METABOLITES OF BACTERIAL-PHYTOPLANKTON INTERACTIONS IN THE SURFACE OCEAN

by

COURTNEY M. THOMAS

(Under the Direction of Mary Ann Moran)

ABSTRACT

Marine dissolved organic matter (DOM) is a mixture of thousands of molecules that impact ocean life. Characterization of the most rapidly-cycling DOM, consisting of metabolites produced by marine microbes, has lagged behind that of other chemical compounds of this marine organic matter reservoir. However, identification of these compounds is important for understanding the flux of recently-fixed carbon and revealing interactions occurring between members of the ocean microbiome. Previous research identified transporter operons in the bacterium *Ruegeria pomeroyi* that had enriched expression levels when co-cultured with the diatom *Thalassiosira pseudonana*. The research presented here aims to identify the substrates that trigger expression of these operons and to quantify the abundance of these genes in the global ocean. Two of the targeted transporter operons increased in expression following the addition of acetate and taurine. Further, several of the operons could be associated with diverse groups of bacterial families in the surface ocean.

INDEX WORDS: Metabolites, Transporter, Microbial Interactions

# METABOLITES OF BACTERIAL-PHYTOPLANKTON INTERACTIONS IN THE SURFACE OCEAN

by

COURTNEY M. THOMAS

B.S., Northwest Missouri State University, 2015

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

MASTER OF SCIENCE

ATHENS, GEORGIA

2017

© 2017

Courtney M. Thomas

All Rights Reserved

# METABOLITES OF BACTERIAL-PHYTOPLANKTON INTERACTIONS IN THE SURFACE OCEAN

by

COURTNEY M. THOMAS

Major Professor: Mary Ann Moran
Committee: William B. Whitman
Brian Hopkinson

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2017

ACKNOWLEDGEMENTS

I would first like to thank my thesis advisor Dr. Mary Ann Moran of the Marine Science Department at the University of Georgia for her support while I pursued my academic career goals. Her extensive knowledge and guidance during my research allowed me to develop a dynamic and exciting research project. I sincerely appreciate everything she has taught me while I worked as her graduate research assistant. I am also deeply appreciative to my co-authors, Shalabh Sharma and Christa Smith, without their advice and hard work this research would not have been possible. Thank you to my committee members Dr. William Whitman and Dr. Brian Hopkinson for your valuable input during this research.

I would like to thank my friends for their support and encouragement during my research, especially my fellow graduate students and lab mates. Finally, I must express my very profound gratitude to my parents and to my partner for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS ..... iv

LIST OF TABLES ..... vi

LIST OF FIGURES ..... vii

CHAPTER

1. INTRODUCTION AND LITERATURE REVIEW .....1

2. BACTERIAL TRANSPORTERS AS BIOSENSORS OF LABILE DISSOLVED ