvre, D., Catala, P., Van Wambeke, F., & Lebaron, P. (2008). Rapid bacterial mineralization of organic carbon produced during a phytoplankton bloom induced by natural iron fertilization in the Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, *55*(5-7), 777-789. doi:10.1016/j.dsr2.2007.12.005

Passow, U. (2002). Transparent exopolymer particles (TEP) in aquatic environments. *Progress in Oceanography*, *55*(3-4), 287-333. doi:10.1016/S0079-6611(02)00138-6

Paul, C., Barofsky, a, Vidoudez, C., & Pohnert, G. (2009). Diatom exudates influence metabolism and cell growth of co-cultured diatom species. *Marine Ecology Progress Series*, *389*(2007), 61-70. doi:10.3354/meps08162

Pennington, J. T., Mahoney, K. L., Kuwahara, V. S., Kolber, D. D., Calienes, R., & Chavez, F. P. (2006). Primary production in the eastern tropical Pacific: A review. *Progress in Oceanography*, *69*(2-4), 285-317. doi:10.1016/j.pocean.2006.03.012

Pinhassi, J., Havskum, H., Peters, F., Guadayol, Ò., Malits, A., & Cie, I. D. (2004). Changes in Bacterioplankton Composition under Different Phytoplankton Regimens. *Applied and environmental microbiology*, *70*(11), 6753. doi:10.1128/AEM.70.11.6753

Price, N. M., & Ahner, B. A. (1994). The equatorial Pacific Ocean: Grazer-controlled populations in an iron-limited ecosystem. *Limnology and Oceanography*, *39*(3), 520-534.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., & Glöckner, F. O. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic acids research*, *35*(21), 7188-96. doi:10.1093/nar/gkm864

Reyes, A., Semenkovich, N. P., Whiteson, K., Rohwer, F., & Gordon, J. I. (2012). Going viral: next-generation sequencing applied to phage populations in the human gut. *Nature reviews. Microbiology*, *10*(9), 607-17. Nature Publishing Group. doi:10.1038/nrmicro2853

Riemann, L., Steward, G. F., & Azam, F. (2000). Dynamics of Bacterial Community Composition and Activity during a Mesocosm Diatom Bloom. *Applied and environmental microbiology*, *66*(2), 578-587. doi:10.1128/AEM.66.2.578-587.2000.Updated

Rinta-Kanto, J. M., Sun, S., Sharma, S., Kiene, R. P., & Moran, M. A. (2012). Bacterial community transcription patterns during a marine phytoplankton bloom. *Environmental microbiology*, *14*(1), 228-39. doi:10.1111/j.1462-2920.2011.02602.x

Romari, K., Vaulot, D., Curie, M., & Cedex, R. (2004). Composition and temporal variability of picoeukaryote communities at a coastal site of the English Channel from 18S rDNA sequences, *49*(3), 784-798.

Rusch, D. B., Halpern, A. L., Sutton, G., Heidelberg, K. B., Williamson, S., Yooseph, S., Wu, D., et al. (2007). The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS biology*, *5*(3), e77. doi:10.1371/journal.pbio.0050077

Shi, X. L., Marie, D., Jardillier, L., Scanlan, D. J., & Vaulot, D. (2009). Groups without cultured representatives dominate eukaryotic picophytoplankton in the oligotrophic South East Pacific Ocean. *PloS one*, *4*(10), e7657. doi:10.1371/journal.pone.0007657

Shi, Y., McCarren, J., & DeLong, E. F. (2011). Transcriptional responses of surface water marine microbial assemblages to deep-sea water amendment. *Environmental Microbiology*, no-no. doi:10.1111/j.1462-2920.2011.02598.x

Slemons, L.O., Murray, J.W., Resing, J., Paul, B., Dutrieux, P. (2010). Western Pacific coastal sources of iron, manganese, and aluminum to the Equatorial Undercurrent. *Global Biogeochemical Cycles*, *24*, GB3024. doi:10.1029/2009GB003693

Smetacek, V. S. (1985). Role of sinking in diatom life-history cycles: ecological, evolutionary and geological significance. *Marine Biology*, *84*(3), 239-251. doi:10.1007/BF00392493

Steigenberger, S., & Statham, P. J. (2010). The role of polysaccharides and diatom exudates in the redox cycling of Fe and the photoproduction of hydrogen peroxide in coastal seawaters. *Biogeosciences*, *7*(Iii), 109-119.

Sunda, W. G., & Huntsman, S. A. (1995). Iron uptake and growth limitation in oceanic and coastal phytoplankton. *Marine Chemistry*, *50*, 189-206.

Sunda, W. G., Swift, D. G., & Huntsman, S. A. (1991). Low-iron requirement for growth. *Nature*, *351*, 55-57.

Takeda, S., & Tsuda, A. (2005). An in situ iron-enrichment experiment in the western subarctic Pacific (SEEDS): Introduction and summary. *Progress in Oceanography*, *64*(2-4), 95-109. doi:10.1016/j.pocean.2005.02.004

Tringe, S. G., von Mering, C., Kobayashi, A., Salamov, A. a, Chen, K., Chang, H. W., Podar, M., et al. (2005). Comparative metagenomics of microbial communities. *Science (New York, N.Y.)*, *308*(5721), 554-7. doi:10.1126/science.1107851

3.9 Supplementary Materials

3.9.1. Cloned 18S rRNA Gene Sequences

Fe_A01 TGATCCTTCTGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGGACATCTAAGGGCATCACAGACCTGTTATTG
CCCCTATCTTCCTGCGTTTAATAGAACGCACGTCCCTCTAAGAAGTTATAATACCAATGAAAAATCATTAGCAAAAC
TATTTAGCAGGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAGTCGGCCCACGAACTAAGAACGGCCATG
CACCACCACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTTACTATGTCTGGACCTGGTAAGTTTTCCC
GTGTTGAGTCAAATTAAGCCGCAGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTTCTTTAAGTTTCAGCCTTGCG
ACCATACTCCCCCCGGAACCCAAAGACTTGTGATTTCTCATAAGGTGCTGACGGGGTCAATACAACGACCGCCAAT
CCCTTGTCGGCATAGTTTATGGTTAAGACTACGATGGTATCTAATCATCTTCGATCCCCTAACTTTCGTTCTTGATT
AATGAAAACATCCTTGGTAAATGCTTTCGCAGTAGTTCGTCTTTCAGAAATCCAA

Fe_A02 GCTGGAATTCGCCCTTACCTGGTTGATCCTGCCAGTAGTCATACGCTCGTCTCAAAGATTAAGCCATGCATGTCTA
AGTATAAATATTTTACTTTGAAACTGCGAACGGCTCATTATATCAGTTATAGTTTATTTGATAGTCCCTTACTACTT
GGATACCCGTAGTAATTCTAGAGCTAATACATGCGTCAATACCCTTCTGGGGTAGTATTTATTAGATTGAAACCAA
CCCCTTCGGGGTGATGTGGTGATTCATAATAAGCTTGCGGATCGCATGCCTCTGGCGGCGATGGATCATTCAAG
TTTCTGCCCTATCAGCTTTGGATGGTAGGGTATTGGCCTACCATGGCTTTAACGGGTAACGGGAAATTAGGGTTT
GATTCCGGAGAGGGAGCCTGAGAGACGGCTACCACATCCAAGGAAGGCAGCAGGCGCGTAAATTACCCAATCCT
GACACAGGGAGGTAGTGACAATAAATAACAATGCCGGGCCTTCTTAGGTCTGGCAATTGGAATGAGAACAATTT
AAACCCCTTATCGAGTATCAATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTA
TATTAAAGTTGTTGCAGTTAAAAAGCTCGTAGTTGAATTTGTGATGTGTCCAGTCGGCCTTTGCTCTTTGAGTGA
TTGTGCTGTTATTGGTCCGTCATGTTTGGGTGGAATCTGTGTGGCATTTTGTCGTGCAGGGATGCCCATCGTTT
ACTGTGAAAAAATTAAGTGTTCAAAGCAGCTTATGCCGTTGAATATATTACATGGAATAATGATATAGGACCTTG
GTACTATTTTGTTGGTTTGCGCACTAAGGTAATGATGAATA

Fe_A04 TGATCCTTCCGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGACCTGTTATTG
CCCCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTAT
TTAGCAGGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCCACCAACTAAGAACGGCCATGCAC
CACCACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTCACTATGTCTGGACCTGGTAAGTTTTCCCGTG
TTGAGTCAAATTAAGCCGCAGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTTCTTTAAGTTTCAGCCTTGCGACC
ATACTCCCCCCGGAACCCAAAGACTTATGATTTCTCACAAGGTGCTGACGAAGACGAAACGAGACTCCGCCCAATC
CCTTGTCGGCATA

Fe_B02 GATCCTTCTGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGTT
CTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGCG
ACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCGT
TCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAGG
TAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGACCTGTTATTGC
CCCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTATT
TAGCAGGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCCACCAACTAAGAACGGCCATGCACC
ACCACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTCACTATGTCTGGACCTGGTAAGTTTTCCCGTGT
TGAGTCAAATTAAGCCGCAGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTTCTTTAAGTTTCAGCCTTGCGACCA
TACTCCCCCCGGAACCCAAAGACTTATGATTTCTCACAAGGTGCTGACGAAGACGAAACGAGACTCCGCCAATCCC
TTGTCGGCATATTTATGGTTAAGACTACGATGGTATCTAATCATCTTCGATCCCCTAACTTTCGTTCTTGATTAATG
AAAACATCCTTGGTAAATGCTTTCGCAGTAGTTTGTC

Fe_B03 TGATCCTTCTGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGACCTGTTATTG
CCCCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTAT
TTAGCAGGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCCACCAACTAAGAACGGCCATGCAC
CACCACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTCACTATGTCTGGACCTGGTAAGTTTTCCCGTG
TTGAGTCAAATTAAGCCGCAGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTTCTTTAAGTTTCAGCCTTGCGACC
ATACTCCCCCC

Fe_B06 TGATCCTTCCGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGCAACAAATCCCCAATAAAGGAGACTCATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGACTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCATGATCAATAATTTCAATGATCAATCCCTATCACGATGAACGTTCAAAGATTTCCCAGGCCTCTCGGCCAAGG
TAGACTTGTTGCATTCATCAGTGTAGCGCGCGTGCGGCCCAGGACATCTAAGGGCATCACAGACCTGTTATTGCC
CCTATCTTCCTGCGTTTAATAGAACGCACGTCCCTCTAAGAAGTTATAATACCAATGAAAAATCATTAGCAAAACTA
TTTAGCAGGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAGTCGGCCCACGAACTAAGAACGGCCATGC
ACCACCACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTTACTATGTCTGGACCTGGTAAGTTTTCCCG
TGTTGAGTCAAATTAAGCCGCAGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTTCTTTAAGTTTCAGCCTTGCGA
CCATACTCCCCCCGGAACCCAAAGACTTGTGATTTCTCATAAGTGCTGACGGGGTCAATACAACGACCGCCAATCC
CTTGTCGG

Fe_B08 TGATCCTTCTGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGACCTGTTATTG
CCCCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTAT
TTAGCAGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCCACCAACTAAGAACGGCCATGCACC
ACCACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTCACTATGTCTGGACTGGTAAGTTTTCCCGTGTT
GAGTCAAATTAA

Fe_B10 TGATCCTTCTGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGACCTGTTATTG
CCCCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTAT
TTAGCAGCGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCCACCAACTAA

Fe_C01 ACCTGGTTGATCCTGCCAGTAGTCATACGCTCGTCTCAAAGATTAAGCCATGCATGTCTAAGTATAAATATTTTACT
TTGAAACTGCGAACGGCTCATTATATCAGTTATAGTTTATTTGATAGTCCCTTACTACTTGGATACCCGTAGTAATT
CTAGAGCTAATACATGCGTCAATACCCTTCTGGGGTAGTATTTATTAGATTGAAACCAACCCCTTCGGGGTGATGT
GGTGATTCATAATAAGCTTGCGGATCGCATGCCTCTGGCGGCGATGGATCATTCAAGTTTCTGCCCTATCAGCTT
TGGATGGTAGGGTATTGGCCTACCATGGCTTTAACGGGTAACGGGAAATTAGGGTTTGATTCCGGAGAGGGAG
CCTGAGAGACGGCTACCACATCCAAGGAAGGCAGCAGGCGCGTAAATTACCCAATCCTGACACAGGGAGGTAGT
GACAATAAATAACAATGCCGGGCCTTCTTAGTCTGGCAATTGGAATGAGAACAATTTAAACCCCTTATCGAGTATC
AATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTAAAGTTGTTGCAGTT
AAAAAGCTCGTAGTTGGATTTGTGGTGTGTCCAGTCGCCTCTGCTCTTTGAGTGGTTGTGCTGTACTGGTCTGCC
ATGTTTGGGTGGAATCTGTGTGGCATTAAGTTGTCGTGCGGGGAGCCCATCGTTTACTGTG

Fe_D02 TGATCCTTCTGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGACCTGTTATTG
CCCCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTAT
TTAGCAGGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCCACCAACTAAGAACGGCCATGCAC
CACCACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTCACTATGTCTGGACCTGGTAAGTTTTCCCGTG
TTGAGTCAAATTAAGCCGCAGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTTCTTTAAGTTTCAGCCTTGCGACC
ATACTCCCCCCGGAACCCAAAGACTTATGATTTCTCACAAGGTGCTGACGAAGACGAAACGAGACTCCGCCCATCC
CTTGTCGGCATAGTTTATGGTTAAGACTACGATGGTATCTAATCATCTTCGATCCCCTAACTTTCGTTCTTGATTAA
TGAAAACATCC

Fe_D11 TGATCCTTCTGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGACCTGTTATTG
CCCCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGTCTAT
TTAGCAGGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCCACCAACTAAGAACGGCCATGCAC
CACCACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTCACTATGTCTGGACCTGGTAAGTTTTCCCGTG
TTGAGTCAAATTAAGCCGCAGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTTCTTTAAGTTTCAGCCTTGCGACC
ATACTCCCCCCGGAACCCAAAGACTTATGATTTCTCACAAGGTGCTGACGAAGACGAAACGAGACTCCGCCAATCC
CTTGTCGGCATATTTATGGTTAAGACTACGATGGTATCTAATCATCTTCGATCCCCTAACTTTCGTTCTTGATTAAT
GAAAACATCCTTGGTAAATGCTTTC

Fe_D12 TGATCCTTCCGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGAACCTGTTATTG
CCCCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTAT
TTAGCAGGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCCACCAACTAAGAACGGCCATGCAC
CACCACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTCACTATGTCTGGACCTGGTAAGTTTTCCCGTG
TTGAGTCAAATTAAGCCGCAGGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTTCTTTAAGTTTCAGCCTTGCGAC
CATACTCCCCCCGGAACCCAAAGACTTGTGATTTCTCACAAGGTGCTGACGAAGACGAAACGAGACTCCGCCATC
CCTTGTCGGCATAGTTTATGGTTAAGACTACGATGGTATCTAATCATCTTCGATCCCCTAACTTTCGTTCCTGATTA
ATGAAAACATCCTTGGTAAATGCTTTCGCAGTAGTTTG

Fe_E03 ACCTGGTTGATCCTGCCAGTAGTCATACGCTCGTCTCAAAGATTAAGCCATGCATGTCTAAGTATAAATATTTTACT
TTGAAACTGCGAACGGCTCATTATATCAGTTATAGTTTATTTGATAGTCCCTTACTACTTGGATACCCGTAGTAATT
CTAGAGCTAATACATGCGTCAATACCCTTCTGGGGTAGTATTTATTAGATTGAAACCAACCCCTTCGGGGTGATGT
GGTGATTCATAATAAGCTTGCGGATCGCATGCCTCTGGCGGCGATGGATCATTCAAGTTTCTGCCCTATCAGCTT
TGGATGGTAGGGTATTGGCCTACCATGGCTTTAACGGGTAACGGGAAATTAGGGTTTGATTCCGGAGAGGGAG
CCTGAGAGACGGCTACCACATCCAGGAAGGCAGCAGGCGCGTAAATTACCCAATCCTGACACAGGGAGTAGTGA
CAATAAATAACAATGCCGGGCCTTCTTAGTCTGGCAATTGGAATGAAACAATTTAAACCCCTTATCGAGTATCAA

Fe_E05 TGATCCTTCTGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGACCTGTTATTG
CCCCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTAT
TTAGCAGGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCCACCAACTAAGAACGGCCATGCAC
CACCACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTCACTATGTCTGGACCTGGTAAGTTTTCCCGTG
TTGAGTCAAATTAAGCCGCAGGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTTCTTTAAGTTTCAGCCTTGCGAC
CATACTCCCCCCGGAACCCAAAGACTTATGA

Fe_E10 TGATCCTTCCGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGCAACAAATCCCCAATAAAGGAGACTCATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGAACTCGTTGAATGCATCAGTGTAGCGCGCGTGCGGCCCAGGACATCTAAGGGCATCACAGACCTGTTATT
GCCCCTATCTTCCTGCATCTAATAGAATGCACGTCCCTCTAAGAAGATTCACCCAGTGAAATTCACTAGGCAACCTA
TTTAGCAGGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCCACCAACTAAGAACGGCCATGCA
CCACCACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTCACTATGTCTGGACCTGGTAAGTTTTCCCGT
GTTGAGTCAAATTAAGCCGCAGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTTCTTTAAGTTTCAGCCTTGCGAC
CATACTCCCCCCGGAACCCAAAGACTTGTGA

Fe_F02 TGATCCTTCCGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGACCTGTTATTG
CCCCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTAT
TTAGCAGGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCCACCAACTAAGAACGGCCATGCAC
CACCACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTCACTATGTCTGGACCTGGTAAGTTTTCCCGTG
TTGAGTCAAATTAAGCCGCAGGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTTCTTTAAGTTTCAGCCTTGCGAC
CATACTCCCCCCGGAACCCAAAGACTTATGATTTCTCACAAGGTGCTGACGAAGACGAAACGAGACTCCGCCAATC
CCTTGTCGGCATAGTTTATGGTTAAGACTACGATGGTATCTAATCATCTTCATCCCTAACTTTCGTTCTTGATTAAT
GAAAACATCCT

Fe_F05 ACCTGGTTGATCCTGCCAGTAGTCATACGCTCGTCTCAAAGATTAAGCCATGCATGTCTAAGTATAAATATTTTACT
TTGAAACTGCGAACGGCTCATTATATCAGTTATAGTTTATTTGATAGTCCCTTACTACTTGGATACCCGTAGTAATT
CTAGAGCTAATACATGCGTCAATACCCTTCTGGGGTAGTATTTATTAGATTGAAACCAACCCCTTCGGGGTGATGT
GGTGATTCATAATAAGCTTGCGGATCGCATGCCTCTGGCGGCGATGGATCATTCAAGTTTCTGCCCTATCAGCTT
TGGATGGTAGGGTATTGGCCTACCATGGCTTTAACGGGTAACGGGAAATTAGGGTTTGATTCCGGAGAGGGAG
CCTGAGAGACGGCTACCACATCCAAGGAAGGCAGCAGGCGCGTAAATTACCCAATCCTGACACAGGGAGGTAGT
GACAATAAATAACAATGCCGGGCCTTCTTAGTCTGGCAATTGGAATGAGAACAATTTAAACCCCTTATCGAGTATC
AATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTAAAGTTGTTGCAGTT
AAAAAGCTCGTAGTTGAATTTGTGATGTGTCCAGTCGGCCTTTGCTCTTTGAGTGATTGTGCTGTTATTGGTCCG
TCATGTTTGGGTGGAATCTGTGTGGCATTA
Fe_F08 ACCTGGTTGATCCTGCCAGTAGTCATACGCTCGTCTCAAAGATTAAGCCATGCATGTCTAAGTATAGTCTTCGTCA
GCACCTTGTGAGAAATCATAAGTCTTTGGGTTCCGGGGGGAGTATGGTCGCAAGGCTGAAACTTAAAGAAATTG
ACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAAACTTACCAGGTCCAGAC
ATAGTGAGGATTGACAGATTGAGAGCTCTTTCTTGATTCTATGGGTGGTGGTGCATGGCCGTTCTTAGTTGGTG
GAGTGATTTGTCTGGTTAATTCCGTTAACGAACGAGACCCCTGCCTGCTAAATAGTCCGCATAGTGTTTATCACTG
TGTAGGACTTCTTAGAGGGACGTGCGTTCTATTAGACGCAGGAAGATAGGGGCAATAACAGGTCTGTGATGCCC
TTAGATGTTCTGGGCCGCACGCGCGCTACACTGATGCAAGCAACGAGTACTACCTTGGCCGAGAGGCCTGGGCA
ATCTTTTGAACCTGCATCGTGATAGGGATAGATTATTGCAATTATTAATCTTGAACGAGGAATTCCTAGTAAACGC
AGATCATCAATCTGCATTGATTACGTCCCTGCCCTTTGTACACACCGCCCGTCGCACCTACCGATTGAATGGTCCG
GTGAAGCCTCGGGATTGTGATTATCTTCCTTTATTGGATGGTTGTCACGAGAACTTGTCTAAACCTTATCATTTA

Fe_F11 TGATCCTTCTGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGACCTGTTATTG
CCCCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTAT
TTAGCAGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCTACCAACTAAAACGGCCATGCACCAC
CACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTCACTATGTCTGGACCTGGTAAGTTTTCCCGTGTTG
AGTCAAATTAAGCCGCGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTTCTTTAAGTTTCACCTTGCGACCATACT
CCCCCCGGAACCCAAAGACTTATGATTTCTCACAAGGTGCTGACAAACAAACGAGACTCCGCCAATCCCTTGTCGG
CATATTTATGGTTAAGATACATGGTATCTAATCATCTTCATCCCCTAACTTTCTTCTTGATTAATGAAAACATCCTTG
GTAAATGCTTTC
Fe_F12 ACCTGGTTGATCCTGCCAGTAGTCATACGCTCGTCTCAAAGATTAAGCCATGCATGTCTAAGTATAAATATTTTACT
TTGAAACTGCGAACGGCTCATTATATCAGTTATAGTTTATTTGATAGTCCCTTACTACTTGGATACCCGTAGTAATT
CTAGAGCTAATACATGCGTCAATACCCTTCTGGGGTAGTATTTATTAGATTGAAACCAACCCCTTCGGGGTGATGT
GGTGATTCATAATAAGCTTGCGGATCGCATGCCTCTGGCGGCGATGGATCATTCAAGTTTCTGCCCTATCAGCTT
TGGATGGTAGGGTATTGGCCTACCATGGCTTTAACGGGTAACGGGAAATTAGGGTTTGATTCCGGAGAGGGAG
CCTGAGAGACGGCTACCACATCCAAGGAAGGCAGCAGGCGCGTAAATTACCCAATCCTGACACAGGGAGGTAGT
GACAATAAATAACAATGCCGGGCCTTCTTAGGTCTGGCAATTGGAATGAGAACAATTTAAACCCCTTATCGAGTAT
CAATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTAAAGTTGTTGCAGT
TAAAAAGCTCGTAGTTGAATTTGTGATGTGTCCAGTCGGCCTTTGCTCTTTGAGTGATTGTGCTGTTATTGGTCC
GTCATGTTTGGGTGGAATCTGTGTGGCATTAGTTGTCGTGCAGGGATGCCCATCGTTTACTGTGAAAAAATTAA
GTGTTCAAAGCAGCTTATGCCGTTGAATATATTACATGGAATAATGATATAGACTTGGTACTATTTTGTTGGTTTG
CGCACTAA
Fe_G07 CCTTCCGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGTTCTC
GTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGCGACG
GGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCGTTCA
AGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAGGTAG
TACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGACCTGTTATTGCCCC
TATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTATTTA
GCAGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCCACCAACTAAGAACGGCCATGCACCACC
ACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTCACTATGTCTGGACCTGGTAAGTTTTCCCGTG
Fe_G08 ACCTGGTTGATCCTGCCAGTAGTCATACGCTCGTCTCAAAGATTAAGCCATGCATGTCTAAGTATAAATATTTTACT
TTGAAACTGCGAACGGCTCATTATATCAGTTATAGTTTATTTGATAGTCCCTTACTACTTGGATACCCGTAGTAATT
CTAGAGCTAATACATGCGTCAATACCCTTCTGGGGTAGTATTTATTAGATTGAAACCAACCCCTTCGGGGTGATGT
GGTGATTCATAATAAGCTTGCGGATCGCATGCCTCTGGCGGCGATGGATCATTCAAGTTTCTGCCCTATCAGCTT
TGGATGGTAGGGTATTGGCCTACCATGGCTTTAACGGGTAACGGGAAATTAGGGTTTGATTCCGGAAGGGAGC
CTGAGAGACGGCTACCACATCCAGGAAGGCAGCAGCGCGTAAATTACCCAATCCTGACACAGGAGTAGTGACAAT
AAATAACAATGCCGGGCCTTCTTTCTGGCAATTGATGAGAACAATTTAAACCCCTTATCGAGTATC
Fe_G09 TGATCCTTCTGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGCGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCGT
TCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAGG
TAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGCATCACAGACCTGTTATTGCC
CCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTATTT
A

Fe_H03 ACCTGGTTGATCCTGCCAGTAGTCATACGCTCGTCTCAAAGATTAAGCCATGCATGTCTAAGTATAAATATTTTACT
TTGAAACTGCGAACGGCTCATTATATCAGTTATAGTTTATTTGATAGTCCCTTACTACTTGGATACCCGTAGTAATT
CTAGAGCTAATACATGCGTCAATACCCTTCTGGGGTAGTATTTATTAGATTGAAACCAACCCCTTCGGGGTGATGT
GGTGATTCATAATAAGCTTGCGGATCGCATGCCTCTGGCGGCGATGGATCATTCAAGTTTCTGCCCTATCAGCTT
TGGATGGTAGGGTATTGGCCTACCATGGCTTTAACGGGTAACGGGAAATTAGGGTTTGATTCCGGAGAGGGAG
CCTGAGAGACGGCTACCACATCCAAGGAAGGCAGCAGGCGCGTAAATTACCCAATCCTGACACAGGGAGGTAGT
GACAATAAATAACAATGCCGGGCCTTCTTAGGTCTGGCAATTGGAATGAGAACAATTTAAACCCCTTATCGAGTAT
CAATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTAAAGTTGTTGCAGT
TAAAAAGCTCGTAGTTGAATTTGTGATGTGTCCAGTCGGCCTTTGCTCTTTGAGTGATTGTGCTGTTATTGGTCC
GTCATGTTTGGGTGGAATCTGTGTGGCATTAGTTGTCGTGCAGGGATGCCCATCGTTTACTGTGAAAAAATTAA
GTGTTCAAAGCAGCTTATGCCGTTGAATATATTACATGGATAATGATATAGACCTTGGTACTATTTTGTTGGTTTG
CGCACTAGTTGGGGGGTATTTGTA

Fe_H05 TGATCCTTCTGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGACCTGTTATTGC
CCCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTA

Fe_H09 TGATCCTTCTGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGACCTGTTATTG
CCCCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTAT
TTAGCAGGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCCACCAACTAAGAACGGCCATGCAC
CACCACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTCACTATGTCTGGACCTGGTAAGTTTTCCCGTG
TTGAGTCAAATTAAGCCGCAGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTTCTTTAAGTTTCAGCC

Fe_H11 TGATCCTTCTGCAGGTTCACCTACGGAAACCTTGTTACGACTTCACCTTCCTCTAAATGATAAGGTTTAGACAAGT
TCTCGTGACAACCATCCAATAAAGGAAGATAATCACAATCCCGAGGCTTCACCGGACCATTCAATCGGTAGGTGC
GACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAATGCAGATTGATGATCTGCGTTTACTAGGAATTCCTCG
TTCAAGATTAATAATTGCAATAATCTATCCCTATCACGATGCAGGTTCAAAAGATTGCCCAGGCCTCTCGGCCAAG
GTAGTACTCGTTGCTTGCATCAGTGTAGCGCGCGTGCGGCCCAGAACATCTAAGGGCATCACAGACCTGTTATTG
CCCCTATCTTCCTGCGTCTAATAGAACGCACGTCCCTCTAAGAAGTCCTACACAGTGATAAACACTATGCGGACTAT
TTAGCAGCAGGGGTCTCGTTCGTTAACGGAATTAACCAGACAAATCACTCCACCAACTAAGAACGGCCATGCACC
ACCACCCATAGAATCAAGAAAGAGCTCTCAATCTGTCAATCCTCACTATGTCTGG

Chapter 4

Microbial Community Dynamics in Response to Changing Iron Availability in the North Pacific Subtropical Gyre

4.1. Abstract

Iron is essential to microbial life. As a co-factor of key enzymes, for example those controlling nitrogen fixation and photosynthesis, iron availability regulates key biochemical reactions that drive global biogeochemical cycles. Yet, in the ocean, iron is present in minute concentrations and may potentially limit microbial metabolism. Over 99% of dissolved iron in seawater is complexed with organic ligands of unknown structure and origin. Previous studies have suggested that these organic ligands may act as siderophores, increasing the microbial community's access to iron. We aimed to explore how sea-surface microbial assemblages respond to additions of iron and naturally occurring iron-binding ligands. By following changes in microbial community structure and patterns of gene expression, we aimed to reveal pathways involved in iron acquisition and utilization in the most abundant autotrophs and heterotrophs. Ship-board incubations were set up with surface seawater from Hawai'i Ocean Time Series Station ALOHA and amended with Fe(III), organic matter (containing iron-binding ligands), or Fe(III) + organic matter. A combination of flow-cytrometric, metagenomic and metatranscriptomic measurements revealed that both autotrophs and heterotrophs were not significantly affected by the addition of iron. The dominant autotroph, *Prochlorococcus* showed differential physiological and metatranscriptomic responses to the combined addition of organic matter and iron (relative to the organic matter alone addition). However, we suggest that this was more likely due to shifts in the type of limiting nutrient caused by an influx of organic nutrients provided in the form of the organic matter than to the increased availability of iron. To support this hypothesis, we showed how the overall microbial community was dramatically affected by the addition of organic matter both in terms of composition and gene expression patterns.

With Kathleen Munson[1], Mar Nieto-Cid[1], Erin Bertrand[1], Curtis R. Young[2], Dan Repeta[1], Mak Saito[1], Edward F. DeLong[2,3]

[1]Marine Chemistry and Geochemistry Department, WHOI, Woods Hole, MA, USA
[2]Civil and Environmental Engineering Department, MIT, Cambridge, MA, USA
[3]Biological Engineering Department, MIT, Cambridge, MA, USA

4.2. Background

Iron is present in minute concentrations in large regions of the world's oceans, thus limiting microbial processes of global importance, such as primary production and nitrogen fixation (Morel & Price 2003; M. Moore et al. 2009). Understanding what forms of iron may be biologically available to the microbial community is crucial in improving our knowledge of microbial functioning in the ocean ecosystem. The discovery almost 20 years ago that 99% of the dissolved iron pool is complexed by organic ligands (Gledhill & van den Berg 1994) changed our fundamental understanding of iron speciation, and in turn of iron bio-availability in seawater. Up until then, it was thought that unchelated inorganic ferric iron (Fe^{3+}) was the main source of bio-available iron, the Fe' model (Hudson & Morel 1990). Because 99% of the dissolved iron pool is complexed by these ligands (Gledhill & van den Berg 1994; Rue & Bruland 1995) and because ligand concentrations correlate well with increased dissolved iron concentration (Boye et al. 2003), it is now a widely accepted hypothesis that these ligands may act to maintain iron in the dissolved fraction, thereby potentially providing a greater source of bioavailable iron. However, the biological role of these iron-complexing organic ligands remains to be elucidated.

Since their discovery, iron-binding organic ligands have been detected in a variety of ocean basins, including in the Northeast Atlantic (Gledhill & van den Berg 1994), the Western Mediterranean (van den Berg 1995), the North Pacific (Rue & Bruland 1995), the Northwest Atlantic (Wu & Luther 1995), the Equatorial pacific (Rue & Bruland 1997) as well as the Bering Sea (Buck & Bruland 2007). Despite this ubiquitous distribution in marine waters, their exact chemical composition as well as their origin remain unknown. Based on their iron-binding affinity, two classes of iron-complexing ligands have consistently been identified: a strong ligand, referred to as L_1 with concentrations in surface waters of ~0.5nM and conditional stability constant of (K_{L1}^{cond}, Fe') ~10^{12} M^{-1}, and a weaker ligand pool, L_2 with a conditional stability constant of (K_{L1}^{cond}, Fe') ~10^{11} M^{-1} (Rue & Bruland 1997; Boyd et al. 2000). Based on these properties, a few compounds have been suggested as likely components of the ligand pool. Because the L_1 pool has a conditional stability constant similar to that of strong siderophores in seawater, it has been suggested that L_1 likely comprises biologically produced siderophores (Macrellis et al. 2001; Rue & Bruland,

1997). Other candidate compounds include iron-containing porphyrins, which could be released into seawater after cell lysis (Hutchins et al. 1999), algal exudates including saccharides (Hassler et al. 2011), and humic substances believed to be an important component of the iron-binding ligand pool especially in coastal and deep-sea waters (Laglera & van den Berg 2009). Despite the uncertainty surrounding their chemical composition, experiments showing increased ligand concentration upon addition of iron to either natural plankton assemblages (Rue & Bruland 1997) or cultures of *Emiliana huxleyi* (Boye & van den Berg 2000) suggest that these ligands are biologically produced.

Despite these clues, many uncertainties remain: it is still unclear which microorganisms produce these ligands, whether these are by-products of metabolism, or whether they are produced with the specific goal of iron acquisition, and finally whether these ligand in fact enhance biological availability of iron, as it has been suggested many times. A number of experiments have been performed on cultured isolates and on natural microbial assemblages to determine whether organic ligands enhance iron bioavailability. Most of these quantified the increased bioavailability by measuring iron uptake rates (using radiolabeled Fe) and by using model ligands such as the siderophores ferrioxamines, ferrichrome, rhodotorulic acid, enterobactin (Hutchins et al. 1999; Weaver et al. 2003; Maldonado & Price 1999), the synthetic metal chelators EDTA and DTPA (Maldonado 2005), porphyrins (Hutchins et al., 1999; Weaver et al. 2003), phytic acid (Weaver et al. 2003; Maldonado 2005), or more recently saccharides (Hassler et al. 2011). Taken together, these experiments show mixed results, with increases in iron bioavailability depending both on the type of model ligand used and the composition of the community (identified in by size-fractionation and pigment analyses) responsible for the uptake. Until more specific chemical characterization of this pool of organic ligands is available, connecting the results of such experiments to what may be happening *in situ* will be difficult. Here, we set out to investigate whether organic matter directly obtained from seawater samples enhances the iron-specific response of a natural microbial assemblage from the North Pacific Subtropical Gyre (NPSG).

The North Pacific Subtropical Gyre is one of the largest habitats on Earth and one of the most expansive subtropical gyres of the ocean system (Sverdrup et al., 1946). Permanently stratified, with low

standing stocks of biomass and nutrients, the NPSG was long thought to be a desert of biological activity (Karl, 2010). However, long-term monitoring of the gyre has revealed a diverse and highly active plankton assemblage dominated by small-celled bacterial and archaeal species (Karl and Lukas, 1996). Dissolved iron concentrations (≤0.4 μm; now ≤0.2 μm) measured in surface waters at the Hawai'i Ocean Time series (HOT) station ALOHA (22°45' N, 158° W) fluctuate seasonally between 0.2 nmol kg^{-1} (in mid-January) and 0.7 nmol kg^{-1} (in late April) due to temporally variable eolian dust deposition (Boyle et al 2005). While the NPSG is not considered to be one of the highly iron-limited HNLC regions, it is conceivable that during (or right after) the low-dust input period that extends from July to January surface iron concentrations may reach levels below 0.5 nmol kg^{-1}, which are potentially limiting to some picocyanobacteria as well as diazotrophs (Blain et al. 2002; Moore et al., 2009). In fact, recent studies suggest that iron-laden dust episodic deposition might be one of the factors driving the development of phytoplankton summer blooms (Calil et al. 2011). Situated in the NPSG, the present study also aimed to explore the *in situ* iron-specific response of the autotrophic as well as heterotrophic community indigenous to this low-nutrient low-chlorophyll regime.

While the effects of iron on marine microbial communities have mostly been studied with respect to iron's potential in limiting the growth of large eukaryotic phytoplankton and primary production (Martin et al. 1994; Boyd et al. 2007), the probable impact of iron on the heterotrophic community cannot be ignored. Indeed, the requirement for iron is pervasive across the microbial tree of life as it acts as a co-factor to enzymes participating in pathways as diverse as respiration (as part of Fe-S clusters in the respiratory electron transport chain), the TCA cycle (as part of the [4Fe-4S] cluster in the aconitase enzyme), or DNA biosynthesis (in the iron-dependent ribonucleotide reductase) (Andrews et al. 2003; Beinert et al. 1996). In fact, marine heterotrophs have been found to assimilate 50% of the total dissolved iron in the Subarctic North Pacific (Tortell et al. 1999) and their abundance as well as their activity (measured by leucine incorporation rates) both increased during the IronEx II study when iron amendments were performed to the waters of the Equatorial Pacific (Cochlan 2001). In the North Pacific Subtropical Gyre, microbial heterotrophs play a major role in sustaining biological productivity. By engaging

in the rapid remineralization of organic matter, they supply primary producers with a precious source of cycled nutrients in an oceanic regime where external nutrient inputs are otherwise rare (Cotner & Biddanda, 2002; Karl, 2002; Karl & Lukas, 1996). Because of the high variability in surface iron concentrations and because microbial heterotrophs have a central role in sustaining nutrient cycling in the NPSG, understanding how iron's availability may affect the growth and activity of these is key if we are to determine the factors governing and driving biological variability in the NPSG.

Using an experimental meta-transcriptomic approach allowing for the exploration of rapid changes in gene expression patterns across the entire microbial community (Frias-Lopez et al. 2008; McCarren et al. 2010; Shi et al. 2011), we set out to explore how changes in iron bioavailability affect the community structure and gene expression patterns of the natural microbial assemblage at station ALOHA. **We aimed to understand whether organic matter enhances the iron-specific gene expression response of the microbial community,** a step towards testing the hypothesis that iron-binding ligands render iron more bioavailable. Furthermore, this study could be used as a platform to **explore the response of the community to mimicked episodic iron inputs by atmospheric deposition**, a seemingly important process in the development of summer blooms. Specifically we asked: **1) Is there an iron-specific *in situ* response of the microbial community, with particular focus on the most abundant autotrophic and heterotrophic members? 2) Does the addition of (Fe-complexed) organic matter enhance the iron-specific response?**

4.3. Materials and Methods

4.3.1 Preparation of Amendments

Organic matter was collected aboard the *R/V Knorr* during a cruise to the Costa Rica Upwelling Dome in July 2005 (cruise number KN182-50). About ~3,000 L of surface seawater from seven different stations (see Table 4.1) was pumped through Teflon tubing (Cole-Parmer, Vernon Hills, IL, USA) via a Kynar Lutz pump (Wertheim, Germany) into acid cleaned barrels. Water was pumped via a Masterflex pumping system (Cole-Parmer) and Teflon tubing (Cole-Parmer) onto 3 parallel columns (re-utilized at each station) packed with XAD-16 Amberlite resin (Sigma-Aldrich, Saint-Louis, MO, USA). The resin from each column was stored at -20°C in HDPE bottles until further processing.

Table 4.1. Coordinates of stations where organic matter was obtained.

Date of collection	Latitude	Longitude
18 July 2005	8° 27' 675" N	90° 00' 41" W
19 July 2005	10° 00' 37" N	90° 00' 57" W
20 July 2005	9° 33' 96" N	86° 11' 88" W
21 July 2005	8° 42' 57" N	86° 29' 33" W
22-23 July 2005	8° 42' 12" N	87° 28' 80" W
24 July 2005	9° 30' 77" N	89° 00' 04" W
25 July 2005	9° 30' 15" N	92° 19' 72" W

Organic matter was eluted from the resin obtained from one of the three columns with 100% methanol (Sigma-Aldrich, Saint-Louis, MO, USA), and 58.4 mg (50.4mL) of eluted sample was rotovapped to remove residual methanol, then passed through a column filled with iron-chelating Toyopearl resin (previously washed in pH2 water then rinsed with MilliQ water) (Tosoh Bioscience, King of Prussia, PA, USA) to remove residual iron. The organic matter sample was loaded three times onto the resin, then rinsed with pH2 water and rinsed thoroughly with MilliQ water until a pH read of 7.

Organic matter was transported at room temperature to the sampling site, in two separate bottles that had previously been soaked in acid and rinsed with pH2 water. A stock for the organic matter (+OM) amendment was prepared by adding 17 µL of pH2 water to one of these two bottles. The stock for the

organic matter and iron (OM+Fe) amendment was similarly prepared by adding 17 μL of a solution of 3.24 mM $FeCl_3$ (Thermo Fisher Scientific, Waltham, MA, USA) in pH 2 water to the other bottle. A solution of 18 μM $FeCl_3$ in pH2 served as the stock solution for the iron (+Fe) amendment.

4.3.2 Experimental Set-up and Sample Collection

Surface seawater (0-5 m) for microcosm incubations was collected at Hawai'i Ocean Time Series (HOT) station ALOHA (22°45.793'N , 157°57.901'W) aboard the *R/V Kilo Moana* on September 19, 2009, at 3 am (HST) using a trace metal clean system consisting of a trace metal clean Teflon tubing (Cole-Parmer) connected to Wilden P100 diaphragm pumping system (Wilden Pump and Engineering, Grand Terrace, CA, USA). The tubing arrived directly into an on-deck positive-pressure trace-metal clean chamber maintained by an AirClean filtration system (AirClean, Rayleigh, NC, USA), directly filling a trace-metal clean 20 L Nalgene® polycarbonate bottle (Thermo Fisher Scientific). This bottle was used to homogenize the water being sampled before being used to rinse and fill the microcosm bottles. The 9-liter trace-metal clean Nalgene® polypropylene microcosm bottles (Thermo Fisher Scientific) were rinsed three times and filled with seawater. Amendments (organic matter and iron) were added to the appropriate incubation bottles once all bottles were filled with water. Amendments were added in the following order +Fe, +OM, +OM+Fe with intentional 20 minute scatter to allow for rapid sampling time between bottles. A 6.66 mL volume of the (+OM) stock was added to the corresponding three 9 L microcosm bottles, 6.66 mL of (OM+Fe) stock was added to the (OM+Fe) 9L microcosm bottles, and 1mL of the (+Fe) stock was added to the three (+Fe) 9 L microcosm bottles.

Prior to the cruise, all incubation (polycarbonate and polypropylene) and sampling bottles (LDPE) were subjected to rigorous acid-washing to remove any residual trace metals. The procedure followed was based on the cleaning procedure utilized in the Saito lab. Briefly, new bottles were rinsed three times with MilliQ water, then filled and soaked with ~0.5-1% Citranox (Alconox, White Plains, NY, USA) for one week. Bottles were then rinsed seven times with MilliQ water and filled and soaked with 10% Baker Instra-analyzed hydrochloric acid (Mallinckrodt Baker, Phillipsburg, NJ, USA) for two weeks (after one week, the bottles were flipped). The bottles were then rinsed three times with pH2 water. Polycarbonate (0.2 μm)

filters (GE Osmonics, Minnetonka, MN, USA) for particulate iron measurements were soaked in 10% trace-metal grade hydrochloric acid for 12 hours, rinsed with pH2 water and then with MilliQ water until pH 7 was restored

At each sampling time point (T0h, T3h, T12h, T36h), incubation bottles were wiped outside the positive-pressure trace-metal clean chamber then carefully opened inside the clean chamber, making sure that no water lodged in grooves of the cap could drop inside the bottle. Inside the clean chamber, subsamples were taken for total and particulate iron measurements by pouring 250 mL in LPDE sampling bottles. Filtration for particulate iron measurements was performed on a custom-made filtration apparatus (provided by the Saito lab, following the model developed in the Boyle lab). For RNA samples, an aliquot of 1.75 liters was poured into a separate container and immediately filtered through a 1.6μm glass-fiber pre-filter (GF/A, Whatman) onto a 0.2μm polyvinyl difluoride (PVDF) durapore filter (Millipore, Billerica, MA, USA) using a Masterflex peristaltic pump (Cole Parmer Instrument Company, Vernon Hills, IL, USA). The durapore filter was then placed into *RNALater* (Applied Biosystems, Foster City, CA, USA) and stored at -80°C until further processing. At time initial, DNA was collected by filtering 10 liters of seawater through a 1.6 μm glass-fiber pre-filter (GF/A, Whatman) onto a 0.2 μm PVDF durapore filter (Millipore, Billerica, MA, USA) using a Masterflex peristaltic pump (Cole Parmer Instrument Company, Vernon Hills, IL, USA). Filters were stored at -80°C in cryovials filled with 400 μL sterile storage buffer (0.73 M sucrose, 40 mM EDTA, 50 mM Tris, pH 8.3) until further processing. Any water remaining in the incubation bottles at the end of the experiment (between 1.2 and 1.75 liters) was handled similarly to collect T36h DNA. Refer to Figure 4.1 for summary of experimental set-up and sampling.

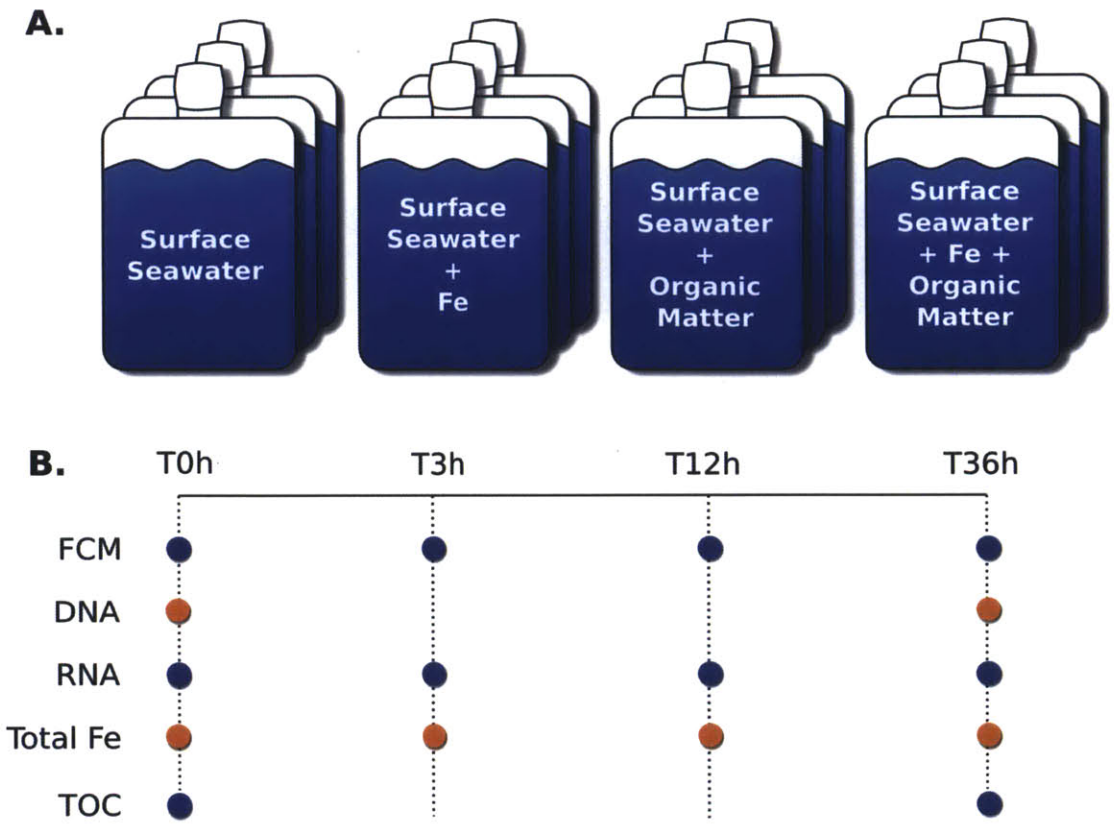


Figure 4.1. Experimental set-up and sampling scheme. A. Experimental Set-up. Four microcosm incubations were performed in triplicates. B. Sampling scheme. Colored dots indicate which samples were taken at each time-point. FCM: flow-cytometry, Total Fe: Total iron, TOC: Total organic carbon. T0h: 4:45am HST, T36h: 5:00pm-6pm HST.

4.3.3 Iron Measurements

All procedures are based on (Saito & Schneider 2006). All samples were processed in a laminar flow hood in a class 100 trace-metal clean room to prevent contamination. Pipette tips were rinsed in 10% Baseline® hydrochloric acid (Seastar Chemicals Inc., Sidney, British Columbia, Canada), thereafter referred to as clean acid, and pH 2 water three times prior to handling of any reagent. Briefly, total Fe samples were acidified to 0.5% volume/volume with concentrated clean hydrochloric acid in January 2010. Fe concentrations were measured in July 2010. A stock of ^{57}Fe spike was prepared by dissolving ^{57}Fe metal (Oak Ridge National Laboratory, TN, USA) in concentrated Baseline® nitric acid (Seastar Chemicals Inc., Sidney, BC, Canada) followed by overnight heating in an acid washed Teflon bottle (Nalgene, Thermo Fisher Scientific) in a 110°C oven. A ~100 nM ^{57}Fe spike working stock was prepared by dilution in 5% clean nitric acid.

Between 12-14 mL of each sample were poured into 15 mL Falcon tubes (BD Biosciences, Franklin Lakes, NJ, USA) (previously acid cleaned by soaking in 10% clean hydrochloric acid and rinsing with pH2 water). Each sample was spiked with 100 µL of ^{57}Fe and allowed to equilibrate overnight (the ^{57}Fe is used correct for variation in precipitation). Fe was precipitated with 120-145 µL Baseline® ammonium hydroxide (Seastar Chemicals Inc., Sidney, British Columbia, Canada) and centrifuged for 3 minutes at 4,800 rpm in a an Eppendorf 5810R centrifuge. The supernatant was poured off and the tube was flicked to remove excess supernatant. The pellet was centrifuged again for 3 minutes at 4,800 rpm. Remaining supernatant was removed by flicking the tube. Pellets were resuspended in 500 mL of 5% clean nitric acid with 1 ppb indium added as an internal standard to monitor instrument sensitivity (In used for blank and recovery corrections, to correct for variation in instrument sensitivity). Measurements were performed on an Element II inductively-coupled plasma mass spectrometer (ICPMS) (Thermo Fisher Scientific, Waltham, MA, USA), downstream of an Aridus II™ desolvating nebulizing system (CETAC Technologies, Omaha, NE, USA). Final Fe concentrations were calculated from the ^{56}Fe: ^{57}Fe ratios using an isotope dilution equation relative to a standard curve of 1 ppb, 2 ppb, 3 ppb, and 4 ppb Fe (iron), and Mn (manganese). ^{56}Fe and ^{57}Fe concentrations were determined relative to the spiked In.

4.3.4 Total Organic Carbon Measurements

Approximately 10 mL of seawater were collected in pre-combusted (450°C, 12 h) glass ampoules for TOC analysis at T0h and T36h. Phosphoric acid H_3PO_4 was added to acidify the sample down to pH <2. The ampoules were heat-sealed and stored at 4°C. TOC was measured by high-temperature catalytic oxidation with a TOC-V Series organic carbon analyser (Shimadzu, Kyoto, Japan). After decarbonation of the sample by vigorous stirring with high-purity synthetic air for 15 min, 150 µL were injected into the vertical furnace (a quartz tube) of the analyzer, filled with a 0.5% Pt-coated Al_2O_3 catalyst at 680°C. Quantitative production of CO_2 occurs from the TOC in the sample; and this CO_2 was measured in the Shimadzu Infrared Gas Analyzer. Three to 5 replicate injections were performed per sample and the system was standardized daily with potassium hydrogen phthalate in Milli-Q water. The concentration of TOC was determined by subtracting the average peak area from the instrument blank area and dividing by the slope of the standard curve. The system blank, obtained by frequent injection (every 5 samples) of UV-Milli-Q water, was equivalent to 5-7 µmol C L^{-1}.The precision of measurements was ±0.8 µmol C L^{-1}.

4.3.5 Flow-Cytometry Measurements

At each time point, a 1 mL aliquot was preserved in 0.125% glutaraldehyde (final concentration) (Sigma-Aldrich) and stored at -80°C until measurements were made. Samples were run on an Influx™ flow cytometer (Becton Dickinson, Franklin Lakes, NJ, USA). Each sample was split in two fractions: a. fraction 1 stained for 15 minutes with the nucleic acid stain SYBR Green (Invitrogen, Carlsbad, CA), and b. fraction 2 unstained. Two micron-diameter fluoresbrite beads (Polysciences Inc. Warrington, PA, USA) were used for normalizing forward-angle light scatter and red fluorescence (Polysciences Inc. Warrington, PA, USA).

4.3.6 Nucleic Acids Processing and Sequencing

DNA extraction and purification

DNA extraction and purification was performed as previously described (DeLong et al., 2006). Briefly, filters were thawed on ice. Cells on the filter were lysed after addition of 40 µL of lysozyme buffer (50 mg lysozyme into 1 mL of an aqueous solution of 0.73 M sucrose, 40 mM EDTA, 50 mM Tris, pH 8.3)

and incubating at 37°C for 45 min. Proteins were degraded after treatment with 100 µL of proteinaseK and 100 µL of MDT (Autogen, Holliston, MA, USA) at 55°C for 55 min. The filter was then removed and 2 mL of LDT (Autogen) solution was added to the lysate and incubated at 55°C for 15 min. About 2.7 mL of 99% ethanol was added and the mixture was vortexed at high speed for 15 seconds. DNA was purified with an automated DNA purification system (AutoGen). Quantification of the extracted DNA was performed with Invitrogen's Quant-iT™ Picogreen® dsDNA Assay kit (Life Technologies, Carlsbad, CA, USA). Roughly 8 µg of DNA was extracted for the T0h sample and ~500-800 ng for each of the four T36h samples (see Table 4.4).

RNA extraction and purification

RNA extraction, purification and DNAse treatment were performed as previously described (Frias-Lopez et al. 2008) on RNA samples from one of the triplicate bottles for times T0h, T12h and T36h for a total of nine samples. Briefly, RNA was extracted using the Ambion *mir*Vana™ miRNA isolation kit (Life Technologies, Carlsbad, CA, USA). DNAse treatment was performed following directions of the Ambion TURBO DNA-free™ kit (Life Technologies, Carlsbad, CA USA). Purification of DNAse treated RNA was performed following the RNease MinElute Cleanup kit (Qiagen, Germantown, MD, USA). Quantification of the extracted RNA was performed with Invitrogen's Quant-iT™ Ribogreen® RNA Assay kit (Life Technologies). Each of the nine samples yielded between 100 to 600ng of total RNA (see Table 4.4).

Sample specific subtraction of ribosomal RNA

Ribosomal RNA was subtracted from RNA samples by hybridization of sample-specific biotinylated probes as previously described (Ottesen et al., 2011; Shi et al., 2011; Stewart et al., 2010). Briefly, ribosomal RNA genes were PCR amplified from T0h and T36h DNA samples using the Stratagene Herculase®II Fusion polymerase (Agilent Technologies, Santa Clara, CA, USA) and T7-promoter-appended eubacterial and archaeal 16S or 23S primers (see Table 4.2 for detailed description of primers and Table 4.3 for detailed description of DNA templates used). Ribosomal RNA gene amplicons were purified per manual's instruction of the QIAquick PCR purification kit (Qiagen, Germantown, MD, USA). In-vitro transcription of each amplicon (300 ng) was then performed using the Invitrogen™ MEGAscript™ High

Yield Transcription kit (Life Technologies, Carlsbad, CA, USA) with T7-RNA polymerase and biotin-labeled cytosine and uracil nucleotides (Roche, Indianapolis, IN, USA). Synthesized biotinylated antisense RNA (aRNA) was purified using the Ambion® MEGAclear™ kit (Life Technologies, Carlsbad, CA, USA) yielding between ~27 and 85 μg of product per targeted ribosomal gene set (see Table 4.3). Biotinylated aRNA probes were then pooled and hybridized to total RNA extracted from samples in a 2:1 probe-to-RNA ratio. The hybridization reaction in 1X sodium chloride-citrate (SSC) and 20% formamide was carried out for 5 min at 70°C, followed by a set of 5°C increment rampdown steps (1 min each) to 25°C. The biotin-labeled dsRNA was then removed via interaction with streptavidin-coated magnetic beads (NEB, Ipswich, MA, USA). Probe set 1 (PS1) was generated after amplification of the eubacterial 16S, 23S and archaeal 16S and 23S ribosomal RNA genes from the T0h_DNA sample (Table 4.3). PS1 was used to subtract ribosomal RNA from the following RNA samples: T0h_RNA, T12h_Control_RNA, T12h_+Iron_RNA, T36h_Control_RNA, T36h_+Iron_RNA. Probe set 2 (PS2) was generated after amplification of the eubacterial 16S, 23S and archaeal 16S and 23S ribosomal RNA genes from a 1:1 mix of the T36h_+Organic_Matter_DNA and T36h_+Organic_Matter_+Iron_DNA samples. PS2 was used to subtract ribosomal RNA from the following RNA samples: T12h_+OM_RNA, T12h_+OM+Fe_RNA, T36h_+OM_RNA, T36h_+OM+Fe_RNA (see Table 4.3). Refer to Table 4.4 for amounts retrieved post subtraction.

Linear amplification of RNA and cDNA synthesis

cDNA was generated from all RNA samples as previously described (Frias-Lopez et al. 2008; Shi et al. 2009). Based on the Invitrogen™ MessageAmp™II-Bacteria kit (Life Technologies, Carlsbad, CA, USA), the 3'-end of rRNA-subtracted RNA was first polyadenylated, then reverse transcribed to single-stranded cDNA by the ArrayScript reverse transcriptase (Roche, Indianapolis, IN, USA), which is primed by the polyT oligo in the primer T7BpmIdT$_{16}$VN. Second-strand synthesis was then performed by DNA polymerase and purified, as stipulated in the MessageAmp™II-Bacteria kit. The purified double stranded cDNA was then linearly amplified by random-primed T7-RNA polymerase in-vitro transcription, following instructions in SuperScript® double stranded cDNA synthesis kit (Life Technologies, Carlsbad, CA, USA).

The amplified RNA (aRNA) was purified using the MessageAmp™II-Bacteria kit. After purification of aRNA it was reverse-transcribed to double stranded cDNA, which was purified using the QIAquick® pcr purification kit (Qiagen, Germantown, MD, USA). The polyA tail used to prime reverse-transcription was removed by BpmI restriciton enzyme digestion.

cDNA and DNA sequencing

The produced cDNA was sequenced using a GS FLX Titanium pyrosequencing platform (454 Life Sciences, Roche, Branford CT, USA). For gDNA samples, library preparation was performed per the manufacturer's recommendation in the Titanium Rapid Library Preparation protocol (454 Life Sciences). For cDNA samples, library preparation was performed with a modified size-selection protocol optimized to capture smaller cDNA fragments (300 bp and up). The selection of smaller fragments was achieved by performing the size-selection step using AMPure XP beads with undiluted adaptor-ligated sample. Except for this slight modification, the remainder of the library preparation followed manufacturer recommendations. Libraries were then quantified using the Slingshot™ kit (Fluidigm, South San Francisco, CA, USA) and added to beads for emulsion PCR at a concentration of roughly 0.1 molecules per bead. Sequencing and quality control were performed following the manufacturer's recommendation. All nine cDNA samples were sequenced on a half plate. DNA time-initial was sequenced on a full plate of a Titanium 454 Sequencer (454 Life Sciences, Roche, Branford, CT, USA). Each of the four time-final DNA samples was sequenced on a half-plate of a Titanium 454 Sequencer. Sequencing data statistics are summarized in Table 4.6.

Table 4.2. Set of primers used for processing RNA samples. The underlined portion of the primers indicates the T7-RNA polymerase promoter site. See (DeLong et al. 1999) for design of T7-appended (promoter underlined above) primers and (Hunt et al. 2006) for bacterial 23S primers.

Step	Primer Description	Primer
PCR amplification of ribosomal RNA genes in the DNA sample	Eubacterial 16S Forward	*AGAGTTTGATCCTGGCTCAG*
	Eubacterial 16S Reverse: includes T7 RNA polymerase promoter site.	*GCCAGTGAATTGTAATACGACTCACTAT AGGGACGGCTACCTTGTTACGACTT*
	Eubacterial 23S Forward	*GAASTGAAACATCTHAGTA*
	Eubacterial 23S Reverse: includes T7 RNA polymerase promoter site.	*GCCAGTGAATTGTAATACGACTCACTAT AGGGCGACATCGAGGTGCCAAAC*
	Archaeal 16S Forward	*TCCGGTTGATCCYGCCGG*
	Archaeal 16S Reverse: includes T7 RNA polymerase promoter site.	GCCAGTGAATTGTAATACGACTCACTAT AGGGGGYYACCTTGTTACGACTT
	Archaeal 23S Forward	*ASAGGGTGAHARYCCCGTA*
	Archaeal 16S Reverse: includes T7 RNA polymerase promoter site.	*GCCAGTGAATTGTAATACGACTCACTAT AGGGCTGTCTCRCGACGGTCTRAACCCA*
Reverse transcription of poly-adenylated RNA	T7BpmIdT$_{16}$VN: includes T7 RNA polymerase promoter site, BpmI restriction enzyme recognition site, poly(T) tail.	*5'-GCC AGT GAA TTG TAA TAC GAC TCA CTA TAG GGG CGA CTG GAG TTT TTT TTT TTT TTT TVN-3'*

Table 4.3. Summary of Probe Sets Used for the Sample Specific rRNA Subtraction

	DNA template	Gene targeted	Amount of DNA template	Amount of rRNA gene amplicons	Amount of RNA probe synthesized
Probe Set 1 (PS1)	T0h	Eubacterial 16S rRNA	4 reactions x 100ng; pooled during PCR purification	12.5 µg	84.5 µg
		Eubacterial 23S rRNA	4 reactions x 100ng; pooled during PCR purification	11.2 µg	77.7 µg
		Archaeal 16S rRNA	4 reactions x 100ng; pooled during PCR purification	11.5 µg	75.6 µg
		Archaeal 23S rRNA	4 reactions x 100ng; pooled during PCR purification	12.4 µg	84.2 µg
Probe Set 2 (PS2)	4 templates: T36_+OM (A and C replicates), T36_OM+Iron (A and C replicates)	Eubacterial 16S rRNA	4 reactions x 100ng; each reaction with one of the 4 templates; pooled during PCR purification	10.1 µg	81.1 µg
		Eubacterial 23S rRNA	4 reactions x 100ng; each reaction with one of the 4 templates; pooled during PCR purification	3.7 µg	100.2 µg
		Archaeal 16S rRNA	4 reactions x 100ng; each reaction with one of the 4 templates; pooled during PCR purification	4.9 µg	27.4 µg
		Archaeal 23S rRNA	4 reactions x 100ng; each reaction with one of the 4 templates; pooled during PCR purification	7.9 µg	78.9 µg

Table 4.4. Summary of amounts of DNA/RNA retrieved or used at each step.

Sample	Total DNA extracted	Total RNA extracted	RNA used for subtraction	Probe Set used for subtraction	aRNA post linear amplification	aRNA used for cDNA synthesis	Total cDNA synthesized
T0h	8.1 µg	154 ng	100 ng	PS1	38.5 µg	5 µg	1.97 µg
T12h- Control	n/a	216 ng	130 ng	PS1	60.5 µg	5 µg	1.22 µg
T12h_+Iron	n/a	163 ng	100 ng	PS1	36.9 µg	5 µg	1.29 µg
T12h_+OM	n/a	303 ng	130 ng	PS2	70 µg	5 µg	495 ng
T12h_+OM+Iron	n/a	224 ng	130 ng	PS2	68 µg	5 µg	1.58 µg
T36h- Control	586 ng	175 ng	75 ng	PS1	47.2 µg	5 µg	1.09 µg
T36h_+Iron	730 ng	162 ng	100 ng	PS1	51.4 µg	5 µg	1.17 µg
T36h_+OM	802 ng	628 ng	130 ng	PS2	10 µg	5 µg	291 ng
T36h_+OM+Iron	946 ng	357 ng	75 ng	PS2	20 µg	10 µg	299 ng

4.3.7 Bioinformatic Analysis

Removal of duplicate reads

Prior to analysis, duplicate reads thought to be artifacts of the pyrosequencing technology (Huse et al. 2007) were removed. The software CD-HIT (Li & Godzik 2006) was used to identify duplicate sequences (100% identity over 90% to 100% length) in both the DNA and cDNA datasets.

Identification and removal of ribosomal RNA (rRNA) reads

Ribosomal RNA reads were identified in both the cDNA and DNA datasets by BLASTn against the ARB SILVA SSU and LSU release 104 (October 2010) database (Pruesse et al. 2007) comprising available gene sequence information for eukaryotic, bacterial and archaeal small and large subunits (16S, 23S, 18S, 28S). Reads producing an alignment with a bit score greater than 50 were considered to be a ribosomal RNA sequence and removed from the DNA and cDNA datasets. Many of the following analyses were adapted from previously described approaches (Ottesen et al. 2011; Ottesen et al. 2012).

Identification of functional genes

Non-rRNA reads were compared to NCBI's non-redundant (nr) protein database (downloaded Jan. 5th, 2011) using BLASTx (Altschul et al. 1990) for functional gene analyses. Reads (query) generating an alignment with a bit score of 50 or greater were retained for further analysis. When a query mapped to multiple hits in the database with a bit score >50, the read was assigned to the hit with the highest bit score (top hit). When a query had multiple top hits with the same bit score (equal top hits), the read was assigned to the hit with the more frequent representation across all datasets (all samples). This single-top-hit file was utilized in the "Identification of organism specific reads" section as well as in.

Taxonomy assignments

Taxonomy assignments to profile the change in community composition across treatments and time were performed with taxonomic information associated to protein-coding genes. The taxonomic information (NCBI tax ID) associated with all top hits and all equal top hits was retained and counts were compiled for the different levels of the NCBI taxonomy (superkingdom down to the species level). When

one read had equal top hits, one count was assigned to the highest common taxonomic level between the hits. For example, if one read had three equal top hits, one with associated taxonomic information to *Prochlorococcus marinus* MIT 9515 and two to *Synechococcus elongatus* PCC6301, one count would be added to the Cyanobacteria level. For relative representation of taxonomic groups (Figure 4.5), counts were normalized to the total number of reads with a significant hit to the NCBI nr database. The community composition is established based on the taxonomy assignments of the protein coding genes identified, thus possibly missing groups that are not well represented in the NCBI database. However, previous studies performed at station ALOHA have shown that this approach gives very similar results to one based on the identification of 16S ribosomal RNA reads (Shi et al. 2011).

Identification of Organism Specific Reads

In addition to a community overview, an organism centric approach was taken to analyze the sequencing data. Given an organism of interest (identified by its NCBI taxonomy ID), all reads with at least one top hit (in the case of reads with equal top hits) to the organism of interest were retrieved. This approach was taken to retrieve reads mapping to *Prochlorococcus* (genus level taxID: 1218), SAR11 cluster (taxID: 54526), *Alteromonas* (genus level, taxID: 226), *Pseudoalteromonas* (genus level, taxID: 53246), SAR116 cluster (taxID: 62654), *Psychroflexus torquis* (species level: 57029).

Generation of Organism Specific Clusters of Orthologous Genes

In a simple BLAST against the NCBI non-redundant protein database, reads map to the closest sequenced relative in the database. In reality, meta-omics samples are composed of organisms whose genomes are likely to be combinations of these sequenced genomes. Hence, the abundance of a transcript in a sample may be underestimated because counts are spread across the sequenced relatives. To account for this, clusters of orthologous genes were generated as previously described (Ottesen et al. 2012) for the most abundant taxon groups: *Prochlorococcus*, SAR11 cluster, *Alteromonas, Pseudoalteromonas,* SAR116 cluster, *Psychroflexus torquis.* Translated protein-coding sequences from all available genomes for a taxon group of interest (ie. 13 available genomes for *Prochlorococcus*) were retrieved and compared by BLASTp (all against all). All hits with an e-value $\leq 10^{-5}$, at least 30% amino acid identity and an alignment length of 80% of the

longest sequence were considered to be significant. One cluster consisted of the reciprocal best hits for all pairwise genome comparisons. Using the automated KEGG annotation server, KAAS (http://www.genome.jp/tools/kaas/), protein-coding genes for each genome were associated with a KEGG pathway for further analysis (Kanehisa & Goto 2000). Table 4.5 summarizes which genomes were used for the generation of orthologous clusters.

Table 4.5 Summary of Organism Centric Clusters Generation.

Organism	Genomes used to generate clusters	NCBI Accession number for genome	Number of protein coding genes	Number of unique genes	Number of Clusters of Orthologous Genes*	Number of Clusters of Orthologous genes with K0 assignment
Alteromonas (genus-level)	*Alteromonas macleodii* ATCC 27126	NZ_ABQB01000000	4,396	1,492	7,215	2,651
	Alteromonas macleodii Deep Ecotype	NC_011138.2	4,084	1,022		
	Alteromonas sp. SN2	NC_015554.1	4,371	1,365		
Pseudo-alteromonas (genus level)	*Pseudoalteromonas sp.* SM9913	NC_014803.1	3,712	684	10,573	2,726
	Pseudoalteromonas atlantica T6c	NC_008228.1	4,281	1736		
	Pseudoalteromonas haloplanktis TAC125	NC_007481.1	3,485	513		
	Pseudoalteromonas haloplanktis ANT505	NZ_ADOP00000000	4,127	884		
	Pseudoalteromonas tunicata D2		4,504	3,201		
SAR11- cluster	Alpha proteobacterium HIMB114	NZ_ADAC00000000.1	1,425	348	2,762	1,350
	Candidatus Pelagibacter ubique HTCC1002	NZ_AAPV00000000.1	1,393	93		
	Candidatus Pelagibacter ubique HTCC1062	NC_007205.1	1,354	50		
	Candidatus Pelagibacter sp. HTCC7211	NZ_ABVS00000000.1	1,447	310		
	Candidatus Pelagibacter sp. IMCC9063	NC_015380.1	1,447	408		
SAR116-cluster	*Candidatus puniceispirullum marinum* IMCC1322	NC_014010.1	2,543	N/A	2,543	1,465
Psychroflexus torquis	*Psychroflexus torquis* ATCC 700755	AAPR00000000	3,788	N/A	N/A	716
Prochlorococcus	*Prochlorococcus marinus str.* AS9601	NC_008816.1	1,920	85	6,041	1,450
	Prochlorococcus marinus str.	NC_009091.1	1,906	95		

	MIT9301					
	Prochlorococcus marinus str. NATL1A	NC_008819.1	2,193	162		
	Prochlorococcus marinus str. NATL2A	NC_007335.2	2,162	140		
	Prochlorococcus marinus str. MIT9211	NC_009976.1	1,854	239		
	Prochlorococcus marinus str. CCMP1375 / SS120	NC_005042.1	1,883	203		
	Prochlorococcus marinus str. MIT9515	NC_008817.1	1,905	170		
	Prochlorococcus marinus str. CCMP1986/ MED4	NC_005072.1	1,717	78		
	Prochlorococcus marinus str. 9312	NC_007577.1	1,810	101		
	Prochlorococcus marinus str. MIT9202	NZ_ACDW00000000.1	1,890	209		
	Prochlorococcus marinus str. MIT9215	NC_009840.1	1,982	126		
	Prochlorococcus marinus str. MIT 9313	NC_005071.1	2,269	339		
	Prochlorococcus marinus str. MIT9303	NC_008820.1	2,997	1,015		

*The number of orthologous genes corresponds to the number of unique ortholog clusters identified by the reciprocal best blast hit approach described in the text.

Table 4.6. Summary of sequencing statistics for gDNA and cDNA samples.

Type	Sample	Total # of reads	Average read length	% duplicates	% rRNA	% non-rRNA reads to NR
gDNA	T0h	1,408,765	404	0.15	0.43	80
	T36h Control	500,370	310	0.02	0.43	71
	T36h +Fe	541,049	323	0.03	0.35	70
	T36h +OM	781,710	437	0.33	0.56	79
	T36h +OM+Fe	586,879	405	0.18	0.48	75
cDNA	T0h	565,124	320	1.4	33	67
	T12h Control	662,379	316	1.4	17	66
	T12h +Fe	562,195	321	2.1	27	65
	T12h +OM	636,308	340	0.8	14	72
	T12h +OM+Fe	639,023	307	0.7	13	70
	T36h Control	546,764	346	4.1	36	61
	T36h +Fe	473,815	297	0.5	25	71
	T36h +OM	486,980	307	2.1	50	68
	T36h +OM+Fe	486,980	307	2.1	50	68

4.3.8 Statistical Analysis

In this analysis we ask three questions: (i) How does the addition of iron alone impact microbial community gene expression? (ii) How does the addition of organic matter alone impact microbial community gene expression and (iii) What is the interaction (synergistic or antagonistic) effect on microbial community gene expression of the combined addition of iron and organic matter? We use an empirical Bayesian approach, baySeq (Hardcastle & Kelly 2010), to determine what genes/ortholog clusters (defined in Section 4.3.7) are significantly differentially expressed (DE) in the 4 different treatments (1: control, 2: +Fe, 3: +OM, 4: OM + Fe). Because we are working with metatranscriptomic data, 'gene expression' formally refers to relative transcript abundance and 'differential gene expression' to 'differences in relative transcript abundances'.

Assuming that the sequencing count data N observed for each gene transcript/ortholog cluster *i* in each treatment *j* follows a Poisson distribution with mean λ_{ij},: $N_{ij} \sim Poisson(\lambda_{ij})$ (Marioni et al. 2008), we use a loglinear Poisson regression model to determine the extent of the interaction between iron and organic matter (synergy or antagonism) and the significance of the marginals (iron effect, organic matter effect) for the following 2x2 contingency table (Rodriguez 2007):

		Organic Matter -	Organic Matter +
Iron	-	Treatment 1: Control	Treatment 3: OM
	+	Treatment 2: Fe	Treatment 4: OM + Fe

Here, we define 15 models (Table 4.7) that describe how parameters λ_1, λ_2, λ_3, λ_4 each describing the Poisson distribution of count data for a given gene and treatment (treatment 1: control, treatment 2: +Fe, treatment 3: +OM, treatment 4: +OM + Fe) differ from each other. For example, in the model DE. 12, parameters λ_1, and λ_2, describing the distributions of count data for a gene/cluster in treatments 1 and 2 are assumed to be similar to each other but different from λ_3 and λ_4, describing the distributions of count data for a gene/cluster in treatments 3 and 4 (Table 4.7). Using baySeq, we can calculate the posterior

likelihood that the observed count data for a gene is described by one of the 15 models. The baySeq method empirically derives prior probabilities from the data (100,000 iterations). The baySeq analysis was performed to determine significant differential expression between treatments, separately for each time point and for organism specific clusters (defined above).

Table 4.7. Models of differential expression (DE) for which posterior probabilities were computed with baySeq. Samples are referred to by treatment: 1: Control, 2: with iron, 3: with organic matter, 4: with organic matter and iron. Models are defined by the equality or inequality of the parameters λ that describe the distribution of the count data for a given gene/ortholog across treatments. The third column is a visual representation of these parameter comparisons.

Model of DE	Parameter comparison	Ctrl	Fe	OM	OM +Fe
Non-DE	$\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4$				
DE.1	$\lambda_1 \neq \lambda_2 = \lambda_3 = \lambda_4$				
DE.2	$\lambda_2 \neq \lambda_1 = \lambda_3 = \lambda_4$				
DE.3	$\lambda_3 \neq \lambda_1 = \lambda_2 = \lambda_4$				
DE.4	$\lambda_4 \neq \lambda_1 = \lambda_2 = \lambda_3$				
DE.12	$\lambda_1 = \lambda_2 \neq \lambda_3 = \lambda_4$				
DE.23	$\lambda_2 = \lambda_3 \neq \lambda_1 = \lambda_4$				
DE.13	$\lambda_1 = \lambda_3 \neq \lambda_2 = \lambda_4$				
DE.1.2.34	$\lambda_1 \neq \lambda_2 \neq \lambda_3 = \lambda_4$				
DE.2.3.14	$\lambda_2 \neq \lambda_3 \neq \lambda_1 = \lambda_4$				
DE.1.3.24	$\lambda_1 \neq \lambda_3 \neq \lambda_2 = \lambda_4$				
DE.3.4.12	$\lambda_3 \neq \lambda_4 \neq \lambda_1 = \lambda_2$				
DE.1.4.23	$\lambda_1 \neq \lambda_4 \neq \lambda_2 = \lambda_3$				
DE.2.4.13	$\lambda_2 \neq \lambda_4 \neq \lambda_1 = \lambda_3$				
DE.1.2.3.4	$\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_4$				

Using the loglinear model of Poisson distributed data, we can sum the posterior probabilities obtained with baySeq to gain biological insight, namely whether the pattern of gene expression for each gene/ortholog cluster studied significantly reflects an iron effect, an organic matter effect, or a compounded iron and organic matter (interaction) effect (Table 4.7 and description of sum of models

below). Clusters with a summed posterior probability greater than 0.9 for a given effect were considered to be significantly impacted by that effect.

The loglinear Poisson regression model states that

$$\ln(\lambda_i) = \alpha_0 X_{i0} + \alpha_1 X_{i1} + \alpha_2 X_{i2} + \alpha_3 X_{i3} \quad (1)$$

where λ_i is the mean of count data for treatment i (i=1: control, 2: +Fe, 3: +OM, 4: +OM+Fe);

α_0, α_1, α_2, α_3 is the regression coefficient for the effect (α_0 describes the control effect, α_1 describes the iron effect, α_2 describes the organic matter effect and α_3 describes the interaction effect).

X is a design matrix.

Hence, we can describe each parameter λ_1, λ_2, λ_3, λ_4, as:

$$\begin{aligned} \ln(\lambda_1) &= \alpha_1 \\ \ln(\lambda_2) &= \alpha_1 + \alpha_2 \\ \ln(\lambda_3) &= \alpha_1 + \alpha_3 \\ \ln(\lambda_4) &= \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 \end{aligned} \quad (2)$$

If $\lambda_1 = \lambda_2$, then we can assume that for the gene of interest the distribution of count data in the iron treatment is not different from the distribution of count data in the control treatment. Hence, if $\lambda_1 = \lambda_2$, then $\log(\lambda_1)=\log(\lambda_2)$, which implies that $\alpha_2 = 0$. Therefore, an iron effect is described by $\alpha_2 \neq 0$. Similarly, we can conclude that an organic matter effect is described by $\alpha_3 \neq 0$, and an interaction effect by $\alpha_4 \neq 0$.

By solving the system of equations (2) for the constraints on λ listed in table X, we can determine whether each DE model listed contributes to an iron effect ($\alpha_2 \neq 0$), an organic matter effect ($\alpha_3 \neq 0$) or an interaction effect ($\alpha_4 \neq 0$). The posterior probabilities calculated by baySeq for the 15 DE models can then be summed according to the scheme presented in Table 4.8 to infer biological meaning. Summed posterior probabilities greater than 0.9 are considered to be significant for that effect.

Table 4.8. Summing of models. To determine whether one of four effects, iron effect, organic matter effect, iron and organic matter interaction effect or no effect drives the pattern of differential expression for a given ortholog, the posterior probabilities obtained for the different models tested with baySeq were summed according to the table below. Crosses in each column indicate which of the posterior probabilities from the 15 models need to be summed to obtain the probability that an ortholog is differentially expressed according to the corresponding scenario in the column header.

	Iron Effect	Organic Matter Effect	Interaction Effect	No Effect
Non-DE				X
DE.1	X	X	X	
DE.2	X		X	
DE.3		X	X	
DE.4			X	
DE.12		X		
DE.23	X	X	X	
DE.13	X			
DE.1.2.34	X	X	X	
DE.2.3.14	X	X	X	
DE.1.3.24	X	X	X	
DE.3.4.12		X	X	
DE.1.4.23	X	X	X	
DE.2.4.13	X		X	
DE.1.2.3.4 (DE.all)	X	X	X	

Expressing α_2 (iron effect), α_3 (organic matter effect), α_4 (interaction effect) in terms of $\ln(\lambda_1)$ allows us to calculate the extent of each effect (iron, organic matter or interaction) with respect to the control (see Figure 4.6):

$$\alpha_2 = \ln(\lambda_2/\lambda_1)$$

$$\alpha_3 = \ln(\lambda_3/\lambda_1)$$

$$\alpha_4 = \ln(\lambda_4) - \ln(\lambda_1) - \ln(\lambda_2/\lambda_1) - \ln(\lambda_3/\lambda_1)$$

where λ_i is the mean for the count data of a given gene/cluster in treatment *i* (number of hits to the cluster normalized by the library size for that treatment. Library size is the total number of organism specific hits retrieved in the sample).

4.4. Results

4.4.1 Iron and Organic Matter Affect Growth and Overall Community Structure in Different Ways

In this experiment, the addition of iron alone, which resulted in a 6-fold increase in the total iron concentration, from 0.25 nM to 1.5 nM (Figure 4.2) seemed to have little effect on microbial growth and the overall structure of the microbial community. Iron did not elicit detectable growth stimulation since no increase in total cell numbers was observed in either the (+Fe) treatment nor the (OM+Fe) treatment (Figure 4.4). Furthermore, the microbial community composition at T0h, very typical of what is usually found in surface waters at HOT station ALOHA, dominated by small-celled picocyanobacteria of the genus *Prochlorococcus* (here representing between 20% and 45% of the community) and bacterioplankton of the genus Ca. *Pelagibacter* (here representing between ~15 and 25% of the community) (DeLong et al., 2006), remained similar in both the (Ctrl) and the (+Fe) treatment at T36h (Figure 4.5). Interestingly, the microbial community composition in the (Ctrl) and (+Fe) treatments was the same regardless of whether it was assessed by genomic (gDNA) or transcriptomic (cDNA) reads. This suggests that the relative abundance of transcripts parallels the relative abundance of genes, and opens the possibility to explore nuanced changes in gene expression patterns between control and iron treatments without the challenge of accounting for drastic changes in abundances of specific microorganisms.

In contrast, organic matter dramatically affected the microbial community, both in terms of microbial growth and community composition. The addition of organic matter resulted in a 2-fold increase in total organic carbon concentration, from ~90 μM to ~210 μM (Figure 4.3). Such an influx of organic carbon (presumably accompanied by other organic nutrient sources) is likely to have notable effects on the nutrient-starved microbial community (Eilers et al. 2000). Indeed, an overwhelming response was observed both in terms of changes in bulk cell numbers as well as in terms of changes in community composition. The number of SYBR® stained cells (as a proxy for the number of heterotrophic cells) doubled in both the organic matter (+OM) and organic matter + iron (+OM+Fe) treatments (Figure 4.4). The growth stimulation was not observed until at least 12 hours of incubation, but it accelerated greatly thereafter (Figure 4.4). These results are mirrored by a drastic change in community composition observed in the

(+OM) and (+OM + Fe) cDNA and gDNA T36h samples (Figure 4.5), where reads mapping to the *Pseudoalteromonas, Alteromonas* and *Alteromadales* genera (combined) accounted for up to 25% (gDNA) of the community and 55% of the transcripts (cDNA T36h OM treatment) as opposed to less than 2% in the control (Ctrl) and iron (+Fe) samples. These results agree with previous observations performed at station ALOHA, where addition of dissolved organic matter (DOM) or nutrient-rich deep-seawater resulted in an enrichment of the opportunistic, fast-growing copiotrophs of the genus *Alteromonas* (McCarren et al., 2010; Shi et al., 2011). One notable difference between the findings presented here and those from these previous experiments is the magnitude of the enrichment as well as the enrichment specifically of *Pseudoalteromonas.*

While iron did not have any effect on the overall microbial community structure and composition (Figure 4.4 and Figure 4.5), we examined whether changes occurred at the gene expression level with an in-depth exploration of gene expression patterns at the organism level, examining both the most abundant autotrophs (Section 4.4.2) and heterotrophs (Section 4.4.3).

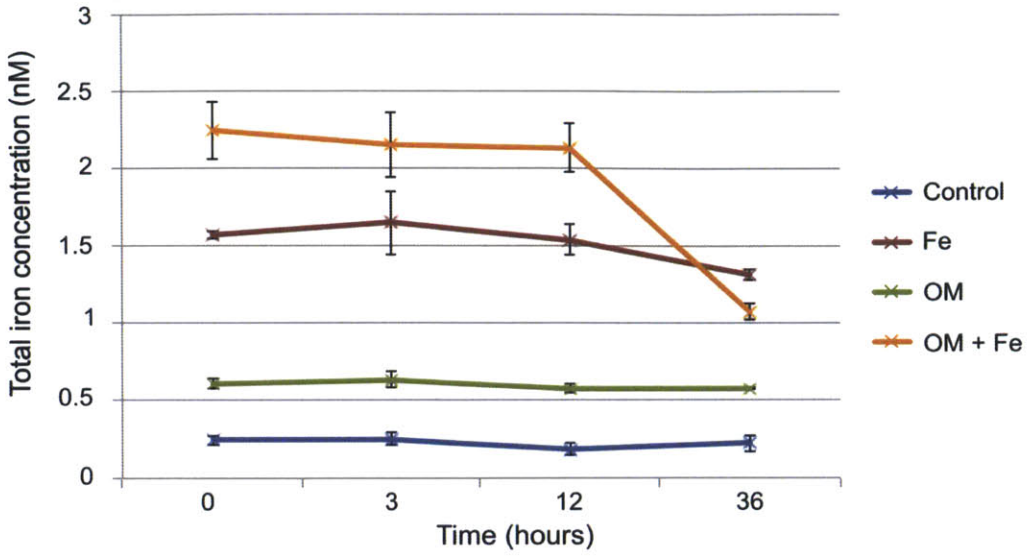


Figure 4.2. Total Iron Concentrations Shows No Contamination of Samples. Total iron concentrations were measured as described in Section 4.3. Error bars represent one standard deviation away from the average value of 3 biological replicates. Measured concentrations are consistent with what would be expected from the experimental set up: the highest iron concentrations are observed for the (+Fe) and the (OM+Fe) samples. The added organic matter seems to carry over a residual ~0.25nM Fe. The sudden drop at T36h in the OM+Fe sample could be attributed to adsorption onto the walls of the incubation bottle.

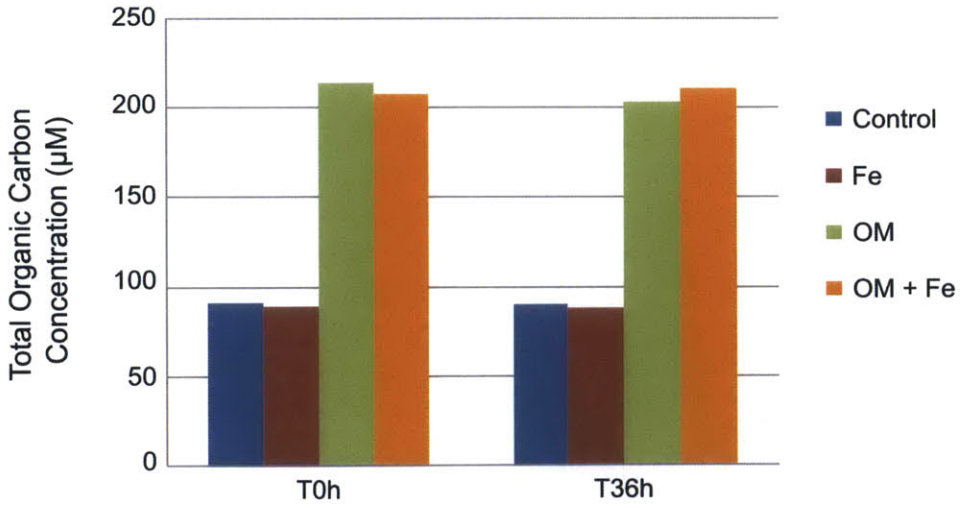


Figure 4.3. Addition of organic matter results in a two-fold increase in total organic carbon concentration (TOC). TOC concentrations were measured as described in Section 4.3. Addition of organic matter leads to a two-fold increase in TOC.

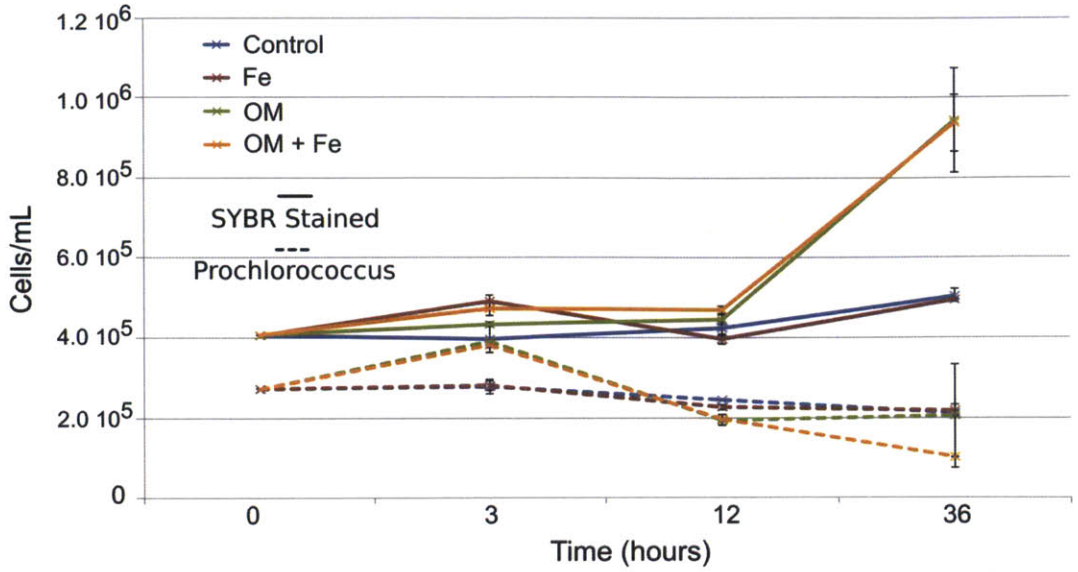


Figure 4.4. Organic matter addition leads to doubling of SYBR® stained cells. Bulk community cell counts as measured by flow-cytometry. Control: no addition, Fe: with iron, OM: with organic matter, OM + Fe: with organic matter and iron. The continuous lines represent counts for cells stained with the universal DNA stain, SYBR® Gold. The dotted lines represent counts for cells identified as *Prochlorococcus* on the red fluorescence channel. Counts were normalized to standard bead fluorescence. Counts for *Prochlorococcus* and *Synechococcus* (not shown) were subtracted from the total number of stained cells counts. Error bars represent one standard deviation from the mean for biological triplicates.

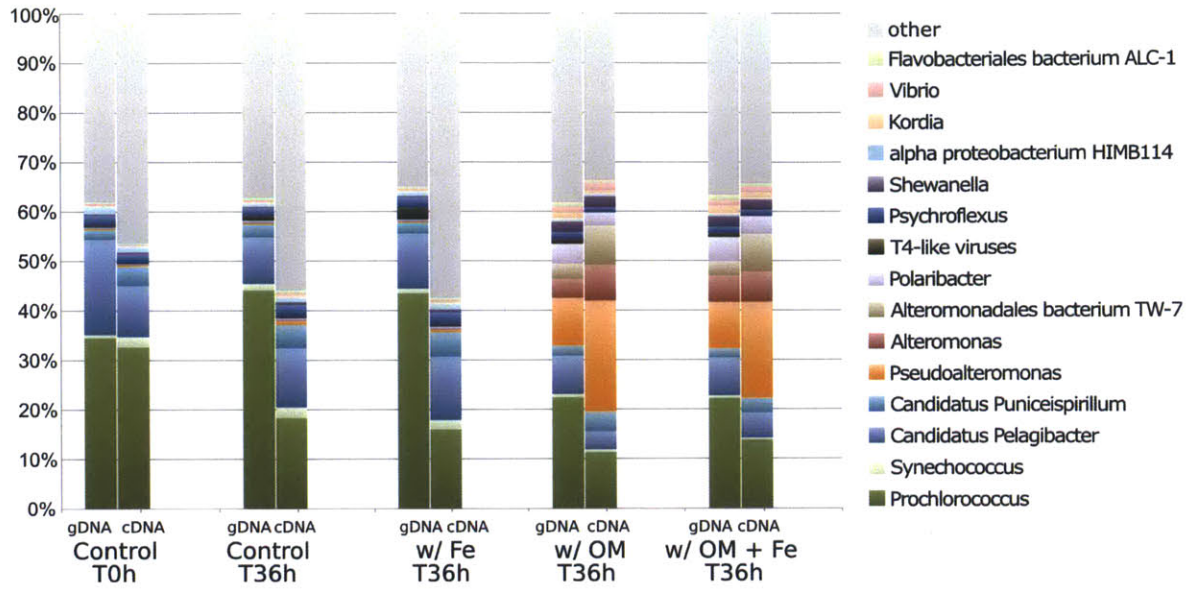


Figure 4.5. Organic matter addition leads to drastic changes in community composition. Genus-level microbial community composition assessed by protein coding sequences in both gDNA and cDNA datasets. Reads with significant matches (bit score > 50) to the NCBI nr-protein database (blastx) were associated to their corresponding taxonomy ID and binned at the appropriate genus level. The genera with at least 1% representation in at least one of the datasets are indicated, all others are grouped in "other".

4.4.2. Organism Specific Patterns of Gene Expression

Organism specific clusters of orthologous genes were generated as described in Section 4.3 for the most abundant members of the community (*Prochlorococcus marinus, SAR11* cluster including Ca. *Pelagibacter ubique*, *SAR116* cluster including Ca. *Puniceispirillum*, *Psychroflexus torquis, Pseudoalteromonas* genus, *Alteromonas* genus- see Figure 5). cDNA reads were assigned to each cluster of orthologous genes on the basis that the read had a significant hit (bit score >50) to one of the genes belonging to the cluster of interest. The statistical software package baySeq (Hardcastle & Kelly 2010) was used to identify the organism specific clusters that displayed a differential expression pattern indicative of a statistically significant organic matter effect, iron effect or a combined organic matter and iron effect (interaction effect). This statistical approach applied separately to each of the most abundant microorganisms revealed that the number of orthologous gene clusters significant for a given effect was highest for the organic matter effect (Figure 4.6). We also notice that most organisms displayed an interaction effect and to a much lesser extent an iron alone effect (as evaluated by the number of clusters for which the pattern of differential expression is significantly influenced by a type of effect).

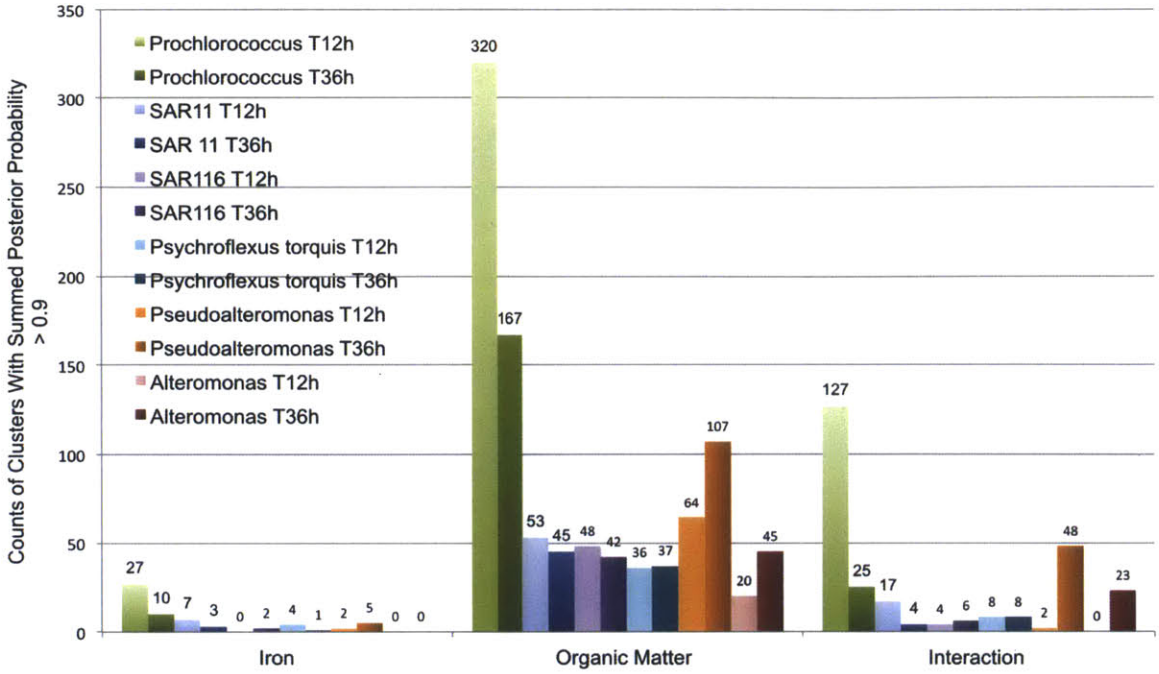


Figure 4.6. Organism-specific orthologous gene clusters with a significant iron effect, organic matter effect or interaction effect. Significance is set as a summed posterior probability greater than 0.9. Calculations of posterior probabilities were performed for 15 different models of differential expression using baySeq (see Section 4.3). Sums to obtain the 4 different effects were performed according to Table 4.8 (see Section 4.3). Numbers at the top of each bar indicate the number of orthologous gene clusters per organism per time point for which the pattern of gene expression across treatments is significant for the effect indicated on the X-axis. The strongest effect on overall gene expression seems to be due to the organic matter addition, as evaluated by the number of clusters for which the pattern of differential expression is significantly influenced by a type of effect. These numbers do not differentiate between the magnitude of the effect nor whether genes are up or down-regulated.

4.4.2.1 Prochlorococcus Displays a Subtle Response to Iron Despite an Overwhelming Response To Organic Matter

Iron Effect –While the addition of iron (both in the (+Fe) and the (OM+Fe) treatments) did not lead to a measurable effect on *Prochlorococcus* cell numbers and relative abundance of metagenomic sequence reads, we suspected that a detailed analysis of *Prochlorococcus* gene expression could nonetheless reveal iron-specific effects. Indeed, the effect of iron on *Prochlorococcus* gene expression, recently investigated in cultures of the high-light adapted *P. marinus* MED4 and of the low-light adapted *P. marinus* MIT9313 (Thompson et al. 2011), revealed that iron-stress (and rescue) affected the expression of over 100 genes in each strain. Furthermore, because the "total cell numbers" parameter as measured by flow-cytometry reflects the net effect of growth and death processes (Landry et al. 2000), we cannot discount the possibility that cell growth was nonetheless occurring despite a constant total number of *Prochlorococcus* cells measured across time in the (+Fe) treatment (Figure 4.4).

The statistical analysis performed revealed that a significant iron effect impacted the relative transcript abundance level of 27 *Prochlorococcus* orthologous gene clusters at T12h (9 increased, 18 decreased) and 10 (5 increased, 5 decreased) at T36h (Figure 4.6). Table 4.9 and Table 4.10 list these clusters. Inspection of the predicted functions of these clusters shows that clusters significantly affected by the iron amendment belong to a number of different pathways: nitrogen and phosphorus transport, oxidative stress response, photosynthesis. Interestingly, not one single pathway is overly represented. None of the listed *Prochlorococcus* clusters have previously been described as being affected by changes in iron concentration (Thompson et al. 2011). This could suggest that the iron-specific *Prochlorococcus* response *in situ* is very different than in culture.

Table 4.9. List of *Prochlorococcus* orthologous gene cluster for which the pattern of gene expression is significantly impacted by iron at T12h. The list comprises the 27 orthologous gene clusters for which the pattern of gene expression across all 4 samples is indicative of an iron effect as assessed by the baySeq analysis described in Section 4.3. Significance is determined for a sum of posterior probabilities greater than 0.9. α_2 denotes the magnitude of the iron effect- positive means that the relative abundance of transcripts that belong to a specific cluster increases in the (+Fe) treatment, negative means that it decreases in the (+Fe) treatment. The α_2 value is calculated as the log ratio of the relative transcript abudance value in the (+Fe) sample to the relative transcript abundance value in the (Ctrl) sample.

Orthologous cluster #	List of Accession Numbers of Genes Populating the Cluster	Gene Name and Description	α_2
Cluster 219	NP_892628.1 YP_001008960.1 YP_001010890.1 YP_001090761.1 YP_001483793.1 YP_397008.1 ZP_05137812.1	*rnhA*, RNase HI	1.37
Cluster 566	NP_892628.1 YP_001008960.1 YP_001010890.1 YP_001090761.1 YP_001483793.1 YP_397008.1 ZP_05137812.1	Putative RNA-dependent reverse transcriptase	0.79
Cluster 961	NP_875331.1 NP_893014.1 NP_894536.1 YP_001009356.1 YP_001011292.1 YP_001014809.1 YP_001017526.1 YP_001091187.1 YP_001484197.1 YP_001550752.1 YP_291508.1 YP_397400.1 ZP_05138559.1	DnaJ2 molecular chaperone	0.47
Cluster 604	NP_874943.1 NP_892667.1 NP_895036.1 YP_001008998.1 YP_001010929.1 YP_001014430.1 YP_001016823.1 YP_001090799.1 YP_001483831.1 YP_001550436.1 YP_293069.1 YP_397046.1 ZP_05139048.1	Carboxysome shell protein CsoS1	0.43
Cluster 4983	NP_875924.1 NP_875966.1 NP_892828.1 NP_894341.1 NP_894824.1 YP_001009167.1 YP_001011533.1 YP_001014975.1 YP_001017122.1 YP_001017766.1 YP_001090998.1 YP_001091463.1 YP_001484009.1 YP_001550570.1 YP_291636.1 YP_397218.1 ZP_05138012.1 AAC45381	*pstS*, phosphate ABC transporter- periplasmic protein	0.36
Cluster 5855	YP_001015925.1	Hypothetical protein NATL1_21051	0.31
Cluster 286	NP_874689.1 NP_892382.1 NP_895680.1 YP_001008680.1 YP_001010612.1 YP_001014170.1 YP_001018476.1 YP_001090510.1 YP_001483490.1 YP_001550175.1 YP_292820.1 YP_396762.1 ZP_05138137.1 AAF15904 ABE11504	High affinity ammonium transporter	0.19
Emb|CAX32370.1|	Emb|CAX32370.1|	Hypothetical protein PMT_2851	0.18
Cluster 1816	NP_876162.1 NP_893728.1 NP_893993.1 YP_001010211.1 YP_001012113.1 YP_001015883.1 YP_001016217.1 YP_001092028.1 YP_001485084.1 YP_001551621.1 YP_292381.1 YP_398201.1 ZP_05138829.1	*thiC*, Thiamine biosynthesis protein	0.12
Cluster 1319	NP_875635.1 NP_893265.1 NP_894997.1 YP_001009714.1 YP_001011627.1 YP_001015418.1 YP_001016875.1 YP_001091562.1 YP_001484554.1 YP_001551107.1 YP_291950.1 YP_397740.1 ZP_05138058.1 ABE11002	DNA-binding domain containing protein	-0.13
Cluster 1488	NP_875759.1 NP_893411.1 NP_894173.1 YP_001009883.1 YP_001011769.1 YP_001015534.1 YP_001017980.1 YP_001091703.1 YP_001484721.1 YP_001551226.1 YP_292055.1 YP_397886.1 ZP_05139142.1	Putative nickel containing superoxide dismutase precursor	-0.21
Cluster 1331	NP_875647.1 NP_893275.1 NP_895010.1 YP_001009726.1 YP_001011639.1 YP_001015430.1 YP_001016858.1 YP_001091574.1 YP_001484565.1 YP_001551119.1 YP_291962.1 YP_397752.1 ZP_05138193.1 ABE10996	*psbC*, Photosystem II PsbC protein	-0.21
Cluster 875	NP_892928.1 YP_001009269.1 YP_001011085.1 YP_001015225.1 YP_001091099.1 YP_001484109.1 YP_291794.1 YP_397316.1 ZP_05138816.1	Hypothetical protein	-0.26
Cluster 503	NP_874837.1 NP_892565.1 NP_895167.1 YP_001008897.1 YP_001010825.1 YP_001014328.1 YP_001016661.1 YP_001090695.1 YP_001483728.1 YP_001550330.1 YP_292969.1 YP_396943.1 ZP_05138518.1 ABE11336	Cytochrome c oxidase subunit XV assembly protein	-0.28
Cluster 363	YP_001008757.1 YP_001090586.1 YP_001483563.1 ZP_05138726.1	Hypothetical protein	-0.28
Cluster 2447	NP_875146.1 NP_893060.1 NP_894477.1 YP_001009309.1 YP_001011339.1 YP_001014760.1 YP_001017595.1 YP_001091140.1 YP_001484149.1 YP_001550587.1 YP_291463.1 YP_397353.1 ZP_05137613.1	*Ruv, holliday junction DNA helicase*	-0.29
Cluster 1536	NP_875812.1 NP_893457.1 NP_895243.1 YP_001009931.1 YP_001011815.1 YP_001015587.1 YP_001016563.1 YP_001091750.1 YP_001484768.1 YP_001551279.1 YP_292105.1 YP_397934.1 ZP_05138569.1 ABE11298 ABE11402	*purD*, phosphoribosylamine-glycine ligase	-0.32
Cluster 1615	NP_893517.1 YP_001010010.1 YP_001015624.1 YP_001484835.1 YP_292138.1 YP_397989.1 ZP_05137944.1	Hypothetical protein	-0.32

Cluster 2447	NP_875146.1 NP_893060.1 NP_894477.1 YP_001009309.1 YP_001011339.1 YP_001014760.1 YP_001017595.1 YP_001091140.1 YP_001484149.1 YP_001550587.1 YP_291463.1 YP_397353.1 ZP_05137613.1	*Ruv, holliday junction DNA helicase*	-0.29
Cluster 1536	NP_875812.1 NP_893457.1 NP_895243.1 YP_001009931.1 YP_001011815.1 YP_001015587.1 YP_001016563.1 YP_001091750.1 YP_001484768.1 YP_001551279.1 YP_292105.1 YP_397934.1 ZP_05138569.1 ABE11298 ABE11402	*purD*, phosphoribosylamine-glycine ligase	-0.32
Cluster 1615	NP_893517.1 YP_001010010.1 YP_001015624.1 YP_001484835.1 YP_292138.1 YP_397989.1 ZP_05137944.1	Hypothetical protein	-0.32
Cluster 979	NP_875350.1 NP_892996.1 NP_894560.1 YP_001009374.1 YP_001011274.1 YP_001014828.1 YP_001017501.1 YP_001091205.1 YP_001484215.1 YP_001550771.1 YP_291527.1 YP_397418.1 ZP_05138108.1	*ilvE*, branched chain amino acid aminotransferase	-0.35
Cluster 906	NP_875137.1 NP_893068.1 NP_894486.1 YP_001009300.1 YP_001011348.1 YP_001014751.1 YP_001017586.1 YP_001091131.1 YP_001484141.1 YP_001550578.1 YP_291454.1 YP_397344.1 ZP_05138113.1	*carA*, carbamoyl-phosphate synthase small sub.	-0.41
Cluster 3988	NP_875242.1 NP_892895.1 NP_894390.1 YP_001009230.1 YP_001011122.1 YP_001014638.1 YP_001017704.1 YP_001091061.1 YP_001484072.1 YP_001550870.1 YP_291378.1 YP_397281.1 ZP_05137880.1	Hypothetical protein	-0.44
Cluster 404	YP_001008798.1 YP_001011894.1 YP_001090628.1 YP_001483623.1 ZP_05139130.1 emb|CAP16335.1| ABE10755	Hypothetical protein	-0.53
Cluster 4781	NP_875863.1 YP_001008874.1 YP_001009698.1 YP_001011202.1 YP_001015202.1 YP_001090672.1 YP_001091545.1 YP_001484537.1 YP_001551327.1 YP_291785.1 YP_396921.1 YP_397726.1 YP_398003.1 ZP_05137533.1 ZP_05139107.1	Hypothetical protein	-0.59
Cluster 1109	YP_001009504.1 YP_001011420.1 YP_001091336.1 YP_001484343.1 YP_654189.1 ZP_05137499.1 ABS83162	Hypothetical protein	-0.61
Cluster 2230	YP_001010358.1;ZP_05139173.1	Hypothetical protein	-0.92
Cluster 263	NP_874664.1 NP_892359.1 NP_895647.1 YP_001008657.1 YP_001010589.1 YP_001014148.1 YP_001018435.1 YP_001090487.1 YP_001483466.1 YP_001550149.1 YP_292798.1 YP_396739.1 ZP_05137796.1	tRNA guanine-N(7)-methyltransferase	-0.98
Cluster 4506	NP_874602.1 NP_892304.1 NP_895891.1 YP_001008596.1 YP_001010528.1 YP_001014088.1 YP_001018736.1 YP_001090427.1 YP_001483406.1 YP_001550086.1 YP_292742.1 YP_396683.1 ZP_05137689.1	*dgkA*, diacylglycerol kinase	-0.98

Table 4.10. List of *Prochlorococcus* orthologous gene clusters that show a significant iron effect at T36h. The list comprises the 10 orthologous gene clusters for which the pattern of gene expression across all 4 samples is indicative of an iron effect as assessed by the baySeq analysis described in Section 4.3. Significance is determined for a sum of posterior probabilities greater than 0.9. α_2 denotes the magnitude of the iron effect- positive means that the relative abundance of transcripts that belong to a specific cluster increases in the (+Fe) treatment, negative means that the relative abundance of transcripts that belong to a specific cluster decreases in the (+Fe) treatment. It is calculated as the log ratio of the relative expression value in the (+Fe) sample to the relative expression value in the (Ctrl) sample.

Orthologous cluster #	List of Accession Numbers of Genes Populating the Cluster	Gene Name and Description	α_2
Cluster 1615	NP_893517.1 YP_001010010.1 YP_001015624.1 YP_001484835.1 YP_292138.1 YP_397989.1 ZP_05137944.1	Hypothetical protein	0.33
Cluster 1261	YP_001009656.1 YP_001091497.1 YP_001484508.1 YP_397703.1 ZP_05137560.1 ABE10907	Hypothetical protein	0.27
Cluster 404	YP_001008798.1 YP_001011894.1 YP_001090628.1 YP_001483623.1 ZP_05139130.1 emb|CAP16335.1| ABE10755	Hypothetical protein	0.21
Cluster 1331	NP_875647.1 NP_893275.1 NP_895010.1 YP_001009726.1 YP_001011639.1 YP_001015430.1 YP_001016858.1 YP_001091574.1 YP_001484565.1 YP_001551119.1 YP_291962.1 YP_397752.1 ZP_05138193.1 ABE10996	Photosystem II PsbC protein	0.15
Cluster 1248	NP_874646.1 NP_892343.1 NP_894252.1 NP_895359.1 YP_001008639.1 YP_001009643.1 YP_001010569.1 YP_001010571.1 YP_001014132.1 YP_001015148.1 YP_001015394.1 YP_001016426.1 YP_001017875.1 YP_001090469.1 YP_001484502.2 YP_001550131.1 YP_291752.1 YP_291929.1 YP_292783.1 YP_396723.1 YP_397639.1 ZP_05137720.1 ZP_05137853.1 emb|CAA89062.1| AAT98419 AAT98420 AAN77566 AAT98416	Photosystem II PsbA protein	0.11
Cluster 1757	NP_876100.1 NP_893672.1 NP_895562.1 YP_001010152.1 YP_001012051.1 YP_001015820.1 YP_001018305.1 YP_001091970.1 YP_001485026.1 YP_001551560.1 YP_292319.1 YP_398144.1 ZP_05138722.1	*rplB*, 50S ribosomal protein L2	-0.53
Cluster 363	YP_001008757.1 YP_001090586.1 YP_001483563.1 ZP_05138726.1	Hypothetical protein	-0.59
Cluster 955	NP_875325.2 NP_893020.1 NP_894527.1 YP_001009350.1 YP_001011298.1 YP_001014801.1 YP_001017535.1 YP_001091181.1 YP_001484191.1 YP_001550743.1 YP_291500.1 YP_397394.1 AAP99977 ABE11464	*rpmB*, 50S ribosomal protein L28	-0.73
Cluster 1183	NP_875508.1 NP_893198.1 NP_894902.1 YP_001009578.1 YP_001011485.1 YP_001015323.1 YP_001016993.1 YP_001091412.1 YP_001484418.1 YP_001550991.1 YP_291863.1 YP_397587.1 ZP_05138362.1	*malQ*, putative 4 alpha glucanotransferase	-1.16
Cluster 489	NP_874823.1 NP_892551.1 NP_895186.1 YP_001008883.1 YP_001010811.1 YP_001014316.1 YP_001016640.1 YP_001090681.1 YP_001483714.1 YP_001550315.1 YP_292956.1 YP_396929.1 ZP_05138003.1 ABE11251	*birA*, putative biotin operon repressor	-1.12

To complement this statistical analysis aimed at identifying *Prochlorococcus*-specific genes affected by iron, we explored the pattern of expression of *Prochlorococcus* genes previously shown to be iron-responsive in laboratory studies of pure cultures. Using microarrays, Thompson et al. investigated the changes in gene expression induced by iron-stress in cultures of *Prochlorococcus marinus* MED4 and MIT9313 (Thompson et al. 2011). Because ~95% of the *Prochlorococcus* specific reads in our experiment mapped to high-light adapted strains (25-30%-MIT9301; ~30% AS9601; 10% MIT9312; 5-10% MIT9315; 5-10% MED4; 3-5% MIT9202; ~2% MIT9515), we focused our comparison on the results obtained by Thompson et al. for the high-light strain MED4 only. Since our experiment relied on the addition of iron to the *in-situ* community, while Thompson's experiment investigated the *Prochlorococcus* response to iron limitation, we expected that genes upregulated under iron-limitation in Thompson's experiment might instead show a decrease in relative transcript abundance in our experiment. Comparison of the log-ratios of gene expression (or relative transcript abundance) for the treatments (iron addition or iron limitation) to the control revealed that, indeed, for a number of orthologous gene clusters the gene expression pattern in response to iron addition seemed to be the opposite than that observed in response to iron limitation (Figure 4.7). The gene expression pattern in response to iron addition was, however not as pronounced and not deemed statistically significant by our baySeq approach. This could also indicate that *Prochlorococcus* at HOT station ALOHA may not be iron-limited.

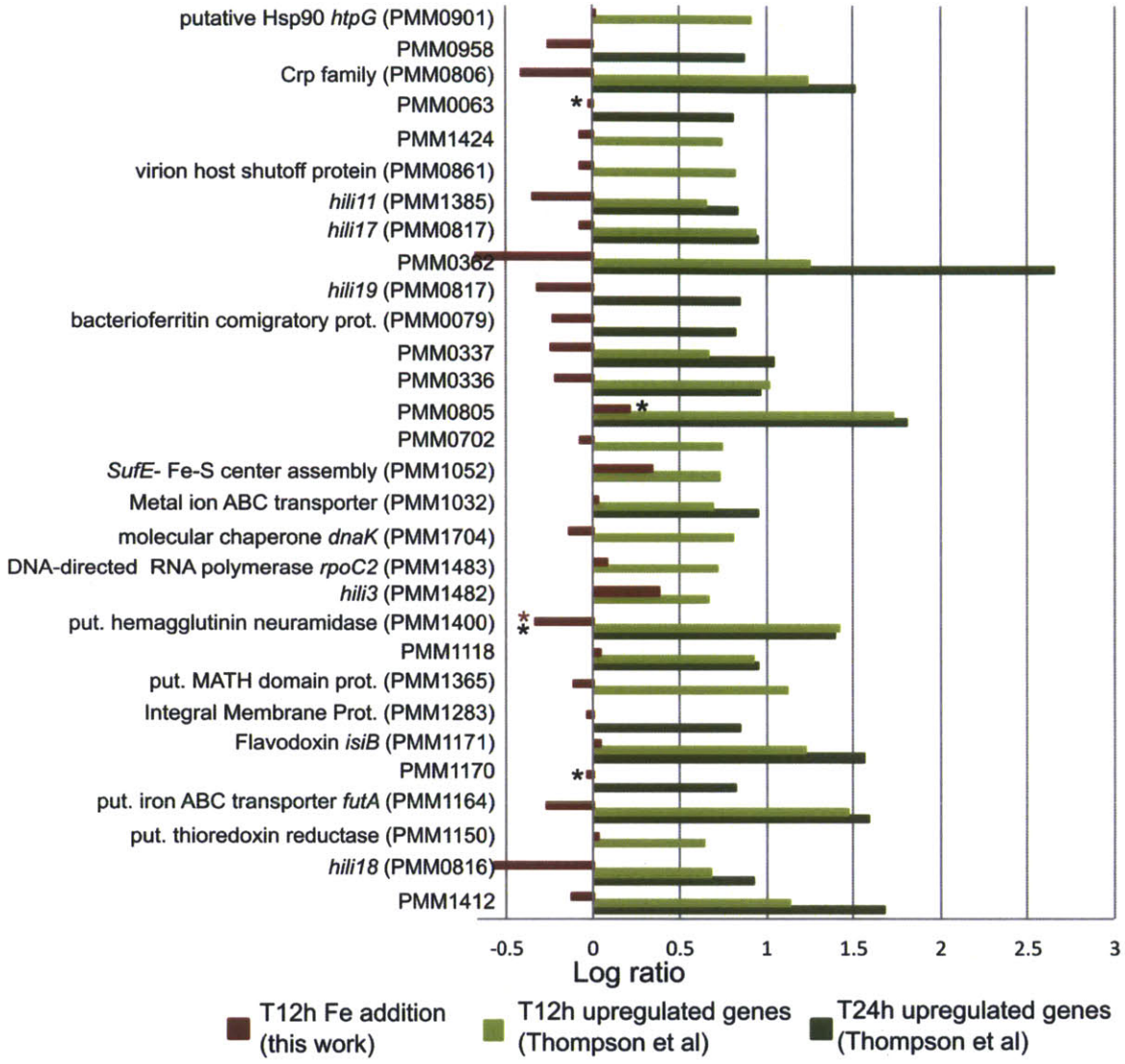


Figure 4.7. Comparison of *Prochlorococcus* gene expression in response to iron addition (this work) and to iron limitation (Thompson et al. 2011)**.** Log ratio of gene expression (treatment vs. control) for *Prochlorococcus MED4* genes upregulated under iron limitation (Thompson et al., 2001) compared to log ratio (relative transcript abundance (+Fe)/relative transcript abundance in (Ctrl)) for the corresponding *Prochlorococcus* orthlogous gene cluster in the present experiment. Asterisks show a significant iron (red) or interaction effect (black) as determined by baySeq. Significant organic matter effects are not displayed.

Organic Matter Effect –While *Prochlorococcus* cell counts remained fairly constant and comparable in all 4 treatments throughout the experiment (Figure 4.4), the addition of organic matter seemed to have an impact on the physiology of *Prochlorococcus*. Indeed, in both the (+OM) and the (+OM+Fe) treatments, we observed a close to 5x increase in both average cell size and average chlorophyll content per cell compared to the (Ctrl) and (+Fe) treatments (Figure 4.8). Interestingly, while the addition of iron alone did not seem to impact neither total *Prochlorococcus* cell counts nor *Prochlorococcus* cell physiology, at T36h, these physiological changes were enhanced in the (OM+Fe) treatment compared to the (+OM) treatment. This differential effect suggests that addition of iron together with organic matter has a compounded effect on *Prochlorococcus* that is not observed in the iron alone treatment. It also suggests that the community's iron-specific response may be identified not only by performing a comparative analysis of the (+Fe) and the (Control) samples but also by comparing the (+OM+Fe) samples to the (+OM) samples.

Analysis of global patterns of gene expression of *Prochlorococcus* specific reads confirmed the impact of organic matter on *Prochlorococcus* physiology (Figure 4.9). Indeed, comparison of the relative transcript abundance for each *Prochlorococcus* orthologous cluster in each of the four treatments at T12h indicates a pattern that lies along an OM-OM+Fe axis (Figure 4.9). The same analysis comparing relative transcript abundance of *Prochlorococcus* orthologous clusters across treatments at T36h (not shown) revealed a similar pattern. However, the differential effect of iron on the physiology of *Prochlorococcus* cells observed when comparing flow-cytometry data from the (+OM) and the (OM+Fe) samples (Figure 4.8) was not mirrored in this analysis of global gene expression patterns. It is possible that only a few genes are responsible for this observed effect and that, hence their differential expression does not contribute to driving the global pattern of gene expression.

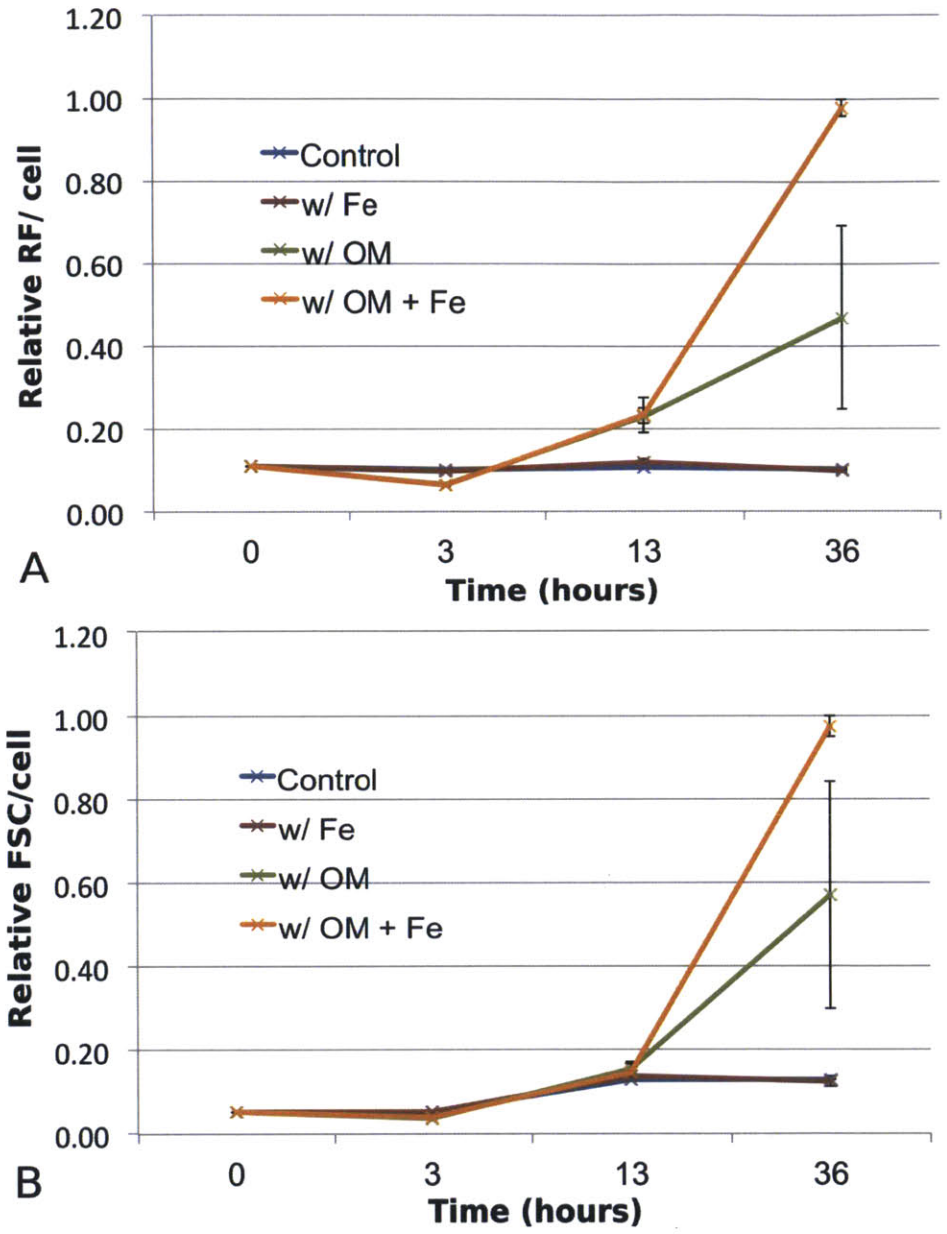


Figure 4.8. Organic matter and iron addition relieves *Prochlorococcus* physiological stress. Panel A: Relative red fluorescence (RF) per cell. Panel B: Relative forward scatter (FSC) per cell. Both properties were assessed by flow-cytometric measurements of the four incubations (Control: seawater only, w/ Fe: seawater with $FeCl_3$, w/ OM: seawater with organic matter, w/OM + Fe: seawater with organic matter and $FeCl_3$). Parameters are plotted for the *Prochlorococcus* population only. Plotted is the mean red fluorescence per cell averaged over three biological replicates and normalized to the maximum value. The error bars represent one standard deviation from the mean. Note that the large error bar at T36hours, in the OM treatment is due to one of the three bottles with unusually high cell counts. We plot FSC as a proxy for cell-size and red-fluorescence per cell as a proxy for chlorophyll content per cell, as previously established (Cavender-Bares et al. 1999).

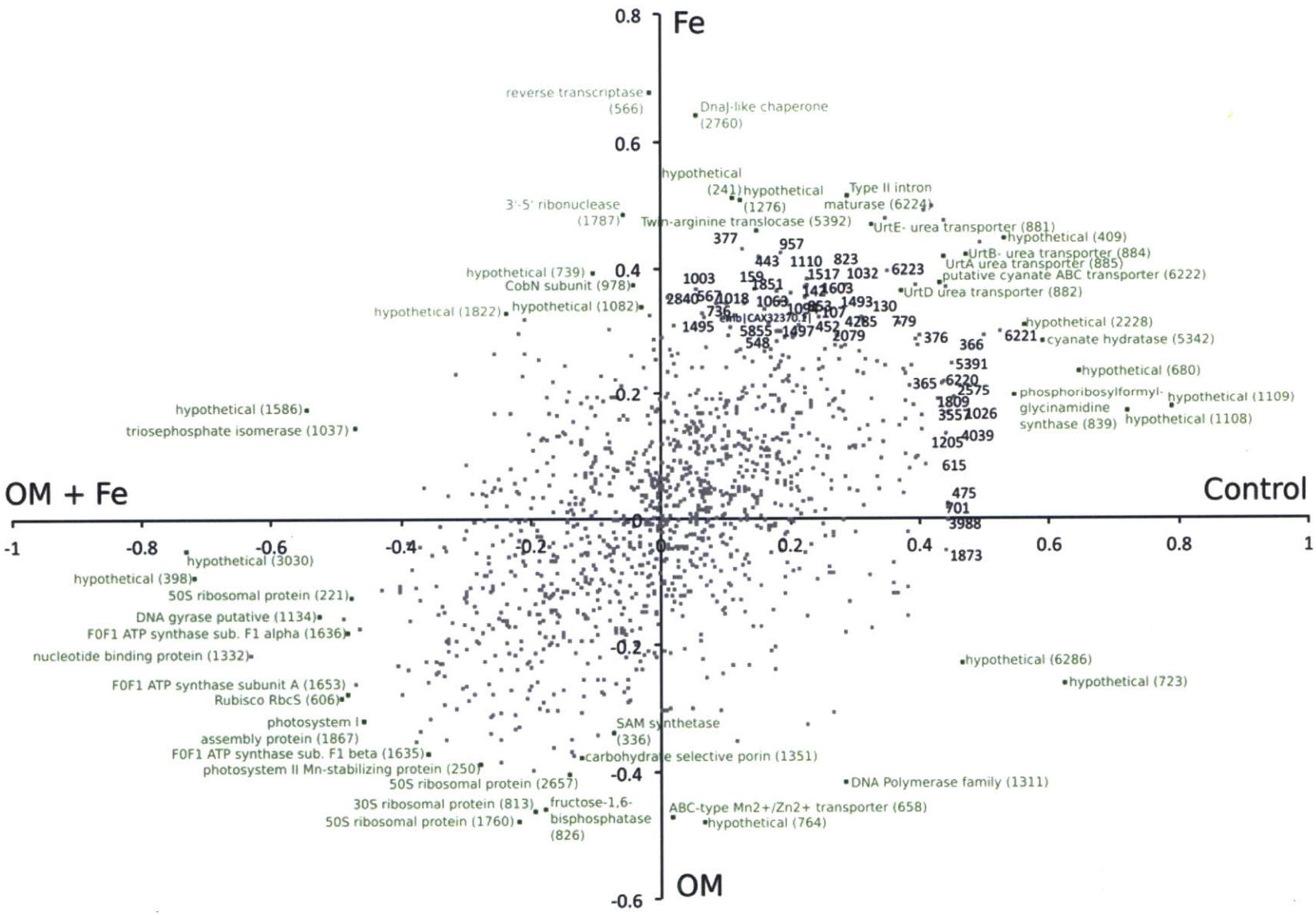
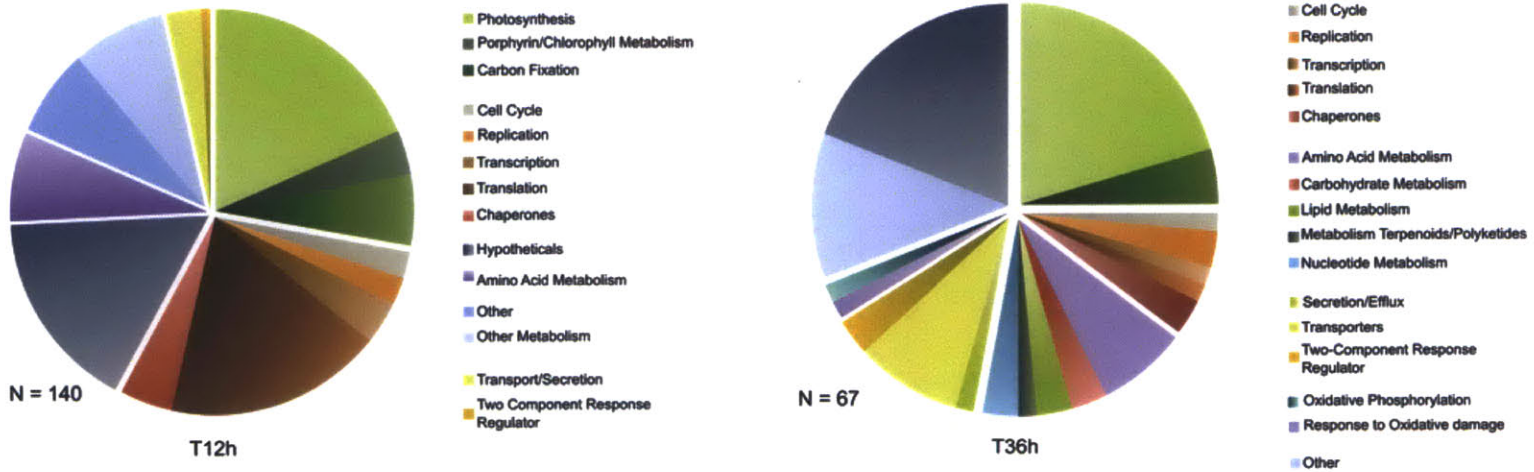

Fe
OM
OM + Fe
Control
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
reverse transcriptase (566)
DnaJ-like chaperone (2760)
hypothetical (241)
hypothetical (1276)
Type II intron maturase (6224)
3'-5' ribonuclease (1787)
Twin-arginine translocase (5392)
UrtE- urea transporter (881)
hypothetical (409)
UrtB- urea transporter (884)
UrtA urea transporter (885)
putative cyanate ABC transporter (6222)
UrtD urea transporter (882)
hypothetical (739)
CobN subunit (978)
hypothetical (1822)
hypothetical (1082)
hypothetical (2228)
cyanate hydratase (5342)
hypothetical (680)
hypothetical (1109)
hypothetical (1108)
phosphoribosylformyl-glycinamidine synthase (839)
hypothetical (1586)
triosephosphate isomerase (1037)
hypothetical (3030)
hypothetical (398)
50S ribosomal protein (221)
DNA gyrase putative (1134)
F0F1 ATP synthase sub. F1 alpha (1636)
nucleotide binding protein (1332)
F0F1 ATP synthase subunit A (1653)
Rubisco RbcS (606)
photosystem I assembly protein (1867)
F0F1 ATP synthase sub. F1 beta (1635)
photosystem II Mn-stabilizing protein (250)
50S ribosomal protein (2657)
30S ribosomal protein (813)
50S ribosomal protein (1760)
fructose-1,6-bisphosphatase (826)
SAM synthetase (336)
carbohydrate selective porin (1351)
DNA Polymerase family (1311)
ABC-type Mn2+/Zn2+ transporter (658)
hypothetical (764)
hypothetical (6286)
hypothetical (723)

Figure 4.9. Influence of organic matter on *Prochlorococcus* is observed in gene expression pattern at T12h. Clusters of *Prochlorococcus* orthologous genes are plotted with respect to their relative representation in each of the four treatments. Only the clusters with at least 10 counts in at least one of the four datasets were considered for the analysis. Values for 1,260 out of 6,041 clusters that meet this criterion are shown. For each orthologous cluster, the relative transcript abundance value was normalized to the sum of the relative transcript abundance values in the 4 treatments. X coordinates: (normalized control - normalized OM + Fe). Y coordinates: (normalized Fe - normalized OM). The pattern of data points follows a distribution along the Fe/Control to OM/OM+Fe axis. A similar distribution is observed for the T36h data (not shown). The descriptions of some clusters on the edges (for ease of reading) are indicated in green. Parentheses indicate the cluster id number.

The effect of organic matter on the global expression pattern of *Prochlorococcus* genes is further supported by the large number of *Prochlorococcus* orthologous gene clusters for which a statistically significant organic matter effect could be identified (Figure 4.6). At T12h, 320 *Prochlorococcus* orthologous gene clusters (that is ~5% of all *Prochlorococcus* orthologous gene clusters inspected) displayed a differential expression pattern indicative of an organic matter effect, while at T36h, the effect was observed for 127 *Prochlorococcus* orthologous gene clusters (Figure 4.6). Grouping these clusters by pathway provides insight into how *Prochlorococcus* responded to the effects of added organic matter. Of the 320 orthologous clusters that showed an organic matter effect at T12h, 140 showed an increase in relative transcript abundance (positive $\alpha 3$) while 180 showed a decrease in relative transcript abundance (negative $\alpha 3$) (see Figure 4.10). Of note is the increase in relative transcript abundance of genes that are part of photosynthesis (mostly F0F1 ATPase and photosystem I and II gene clusters) and related pathways (porphyrin/chlorophyll metabolism and carbon fixation) as well as the increase in relative transcript abundance of genes involved in transcription and translation (mostly ribosomal proteins) processes. Decrease in relative transcript abundance was observed for transporters (ammonium/urea/cyanate transporters) and repair mechanisms (including DNA polymerase III and RecA). At T36h the number of *Prochlorococcus* orthologous gene clusters that were significantly impacted by an organic matter effect dropped to 166 with 67 displaying an increase in relative transcript abundance and 99 displaying a decrease in relative transcript abundance (Figure 4.10). We found that as was the case at T12h, a significant increase in relative transcript abundance was observed at T36h for clusters of orthologous genes that code for photosynthesis proteins (Figure 4.10). However, while an increase in relative transcript abundance of *Prochlorococcus* ribosomal proteins could be noticed at T12h, it was not as widespread at T36h (Figure 4.10).

A. Increase in relative transcript abundance in OM



B. Decrease in relative transcript abundance in OM

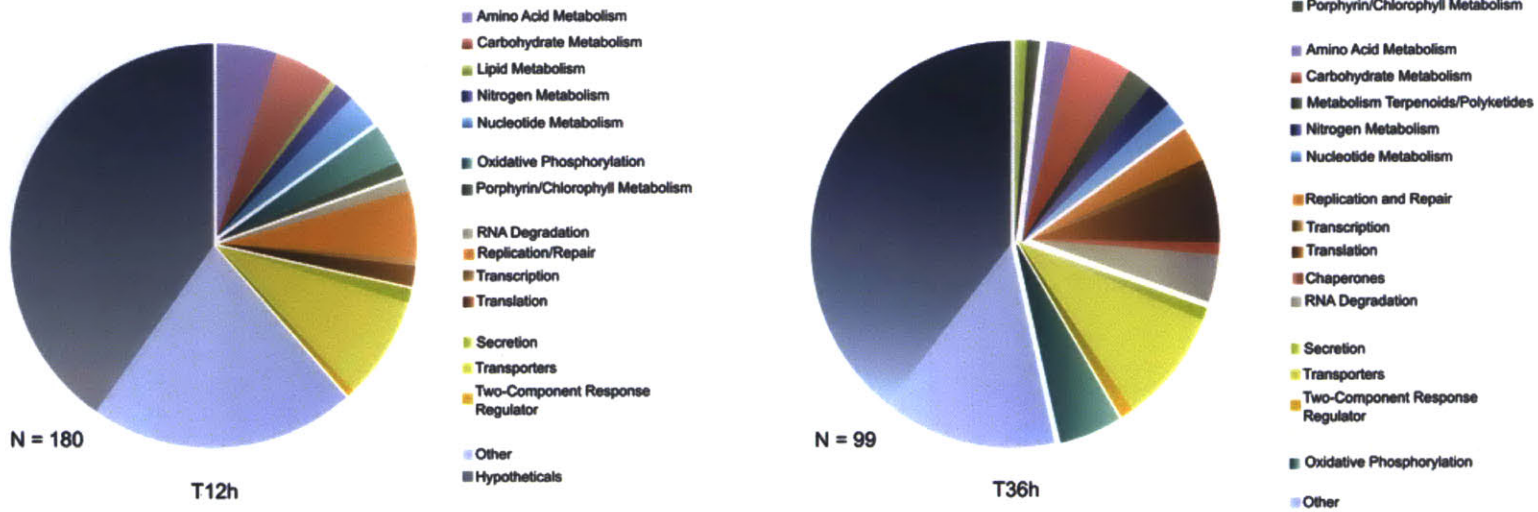


Figure 4.10. Significant Organic Matter Effect on *Prochlorococcus* orthologous gene clusters at T12h and T36h. We group by pathway the gene clusters for which differences in relative transcript abundance values are indicative of a significant organic matter effect. Panel A shows the pathway grouping (manual and level 2 of the KEGG hierarchy) of the gene clusters for which relative transcript abundance is enhanced (140 at T12h and 68 at T36h). Panel B shows the pathway grouping of the orthologous gene clusters for which relative transcript abundance is significantly decreased (180 at T12h, 99 at T36h). N refers to the total number of orthologous gene clusters that are plotted for each pie chart.

Because the physiological response of *Prochlorococcus* cells observed in the two organic matter treatments may be indicative of the relief of nutrient stress (Kjellegerg 1993; Mann & Chisholm 2000; Moore et al. 2008), we proceeded to specifically explore the changes in relative transcript abundance of nitrogen and phosphorus acquisition genes in the +OM sample compared to the Ctrl sample. Indeed, at T36h, 4 of the orthologous gene clusters displaying the greatest difference in relative transcript abundance are related to phosphorus acquisition (*phoA* coding for an alkaline phosphatase-like protein α3=4.4/ *pstS* α3=4.8/ *phnD* coding for phosphonate ABC-transporter periplasmic protein α3=3.3 / and *pho*R part of the pho-regulon α3=2.8). Upregulation of these genes, especially of *pstS* has been recorded for *Prochlorococcus* cells grown in culture under phosphorus stress (Martiny et al. 2006). This phosphorus-specific response is very different from what was observed at T12h, where no significant difference in the relative transcript abundance of these genes was observed between the organic matter treatments and the (Ctrl) and (+Fe) treatments (Figure 4.11). Furthermore, at both T12h and T36h, the relative transcript abundance of urea and cyanate acquisition and utilization orthologous gene clusters significantly decreased in the organic matter treatments (Figure 4.12). These observations suggest that the addition of organic matter provided a nitrogen source that relieved nitrogen limitation. Indeed, the *Prochlorococcus marinus* MED4 genes *urtB* (PMM0971), *urtA* (PMM0970) and *cynA* (PMM0370) were upregulated between ~5 fold (*urtB*) and ~50 fold (*cynA*) when grown in culture under nitrogen limiting conditions (Tolonen et al. 2006).

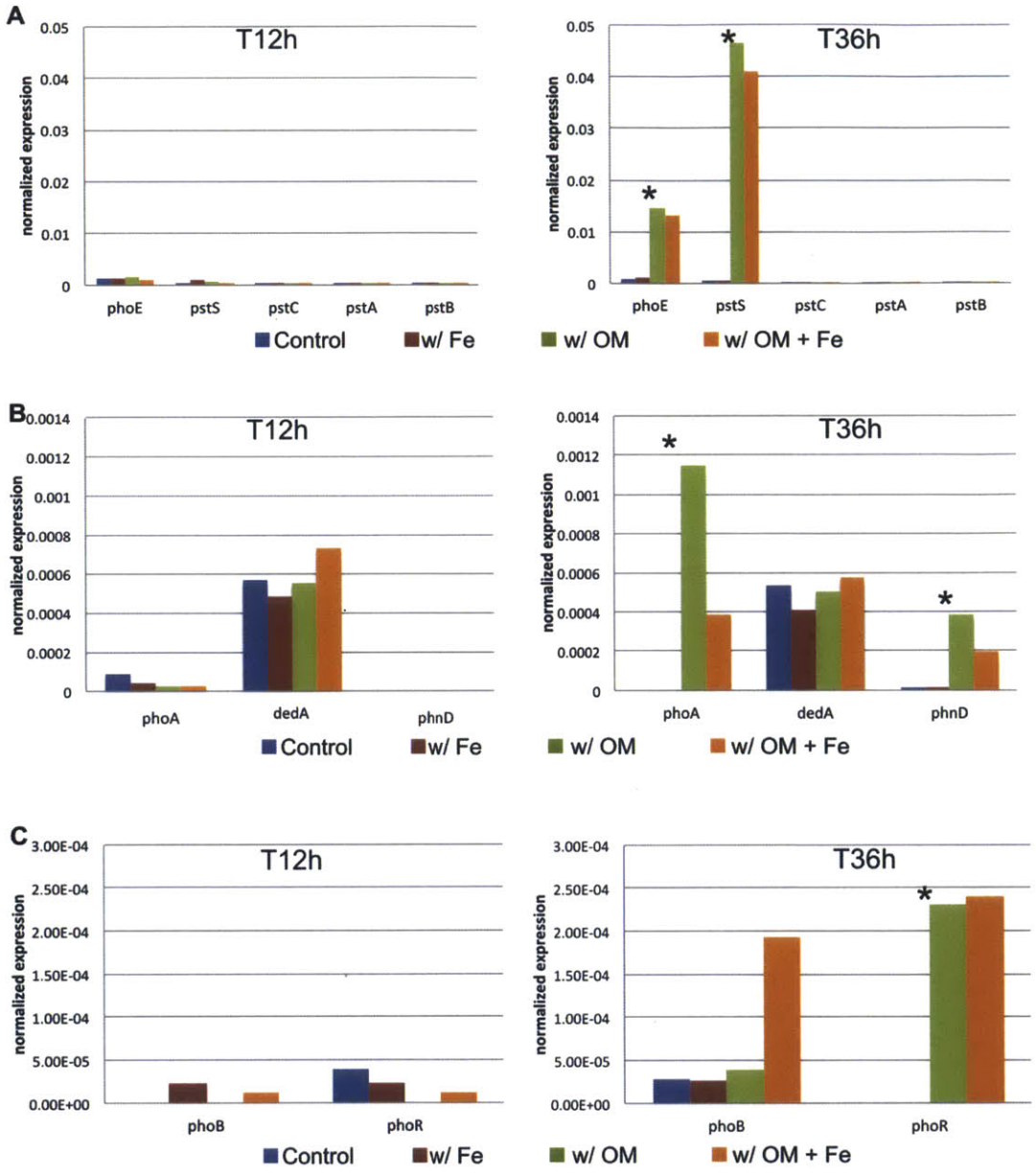


Figure 4.11. Addition of organic matter leads to phosphate limitation after 36 hours. Relative transcript abundance across treatments and time points indicative of inorganic phosphate (A) and organic phosphate (B) acquisition genes and phosphorus regulation (C). Relative transcript abundance refers to the number of read counts included in each *Prochlorococcus* orthologous gene clusters normalized to total number of reads with a significant hit to *Prochlorococcus*. Asterisks indicate that the difference in relative transcript abundance levels observed between the 4 treatments is significant and indicative of an organic matter effect (sum of baySeq posterior probabilities > 0.9). Note that the relative transcript abundance levels displayed are those obtained for the orthologous gene cluster. Most gene designations are based on the annotations of the MED4 genes that belong to the different clusters (phoE: PMM0709, pstS: PMM0710, pstC: PMM0723, pstA: PMM0724, pstB: PMM0725, phoA: PMM0708, dedA: PMM1624, phoB: PMM0705, phoR: PMM0706). One exception is *phnD* for which the cluster is composed of genes only from strains MIT9303 and MIT9301, both originally isolated from the P-limited Sargasso Sea (*phnD*: P9303_1129).

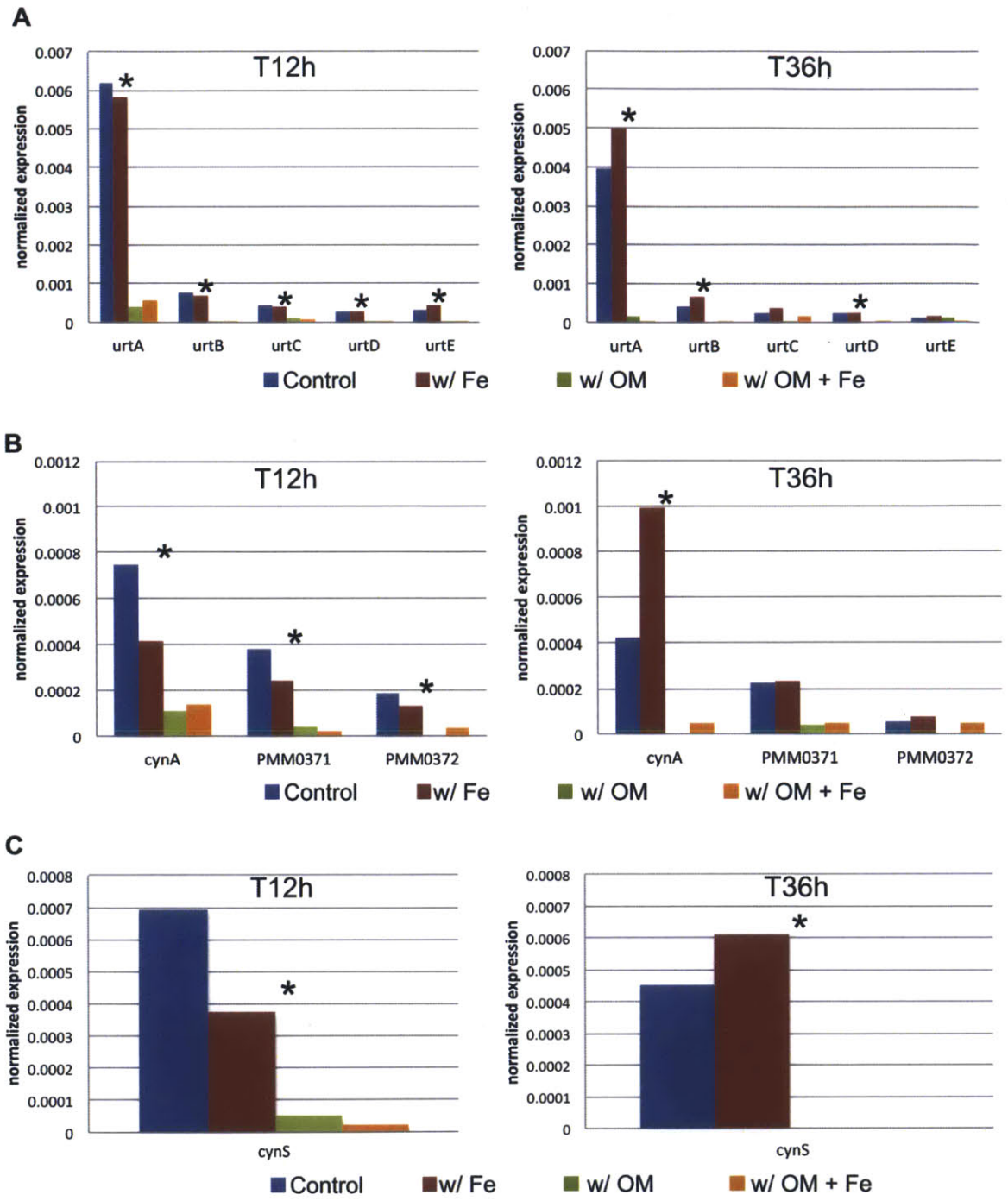


Figure 4.12. Addition of organic matter relieves nitrogen limitation. Relative transcript abundance across treatments and time points for urea (A) and cyanate (B) acquisition and (C) cyanate utilization genes. Relative transcript abundance refers to the number of read counts included in each *Prochlorococcus* orthologous gene clusters normalized to total number of reads with a significant hit to *Prochlorococcus*. Asterisks indicate a significant organic matter effect (sum of baySeq posterior probabilities > 0.9). As for Figure 4.11, the relative transcript abundance levels displayed are those obtained for the orthologous gene cluster. Most gene designations are based on the annotations of the MED4 genes that belong to the different clusters (urtA: PMM0970, urtB: PMM0971, urtC: PMM0972, urtD: PMM0973, urtE: PMM0974, cynA: PMM0370, cynS: PMM0373).

Interaction Effect –Despite this overwhelming response to the organic matter amendment, physiological data suggests that iron addition together with organic matter had a specific impact on *Prochlorococcus* (Figure 4.8). This is likely to be mirrored in patterns of differential transcript abundance. Indeed, at T12h, 16 *Prochlorococcus* orthologous gene clusters agreed with the specific differential expression model 12.3.4 (meaning that transcript abundance pattern in the control and (+Fe) treatment is similar, but different from the pattern in the (+OM) treatment, which itself differs from pattern in the (+OM+Fe) treatment – see Table 4.7 in Section 4.3). The orthologous gene clusters that displayed this pattern of differential expression (Figure 4.13) may in part explain the physiological differences observed for *Prochlorococcus* cells at T36h (Figure 4.8): the increase in relative abundance of the 50S ribosomal L10 transcript could explain an increase in cell size, the increase in relative abundance of the *psaB* transcript (photosystem I P700 chlorophyll a apoprotein) could explain the increase in cellular chlorophyll content. Interestingly, at T36h (the time-point at which the differential effect of iron on *Prochlorococcus* physiology was most pronounced), no orthologous gene cluster showed a differential transcript abundance pattern agreeing with the 12.3.4 model. This suggests that there was a (surprisingly long) delay between the onset of changes in gene expression and the time at which the changes could be observed at the physiological level.

This interesting differential effect between the OM and the OM+Fe treatments prompted us to explore in more detail which *Prochlorococcus* orthologous gene clusters might be significantly impacted by an interaction effect. At T12h, we found that there were 127 such clusters, and 25 at T36h (Figure 4.6). Interestingly, at T12h ~75% of these orthologous gene clusters were associated with a positive combined iron-organic matter effect (positive $\alpha 4$) (Figure 4.14). This means that relative transcript abundance was enhanced in the OM+Fe treatment compared to the control treatment. Clustering of the orthologous gene clusters per their $\alpha 2$ (iron effect), $\alpha 3$ (organic matter effect) and $\alpha 4$ (combined effect- interaction) values revealed 4 distinct groups at T12h (Figure 4.14) and 2 groups at T36h (Figure 4.15).

Group 1 consists of genes for which patterns of relative transcript abundance are positively affected by both an organic matter alone effect and an interaction affect. Genes that belong to this group show an increase in relative transcript abundance in both the +OM sample and in the +OM+Fe sample

and for a number of those genes, the relative transcript abundance is enhanced in the +OM+Fe sample compared to the +OM sample. Furthermore, genes that belong to this group do not show a notable change in relative transcript abundance in the +Fe sample. Therefore, we propose that genes that belong to this group parallel the response pattern observed for *Prochlorococcus* cell size and chlorophyll content (Figure 4.8): cells do not respond to iron alone because they are limited by other macronutrients, upon addition of organic matter the macronutrient limitation is relieved, which then leads to iron limitation (also known as secondary limitation). Interestingly, a number of genes coding for photosynthetic proteins (PsaL-cluster 1718) that are usually down-regulated under iron limitation (and upregulated upon relief of iron limitation) are found as part of group 1 (cluster 1718: PsaL, cluster 518: PetA, cluster: 518) (Thompson et al. 2011). Other photosynthetic genes (cluster 323: PsbL, cluster 1722: PsaB, cluster 274: PsbH, cluster 2105: PsaJ) are also part of this group (Figure 4.14).

Group 2 consists of genes for which patterns of relative transcript abundance are positively affected by an organic matter effect but negatively affected by an interaction effect. If we follow the logic presented above, this group should include genes that are downregulated in response to relief of iron limitation. None of the genes previously shown to be downregulated in response to relief of iron limitation in *Prochlorococcus* cultures are found in group 2 (Thompson et al. 2011). Groups 3 and 4 consist of genes for which patterns of relative transcript abundance are negatively impacted by an iron alone effect and an organic matter alone effect, but positively impacted by an interaction effect. The difference between these two groups is the magnitude of the different effects.

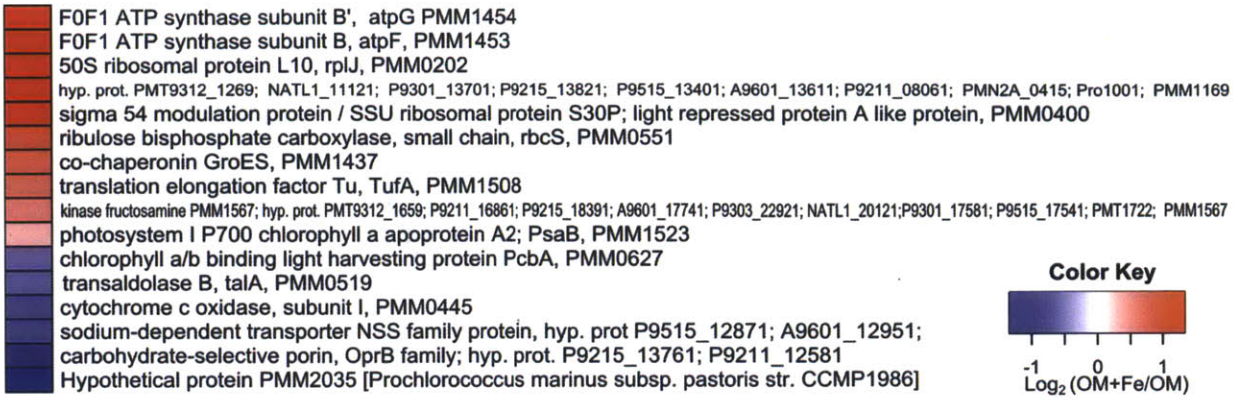


Figure 4.13. Comparison of the relative transcript abundance of *Prochlorococcus* clusters in OM+Fe with clusters in OM at T12h. Log_2 ratio (relative transcript abundance in OM+Fe/ relative transcript abundance in OM) for the 16 *Prochlorococcus* orthologous gene clusters that fit the 12.3.4 model of differential gene expression. Some of these clusters can explain the differential effect between OM and OM +Fe observed on the physiology of *Prochlorococcus* cells. At T36h no *Prochlorococcus* orthologous gene clusters show a pattern of gene expression that fits the 12.3.4 model.

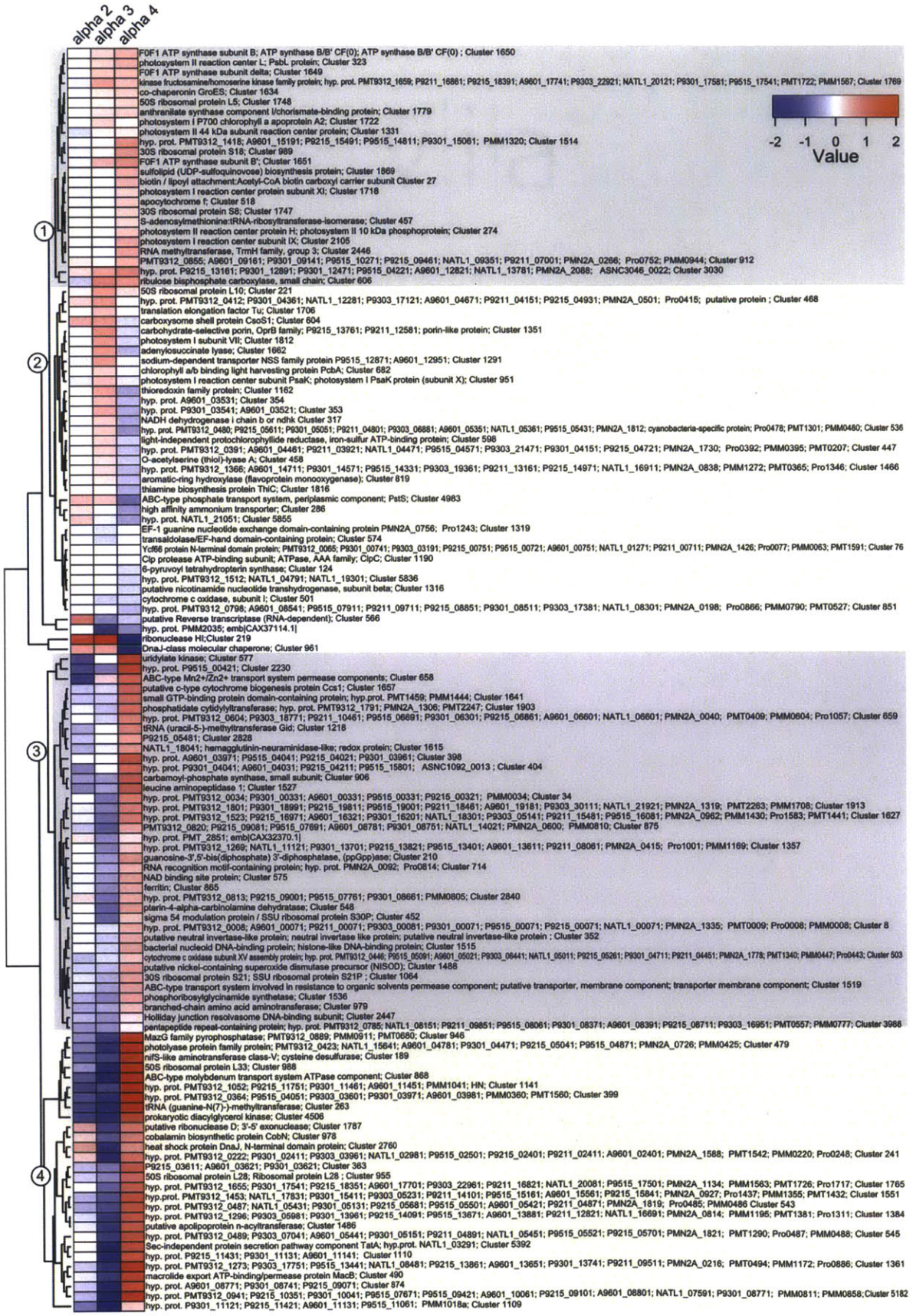

alpha 2
alpha 3
alpha 4
-2 -1 0 1 2
Value
1
2
3
4
F0F1 ATP synthase subunit B; ATP synthase B/B' CF(0); ATP synthase B/B' CF(0) ; Cluster 1650
photosystem II reaction center L; PsbL protein; Cluster 323
F0F1 ATP synthase subunit delta; Cluster 1649
kinase fructosamine/homoserine kinase family protein; hyp. prot. PMT9312_1659; P9211_16861; P9215_18391; A9601_17741; P9303_22921; NATL1_20121; P9301_17581; P9515_17541; PMT1722; PMM1567; Cluster 1769
co-chaperonin GroES; Cluster 1634
50S ribosomal protein L5; Cluster 1748
anthranilate synthase component I/chorismate-binding protein; Cluster 1779
photosystem I P700 chlorophyll a apoprotein A2; Cluster 1722
photosystem II 44 kDa subunit reaction center protein; Cluster 1331
hyp. prot. PMT9312_1418; A9601_15191; P9215_15491; P9515_14811; P9301_15061; PMM1320; Cluster 1514
30S ribosomal protein S18; Cluster 989
F0F1 ATP synthase subunit B'; Cluster 1651
sulfolipid (UDP-sulfoquinovose) biosynthesis protein; Cluster 1869
biotin / lipoyl attachment:Acetyl-CoA biotin carboxyl carrier subunit Cluster 27
photosystem I reaction center protein subunit XI; Cluster 1718
apocytochrome f; Cluster 518
30S ribosomal protein S8; Cluster 1747
S-adenosylmethionine:tRNA-ribosyltransferase-isomerase; Cluster 457
photosystem II reaction center protein H; photosystem II 10 kDa phosphoprotein; Cluster 274
photosystem I reaction center subunit IX; Cluster 2105
RNA methyltransferase, TrmH family, group 3; Cluster 2446
PMT9312_0855; A9601_09161; P9301_09141; P9515_10271; P9215_09461; NATL1_09351; P9211_07001; PMN2A_0266; Pro0752; PMM0944; Cluster 912
hyp. prot. P9215_13161; P9301_12891; P9301_12471; P9515_04221; A9601_12821; NATL1_13781; PMN2A_2088; ASNC3046_0022; Cluster 3030
ribulose bisphosphate carboxylase, small chain; Cluster 606
50S ribosomal protein L10; Cluster 221
hyp. prot. PMT9312_0412; P9301_04361; NATL1_12281; P9303_17121; A9601_04671; P9211_04151; P9215_04931; PMN2A_0501; Pro0415; putative protein ; Cluster 468
translation elongation factor Tu; Cluster 1706
carboxysome shell protein CsoS1; Cluster 604
carbohydrate-selective porin, OprB family; P9215_13761; P9211_12581; porin-like protein; Cluster 1351
photosystem I subunit VII; Cluster 1812
adenylosuccinate lyase; Cluster 1662
sodium-dependent transporter NSS family protein P9515_12871; A9601_12951; Cluster 1291
chlorophyll a/b binding light harvesting protein PcbA; Cluster 682
photosystem I reaction center subunit PsaK; photosystem I PsaK protein (subunit X); Cluster 951
thioredoxin family protein; Cluster 1162
hyp. prot. A9601_03531; Cluster 354
hyp. prot. P9301_03541; A9601_03521; Cluster 353
NADH dehydrogenase i chain b or ndhk Cluster 317
hyp. prot. PMT9312_0480; P9215_05611; P9301_05051; P9211_04801; P9303_06881; A9601_05351; NATL1_05361; P9515_05431; PMN2A_1812; cyanobacteria-specific protein; Pro0478; PMT1301; PMM0480; Cluster 536
light-independent protochlorophyllide reductase, iron-sulfur ATP-binding protein; Cluster 598
hyp. prot. PMT9312_0391; A9601_04461; P9211_03921; NATL1_04471; P9515_04571; P9303_21471; P9301_04151; P9215_04721; PMN2A_1730; Pro0392; PMM0395; PMT0207; Cluster 447
O-acetylserine (thiol)-lyase A; Cluster 458
hyp. prot. PMT9312_1366; A9601_14711; P9301_14571; P9515_14331; P9303_19361; P9211_13161; P9215_14971; NATL1_16911; PMN2A_0838; PMM1272; PMT0365; Pro1346; Cluster 1466
aromatic-ring hydroxylase (flavoprotein monooxygenase); Cluster 819
thiamine biosynthesis protein ThiC; Cluster 1816
ABC-type phosphate transport system, periplasmic component; PstS; Cluster 4983
high affinity ammonium transporter; Cluster 286
hyp. prot. NATL1_21051; Cluster 5855
EF-1 guanine nucleotide exchange domain-containing protein PMN2A_0756; Pro1243; Cluster 1319
transaldolase/EF-hand domain-containing protein; Cluster 574
Ycf66 protein N-terminal domain protein; PMT9312_0065; P9301_00741; P9303_03191; P9215_00751; P9515_00721; A9601_00751; NATL1_01271; P9211_00711; PMN2A_1426; Pro0077; PMM0063; PMT1591; Cluster 76
Clp protease ATP-binding subunit; ATPase, AAA family; ClpC; Cluster 1190
6-pyruvoyl tetrahydropterin synthase; Cluster 124
hyp. prot. PMT9312_1512; NATL1_04791; NATL1_19301; Cluster 5836
putative nicotinamide nucleotide transhydrogenase, subunit beta; Cluster 1316
cytochrome c oxidase, subunit I; Cluster 501
hyp. prot. PMT9312_0798; A9601_08541; P9515_07911; P9211_09711; P9215_08851; P9301_08511; P9303_17381; NATL1_08301; PMN2A_0198; Pro0866; PMM0790; PMT0527; Cluster 851
putative Reverse transcriptase (RNA-dependent); Cluster 566
hyp. prot. PMM2035; emb|CAX37114.1|
ribonuclease HI;Cluster 219
DnaJ-class molecular chaperone; Cluster 961
uridylate kinase; Cluster 577
hyp. prot. P9515_00421; Cluster 2230
ABC-type Mn2+/Zn2+ transport system permease components; Cluster 658
putative c-type cytochrome biogenesis protein Ccs1; Cluster 1657
small GTP-binding protein domain-containing protein; hyp.prot. PMT1459; PMM1444; Cluster 1641
phosphatidate cytidylyltransferase; hyp. prot. PMT9312_1791; PMN2A_1306; PMT2247; Cluster 1903
hyp. prot. PMT9312_0604; P9303_18771; P9211_10461; P9515_06691; P9301_06301; P9215_06861; A9601_06601; NATL1_06601; PMN2A_0040; PMT0409; PMM0604; Pro1057; Cluster 659
tRNA (uracil-5-)-methyltransferase Gid; Cluster 1218
P9215_05481; Cluster 2828
NATL1_18041; hemagglutinin-neuraminidase-like; redox protein; Cluster 1615
hyp. prot. A9601_03971; P9515_04041; P9215_04021; P9301_03961; Cluster 398
hyp. prot. P9301_04041; A9601_04031; P9215_04211; P9515_15801; ASNC1092_0013 ; Cluster 404
carbamoyl-phosphate synthase, small subunit; Cluster 906
leucine aminopeptidase 1; Cluster 1527
hyp. prot. PMT9312_0034; P9301_00331; A9601_00331; P9515_00331; P9215_00321; PMM0034; Cluster 34
hyp. prot. PMT9312_1801; P9301_18991; P9215_19811; P9515_19001; P9211_18461; A9601_19181; P9303_30111; NATL1_21921; PMN2A_1319; PMT2263; PMM1708; Cluster 1913
hyp. prot. PMT9312_1523; P9215_16971; A9601_16321; P9301_16201; NATL1_18301; P9303_05141; P9211_15481; P9515_16081; PMN2A_0962; PMM1430; Pro1583; PMT1441; Cluster 1627
PMT9312_0820; P9215_09081; P9515_07691; A9601_08781; P9301_08751; NATL1_14021; PMN2A_0600; PMM0810; Cluster 875
hyp. prot. PMT_2851; emb|CAX32370.1|
hyp. prot. PMT9312_1269; NATL1_11121; P9301_13701; P9215_13821; P9515_13401; A9601_13611; P9211_08081; PMN2A_0415; Pro1001; PMM1169; Cluster 1357
guanosine-3',5'-bis(diphosphate) 3'-diphosphatase, (ppGpp)ase; Cluster 210
RNA recognition motif-containing protein; hyp. prot. PMN2A_0092; Pro0814; Cluster 714
NAD binding site protein; Cluster 575
ferritin; Cluster 865
hyp. prot. PMT9312_0813; P9215_09001; P9515_07761; P9301_08661; PMM0805; Cluster 2840
pterin-4-alpha-carbinolamine dehydratase; Cluster 548
sigma 54 modulation protein / SSU ribosomal protein S30P; Cluster 452
hyp. prot. PMT9312_0008; A9601_00071; P9211_00071; P9303_00081; P9301_00071; P9515_00071; P9215_00071; NATL1_00071; PMN2A_1335; PMT0009; Pro0008; PMM0008; Cluster 8
putative neutral invertase-like protein; neutral invertase like protein; putative neutral invertase-like protein ; Cluster 352
bacterial nucleoid DNA-binding protein; histone-like DNA-binding protein; Cluster 1515
cytochrome c oxidase subunit XV assembly protein; hyp. prot. PMT9312_0446; P9515_05091; A9601_05021; P9303_06441; NATL1_05011; P9215_05261; P9301_04711; P9211_04451; PMN2A_1778; PMT1340; PMM0447; Pro0443; Cluster 503
putative nickel-containing superoxide dismutase precursor (NISOD); Cluster 1488
30S ribosomal protein S21; SSU ribosomal protein S21P ; Cluster 1064
ABC-type transport system involved in resistance to organic solvents permease component; putative transporter, membrane component; transporter membrane component; Cluster 1519
phosphoribosylglycinamide synthetase; Cluster 1536
branched-chain amino acid aminotransferase; Cluster 979
Holliday junction resolvasome DNA-binding subunit; Cluster 2447
pentapeptide repeat-containing protein; hyp. prot. PMT9312_0785; NATL1_08151; P9211_09851; P9515_08061; P9301_08371; A9601_08391; P9215_08711; P9303_16951; PMT0557; PMM0777; Cluster 3988
MazG family pyrophosphatase; PMT9312_0889; PMM0911; PMT0680; Cluster 946
photolyase protein family protein; PMT9312_0423; NATL1_15641; A9601_04781; P9301_04471; P9215_05041; P9515_04871; PMN2A_0726; PMM0425; Cluster 479
nifS-like aminotransferase class-V; cysteine desulfurase; Cluster 189
50S ribosomal protein L33; Cluster 988
ABC-type molybdenum transport system ATPase component; Cluster 868
hyp. prot. PMT9312_1052; P9215_11751; P9301_11461; A9601_11451; PMM1041; HN; Cluster 1141
hyp. prot. PMT9312_0364; P9515_04051; P9303_03601; P9301_03971; A9601_03981; PMM0360; PMT1560; Cluster 399
tRNA (guanine-N(7)-)-methyltransferase; Cluster 263
prokaryotic diacylglycerol kinase; Cluster 4506
putative ribonuclease D; 3'-5' exonuclease; Cluster 1787
cobalamin biosynthetic protein CobN; Cluster 978
heat shock protein DnaJ, N-terminal domain protein; Cluster 2760
hyp. prot. PMT9312_0222; P9301_02411; P9303_03961; NATL1_02981; P9515_02501; P9215_02401; P9211_02411; A9601_02401; PMN2A_1588; PMT1542; PMM0220; Pro0248; Cluster 241
P9215_03611; A9601_03621; P9301_03621; Cluster 383
50S ribosomal protein L28; Ribosomal protein L28 ; Cluster 955
hyp. prot. PMT9312_1655; P9301_17541; P9215_18351; A9601_17701; P9303_22961; P9211_16821; NATL1_20081; P9515_17501; PMN2A_1134; PMM1563; PMT1726; Pro1717; Cluster 1765
hyp.prot. PMT9312_1453; NATL1_17931; P9301_15411; P9303_05231; P9211_14101; P9515_15161; A9601_15561; P9215_15841; PMN2A_0927; Pro1437; PMM1355; PMT1432; Cluster 1551
hyp. prot. PMT9312_0487; NATL1_05431; P9301_05131; P9215_05681; P9515_05501; A9601_05421; P9211_04871; PMN2A_1819; Pro0485; PMM0486 Cluster 543
hyp. prot. PMT9312_1296; P9303_05981; P9301_13961; P9215_14091; P9515_13671; A9601_13881; P9211_12821; NATL1_16691; PMN2A_0814; PMM1195; PMT1381; Pro1311; Cluster 1384
putative apolipoprotein n-acyltransferase; Cluster 1486
hyp. prot. PMT9312_0489; P9303_07041; A9601_05441; P9301_05151; P9211_04891; NATL1_05451; P9515_05521; P9215_05701; PMN2A_1821; PMT1290; Pro0487; PMM0488; Cluster 545
Sec-independent protein secretion pathway component TatA; hyp.prot. NATL1_03291; Cluster 5392
hyp. prot. P9215_11431; P9301_11131; A9601_11141; Cluster 1110
hyp. prot. PMT9312_1273; P9303_17751; P9515_13441; NATL1_08481; P9215_13861; A9601_13651; P9301_13741; P9211_09511; PMN2A_0216; PMT0494; PMM1172; Pro0886; Cluster 1361
macrolide export ATP-binding/permease protein MacB; Cluster 490
hyp. prot. A9601_08771; P9301_08741; P9215_09071; Cluster 874
hyp. prot. PMT9312_0941; P9215_10351; P9301_10041; P9515_07671; P9515_09421; A9601_10061; P9215_09101; A9601_08801; NATL1_07591; P9301_08771; PMM0811; PMM0858;Cluster 5182
hyp. prot. P9301_11121; P9215_11421; A9601_11131; P9515_11061; PMM1018a; Cluster 1109

Figure 4.14. *Prochlorococcus* **orthologous gene clusters for which the pattern of gene expression is significantly impacted by an interaction effect at T12h.** We represent the 127 gene clusters that show an interaction effect. As described in the methods, $\alpha 2$ is indicative of an iron effect, $\alpha 3$ of an organic matter effect and $\alpha 4$ of an interaction effect (note: here the α values are in $\log_{10}$ scale). Clustering of the orthologous gene clusters was performed using the heatmap.2 command in the statistical package R (Computing 2012). The numbers on the dendogram distinguish between the 4 different types of patterns observed. These are discussed in the text.

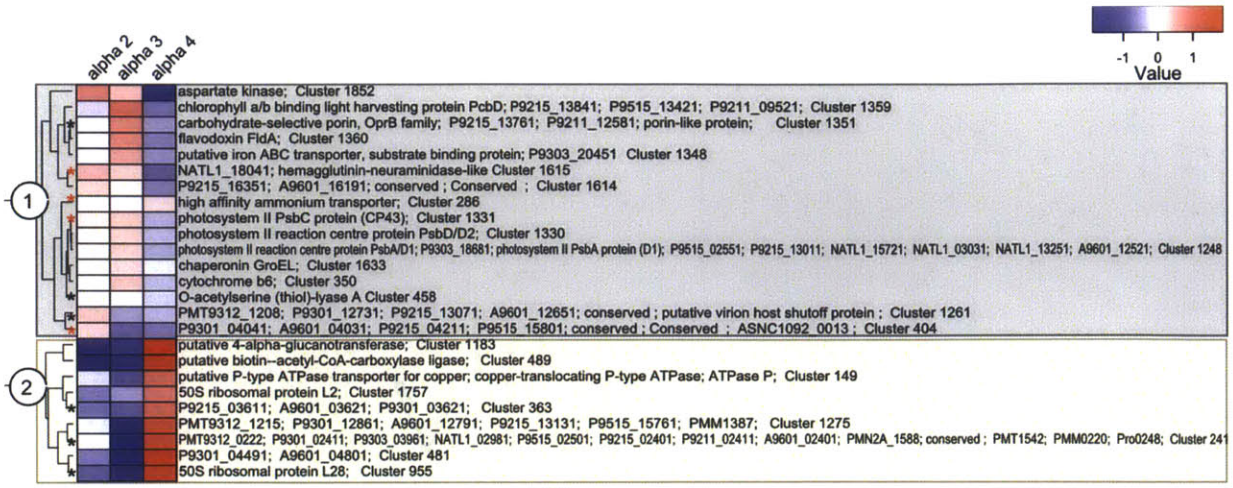


Figure 4.15. *Prochlorococus* **clusters of orthologous genes significantly impacted by an interaction effect at T36h.** We represent the 25 gene clusters that show an interaction effect at T36h. As described in the methods, α2 is indicative of an iron effect, α3 of an organic matter effect and α4 of an interaction effect (note: here the α values are in $\log_{10}$ scale). Clustering of the orthologous gene clusters was performed using the heatmap.2 command in the statistical package R (Computing 2012). The numbers on the dendogram distinguish between the 2 different types of patterns observed. The asterisks indicate the clusters that show an interaction effect at T12h as well (see Figure 13). Red asterisk indicate the clusters for which the value of α4 is of the opposite sign at T12h.

4.4.2.2 Organism Specific Patterns of Gene Expression of Heterotrophs Show no Iron-Specific Response

As was the case for *Prochlorococcus,* the addition of iron alone did not impact the growth of heterotrophs. Because of the widespread impact of iron on the regulation of bacterial genes and because lack of growth does not necessarily mean lack of activity (as mentioned in Section 4.4.2.1), we nonetheless expected a measurable change in gene expression. We found that for all the heterotrophic organisms investigated, the number of orthologous gene clusters significantly impacted by an iron effect was low (Figure 4.6), making it difficult to infer a pathway-level understanding of the iron effect. Table 4.11 lists the organism specific clusters of orthologous genes for which the pattern of gene expression was driven by a significant iron effect (as established by the baySeq analysis). Interestingly, across the five most abundant groups investigated, none of the listed genes are orthologous to each other, suggesting that the iron effect would be organism-specific.

The SAR11 cluster represents the most abundant group of heterotrophs indigenous to surface waters of the North Pacific Subtropical Gyre, at times constituting as much as 30% of the bacterioplankton (Shi et al. 2011). Well-adapted to life at low-nutrient conditions, their abundance as well as their demonstrated ability to utilize various substrates found in dissolved organic matter such as proteins, amino acids, dimethylsulfonopropionate and glucose suggest their potentially important contribution to the cycling of nutrients in the gyre (Alonso & Pernthaler 2006; Malmstrom et al. 2004; Malmstrom et al. 2005; Vila-Costa et al. 2007). Interestingly, we find that addition of iron led to a slight increase in the relative abundance of the transcript coding for the molecular chaperone GroEL (Table 4.11). This could be explained by the fact that increased iron could lead to an oxidative stress response for cells of the SAR11 clade. Indeed GroEL is one of the better characterized microbial chaperone systems and is known – together with its co-chaperone GroES –to be involved in various stress responses, including oxidative stress response (Melkani et al. 2002; Melkani et al. 2008). However, the expression level and abundance of GroEL usually increase in response to iron-limitation (Hennequin et al. 2001). This has also recently been shown in a recent study where GroES protein abundances increased 2-fold in iron-limited cultures of Ca. *Pelagibacter*

ubique, a cultured representative of the SAR11 clade (Smith et al. 2010). This suggests that both iron-limitation and 'iron overload' could lead to a similar oxidative stress response.

Initially identified thanks to molecular techniques, members of the SAR116 clade have been found to be widespread in surface waters of the global ocean and representative members have just recently been brought into culture (Mullins et al. 1995; Rappé et al. 1997; Acinas et al. 1999; Rappé et al. 2000; Suzuki et al. 2001; Grote et al. 2011). While it is still unclear whether microorganisms belonging to this clade are motile, we found that at both T12h and T36h the relative transcript abundance of two genes (*flgE* at T12h, and *fliQ* at T36h) coding for flagellum proteins significantly decreased in the iron treatment (Table 4.11). This finding is interesting, as iron has been shown in the past to up-regulate flagellar-assembly processes and chemotaxis in the gamma-proteobacterium *Vibrio parahaemolyticus* (McCarter & Silverman 1989). The authors of this study rationalized that in an iron-poor environment, the cell would increase its motility in order to maintain itself in an iron-rich environment. It is unclear why in the present experiment, flagellar proteins would be potentially down-regulated in response to iron addition.

The increase in the relative transcript abundance of *Pseudoalteromonas* cysteine desulfurase and isocitrate lyase in the iron addition is interesting. The cysteine desulfurase protein IscS is an iron-binding protein that is involved in the assembly of Fe-S clusters (Flint 1996). However, it is generally thought to take on this role when cells are subjected to oxidative stress or iron starvation (Fontecave et al. 2008), as is observed with the close to 2-fold upregulation of SufC, the SAR11 IscS homolog in iron-limited cultures of *Ca. P. ubique* (Smith et al. 2010).

The anecdotal nature of these observations indicate that our current understanding of how iron may regulate certain genes is limited, making a sound interpretation of these signals difficult. Furthermore, the low number of orthologous gene clusters for which a significant iron effect is detected could suggest that the addition of iron alone may have little effect on the heterotrophic picoplankton community.

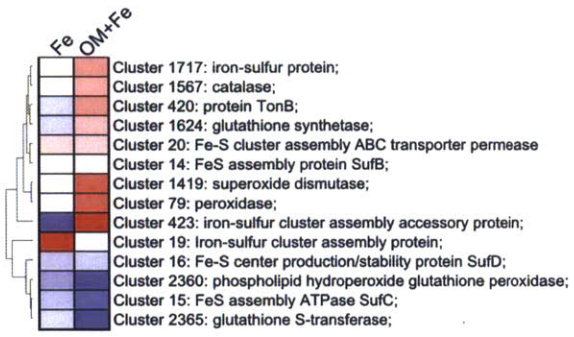
As a complementary approach, we tracked the response of the orthologous gene clusters for which an iron response was expected (Figure 4.16). These orthologous gene clusters include those with the

following annotation descriptions: TonB-dependent transporters, siderophore, Fe^{3+} ABC transporters, Fe-S clusters, and genes involved in responding to oxidative damage, namely: superoxide dismutase, catalase, peroxidase, carotenoids, and glutathione. In Figure 4.16, two values are represented for each organism specific orthologous cluster: the $\log_2$ value of the ratio of the relative transcript abundance in the (+Fe) sample to the relative transcript abundance in the (Control) sample and the $\log_2$ value of the ratio of the relative transcript abundance in the (+OM+Fe) sample to the relative transcript abundance in the (+OM) sample. We found that for a given organism orthologous gene clusters with the same predicted function (ie. iron transporters) behave differently (ie. Clusters 1254 and 73 in SAR11). While some of the orthologous clusters represented in Figure 4.16 show important changes in relative transcript abundance between the +Fe sample compared to the Control sample and in the OM+Fe sample compared to the +OM sample, these were not deemed to be indicative of a statistically significant iron effect per the statistical baySeq test used. This suggests that the effect of iron alone may not have a significant effect on the heterotrophic community.

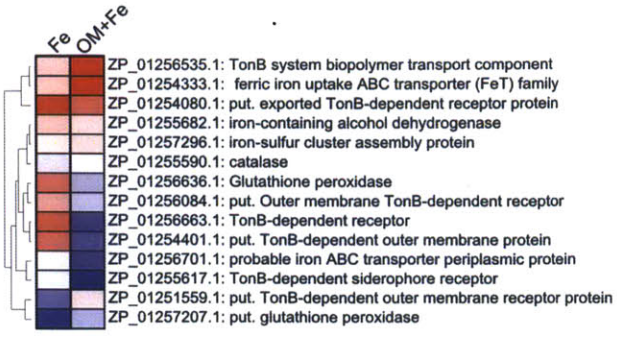
Table 4.11. List of orthologous gene clusters significantly impacted by an iron effect (sum of probabilities >0.9) for the most abundant heterotrophs. Bolded are the clusters for which the iron-responsiveness is discussed in the text. α2 indicates the magnitude of the iron effect and is calculated for each orthologous gene cluster as the $\log_{10}$ ratio of the relative transcript abundance in the (+Fe) sample to the relative transcript abundance in the (Ctrl) sample.

Organism	Orthologous Cluster #	Accession numbers of protein coding genes that form the cluster	Description of Cluster	α2 (log_{10})
SAR11 cluster (T12h)	**Cluster 1187**	**YP_004357857.1 YP_265587.1 ZP_01264674.1 ZP_05069866.1 ZP_06054579.1 ABR27782 ABR27777 ABR27776**	**Chaperonin GroEL**	**0.180**
	Cluster 1296	YP_004358204.1 YP_266038.1 ZP_01264216.1 ZP_05069975.1 ZP_06055235.1	ispB; octaprenyl diphosphate synthase	-0.209
	Cluster 1302	YP_004357946.1 YP_265510.1 ZP_01264751.1 ZP_05069981.1 ZP_06054648.1	folB; dihydroneopterin aldolase	-0.949
	Cluster 2218	YP_265684.1 ZP_01264576.1	hypothetical protein PU1002_05061; SAR11_0259;	-0.984
	Cluster 1031	YP_004358296.1 YP_266094.1 ZP_01264166.1 ZP_05069710.1 ZP_06055035.1	S-adenosyl-L-methionine dependent methyltransferase	-0.569
	Cluster 559	YP_004358452.1 YP_266537.1 ZP_01265111.1 ZP_05069238.1 ZP_06054874.1	tufA; elongation factor EF-Tu	0.164
	Cluster 636	YP_004357981.1 YP_265445.1 ZP_01264817.1 ZP_05069315.1 ZP_06054680.1	outer membrane assembly lipoprotein YfiO; putative competence lipoprotein ComL; DNA uptake lipoprotein	-0.950
SAR116 (T12h)	**Cluster 866**	**YP_003551194.1**	***flgE* flagellar hook protein;**	**-1.06**
Psychroflexus torquis (T12h)	None	ZP_01257225.1	DNA-directed RNA polymerase, beta' subunit	-1.00
	None	ZP_01256203.1	hypothetical protein P700755_26587	0.74
	None	ZP_01257149.1	hypothetical protein P700755_26060	-0.96
	None	ZP_01257080.1	putative adenylate cyclase (iron-containing protein)	1.07
SAR11 (T36h)	Cluster 1274	YP_004357927.1 YP_265528.1 ZP_01264733.1 ZP_05069953.1 ZP_06054631.1	RNA methyltransferase, TrmH family, group 2	-1.44
	Cluster 936	YP_266631.1 ZP_01264985.1 ZP_05069615.1	sarcosine dehydrogenase	0.54
	Cluster 1045	YP_004357552.1 YP_265815.1 ZP_01264442.1 ZP_05069724.1 ZP_06055548.1	30S ribosomal protein S15	1.28
SAR116 (T36h)	Cluster 820	YP_003551148.1	***fliQ* flagellar biosynthetic protein FliQ**	-1.204
	Cluster 1163	YP_003551491.1	pbuG; putative MFS transporter, xanthine/uracil permease	1.022
P. torquis (T36h)	None	ref|ZP_01256273.1|	hypothetical protein P700755_32769	-0.701
Pseudoalteromonas (T36h)	Cluster 1004	YP_004069937.1 YP_341288.1 YP_660591.1 ZP_08409059.1 EAR26349.1	30S ribosomal subunit protein S13, rpsM gene product	-1.33
	None	ZP_01136100	preprotein translocase, membrane component	-0.97
	Cluster 1208	**YP_004065013.1 YP_341558.1 YP_660795.1 ZP_01131853.1 ZP_08408920.1 EAR30219.1**	**aceA; isocitrate lyase**	**0.71**
	Cluster 1231	**YP_004067530.1 YP_341161.1 YP_660818.1 ZP_01133138.1 ZP_08408231.1 EAR29926.1**	**iscS, NFS1; cysteine desulfurase**	**0.45**
	Cluster 1568	YP_004068113.1 YP_339467.1 YP_661155.1 ZP_01135161.1 ZP_08411061.1 EAR27538.1	tRNA (guanine37-N1) –methyltransferase	0.93

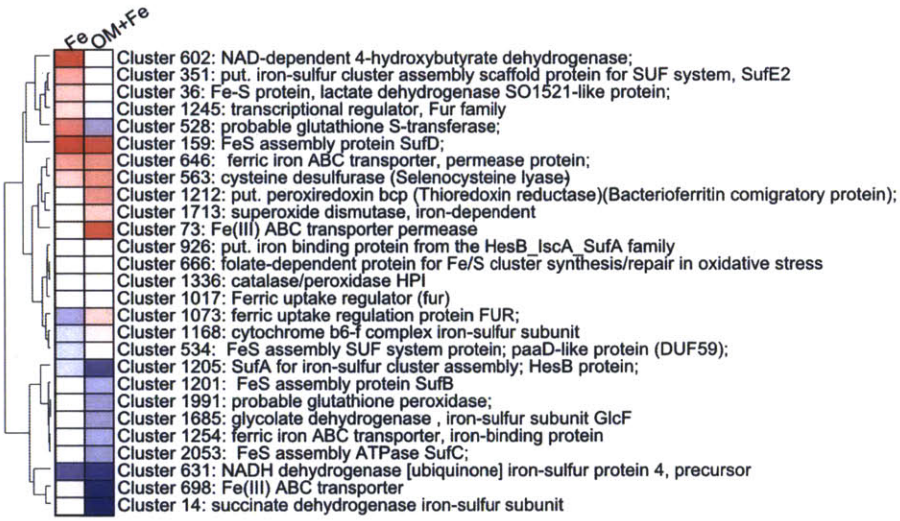
A. SAR116 T36h



B. Psychroflexus torquis T36h



C. SAR11 T36h



D. Pseudoalteromonas T36h

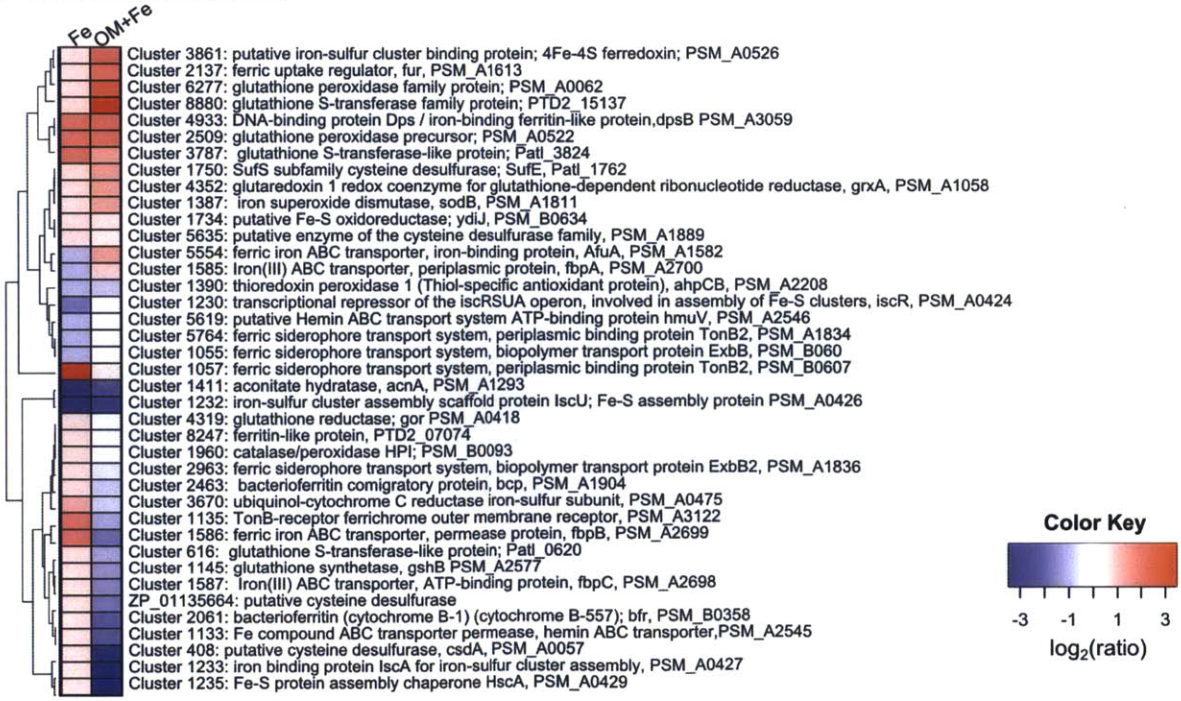


Figure 4.16. Fold change in relative transcript abundance of iron-specific orthologous gene clusters at T36h for organism specific reads. Values represented in the "Fe" column are the $\log_2$ values for the ratio of the relative transcript abundance in the Fe treatment to the relative transcript abundance in the control treatment. Values represented in the "OM+Fe" column are the $\log_2$ values of the ratio of the relative transcript abundance in the OM+Fe treatment to the relative transcript abundance in the OM alone treatment. The values are shown for the clusters of orthologous genes hypothesized to be iron-responsive (per annotation) and with a minimum read count of 5 in at least one of the 4 samples (T36h Ctrl, T36h Fe, T36h OM, T36h OM+Fe) considered. Panel A. SAR116 cluster, encompassing Ca. *Puniceispirullum.* Panel B. *Psychroflexus torquis.* Panel C. SAR11 cluster encompassing Ca. *Pelagibacter ubique.* Panel D. Pseudoalteromonas. None of the clusters of orthologous genes represented show a significant iron effect per the baySeq analysis performed.

Organic Matter Effect – Statistical analysis revealed that changes in patterns of gene expression of the heterotrophs examined (SAR11 cluster, SAR116 cluster, *Psychroflexus torquis*, *Pseudoalteromonas sp.*), were significantly influenced by an organic matter effect (Figure 4.6). The widespread effect of organic matter on the global gene expression patterns of SAR11-specific transcripts confirmed this (Figure 4.17). As was the case with *Prochlorococcus*, the global pattern of expression of orthologous gene clusters attributed to the SAR11 cluster was impacted by an organic matter effect (Figure 4.17). In order to determine the impact of the organic matter on this microbial group, we proceeded to compare the relative transcript abundance values of the main types of transporters found to dominate the genomes of *Pelagibacter ubique* (Figure 4.18) (Giovannoni et al. 2005). We found that the added organic matter led to a decrease in the relative transcript abundance of ammonium and phosphate transporters at T12h, and an increase in the relative transcript abundance of phosphate transporters at T36h. This response is very similar to the one observed for *Prochlorococcus*.

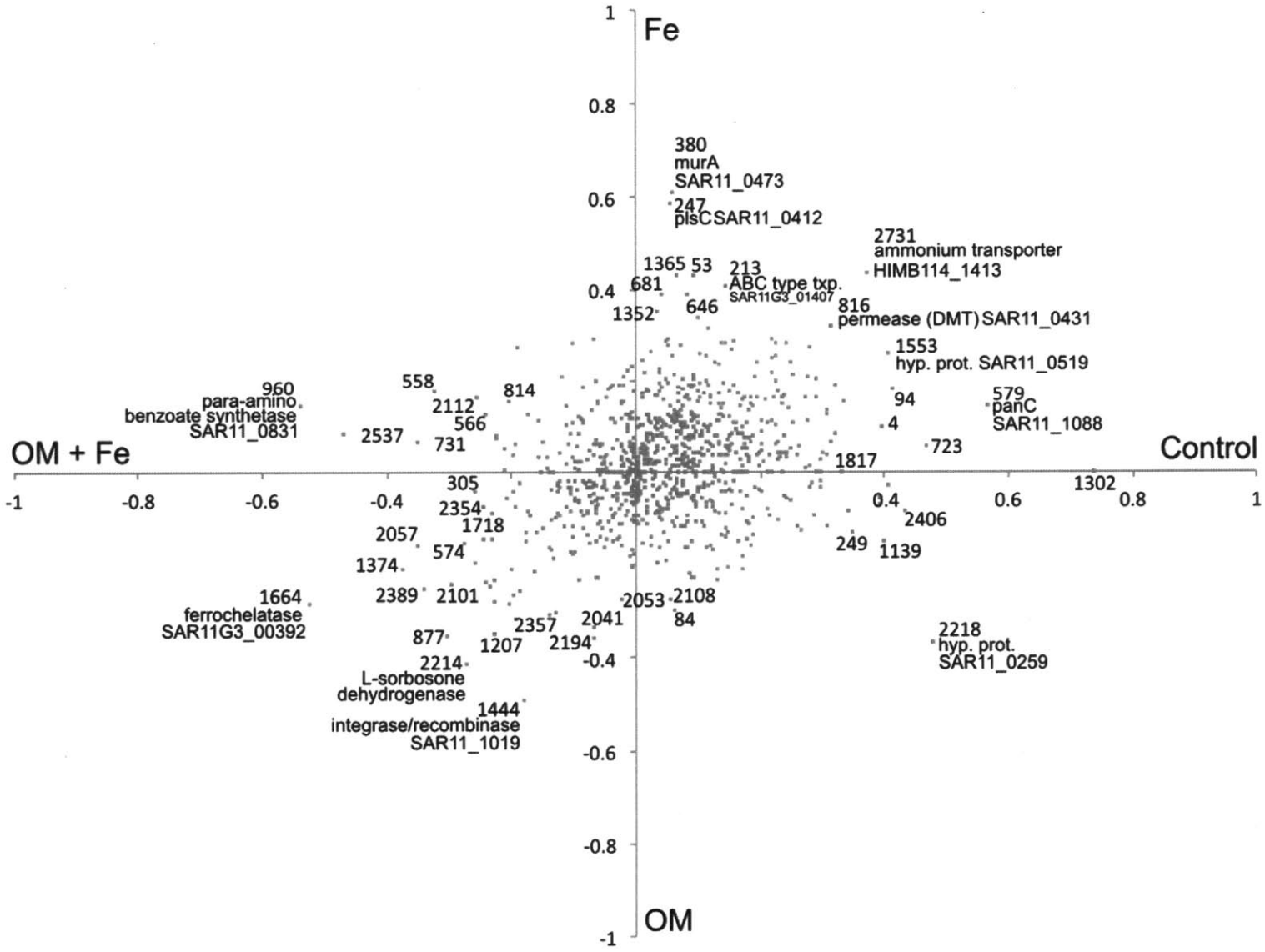

Fe
OM
OM + Fe
Control
1
0.8
0.6
0.4
-0.4
-0.6
-0.8
-1
-1
-0.8
-0.6
-0.4
0.4
0.6
0.8
1
380
murA
SAR11_0473
247
plsC SAR11_0412
2731
ammonium transporter
HIMB114_1413
1365
53
213
ABC type txp.
SAR11G3_01407
681
646
1352
816
permease (DMT) SAR11_0431
1553
hyp. prot. SAR11_0519
960
para-amino
benzoate synthetase
SAR11_0831
558
814
2112
566
2537
731
94
579
panC
SAR11_1088
4
723
1817
1302
305
2354
1718
2057
574
1374
2406
249
1139
1664
ferrochelatase
SAR11G3_00392
2389
2101
2053
2108
84
2357
2041
877
1207
2194
2214
L-sorbosone
dehydrogenase
1444
integrase/recombinase
SAR11_1019
2218
hyp. prot.
SAR11_0259

Figure 4.17. Influence of organic matter on *SAR11* is observed in gene expression pattern at T12h. Clusters of *SAR11* orthologous genes are plotted with respect to their relative representation in each of the four treatments. Only the clusters with at least 10 counts in at least one of the four datasets were considered for the analysis. Hence we represent the values for 874 out of 2,762 clusters. For each orthologous cluster, the relative transcript abundance value was normalized to the sum of the relative transcript abundance values in the 4 treatments. X coordinates: (normalized control - normalized OM + Fe). Y coordinates: (normalized Fe - normalized OM). The pattern of data points follows a distribution along the Fe/Control to OM/ OM+Fe axis. A similar distribution is observed for the T36h data (not shown). The descriptions of some clusters on the edges (for ease of reading) and the Ca. *P. ubique* are indicated.

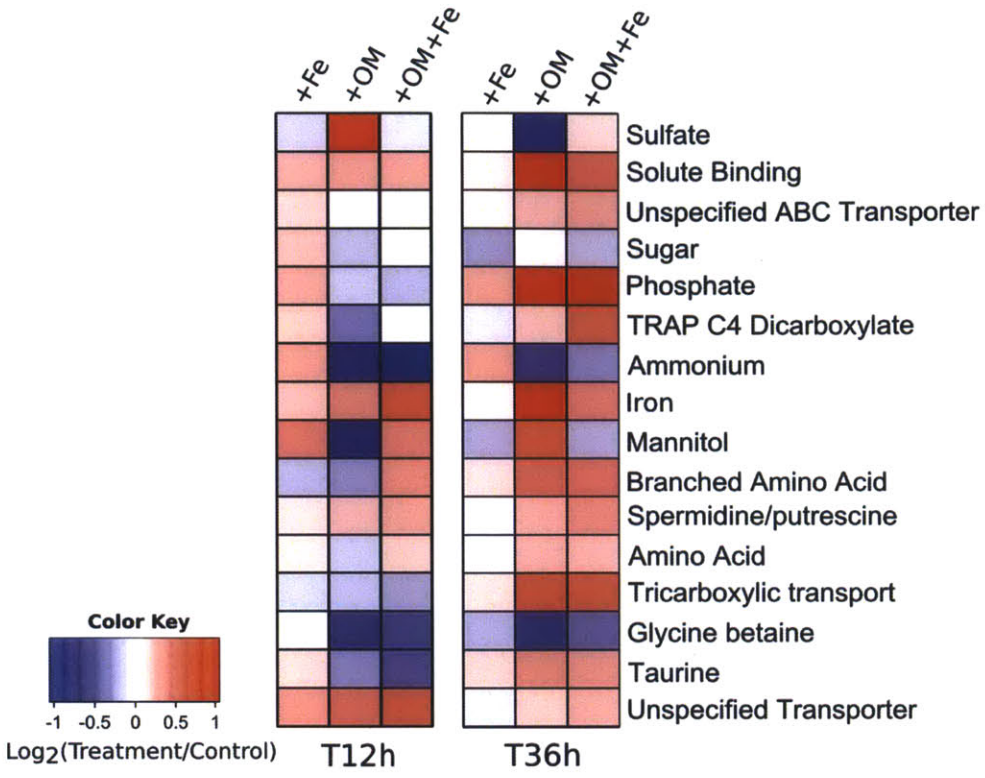


Figure 4.18. Relative transcript abundance of specific transporters in SAR11. We show the log-2 ratio of the relative transcript abundance in a treatment to the relative transcript abundance in the control. Relative transcript abundance refers to the number of hits to each type of transporter normalized to the total number of SAR11 reads in each sample. Transporters were manually grouped per their predicted substrate specificity (Giovannoni et al. 2005). At T12h, relative transcript abundance of both ammonium and phosphate transporters decrease in the OM and OM+Fe treatments, while by T36h, the relative transcript abundance of phosphate transporters increase.

4.5. Discussion

4.5.1 Despite Low Iron Concentrations in the NPSG, Other Nutrients Appear More Limiting

Lack of growth stimulation by the iron alone amendments – Our results suggest that overall the microbial community was not affected by the addition of iron (Figure 4.4). The lack of growth response confirms that the ambient iron concentration at HOT station ALOHA does not limit the growth of both the bacterioplankton and the ambient phytoplankton communities. This is not completely surprising as the lowest dissolved iron concentration recorded to date from surface seawater at station ALOHA is close to ~0.2 nM (Boyle et al. 2005; Noble et al., 2008), likely sufficient to support the iron requirement of small celled heterotrophs, which because they do not need to sustain an iron-rich photosynthetic apparatus are thought to have much lower iron requirements than photosynthetic organisms. Accordingly, our data also indicate that over the time period measured (36 hours), the addition of iron did not stimulate the indigenous autotrophic picoplankton community at station ALOHA (Figure 4.4). A recent study has suggested that the genomes of *Prochlorococcus* species found in extremely iron limited environments had fewer iron-requiring proteins than their coastal counterparts (Rusch et al. 2010). These data may indicate that open ocean *Prochlorococcus* cells are inherently less responsive to changes in iron bioavailability and might explain in part the lack of growth response we observe.

Blooms of diatoms are periodically observed in the NPSG (Villareal et al., 2012). These blooms usually consist of diatom species that harbor diazotroph symbionts, which help sustain the growth of the diatom species in waters with such low nitrate concentrations (Villareal et al., 2012). A number of factors have been hypothesized to drive these blooms (Villareal et al., 2012), one of which could be the deposition of iron-laden dust (Calil et al., 2011). Indeed, both diazotrophs and diatoms have high iron requirements. While our metagenomic and metatranscriptomic approach was not designed to examine responses from large eukaryotic phytoplankton such as diatoms (our pre-filter size excluded them, and the time-course monitored was limited due to volume constraints), we nonetheless did not observe any bloom (usually visible by eye) in our iron-enriched

incubation bottles. One possible explanation is that the duration of our experiment (36 hours) was too short to observe any bloom of large eukaryotic phytoplankton. Indeed, iron addition experiments that have resulted in the relief of iron limitation of larger eukaryotic species have typically been conducted over the course of many days (Boyd et al. 2007), which was beyond the practical limits of our experimental design. A more likely explanation, however is that in the NPSG, iron concentrations may be too high to effectively limit the growth of these diatom species. While one study (Blain et al., 2002) showed that some diatom species could be iron-limited at iron concentrations as high as ~0.33 nM, it is generally accepted that iron concentrations limiting phytoplankton growth are closer to ~0.1 nM as is the case in most HNLC regions.

Lack of iron-specific response in the gene expression pattern of heterotrophs – In addition to the lack of iron-induced growth stimulation of the heterotrophic community, there was little impact of iron on the gene expression patterns of the most abundant heterotrophs (Figure 4.6). Indeed, because of iron's poor solubility but also because of its potential toxicity when generating reactive oxygen species if present in large quantities inside the cell (Touati 2000), iron homeostasis is tightly regulated. Accordingly, in some microbial species iron is known to impact the expression level of as many as ~90 genes, involved in a number of different pathways (Andrews et al. 2003). While an iron-mediated growth enhancement of heterotrophs was not necessarily expected (see reasons mentioned above), we nonetheless did not observe any significant impact of iron alone on the gene expression of the most abundant heterotrophs (Figure 4.6, Table 4.11) is rather surprising. The few orthologous clusters displaying a significant iron effect for which regulation by iron was expected provided confusing results: their gene expression pattern in the (+Fe) treatment was characteristic of expression patterns usually observed under iron limitation (Figure 4.16). Furthermore, because the number of orthologous gene clusters significantly impacted by iron was low, identifying a meaningful biological explanation at the pathway level is difficult (Table 4.11).

Effect of organic matter on growth and gene expression patterns of *Prochlorococcus* – Our findings show that while the addition of iron alone did not affect microbial community structure and composition in microcosms performed over the course of 36 hours (Figure 4.5), the combined addition of iron and organic matter reveals significant changes in expression patterns of *Prochlorococcus* genes (Figures 4.14 and 4.15). A response indicative of relief of iron limitation was observed for *Prochlorococcus* physiology when comparing (+OM) and in the (+OM+Fe) samples (Figure 4.8). This could suggest, as we hypothesized, that organic matter (presumably containing iron-binding ligands) renders iron more available to the microbial community. However, the overwhelming response of the community (both in composition and activity) to the organic matter addition (Figures 4.4 and 4.5) and the down-regulation of *Prochlorococcus* genes diagnostic of nitrogen limitation (Figure 4.12), suggest that a more appropriate explanation is that the organic matter provides a source of otherwise limiting nutrients (here nitrogen).

Indeed, the NPSG is usually thought of as a nitrogen limited environment (Karl et al. 2001). A recent study showing an increase in the synthesis of *Prochlorococcus* RNA upon addition of ammonium to seawater, confirms the nitrogen limitation status of *Prochlorococcus* cells found near station ALOHA (Van Mooy & Devol 2008). Here, we argue that the *Prochlorococcus* population seeding our experiment was nitrogen limited as well, in part explaining why the addition of iron alone had no effect on the growth or the gene expression pattern of *Prochlorococcus* (Figure 4.4, Table 4.9 and 4.10). Upon addition of the organic matter, the nitrogen limitation was relieved, as diagnosed by a down-regulation of urea and cyanate transport genes observed at T12h and T36h (Figure 4.12). Consistent with this hypothesis is the increase in the relative transcript abundance of ribosomal proteins (Figure 4.10) presumably leading to cell growth in the organic matter amendments. By T36h *Prochlorococcus* seemed to have utilized the available phosphate and showed signs of phosphorus limitation, as indicated by the increase in relative transcript abundance of phosphate and organic phosphorus acquisition genes (Figure 4.11). Therefore, the lack of physiological response from *Prochlorococcus* to the addition of iron alone can be explained by the fact that it is more likely limited by nitrogen than by iron.

The physiological response of *Prochlorococcus* cells to organic matter addition (increase in cell size and chlorophyll content per cell, see Figure 4.8) is consistent with these observations. Indeed, reduction of cell size is a general stress response observed of bacteria when starved for nutrients (including nitrogen) (Kjellegerg 1993) and increase in *Prochlorococcus* chlorophyll cellular content has previously been observed upon relief of nitrogen limitation (Moore et al. 2008). Interestingly, both of these physiological responses have also been observed upon relief of iron limitation of *Prochlorococcus* cells in the Equatorial Pacific (Cavender-Bares et al. 1999; Mann & Chisholm, 2000). Since the organic matter utilized in this experiment had a residual iron concentration of 0.25nM (Figure 4.2), it is difficult to tease out which of iron or nitrogen (or both) is responsible for this effect. Because these physiological responses are enhanced at T36h in the OM+Fe treatment relative to the +OM treatment (Figure 4.8), but similar in the two organic treatments and non-existent in the iron alone treatment at T12h, we suggest that addition of organic matter first relieves nitrogen limitation, driving cell growth (as measured by increased cell size and increased expression of ribosomal proteins), which then leads to a depletion of the other nutrients (phosphate and iron). While the data presented here supports this hypothesis, a more complete chemical analysis of the added organic matter to establish the elemental ratio of the added carbon, nitrogen, phosphorus and iron would be needed to further support this scenario.

4.5.2 Organic matter – a poor model of iron-binding ligands

In addition to gaining a better understanding of the impact of iron on the microbial community of the NPSG, we were interested in exploring whether organic matter enhances iron bioavailability. Our data is however inconclusive in that regard. Indeed, we show that members of the community utilize the organic matter as a nutrient source, making it hard to determine whether an interaction effect (between organic matter and iron) can be attributed to higher iron bioavailability or to an interaction between the nutrient source and iron. To address this problem, the prior identification of the fraction that effectively binds iron and adding this fraction instead of the bulk organic matter would be desirable. A technique to do this has recently been perfected in

Dan Repeta's lab (Boiteau et al., *in prep*) and should be used in future experiments. In light of the present experiment and in order to observe any measurable effect of the addition of the obtained iron-binding fraction, a specific nitrogen-source should be added prior the start of the experiment in order to relieve any macronutrient stress and render the system iron-limited. It would also be desirable to perform such studies in a more iron-limited environment.

If the hypothesis that organic matter (containing iron-binding ligands) enhances iron bioavailability by acting as a source of siderophores were true, we would have expected the addition of organic matter to at least have an effect on the gene expression level of transport mechanisms (ie. TonB outer membrane receptors, bacterial surface reductases, ferric and ferrous iron transporters etc). Our analysis of gene expression patterns reveals that this is not the case (Tables 4.9, 4.10 and 4.11). Such expected changes at the transcriptomic-level assume that marine microorganisms would display a regulatory response to iron. Interestingly, it has been suggested that in the case of marine microorganisms, genes involved in iron-metabolism may always be turned on, implying that an iron-specific response may not be detected at the transcriptional level (Marchetti et al., 2012). To confirm that organic matter enhances iron availability, determination of iron-uptake rates could effectively complement our genomic and transcriptomic-based approach. Because of the novelty of this experiment and the difficulty associated with carrying out trace-metal clean incubation experiments with large volumes such as the ones utilized in this experiment, the water budget could not be increased to accommodate this additional measurement, but should in future experiments.

4.5.3 Towards a better understanding of the role of iron in the NPSG

The present experiments reported here show that iron alone does not dramatically impact the heterotrophic and autotrophic microbial community in the NPSG. However, it remains to be investigated how iron impacts the indigenous diazotrophic community. Indeed, the NPSG is permanently stratified with a rather shallow mixed-layer depth, which leads to extreme oligotrophic conditions with chronic nutrient-depletion and low standing stocks of biomass in surface waters

(Karl & Lukas 1996). Despite these conditions, episodic phytoplankton blooms are regularly observed during the summer months (Wilson et al., 2003; Dore et al., 2008), resulting in pulses of particulate matter export to the deep sea, key for sustaining long-term productivity of the gyre (Karl et al. 2012). Recent studies have suggested that these summer-time blooms are actually fueled by diazotrophs that provide a biologically available nitrogen source via N_2 fixation when nitrate and nitrite concentrations are minimal (Calil et al., 2011). Because nitrogen fixation is one of the most iron-intensive microbial processes, with some diazotrophs requiring one order of magnitude more iron than some eukaryotic phytoplankton (Kustka et al. 2003), we can expect that iron availability is likely to play a role in fueling and/or maintaining these blooms (Moore et al., 2009; Dutkiewicz et al., 2012). In fact, some studies performed at station ALOHA have suggested that processes such as aerial iron-laden dust deposition are indeed responsible for fueling those blooms (Calil et al., 2011). Therefore, determining how iron may affect these diazotroph populations would contribute to a better understanding of what factors control these episodic blooms. This would require extending the analysis to cells larger than the ones included in this study. Indeed, most of the diazotrophs found in the NPSG are larger than the pores of the pre-filters used in this study. Of special interest will be the exploration of how iron impacts the *in situ* gene expression patterns of these diazotrophs as well as their nitrogen fixation rates.

4.6. Conclusions and Future Directions

The goals of this study were two-fold: 1) characterize the iron-specific response of the microbial community characteristic of a low-nitrate low-chlorophyll regime, and 2) determine whether the addition of organic matter together with iron would enhance the iron-specific response identified. We showed that addition of iron alone does not impact the growth of either the heterotrophic nor the autotrophic members of the community. While the addition of iron alone did not seem to have a dramatic effect on the gene expression pattern of the most abundant heterotrophs investigated, it had a slight effect on the gene expression pattern of the most abundant autotroph *Prochlorococcus*. Finally, we find that the iron-specific response observed for *Prochlorococcus* is enhanced in the presence of organic matter. We argue, however that this is a result of the added nutrients (organic nitrogen) brought with the organic matter than from an increase in iron bio-availability.

This study is one of the first that takes a meta-transcriptomic approach to understanding the impact of iron on the microbial community composition and function in the North Pacific Subtropical Gyre. While this approach allows for the identification of regulation processes without *a priori* knowledge of the responsive genes, we find that significant changes in gene expression patterns are dominated by relief of nutrient limitation. Having identified and discussed caveats associated with the experimental design, we propose an improved approach to test the hypothesis that iron-binding ligands enhance iron availability. Furthermore, this study opens the door to future studies that could answer some important questions in the field, namely: how does the diazotrophic community, integral to the productivity of the gyre respond to changing iron concentrations?

4.7 Acknowledgements

I thank Erin Bertrand, Mar Nieto Cid, Kathleen Munson for making this project possible by training me in trace metal clean techniques and accompanying me to conduct the experiment at sea. Thanks to Dave Karl and the crew of the R/V Kilo Moana for assistance during the expedition. Many thanks to Elizabeth Ottesen for helping me navigate the bioinformatics. Thanks to Rob Young for his help with the statistical analyses. I also thank Anne Thompson for teaching me how to use the flow cytometer in the Chisholm lab. Finally, thanks to C-MORE EDventures for providing financial support for part of the logistics related to conducting the experiment at sea.

4.8 References

Acinas, S.G., Antón, J. & Rodríguez-Valera, F., 1999. Diversity of free-living and attached bacteria in offshore Western Mediterranean waters as depicted by analysis of genes encoding 16S rRNA. *Applied and environmental microbiology*, 65(2), pp.514-22.

Alonso, C. & Pernthaler, Jakob, 2006. Roseobacter and SAR11 dominate microbial glucose uptake in coastal North Sea waters. *Environmental microbiology*, 8(11), pp.2022-30.

Altschul, S.F. et al., 1990. Basic Local Alignment Search Tool. *Journal of Molecular Biology*, 215, pp. 403-410.

Andrews, S.C., Robinson, A.K. & RodrÃguez-QuiÃ±ones, F., 2003. Bacterial iron homeostasis. *FEMS Microbiology Reviews*, 27(2-3), pp.215-237.

Beinert, H., Kennedy, M.C. & Stout, C.D., 1996. Aconitase as Iron–Sulfur Protein, Enzyme, and Iron-Regulatory Protein. *Chemical Reviews*, 96(7), pp.2335-2374.

van den Berg, C.M.G., 1995. Evidence for organic complexation of iron in seawater. *Marine Chemistry*, 50(1-4), pp.139-157.

Blain, S. et al., 2002. Quantification of algal iron requirements in the Subantarctic Southern Ocean (Indian sector). *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(16), pp.3255-3273.

Boyd, P.W. et al., 2000. A mesoscale phytoplankton bloom in the polar Southern Ocean stimulated by iron fertilization. *Nature*, 407(6805), pp.695-702.

Boyd, P.W. et al., 2007. Mesoscale iron enrichment experiments 1993-2005: synthesis and future directions. *Science (New York, N.Y.)*, 315(5812), pp.612-7.

Boye, M. et al., 2003. Horizontal gradient of the chemical speciation of iron in surface waters of the northeast Atlantic Ocean. *Marine Chemistry*, 80(2-3), pp.129-143.

Boye, M. & van den Berg, C.M.G., 2000. Iron availability and the release of iron-complexing ligands by Emiliania huxleyi. *Marine Chemistry*, 70(4), pp.277-287.

Boyle, E. a. et al., 2005. Iron, manganese, and lead at Hawaii Ocean Time-series station ALOHA: Temporal variability and an intermediate water hydrothermal plume. *Geochimica et Cosmochimica Acta*, 69(4), pp.933-952.

Buck, K.N., Bruland, K.W. 2007. The physicochemical speciation of dissolved iron in the Bering Sea. *Limnology and Oceanography*, 52(5), pp. 1800-08.

Calil, P.H.R. et al., 2011. Episodic upwelling and dust deposition as bloom triggers in low-nutrient, low-chlorophyll regions. *Journal of Geophysical Research*, 116(C6), pp.1-17.

Cavender-Bares, K.K. et al., 1999. Differential response of equatorial Pacific phytoplankton to iron fertilization. *Limnology and Oceanography*, 44(2), pp.237-246.

Chen, M., 2004. Phase partitioning and solubility of iron in natural seawater controlled by dissolved organic matter. *Global Biogeochemical Cycles*, 18(4), pp.1-12.

Cochlan, W., 2001. The heterotrophic bacterial response during a mesoscale iron enrichment experiment (IronEx II) in the eastern equatorial Pacific Ocean. *Limnology and Oceanography*, 46(2), pp.428-435.

Computing, R.F. for S., 2012. R: A Language and Environment for Statistical Computing.

Cotner, J.B. & Biddanda, B.A., 2002. Small Players, Large Role: Microbial Influence on Biogeochemical Processes in Pelagic Aquatic Ecosystems. *Ecosystems*, 5(2), pp.105-121.

DeLong, E.F. et al., 2006. Community genomics among stratified microbial assemblages in the ocean's interior. *Science (New York, N.Y.)*, 311(5760), pp.496-503.

Dutkiewicz, S., Ward, B.A., Monteiro, F., Follows, M.J., 2012. Interconnection of nitrogen fixers and iron in the Pacific Ocean: Theory and numerical simulations. *Global Biogeochemical Cycles*, 26, GB1012.

Eilers, H., Pernthaler, J & Amann, R., 2000. Succession of pelagic marine bacteria during enrichment: a close look at cultivation-induced shifts. *Applied and environmental microbiology*, 66(11), pp.4634-40.

Flint, D.H., 1996. Escherichia coli contains a protein that is homologous in function and N-terminal sequence to the protein encoded by the nifS gene of Azotobacter vinelandii and that

can participate in the synthesis of the Fe-S cluster of dihydroxy-acid dehydratase. *The Journal of biological chemistry*, 271(27), pp.16068-74.
Fontecave, M. & Ollagnier-de-Choudens, S., 2008. Iron-sulfur cluster biosynthesis in bacteria: Mechanisms of cluster assembly and transfer. *Archives of biochemistry and biophysics*, 474(2), pp. 226-37.
Frias-Lopez, J. et al., 2008. Microbial community gene expression in ocean surface waters. *Proceedings of the National Academy of Sciences of the United States of America*, 105(10), pp.3805-10.
Giovannoni, S.J. et al., 2005. Genome streamlining in a cosmopolitan oceanic bacterium. *Science (New York, N.Y.)*, 309(5738), pp.1242-5.
Gledhill, M. & van den Berg, C.M.G., 1994. Determination of complexation of iron(III) with natural organic complexing ligands in seawater using cathodic stripping voltammetry. *Marine Chemistry*, 47(1), pp.41-54.
Grote, J. et al., 2011. Draft genome sequence of strain HIMB100, a cultured representative of the SAR116 clade of marine Alphaproteobacteria. *Standards in Genomic Sciences*, 5(3), pp.269-278.
Hardcastle, T.J. & Kelly, K. a, 2010. baySeq: empirical Bayesian methods for identifying differential expression in sequence count data. *BMC bioinformatics*, 11, p.422.
Hassler, C.S. et al., 2011. Saccharides enhance iron bioavailability to Southern Ocean phytoplankton. *Proceedings of the National Academy of Sciences of the United States of America*, 108(3), pp.1076-81.
Hennequin, C., Collignon, a & Karjalainen, T., 2001. Analysis of expression of GroEL (Hsp60) of Clostridium difficile in response to stress. *Microbial pathogenesis*, 31(5), pp.255-60.
Hudson, R.J.M. & Morel, Francois M M, 1990. Iron Transport in Marine Phytoplankton : Kinetics of Cellular and Medium Coordination Reactions. *Limnology and Oceanography*, 35(5), pp. 1002-1020.
Huse, S.M. et al., 2007. Accuracy and quality of massively parallel DNA pyrosequencing. *Genome biology*, 8(7), p.R143.
Hutchins, D.A. et al., 1999. Competition among marine phytoplankton for different chelated iron species. *Nature*, 400, pp.1-4.
Kanehisa, M. & Goto, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1), pp.27-30.
Karl, D.M., 2002. Nutrient dynamics in the deep blue sea. *Trends in microbiology*, 10(9), pp.410-8.
Karl, D.M. & 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(2-3), pp.129-156.
Kjellegerg, S., 1993. *Starvation in Bacteria*, New York and London: Plenum Presse.
Kustka, A. et al., 2003. A Revised Estimate of the Iron Use Efficiency of Nitrogen Fixation, with Special Reference to the Marine Cyanobacterium Trichodesmium Spp. (Cyanophyta). *Journal of Phycology*, 25, pp.12-25.
Laglera, L.M. & van den Berg, C.M.G., 2009. Evidence for geochemical control of iron by humic substances in seawater. *Limnology and Oceanography*, 54(2), pp.610-619.
Landry, M.R. et al., 2000. Biological response to iron fertilization in the eastern equatorial Pacific (IronEx II). III . Dynamics of phyto- plankton growth and microzooplankton grazing. , 201, pp.57-72.
Li, W. & Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics (Oxford, England)*, 22(13), pp.1658-9.
Macrellis, H.M. et al., 2001. Collection and detection of natural iron-binding ligands from seawater. *Marine Chemistry*, 76(3), pp.175-187.
Maldonado, M.T., 2005. Acquisition of iron bound to strong organic complexes, with different Fe binding groups and photochemical reactivities, by plankton communities in Fe-limited subantarctic waters. *Global Biogeochemical Cycles*, 19(4).
Maldonado, M.T. & Price, Neil M, 1999. Utilization of iron bound to strong organic ligands by plankton communities in the subarctic Pacific Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 46, pp.2447-2473.

Malmstrom, R.R. et al., 2005. Biomass Production and Assimilation of Dissolved Organic Matter by SAR11 Bacteria in the Northwest Atlantic Ocean. , 71(6), pp.2979-2986.
Malmstrom, R.R. et al., 2004. Contribution of SAR11 Bacteria to Dissolved Dimethylsulfoniopropionate and Amino Acid Uptake in the North Atlantic Ocean. , 70(7), pp.4129-4135.
Mann, E.L. & Chisholm, S.W., 2000. Iron limits the cell division rate of Prochlorococcus in the eastern equatorial Pacific. *Limnology and Oceanography*, 45(5), pp.1067-1076.
Marioni, J.C. et al., 2008. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome research*, 18(9), pp.1509-17.
Martin, J. et al., 1994. Testing the iron hypothesis in ecosystems of the equatorial Pacific Ocean. , pp.123-129.
Martiny, A.C., Coleman, M.L. & Chisholm, S.W., 2006. Phosphate acquisition genes in Prochlorococcus ecotypes: evidence for genome-wide adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, 103(33), pp.12552-7.
McCarren, J. et al., 2010. Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. *Proceedings of the National Academy of Sciences*, 107(38), pp.16420-16427.
McCarter, L. & Silverman, M., 1989. Iron regulation of swarmer cell differentiation of Vibrio parahaemolyticus. *Journal of bacteriology*, 171(2), pp.731-6.
Melkani, G.C. et al., 2008. Divalent cations stabilize GroEL under conditions of oxidative stress. *Biochemical and biophysical research communications*, 368(3), pp.625-30.
Melkani, G.C., Zardeneta, G. & Mendoza, J. a, 2002. GroEL interacts transiently with oxidatively inactivated rhodanese facilitating its reactivation. *Biochemical and biophysical research communications*, 294(4), pp.893-9.
Moore, C.M. et al., 2008. Relative influence of nitrogen and phosphorus availability on phytoplankton physiology and productivity in the oligotrophic sub-tropical North Atlantic Ocean. , 53(1), pp.291-305.
Moore, M. et al., 2009. Large-scale distribution of Atlantic nitrogen fixation controlled by iron availability. *Nature Geoscience*, 2(12), pp.867-871.
Van Mooy, B. a. S. & Devol, A.H., 2008. Assessing nutrient limitation of Prochlorococcus in the North Pacific subtropical gyre by using an RNA capture method. *Limnology and Oceanography*, 53(1), pp.78-88.
Morel, F M M & Price, N M, 2003. The biogeochemical cycles of trace metals in the oceans. *Science (New York, N.Y.)*, 300(5621), pp.944-7.
Mullins, T.D. et al., 1995. Genetic comparisons reveal the same unknown bacterial lineages in Atlantic and Pacific bacterioplankton communities. *Limnology and Oceanography*, 40(1), pp. 148-158.
Noble, A.E., Saito, M.A., Maiti, K., Benitez-Nelson, C. 2008. Cobalt, manganese, and iron near the Hawaiian Islands: A potential concentrating mechanism for cobalt within a cyclonic eddy and implications for the hybrid-type trace metals. *Deep-Sea Research II*, 55, pp. 1473-1490.
Ottesen, E. et al., 2012. Rhythm and synchrony in gene expression of sympatric microbial populations. *submitted.*
Ottesen, E. a et al., 2011. Metatranscriptomic analysis of autonomously collected and preserved marine bacterioplankton. *The ISME journal*, pp.1881-1895.
Pruesse, E. et al., 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic acids research*, 35(21), pp.7188-96.
Rappé, M., Vergin, K. & Giovannoni, S., 2000. Phylogenetic comparisons of a coastal bacterioplankton community with its counterparts in open ocean and freshwater systems. *FEMS microbiology ecology*, 33(3), pp.219-232.
Rappé, M.S., Kemp, P.F. & Giovannoni, S.J., 1997. Phylogenetic diversity of marine coastal picoplankton 16s rRNA genes cloned from the continental shelf off Cape Hatteras , North Carolina. *Limnology and Oceanography*, 42, pp.811-826.
Rodriguez, G., 2007. Lecture Notes on Generalized Linear Models.

Rue, E.L. & Bruland, K.W., 1995. Complexation of iron (III) by natural organic ligands in the Central North Pacific as determined by a new competitive ligand equilibration / adsorptive cathodic stripping voltammetric method. *Marine Chemistry*, 50, pp.117-138.

Rue, E.L. & Bruland, K.W., 1997. The role of organic complexation on ambient iron chemistry in the equatorial Pacific Ocean and the response of a mesoscale iron addition experiment. *Limnology and Oceanography*, 42, pp.901-910.

Rusch, D.B. et al., 2010. Characterization of Prochlorococcus clades from iron-depleted oceanic regions. *PNAS*.

Saito, M. a. & Schneider, D.L., 2006. Examination of precipitation chemistry and improvements in precision using the Mg(OH)2 preconcentration inductively coupled plasma mass spectrometry (ICP-MS) method for high-throughput analysis of open-ocean Fe and Mn in seawater. *Analytica Chimica Acta*, 565(2), pp.222-233.

Shi, Y., Tyson, G.W., et al., 2011. Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean. *The ISME journal*, 5(6), pp.999-1013.

Shi, Y., McCarren, J. & DeLong, E.F., 2011. Transcriptional responses of surface water marine microbial assemblages to deep-sea water amendment. *Environmental Microbiology*, p.no-no.

Shi, Y., Tyson, G.W. & DeLong, E.F., 2009. Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column. *Nature*, 459(7244), pp.266-9.

Smith, D.P. et al., 2010. Transcriptional and translational regulatory responses to iron limitation in the globally distributed marine bacterium Candidatus pelagibacter ubique. *PloS one*, 5(5), p.e10487.

Stewart, F.J., Ottesen, E. a & DeLong, E.F., 2010. Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. *The ISME journal*, 4(7), pp.896-907.

Sunda, W.G. & Huntsman, S.A., 1995. Iron uptake and growth limitation in oceanic and coastal phytoplankton. *Marine Chemistry*, 50, pp.189-206.

Suzuki, M.T. et al., 2001. Phylogenetic analysis of ribosomal RNA operons from uncultivated coastal marine bacterioplankton. *Environmental microbiology*, 3(5), pp.323-31.

Thompson, A.W. et al., 2011. Transcriptome response of high- and low-light-adapted Prochlorococcus strains to changing iron availability. *The ISME Journal*, 9313, pp.1-15.

Tolonen, A.C. et al., 2006. Global gene expression of Prochlorococcus ecotypes in response to changes in nitrogen availability. *Molecular systems biology*, 2, p.53.

Tortell, P. et al., 1999. Marine bacteria and biogeochemical cycling of iron in the oceans. *FEMS Microbiology Ecology*, 29(1), pp.1-11.

Touati, D., 2000. Iron and Oxidative Stress in Bacteria. *Archives of biochemistry and biophysics*, 373(1), pp.1-6.

Vila-Costa, M. et al., 2007. An annual cycle of dimethylsulfoniopropionate-sulfur and leucine assimilating bacterioplankton in the coastal NW Mediterranean. *Environmental microbiology*, 9(10), pp.2451-63.

Weaver, R., Kirchman, D. & Hutchins, D., 2003. Utilization of iron/organic ligand complexes by marine bacterioplankton. *Aquatic Microbial Ecology*, 31, pp.227-239.

Wu, J. & Luther, G.W., 1995. Complexation of Fe(III) by natural organic ligands in the Northwest Atlantic Ocean by a competitive ligand equilibration method and a kinetic approach. *Marine Chemistry*, 50, pp.159-177.

Chapter 5

Conclusion

With unique redox properties, iron is a co-factor involved in a large number of electron-transfer reactions, and hence one of the trace micronutrients necessary for life (Crichton, 2001). Given the chronic scarcity of iron in certain regions of the ocean as well as the variable inputs of this micronutrient in others, the success of marine microorganisms depends in part on their ability to adapt to either chronically low iron concentrations or to changing iron availabilities. In this thesis, I explored how changes in iron availability impact the ecology of marine microbial communities. At first, I took a focused perspective characterizing iron acquisition systems amongst specific populations of heterotrophic bacteria with different micro-niche preferences (Chapter 2). I then expanded the approach to the community and functional levels, characterizing iron-induced changes of open-ocean microbial community structure (Chapter 3) and microbial community gene expression dynamics (Chapter 4).

5.1. Characterizing iron acquisition strategies at the population level

5.1.1. Summary

One of the current uncertainties in the field is the specific mechanism by which different marine microorganisms acquire iron. Initial studies have started to reveal that these may not only depend on the identity of the microorganism but also on the form of iron that may be locally present (Morel et al. 2008). While siderophore-based iron acquisition is widespread amongst pathogens and terrestrial microorganisms, the utility of siderophore-based iron acquisition in the marine environment has often been questioned, and it has been hypothesized that such a strategy

may only be useful under conditions of high-cell densities, as could be the case on marine particulate matter (Hutchins & Rueter, 1991; Volker & Wolf-gladrow, 1999). Using a combination of functional screens and genomics I characterized the different iron acquisition strategies – including the prevalence and diversity of siderophores– utilized by a group of closely related vibrio strains constituting distinct ecological populations (Chapter 2). This represents the first study that relates siderophore production not only to phylogenetic diversity but also to micro-niche preference. This study is also unique in that it quantifies the siderophore production trait within a well-characterized bacterial population– a complementary approach to the majority of studies on marine siderophores to date, which focus on structurally characterizing the siderophore produced or identifying the marine bacteria that produce them (refer to Chapter 1 and references therein). Findings of this study have interesting repercussions particularly for our understanding of iron acquisition in the marine environment at the micro-scale and more broadly for social evolutionary dynamics of marine microbes on particles.

5.1.2. Implications of key findings and future work

My analysis of iron-related transport genes (including siderophore synthesis genes) in 69 vibrio genomes provides information on the importance of different iron forms as potential iron sources in these strains' micro-environment. I argue, for example, that inorganic ferrous iron could be an important source of iron for the vibrio species investigated. Another interesting example stems from the detailed analysis of siderophore types, which revealed that siderophores produced by host-associated vibrios may have much higher affinity constants to iron than those produced by non-host associated vibrios. I suggest that this could be an indication that iron in the marine environment may be less tightly bound to competing organic ligands than inside a host. Because this type of analysis is entirely performed *in silico,* such findings would need to be confirmed experimentally, for example by chemically characterizing the different types of siderophores for host-associated and non host-associated vibrios, or by measuring ferrous iron uptake rates. Measuring expression levels or even (but more experimentally challenging) protein abundances of

specific transporters in a given micro-environment could also help confirm these findings. Nonetheless, this type of analysis proves useful to generate hypotheses regarding the potential iron sources utilized by vibrios in various micro-environments. In order to obtain more information on potential sources of iron to specific microorganisms, this type of approach could be applied to other microorganisms and environments beyond the micro-scale, as has been done recently in a study of potential iron transporters in the Global Ocean Survey (GOS) metagenomic dataset (Toulza et al. 2012). As chemical measurements of iron speciation in the environment improve, this type of approach focused on iron-specific transporters could act as a complement, helping to identify the types of microorganisms that may be well poised to utilize the different forms of iron.

Another key finding of this study is that siderophore production within the group of vibrios investigated is not only widespread but it also appears to be an intermediate frequency trait, with roughly 40% of the ~1,000 strains screened producing a siderophore. This finding demonstrates that siderophore-based iron acquisition is indeed an important strategy for these vibrio isolates obtained in a coastal environment, likely to have high iron concentrations. Another observation is that except for a few extreme cases (ie. host-associated vibrio populations for which nearly 100% of the strains tested produce a siderophore), siderophore production also centers around intermediate values (~30-50%) within populations. This implies that while siderophore production may be a superfluous and costly strategy to acquire iron for some individual strains, siderophore production is a stable trait at the population level. In fact, this study shows that some of the non-producing strains appear to be "cheaters", potentially benefitting from the siderophores produced by their neighbors without incurring the costs of siderophore biosynthesis. This opens many avenues for future research, including testing the prevalence of siderophore production amongst other bacterial genera and phyla. Is siderophore production also a stable trait amongst other known siderophore-producing bacterial genera (ie. Marinobacter, Alteromonas, Pseudoalteromonas, Halomonas)? Or do the known siderophore-producing marine isolates represent just small fractions of larger populations? Finally, moving beyond the population level, it would be interesting to test whether this could be true at the community level as well, where the

burden of siderophore production would be on members of the community that do not necessarily belong to the same bacterial genus, phylum or even kingdom of life than the individual microorganisms benefitting from the siderophore. In fact, a recent study demonstrated the existence of such inter-kingdom cooperation: describing the association between a siderophore-producing Marinobacter and a dinoflagellate phytoplankter that did not produce a siderophore but nonetheless benefitted from the siderophore produced by the Marinobacter (Amin et al., 2009). Another study showed the importance of siderophore-producing bacteria in helping the growth of as of yet uncultivated bacteria (D'Onofrio et al., 2010). Such a dynamic could potentially explain recent genomic and metagenomic findings, where while siderophore uptake genes seem to be widespread in marine bacterial genomes (Hopkinson & Barbeau, 2012) and marine metagenomes (Toulza et al. 2012), siderophore synthesis genes are not.

Finally, this study demonstrates that siderophore production and micro-habitat are correlated: a significant negative correlation is described between siderophore production and increasing sampling size fraction. This finding is rather surprising at first, as modeling studies have suggested that siderophore production would be favored in high cell-density environments, such as on particles (Volker & Wolf-gladrow, 1999). A possible explanation could be that iron is more readily available on particles. Because iron adsorbs readily unto particulate matter (Balistrieri et al. 1981), iron is likely to be more concentrated on particles than in the ambient seawater, potentially precluding the need for siderophore-based iron acquisition. Another possible explanation could be that the decrease in the prevalence of siderophore production on large particles could potentially be explained by a rise in the frequency of "cheaters". One suggestion is that this would be made possible by longer attachment times and presumably higher cell concentrations on large particles. This would lead to higher local concentrations of siderophore, a prime condition for the evolution of "cheating" (Nadell et al. 2009). This finding has broad implications, specifically with respect to our understanding of microbial ecology at the micro-scale. Marine particles have long thought to be hotspots of microbial activity in the ocean (Simon et al. 2002), where metabolic activity is higher than in the surrounding seawater (Grossart et al. 2003). This study (Chapter 1) shows that the rise

of "cheating" is accompanied by a loss of specific siderophore synthesis genes, presumably introducing greater genetic diversity within the population. Whether this is also the case for other "public good" type of molecules remains to be determined. However, if this mechanism is indeed widespread, it could suggest that particles, in addition to being hotspots of microbial activity, could also be hotspots of microbial diversity, and potentially reservoirs of genetic diversity.

5.2. Characterizing iron-induced changes in microbial communities

Having described how iron acquisition strategies may differ at the micro-scale, I proceeded to explore how changes in iron availability shape microbial communities. I performed and analyzed two microcosm experiments during which I followed the response of the whole microbial community using meta-omic approaches by investigating changes in community structure (Chapter 3) and gene expression dynamics (Chapter 4).

5.2.1. Metagenomic analysis of iron-induced changes in the Central Equatorial Pacific

5.2.1.1. Summary

Using a metagenomic approach I characterized the changes in microbial community structure induced by iron addition in the Central Equatorial Pacific Ocean (Chapter 3). This study represents the first to use an experimental metagenomic approach to characterize iron-induced changes in microbial community structure in the Central Equatorial Pacific Ocean. The study revealed that iron enrichment of surface waters led to a ~3.5 fold enrichment of eukaryotic metagenomic reads, corresponding to a phytoplankton bloom dominated by the *Pseudo-nitzschia*-like diatom. I argue that, because it is present at low abundances in the non-enriched sample, the *Pseudo-nitzschia* like population is well adapted to the low iron conditions of the Central Equatorial Pacific Ocean but nonetheless well-poised for opportunistic growth if provided with an iron pulse. This study also showed that iron enrichment selected for a functionally distinct Bacteroidetes

population that –I argue– was more likely responding directly to the diatom bloom rather than to the iron addition.

5.2.1.2. Implications and Future Work

This study has some interesting implications for carbon cycling in the Central Equatorial Pacific Ocean. Indeed, in this region the main input of iron to surface waters is thought to occur via upwelling (Gordon, Coale, & Johnson, 1997). Acknowledging that the sole addition of 1 nM $FeCl_3$ differs from an upwelling event –during which other macronutrients and potentially a different microbial assemblage are brought together with the upwelled waters– this study provides some indication of how the ambient microbial community could be affected by such an upwelling event. The results suggest that it would likely lead to a *Pseudo-nitzschia*-dominated bloom, presumably resulting in increased rates of photosynthesis and carbon fixation. The results also suggest that a fraction of the fixed carbon could, however be immediately respired back by diatom-associated bacteria, and specifically by the Bacteroidetes (this work, Chapter 3). Testing these hypotheses would require follow-up experiments, where rates of primary and bacterial production could be measured. Ideally, such measurements would be species or genus specific- but such experiments are quite challenging (Bard & Ward, 1997). Furthermore, it would be interesting to explore the evolution of the iron-induced diatom bloom and its associated Bacteroidetes community over time. Because only end point samples were obtained, this study provides little insight into these dynamics. Yet, many questions specifically pertaining to the dynamics of the bloom could potentially have important repercussions on our understanding of the system: Is the Bacteroidetes population stimulated soon after the onset of the *Pseudo-nitzschia* bloom? Would the Bacteroidetes population persist long after the termination of the bloom? Does a fraction of the diatom-associated Bacteroidetes community sink together with senescing phytoplankton cells, thus remineralizing carbon at depth?

5.2.2. Metatranscriptomic analysis of iron-induced changes in the North Pacific Subtropical Gyre

As this study (Chapter 3) demonstrates, a metagenomic approach alone, while enabling analysis of changes in community structure (both in terms of composition and gene content), does not provide for a detailed understanding of how specific microorganisms may adapt to changing iron availabilities. In Chapter 4, I investigated how whole microbial communities responded to changing iron availability by exploring possible changes at the gene expression level.

5.2.2.1. Summary

Using an experimental metatranscriptomic approach complemented with flow-cytometric measurements of cell abundances and specific physiological characteristics, I characterized the iron-induced changes of the microbial community in the North Pacific Subtropical Gyre (NPSG) (Chapter 4). This study showed that iron enrichment of surface seawater led to subtle iron-specific changes in cell size, chlorophyll content and relative abundance of a few photosynthesis-related transcripts of the picocyanobacterium *Prochlorococcus,* but only when iron was supplied together with other limiting macronutrients (in the form of poorly characterized dissolved organic matter). On the other hand, neither the abundance nor gene expression patterns of dominant heterotrophs seemed to be affected by the addition of iron, with or without the concurrent addition of other limiting macronutrient, at the level of sensitivity of our assays. These findings imply that in the North Pacific Subtropical Gyre, iron alone played a minimal role (in comparison to other limiting macronutrients) in regulating picoplankton growth and activity.

5.2.2.2. Implications and Future Work

Future work should include a better characterization of the iron-induced response of *Prochlorococcus.* The increase in cell size and chlorophyll content per cell had previously been observed upon relief of iron limitation in the severely iron-limited Equatorial Pacific (Cavender-Bares et al. 1999). This is the first time that this response has been observed in the NPSG. However, it is likely that this response is not just specific to the relief of iron limitation, but that it

could also be a response to the resupply of any limiting nutrient. Indeed, this study showed that the sole addition of organic matter –presumably containing nitrogen and phosphorus sources– resulted in this physiological response. However, it is intriguing that the effect was exacerbated when iron was added together with the organic matter. Because the organic matter added in this experiment was poorly characterized, it would be worthwhile to test how *Prochlorococcus* cell-size and chlorophyll content change *in situ* in response to the addition of defined nitrogen sources, alone and together with iron. If the effect is observed upon addition of these nitrogen-sources, but exacerbated by the addition of iron, it would imply that iron also exerts a control on *Prochlorococcus* physiology (and potentially activity) in the NPSG. Because changes in chlorophyll content most likely translate to changes in photosynthetic activity, it would also be interesting to test how iron addition (alone or together with a defined nitrogen source) affects *Prochlorococcus* primary production rate and efficiency. Finally, determining the molecular markers linked to this effect (changes in size and chlorophyll content) could be useful as a diagnostic tool of *Prochlorococcus* relief of nutrient limitation, especially as meta-transcriptomic studies of marine microbial communities become more widespread.

Because the NPSG is chronically N-limited (Karl et al., 2001) and is characterized by low iron concentrations (Boyle et al. 2005; Noble et al. 2008), iron is likely to act as an important control of the diazotrophic population in the NPSG. Unfortunately, our study design, which –in order to reduce the complexity of sequence data associated with eukaryotic genomes excluded cells greater than 1.6 μm– excluded the majority of the diazotrophs in the NPSG. Future work should specifically focus on the response of the diazotrophic population *in situ*, focusing on identifying the specific diazotrophs stimulated by addition of iron (*Trichodesmium* spp., *Crocosphaera* spp. or the diatom-associated *Richelia* spp.?) and the associated changes in rates of nitrogen fixation and in patterns of expression of genes involved in nitrogen fixation. Laboratory studies with cultured isolates of *Trichodesmium* spp. (Berman-Frank et al. 2001) and *Crocosphaera* spp. (Jacq & Ridame, 2012) have shown that iron limits the diazotrophic activity of these microorganisms.

Understanding how iron may control the growth and activity of diazotrophic populations *in situ* in the NPSG and how these changes may cascade through to the microbial community could have important repercussions on our understanding of the ecosystem. Indeed, it is thought that N_2 fixation contributes more than 50% of total new nitrogen needed to support primary production in the gyre (Dore et al. 2002; Karl, et al. 2002). Yet, while important blooms of nitrogen fixers are observed annually during the summer months, the factors controlling their onset and decline remain unresolved (Church et al. 2009; Sohm et al. 2011).

One of the goals of this study (Chapter 4) was to determine whether iron-binding organic ligands directly obtained from the marine environment rendered iron more available to the microbial community. I hypothesized that one of the possible responses could be a difference in the relative abundance of specific iron transport transcripts. As these iron-binding ligands have been hypothesized to be siderophores (Macrellis et al. 2001; Rue & Bruland, 1997), I was also curious to see if changes in transcript levels of siderophore-specific receptors could be observed. Unfortunately, this study revealed inconclusive in this respect: partly because of the low transcript abundance of iron-specific transporters and partly because of the type of iron-binding organic ligands used. Because the pool of iron-binding ligands utilized this experiment was rich in organic material, changes in community structure and gene expression dynamics were overwhelmed by an organic matter specific response. Very recent progress in analytical techniques are indicating that it is now possible to separate the iron-binding fraction from the rest of the organic matter that is directly obtained from seawater (Boiteau et al., *in prep*). This opens the door for future interesting experiments, to not only test whether these iron-binding ligands do in fact enhance iron availability, but whether the ligand-bound iron is more or less available to specific organisms.

5.3. Final thoughts

This thesis discusses specific examples of how iron affects microorganisms both at the population (Chapter 2) and at the community levels (Chapters 3 and 4). In both cases, iron plays an important role in shaping these two levels of ecological organization. The specific processes by which this is achieved are however quite different. At the population level, competition for iron as a scarce resource is at the center of a "public good" game, which seems to drive the diversification of certain populations (Chapter 2). At the community level, the greater requirement for iron by specific communities (ie. phototroph verses heterotroph) seems to drive the succession of communities that represent different trophic levels (Chapter 3). We suggest that iron distribution and availability does not only have an impact on primary production and its repercussions onto higher trophic levels (Chapter 3), but also on population diversity (at least within different micro-habitats) (Chapter 2 and Chapter 3). How these diversification processes at the population level trickle up to the community level and then impact ecosystem function is still unclear at this stage. But it is interesting to keep those in mind, especially as models of global geochemical cycles start to incorporate microbial processes.

5.4. References

Amin, S. A., Green, D. H., Hart, M. C., Kupper, F. C., Sunda, W. G., & Carrano, C. J. (2009). Photolysis of iron–siderophore chelates promotes bacterial–algal mutualism. *Proceedings of the National Academy of Sciences*, *106*(40), 901-910.

Balistrieri, L., Brewer, P., & Murray, J. (1981). Scavenging residence times of trace metals and surface chemistry of sinking particles in the deep ocean. *Deep-Sea Research*, *28*, 101-121.

Bard, D. ., & Ward, B. . (1997). A species-specific bacterial productivity method using immunomagnetic separation and radiotracer experiments. *Journal of Microbiological Methods*, *28*(3), 207-219. doi:10.1016/S0167-7012(97)00987-1

Berman-Frank, I., Cullen, J. T., Shaked, Y., Sherrell, R. M., & Falkowski, P. G. (2001). Iron availability, cellular iron quotas, and nitrogen fixation in Trichodesmium. *Limnology and Oceanography*, *46*(6), 1249-1260. doi:10.4319/lo.2001.46.6.1249

Boyle, E. a., Bergquist, B. a., Kayser, R. a., & Mahowald, N. (2005). Iron, manganese, and lead at Hawaii Ocean Time-series station ALOHA: Temporal variability and an intermediate water hydrothermal plume. *Geochimica et Cosmochimica Acta*, *69*(4), 933-952. doi:10.1016/j.gca.2004.07.034

Cavender-Bares, K. K., Mann, E. L., Chisholm, S. W., Ondrusek, M. E., & Bidigare, R. R. (1999). Differential response of equatorial Pacific phytoplankton to iron fertilization. *Limnology and Oceanography*, *44*(2), 237-246. doi:10.4319/lo.1999.44.2.0237

Church, M. J., Mahaffey, C., Letelier, R. M., Lukas, R., Zehr, J. P., & Karl, D. M. (2009). Physical forcing of nitrogen fixation and diazotroph community structure in the North Pacific subtropical gyre. *Global Biogeochemical Cycles*, *23*(2). doi:10.1029/2008GB003418

Crichton, R. (2001). The importance of iron for biological systems. In R. Crichton (Ed.), *Inorganic Biochemistry of Iron Metabolism: From Molecular Mechanisms to Clinical Consequences* (2nd ed., pp. 17-48). Chichester, NY: John Wiley and Sons Ltd.

Dore, J. E., Brum, J. R., Tupas, L. M., & Karl, D. M. (2002). Seasonal and interannual variability in sources of nitrogen supporting export in the oligotrophic subtropical North Pacific Ocean. *Limnology and Oceanography*, *47*(6), 1595-1607.

D'Onofrio, A., Crawford, J. M., Stewart, E. J., Witt, K., Gavrish, E., Epstein, S., Clardy, J., et al. (2010). Siderophores from neighboring organisms promote the growth of uncultured bacteria. *Chemistry & biology*, *17*(3), 254-64. Elsevier Ltd. doi:10.1016/j.chembiol.2010.02.010

Gordon, R. M., Coale, K. H., & Johnson, K. S. (1997). Iron distributions in the equatorial Pacific: Implications for new production. *Limnology and Oceanography*, *42*(3), 419-431. doi:10.4319/lo.1997.42.3.0419

Grossart, H.-peter, Kiørboe, T., Tang, K., & Ploug, H. (2003). Bacterial Colonization of Particles : Growth and Interactions Bacterial Colonization of Particles : Growth and Interactions. *Applied and environmental microbiology*, *69*(6), 3500-3509. doi:10.1128/AEM.69.6.3500

Hopkinson, B. M., & Barbeau, K. a. (2012). Iron transporters in marine prokaryotic genomes and metagenomes. *Environmental Microbiology*, *14*, 114-128. doi:10.1111/j.1462-2920.2011.02539.x

Hutchins, D. A., & Rueter, J. G. (1991). Siderophore production and nitrogen fixation are mutually exclusive strategies in Anabaena 7120. *Limnology and Oceanography*, *36*, 1-12.

Jacq, V., & Ridame, C. (2012). The influence of iron limitation on the growth and activity of Crocosphaera watsonii , an unicellular diazotrophic cyanobacterium, *14*, 12080.

Karl, D. M., Bjorkman, K. M., Dore, J. E., Fujieki, L., Hebel, D. V., Houlihan, T., Letelier, R. M., 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., Michaels, A., Bergman, B., & Capone, D. (2002). Dinitrogen fixation in the world ' s oceans. *Earth Science*, 47-98.

Macrellis, H. M., Trick, C. G., Rue, E. L., Smith, G., & Bruland, K. W. (2001). Collection and detection of natural iron-binding ligands from seawater. *Marine Chemistry*, *76*(3), 175-187. doi: 10.1016/S0304-4203(01)00061-5

Morel, Francois, Kustka, A, Shaked, Y. (2008). The role of unchelated Fe in the iron nutrition of phytoplankton. *Limnology and Oceanography*, *53*(1), 400-404.

Nadell, C. D., Xavier, J. B., & Foster, K. R. (2009). The sociobiology of biofilms. *FEMS microbiology reviews*, *33*(1), 206-24. doi:10.1111/j.1574-6976.2008.00150.x

Noble, A. E., Saito, M. a., Maiti, K., & Benitez-Nelson, C. R. (2008). Cobalt, manganese, and iron near the Hawaiian Islands: A potential concentrating mechanism for cobalt within a cyclonic eddy and implications for the hybrid-type trace metals. *Deep Sea Research Part II: Topical Studies in Oceanography*, *55*(10-13), 1473-1490. doi:10.1016/j.dsr2.2008.02.010

Rue, E. L., & Bruland, K. W. (1997). The role of organic complexation on ambient iron chemistry in the equatorial Pacific Ocean and the response of a mesoscale iron addition experiment. *Limnology and Oceanography*, *42*, 901-910.

Simon, M., Grossart, H., Schweitzer, B., & Ploug, H. (2002). Microbial ecology of organic aggregates in aquatic ecosystems. *Aquatic Microbial Ecology*, *28*, 175-211. doi:10.3354/ame028175

Sohm, J. a., Subramaniam, A., Gunderson, T. E., Carpenter, E. J., & Capone, D. G. (2011). Nitrogen fixation by Trichodesmium spp. and unicellular diazotrophs in the North Pacific Subtropical Gyre. *Journal of Geophysical Research*, *116*(G3), G03002. doi: 10.1029/2010JG001513

Toulza, E., Tagliabue, A., Blain, S., & Piganeau, G. (2012). Analysis of the Global Ocean Sampling (GOS) Project for Trends in Iron Uptake by Surface Ocean Microbes. (F. Rodriguez-Valera, Ed.)*PLoS ONE*, *7*(2), e30931. doi:10.1371/journal.pone.0030931

Volker, C., & Wolf-gladrow, D. A. (1999). Physical limits on iron uptake mediated by siderophores or surface reductases. *Marine Chemistry*, *65*, 227-244.

Appendix A

Functional Screening of Environmental Fosmid Libraries for Siderophore Production

A.1. Background

Iron is necessary for nearly all microbial life. In order to acquire iron in environments where iron is otherwise scarce, microorganisms are known to produce small molecules called siderophores that strongly chelate ferric iron (Andrews, 2003). In the marine environment, the role of siderophores as a widespread and worthwhile iron acquisition mechanism is debated. Indeed, it has also been argued that siderophores only present an evolutionary advantage in environments where cell-densities are high enough to counter siderophore loss by diffusion (Volker and Wolf-gladrow 1999). Consistent with this idea, the most numerically abundant free-living bacterial species (*Prochlorococcus* and *Pelagibacter*) both lack siderophore biosynthesis genes (Rocap et al. 2003; Giovannoni et al. 2005). However, another likely reason for the lack of siderophore synthesis genes in the *Prochlorococcus* and the *Pelagibacter* genomes is the possibility that these would be too distant from currently known siderophore biosynthetic genes, precluding their discovery via bioinformatic similarity searches alone.

Fosmid libraries of environmental genomic DNA have proved to be extremely useful in accessing and probing the genetic material of uncultivated microorganisms (Martinez et al. 2007, Rondon et al. 2000, Schleper et al. 1997). The screening of such libraries has been instrumental in the discovery of the light-harvesting membrane protein proteorhodopsin (Beja et al. 2000). Recently, the screening of a fosmid library prepared from marine picoplankton genomic DNA in the heterologous host *E.coli*, deficient in phosphonate metabolism has revealed a novel

phosphonate degradation pathway (Martinez et al. 2010). The main advantage in this type of approach is the potential for discovering novel genes without necessary *a priori* knowledge of their sequence. Furthermore, unlike genes identified via bioinformatic similarity searches, functional screening of fosmid libraries enables the discovery of genes whose function can immediately be ascertained. Here, I screened fosmid libraries generated from environmental genomic DNA obtained from different depths of the Hawai'i Ocean Time Series (HOT) station ALOHA (22°45' N 158°W). Unfortunately, the positive hits that were identified in this screen were un-informative– they all turned out to be *E.coli entF* (the very same gene that was knocked out of the heterologous host). I present the approach and methods used, as these could be useful should someone want to repeat these experiments in the future. I also discuss possible reasons that could explain why the screen may have have failed to detect siderophore genes in these environmental fosmid libraries.

A.2. Methods

A.2.1. Fosmid Library Generation

Fosmid libraries were generated for use in previous experiments (DeLong et al. 2006). They were constructed from genomic DNA samples (0.2 μm to 1.6 μm size fraction) obtained at various depths from the water column at the Hawai'i Ocean Time Steries (HOT) station ALOHA (22°45' N 158°W) and cloned into the copy-control pCC1FOS fosmid vector (Epicenter). Fosmid libraries were pooled (according to depth) and fosmid DNA was isolated by alkaline lysis and cesium chloride ultracentrifugation (Sambrook et al. 1989). We used aliquots of each pooled fosmid library to transform into the heterologous *E.coli* host.

A.2.2. Heterologous Host: Escherichia coli BW25113 entF⁻ trfA.

The Keio collection (Baba et al. 2006) is a collection of *E.coli* BW25113 for which an in-frame knockout mutant of every non-essential gene is available. For our functional screen, we utilized the *E.coli* BW25113 *entF:Kan* mutant, because the *entF* gene in *E.coli* codes for 2,3-dihydro-2,3-dihydroxybenzoate synthetase, a key enzyme in the biosynthesis of the enterobactin siderophore (Crosa, Walsh 2002). The pCC1FOS fosmid vector was designed to allow for copy-up

control of the fosmid insert via the activation of the *ori*V high-copy origin of replication (Epicentre). However, the *ori*V is activated by the protein product of the *trfA* gene, which is under pBAD (arabinose promoter) control in the Transformax™ EPI300™ *E.coli* electrocompetent cells that are usually commercialized along with the the pCC1FOS fosmid vector (Epicentre). To utilize the copy-up functionality of the pCC1FOS fosmid vector, we had to construct a new *E.coli* BW25113 *entF:Kan* strain with the *trfA* gene under pBAD promoter control as described previously (Martinez et al. 2010).

Briefly, we transduced the pBAD-*trfA* construct from Replicator v2.0 *E.coli* strain (same pBAD-*trfA* construct than in the EPI300™ cells) (Lucigen) (donor) to the *E.coli* BW215113 *entF* *Kan*^{R} strain (recipient) using P1 phage transduction (Miller 1992). Briefly, a phage lysate of the donor strain (Replicator v2.0) was prepared by incubating an overnight culture of the donor strain in LB broth with 5mM $CaCl_2$, 0.2% (wt/vol) glucose and 50 µg/mL ampicillin at 37°C with shaking. After 1 hour, phage P1 *vir* was added (1:100), and the culture was incubated for another 2 hours at 37°C with vigorous shaking until lysis occurred. The lysate was transferred to a 2 mL microcentrifuge tube, to which two drops of chloroform ($CHCl_3$) were added. The tube was inverted rapidly for 30 seconds, and spun for 10-15 min at max speed in a microcentrifuge to remove cell debris. The supernatant was retrieved and stored with 1 drop of chloroform at 4°C.

An overnight culture of the recipient cells (*E.coli* BW25113 *entF:Kan*) was resuspended 1:1 in MC buffer (solution of 100mM $MgSO_4^{2-}$ and 5mM $CaCl_2$) and incubated at 37°C with slight shaking. After 15 min of incubation, the culture was mixed with the P1 lysate of the donor cell at different ratios (recipient : donor lysate = 1:1, 1:2, 10:1, 100:1) and incubated for 30 min at 37°C. To prevent readsorption of phage particles 0.2 mL of 1 M NaCit (sodium citrate) was added to each recipient-donor lysate mixture before plating on LB agar plates with 25 µg/mL ampicillin and 50 µg/mL kanamycin and incubated overnight at 37°C. Colonies that grew on LB amp^{50} kan^{25} were recipient cells for which the phage transduction had been successful and are hereafter referred to as *E.coli* BW25113 *entF trfA*. *E.coli* BW25113 *entF trfA* were made electrocompetent by being washed

thoroughly of any salts (Sambrook et al. 1989) and stored at -80°C until needed for electroporation.

A.2.3. Fosmid Library Screening for Siderophore Production

Aliquots of the pooled fosmid libraries (1 μL) were transformed into 50 μL of electrocompetent *E.coli* BW25113 *entF trfA* cells by electroporation (2.5 kV cm^{-1}, 200 Ω, 25 μF in 2 mm cuvettes). Electroporated cells were resuspended in 1 mL of sterile SOC medium and incubated at 30°C to allow for expression of antibiotic resistance. Cells were then spread on large CAS agar plates (see below for preparation). Plates were incubated at 30°C and checked regularly (over the course of ~2 weeks) for the appearance of colonies surrounded by an orange halo (indication that they produce a siderophore). Results of this functional screen are summarized in Table 1.

Based on Schwyn & Neilands, 1987, CAS agar plates were prepared as follows:

CAS agar plates were prepared by mixing the CAS-Fe-HDTMA dye with a modified M9 growth medium. For 1 L of CAS-agar, 100 mL of CAS-Fe-HDTMA dye was mixed with 900mL of freshly prepared growth media. The CAS-Fe-HDTMA dye was prepared in advance as follows, for 1 L: 10 mL of a 10 mM ferric chloride (FeCl3) in 100 mM hydrochloric acid (HCl) solution was mixed with 590 mL of a 1 mM aqueous solution of chrome azurol S (CAS). The Fe-CAS solution was then added to 400 mL of a 2 mM aqueous solution of hexadecyl-trimethyl-ammonium bromide (HDTMA). The resulting CAS-Fe-HDTMA solution was autoclaved for 25 min in a polycarbonate bottle that had previously been soaked overnight in 10% HCl then rinsed 5 times with MilliQ water. The CAS-Fe-HDTMA dye was stored at room temperature covered from light until usage.

The modified M9 growth media was prepared as follows, for 1 L of CAS-agar: 30.24 g of 1,4-piperazine-diethanesulfonic acid (PIPES), together with 1 g of ammonium chloride (NH_4Cl), 3 g potassium phosphate (KH_2PO_4), 0.5 g of sodium chloride (NaCl), were dissolved into MilliQ

water by adjusting the pH to 6.8 with 10 M NaOH. Note that the commonly used phosphate buffer disodium phosphate ($Na_2HPO_4 \bullet 7H_2O$) was omitted as phosphate can chelate iron and lead to a discoloration of the CAS dye. As a solidifying agent, 9 g of agar noble (Difco) were added to the solution. We found that the more commonly used solidifying agents, agarose and agar also led to a discoloration of the CAS dye, likely due to higher phosphate content. The volume was adjusted to 860 mL and the solution was autoclaved. After cooling, 30 mL of a sterile 10% (w/v) Casamino acids (Difco) aqueous solution and 8 mL of a sterile 50% (w/v) glycerol aqueous solution were added. Appropriate antibiotics were added to final concentrations of: 12.5 μg/mL chloramphenicol (Chl^R is an antibiotic selection marker on the pCC1FOS vector), 25 μg/mL kanamycin (Kan^R is the antibiotic selection marker for the *E.coli EntF* mutant from the Keio collection), and 50 μg/mL ampicillin (Amp^R is the antibiotic marker for the araC-pBAD-*trfA* construct). In the case of arabinose induction of the copy-up function of the pCC1FOS fosmid, an aqueous solution of L-arabinose was added to the media to a final concentration of 0.01% (w/v). Finally, the 100 mL of CAS-Fe-HDTMA was added to the growth media, which was poured into large sterile petri dishes.

The final concentrations of the CAS-agar components are: 100 mM PIPES, 18 mM NH_4Cl, 22 mM KH_2PO_4, 0.05% (w/v) NaCl, 0.3% casamino acids, 0.4% glycerol, 10 μM $FeCl_3$, 58 μM CAS, 80 μM HDTMA. Unless otherwise noted, all chemicals were obtained from Sigma-Aldrich (St. Louis, MO, USA). Note that we use glycerol instead of glucose because the araC-pBAD promoter (needed for the arabinose-induction of the copy-up functionality of the pCC1FOS fosmid) is inhibited by glucose.

A.2.4. Identification of Siderophore Producing Fosmid Clones.

Colonies that were surrounded by an orange halo were picked and grown overnight in LB with 25 μg/mL kanamycin, 50 μg/mL ampicillin and 12.5 μg/mL chloramphenicol. The fosmid was extracted and purified (Sambrook et al. 1989) and retransformed into a fresh *E.coli* BW25113 *entF trfA* (1 μL of fosmid for 50 μL of electrocompetent cells, 2.5 kV cm^{-1}, 200 Ω, 25 μF in 2 mm

cuvettes). After confirmation that the fosmid clone did indeed complement for the loss of siderophore production by restreaking on CAS agar, all fosmid clones obtained from siderophore-producing colonies were compared to each other by restriction digest with the PstI restriction enzyme.

Transposon Sequencing:

Only unique fosmids (as identified by the restriction digest) were further characterized. To identify the specific genes (within the 40 kb fosmid insert) responsible for rescuing the loss of siderophore production phenotype, transposon mutagenesis was used. Following the manufacturer's recommendations, fosmid clones were subjected to random *in vitro* transposon mutagenesis using the EZ-Tn5 kan-2 insertion kit (Epicentre). The transposition reaction was transformed into EPI300 cells, and clones containing fosmids were selected on LB plates with 25 μg/mL kanamycin and 12.5 μg/mL chloramphenicol. Transposon mutants were pooled. Fosmid DNA was extracted and retransformed into *E.coli* BW25113 *entF⁻ trfA*. Clones were selected on LB with 25 μg/mL kanamycin and 12.5 μg/mL chloramphenicol and then replica plated onto CAS agar containing 25 μg/mL kanamycin and 12.5 μg/mL chloramphenicol.

Clones for which the transposon inserted within the gene originally responsible for rescuing the siderophore production phenotype in *E.coli* BW25113 *entF⁻ trfA* were identified as they no longer produced an orange halo when plated on and CAS agar. These were picked and sequenced in house on an ABI Prism 3700 DNA Analyzer (Applied Biosystems, Foster City, CA, USA) using the KAN-2 FP-1 (5' - ACCTACAACAAAGCTCTCATCAACC - 3') and KAN-2 RP-1 (5' - GCAATGTAACATCAGAGATTTTGAG - 3') sequencing primers, and the BigDye version 3.1 cycle sequencing kit. DNA sequence was curated by using Sequencher version 4.5 (Gene Codes Corporation, Ann Arbor, MI, USA) and analyzed using BLAST (Altschul et al. 1990).

A.3. Results

Production of heterologously produced siderophore was determined by using the well-established CAS assay (Schwyn and Neilands 1987). We transformed the fosmid libraries into an *E.coli* heterologous host deficient in siderophore production (Figure 1). We expect that fosmid clones that possess environmental genomic DNA containing siderophore synthesis genes could complement the loss of siderophore production in the heterologous host utilized. Colonies that produce a siderophore can easily be detected visually because they are surrounded by an orange halo. Because "hits" can visually be identified, we are able to screen through fosmid clones in a high-throughput manner.

Having established that the heterologous host utilized is robust (complete loss of siderophore production as observed in Figures 1 and 2), we screened fosmid libraries generated from environmental genomic DNA obtained at different depth at HOT station ALOHA. A thorough screening of libraries representing an entire depth-profile at HOT station ALOHA (10m, 70m, 130m, 200m, 500m, 770m, and 4000m) was performed (Table 1). We found that a number of fosmid clones rescue the loss of siderophore production phenotype of the heterologous *E.coli entF- trfA* (see Figure 3). Table 1 summarizes the results from this high-throughput screening of fosmid libraries. We then proceeded to identifying the single genes responsible for the rescue. It is possible that these genes are homologs of *entF*, potentially from different organisms. Alternatively these genes could be part of an entirely different siderophore pathway (which would have to be entirely contained in the fosmid insert).

Table A.1. Summary of Results

Fosmid Library	Media	Number of Clones Screened	Number of Clones in Library	Fold coverage	# of 'hits'
HF 70	CAS, 4 plates	63,000	13,896	~5X	0
	CAS + Arabinose, 5 plates	66,325		~5X	0
HF 130	CAS	70,000	12,738	5.5X	11
HF 200	CAS, 5 plates	58,625	18,914	3X	0
	CAS + arabinose, 5 plates	57,750		3X	0
HF 500	CAS, 5 plates	25,500	15,440	1.6X	0
	CAS + arabinose, 4 plates	29,800	15,440	~2X	0
HF 770	CAS, 5 plates	154,000	17,370	8.8X	19
	CAS + arabinose, 5 plates	140,000		8X	9
HF 4000	CAS, 8 plates	168,000	18,258	9X	24
	CAS + arabinose, 10 plates	20,600		1X	6

While our high-throughput screen resulted in 69 "hits" (Table 1), further characterization of the fosmid inserts revealed that those were all the same (Figure 4). In Figure 4 we only represent 9 of the positive clones, however other fosmid clones that displayed different restriction digest patterns (from the HF 4000 fosmid library for example- data not shown), were nevertheless identical as subsequent sequencing of fosmid inserts revealed.

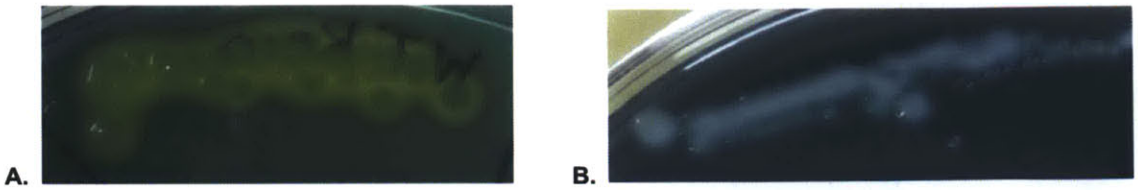


Figure A.1. Behavior of Heterologous *E.coli* host on CAS agar. Panel A. Wild Type *E.coli* BW25113 used to construct the Keio collection clearly produces a siderophore. Panel B. *E.coli* BW25113 *entF-trfA* in which the *entF* gene has been knocked out has lost the ability to produce a siderophore.

Figure A.2. Negative Control. We transformed the *E.coli* BW25113 *entF trfA* with an empty pCC1FOS vector to test as a negative control. When grown on CAS agar, no halo is observed.

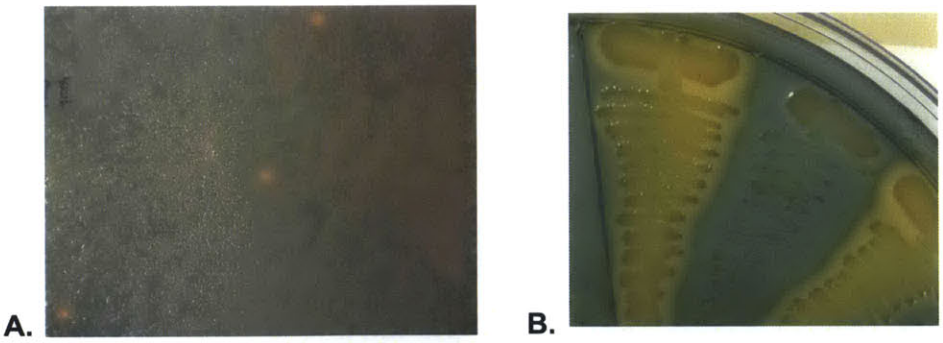


Figure A.3. Functional Screen of Fosmid Libraries. A. Example of a screen where amongst of thousands of transformants 3 clones are surrounded by an orange halo, meaning that the siderophore production has been restored thanks to genes present in the fosmid insert. B. Positive hits, retransformed into fresh heterologous background restreaked next to the negative control, *E.coli* BW25113 *entF trfA* containing the empty pCC1FOS vector.

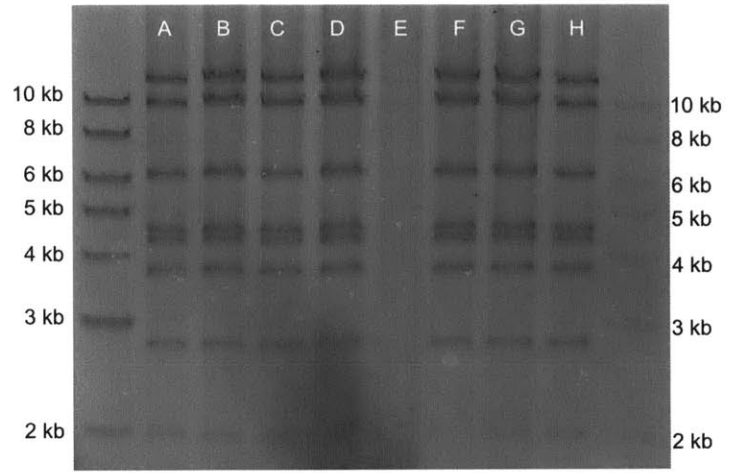


Figure A.4. Restriction Digest Profiles of Fosmid Clones that Complemented the Heterologous Host for Siderophore Production. We show the restriction digest pattern of fosmids obtained from clones (HF770 fosmid library transformed into *E.coli entF⁻ trfA*) that complemented the loss of siderophore production in *E.coli entF⁻ trfA*. The digest pattern is visualized 0.8% agarose gel with SYBR safe stain. A-H represent the restriction patterns of fosmids obtained from 8 distinct "positive hits". The restriction digest patterns show that 7 of the 8 fosmid clones displayed actually have the same identical insert.

Because fosmid inserts are large, ~40 kb, the single genes responsible for rescuing the siderophore production phenotype first need to be identified by transposon mutagenesis. After successfully identifying the transposon mutants that lead to a loss of the rescued phenotype, they were sequenced and their sequence was compared to the NCBI database of non-redundant protein sequences by BLAST. We found that all identified transposon mutants were identical to *entF* encoded by *E.coli* with an E-value of 0 and percent identity of 99-100%. This result is quite surprising as *E.coli* is not a very abundant microorganism in the marine environment, especially not at 4,000 m depth. This suggests that the 'positive' hits identified in our screen could in fact be potential contaminating *E.coli*.

A.4. Conclusion and Future Work

This work shows that screening large-insert fosmid libraries for siderophore synthesis genes is technically feasible. The results we obtained are not those that we were hoping for, which could be explained by a variety of reasons.

The fact that no siderophore synthesis gene was found in the cloned environmental genomic DNA could be due to some inherent limitations in our approach. One of the main limitations of our approach is at the same time its strength: heterologous expression. It is possible that *E.coli,* as a heterologous host, may not be able to express genes that may be too divergent with presumably different codon usage bias requiring different translation machinery (Adzhubei et al. 1996; Kurland and Gallant 1996). Another possibility may be the inability for *E.coli* to export the gene product. This could be envisioned if the genes required for siderophore synthesis were able to complement *entF* or could encode for an entirely different pathway, but that the genes required for siderophore export were either absent from the fosmid insert or just too divergent for *E.coli* to express. We suggest repeating this experiment in a *Pseudomonas*, which are known to recognize a wide variety of exogenously produced siderophores

Another limitation of this approach is the fact that while fosmids allow for large inserts (~40 kb) to be cloned, some siderophore biosynthesis and transport clusters are quite large, and may be truncated when cloned into the fosmid. Indeed, a number of siderophores are synthesized via the non-ribosomal synthetase (NRPS) pathway (Crosa and Walsh 2002). These modular assembly lines represent large gene clusters, as in the case of *E.coli* where enterobactin synthesis and transport genes form a ~22 kb cluster (Crosa and Walsh 2002). In the case of the cyanobacterium *Anabaena* sp. strain PCC 7120, a 76kb gene cluster is involved in siderophore biosynthesis (Jeanjean et al. 2008). In recent years, the NRPS-independent synthesis (NIS) pathway has appeared as another important pathway for siderophore synthesis (Challis 2005). Unlike the NRPS pathway, the NIS pathway relies on a few genes (ie. 4 synthesis and 3 transport genes for

aerobactin) (Lorenzo et al. 1986). In this light, it is surprising that none of these genes have been identified as potential 'hits' in our fosmid screen.

Finally, our results may indeed show that siderophore synthesis genes in the marine environment are scarce. This has been observed in recent metagenomic surveys (Toulza et al. 2012). It is often thought that because of the dilute nature of the marine environment, production and excretion of siderophores make little evolutionary sense for marine microorganisms. We suggest that siderophores may be more abundant on marine particles, where cell densities are higher (Simon et al. 2002). An interesting follow-up to this work would be to screen fosmid libraries generated from environmental genomic DNA obtained from particle-associated marine picoplankton for siderophore biosynthesis. It would also be interesting to take a similar approach and screen fosmid libraries generated from marine picoplankton genomic DNA obtained in other oceanic regions such as in the Equatorial Pacific Ocean or the Subarctic North Pacific, which are regions that are thought to be more iron limited than HOT Station ALOHA.

A.5. Acknowledgements

Many thanks to Asunción Martinez who guided me through the molecular biology. Thanks to Tsultrim Palden for setting up and running the sequencing reactions. Finally, thanks to Jay McCarren who helped me with the CAS assay towards the beginning of this project.

A.6. References

Adzhubei, a a, I a Adzhubei, I a Krasheninnikov, and S Neidle. 1996. "Non-random Usage of 'Degenerate' Codons Is Related to Protein Three-dimensional Structure." *FEBS Letters* 399 (1-2) (December 9): 78-82.

Altschul, Stephen F, Warren Gish, Webb Miller, Eugene Myers, and David Lipman. 1990. "Basic Local Alignment Search Tool." *Journal of Molecular Biology* 215: 403-410.

Andrews, Simon C, Andrea K Robinson, and Francisco RodrÃguez-QuiÃ±ones. 2003. "Bacterial Iron Homeostasis." *FEMS Microbiology Reviews* 27 (2-3) (June): 215-237. doi:10.1016/S0168-6445(03)00055-X.

Baba, Tomoya, Takeshi Ara, Miki Hasegawa, Yuki Takai, Yoshiko Okumura, Miki Baba, Kirill a Datsenko, Masaru Tomita, Barry L Wanner, and Hirotada Mori. 2006. "Construction of Escherichia Coli K-12 In-frame, Single-gene Knockout Mutants: The Keio Collection." *Molecular Systems Biology* 2 (January): 2006.0008. doi:10.1038/msb4100050.

Beja, O., L. Aravind, Eugene Koonin, Marcelino Suzuki, Andrew Hadd, Linh Nguyen, Stevan Jovanovich, et al. 2000. "Bacterial Rhodopsin: Evidence for a New Type of Phototrophy in the Sea." *Science* 289 (5486) (September 15): 1902-1906. doi:10.1126/science.289.5486.1902.

Challis, Gregory L. 2005. "A Widely Distributed Bacterial Pathway for Siderophore Biosynthesis Independent of Nonribosomal Peptide Synthetases." *Chembiochem : a European Journal of Chemical Biology* 6 (4) (April): 601-11. doi:10.1002/cbic.200400283.

Crosa, Jorge H, and Christopher T Walsh. 2002. "Genetics and Assembly Line Enzymology of Siderophore Biosynthesis in Bacteria." *Microbiology and Molecular Biology Reviews* 66 (2): 223-249. doi:10.1128/MMBR.66.2.223.

DeLong, Edward F, Christina M Preston, Tracy Mincer, Virginia Rich, Steven J Hallam, Niels-Ulrik Frigaard, Asuncion Martinez, et al. 2006. "Community Genomics Among Stratified Microbial Assemblages in the Ocean's Interior." *Science (New York, N.Y.)* 311 (5760) (January 27): 496-503. doi:10.1126/science.1120250.

Giovannoni, Stephen J, H James Tripp, Scott Givan, Mircea Podar, Kevin L Vergin, Damon Baptista, Lisa Bibbs, et al. 2005. "Genome Streamlining in a Cosmopolitan Oceanic Bacterium." *Science (New York, N.Y.)* 309 (5738) (August 19): 1242-5. doi:10.1126/science.1114057.

Jeanjean, Robert, Emmanuel Talla, Amel Latifi, Michel Havaux, Annick Janicki, and Cheng-Cai Zhang. 2008. "A Large Gene Cluster Encoding Peptide Synthetases and Polyketide Synthases Is Involved in Production of Siderohores and Oxidative Stress Response in the Cyanobacterium Anabaena Sp. Strain PCC 7120." *Environmental Microbiology* 10 (10): 2574-2585.

Kurland, C, and J Gallant. 1996. "Errors of Heterologous Protein Expression." *Current Opinion in Biotechnology* 7 (5) (October): 489-93.

Lorenzo, Victor D E, Albrecht Bindereif, Barry H Paw, and J B Neilands. 1986. "Aerobactin Biosynthesis and Transport Genes of Plasmid Co1V-K30 in Escherichia Coli K-12." *Journal of Bacteriology* 165 (2): 570-578.

Martinez, Asuncion, Gene W Tyson, and Edward F Delong. 2010. "Widespread Known and Novel Phosphonate Utilization Pathways in Marine Bacteria Revealed by Functional Screening and Metagenomic Analyses." *Environmental Microbiology* 12 (1) (January): 222-38. doi:10.1111/j.1462-2920.2009.02062.x.

Miller, Jeffrey. 1992. *A Short Course in Bacterial Genetics.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Rocap, Gabrielle, Frank W Larimer, Jane Lamerdin, Stephanie Malfatti, Patrick Chain, Nathan a Ahlgren, Andrae Arellano, et al. 2003. "Genome Divergence in Two Prochlorococcus Ecotypes Reflects Oceanic Niche Differentiation." *Nature* 424 (6952) (August 28): 1042-7. doi:10.1038/nature01947.

Sambrook, J, EF Fritsch, and T Maniatis. 1989. *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Schwyn, Bernard, and JB Neilands. 1987. "Universal Chemical Assay for the Detection and Determination of Siderophores." *Analytical Biochemistry* 160: 47-56.

Simon, M, Hp Grossart, B Schweitzer, and H Ploug. 2002. "Microbial Ecology of Organic Aggregates in Aquatic Ecosystems." *Aquatic Microbial Ecology* 28: 175-211. doi:10.3354/ame028175.

Toulza, Eve, Alessandro Tagliabue, Stéphane Blain, and Gwenael Piganeau. 2012. "Analysis of the Global Ocean Sampling (GOS) Project for Trends in Iron Uptake by Surface Ocean Microbes." Ed. Francisco Rodriguez-Valera. *PLoS ONE* 7 (2) (February 17): e30931. doi: 10.1371/journal.pone.0030931.

Volker, Christoph, and Dieter A Wolf-gladrow. 1999. "Physical Limits on Iron Uptake Mediated by Siderophores or Surface Reductases." *Marine Chemistry* 65: 227-244.

Appendix B

Chrome Azurol S Assay.

This protocol was used to screen for siderophore production (See Chapter 2 and Appendix A).

1. CAS-Fe-HDTMA dye

When preparing this I use all plastic containers (Nalgene bottles that can be autoclaved work well) that I first soak in 10% HCl for at least 24hours.

Reagent Name	Sigma Product #	Molecular Weight ($g.mol^{-1}$)
Chrome Azurol S (CAS)	C1018	605.28
Hexadecyltrimethylammonium bromide (HDTMA)	H6269	364.45
Ferric Chloride ($FeCl_3$)	157740	162.20
Concentrated Hydrochloric acid (HCl)	N/A	N/A

a. make Fe solution (10 mM in 100 mM HCl)– keep as stock at room temperature.
   i. add 368 µL of concentrated HCl to 45 mL of dH_2O
   ii. dissolve 0.0729 g of ferric chloride ($FeCl_3$) using the HCl solution.

b. make HDTMA solution– use all right away to make 1 L stock solution of Fe-CAS-HDTMA dye.
   i. dissolve 0.724 g of HDTMA in 400mL of dH_2O.
   ii. mix well until all reagent is in solution

c. make Fe-CAS solution
   i. dissolve 0.6 g of CAS in 590 mL of water
   ii. add 10 mL of 10 mM $FeCl_3$ in 100 mM HCl

d. slowly add the 500 mL of Fe-CAS solution to the 400 mL of HDTMA solution. At this point the dark red CAS dye should turn blue.

e. Autoclave for 25min.

This stock solution of CAS-Fe-HDTMA can be kept for almost 6 months if kept at room temperature, protected from the light.

100 mL of CAS-Fe-HDTMA dye is used to make up 1 L of growth media. The final concentration of reagents in the growth media are as follows: 200 µM HDTMA, 100 µM CAS, 10 nM Fe.

2. M9 modified growth medium

The growth media I use for *E.coli* and *Vibrio* is a modified M9, where the phosphate buffer is removed and replaced by PIPES (could try with MOPS- but I have never tried it). Phosphate chelates iron, so it is important that the phosphate is kept at a minimum.

Reagent Name	Sigma Product #	Molecular Weight ($g.mol^{-1}$)
Potassium Phosphate Monobasic (KH_2PO_4)	P5379	136.09
Ammonium Chloride (NH_4Cl)	A4514	53.49
1,4-piperazine-diethanesulfonic acid (PIPES)	P6757	302.37

a. Mix the following reagents in dH_2O (~850mL)
   i. 30.24 g PIPES
   ii. 1 g ammonium chloride (NH_4Cl)
   iii. 3 g potassium phosphate (KH_2PO_4)
   iv. 20 g of sodium chloride (NaCl) (high salt concentration because of *Vibrio*– bring down to 0.5g NaCl if preparing for *E.coli*).
   v. Adjust pH to 6.8 with 10N NaOH.
   vi. Add 9 g of agar noble (agarose may not be pure enough and may contain compounds that could chelate iron and mess up the assay)
   vii. Adjust volume to 940 mL, autoclave ~25min

b. Wait for media to cool down and add:
   i. 30 mL of a 10% (w/v) filter-sterilized aqueous solution of Casamino acids (Difco)
   ii. 10 mL of a 20% (w/v) filter-sterilized aqueous solution of glucose (for vibrio work)/ or 8 mL of sterile 50 % (wt/wt) glycerol solution (for arabinose induction in *E.coli* work).
   iii. For work presented in Appendix A, when needed, L-arabinose was added to a final concentration of 0.01 % (wt/vol).
   iv. 100 mL of the Fe-CAS-HDTMA
   v. add antibiotics at this stage if needed.
   vi. Pour into sterile petri dishes.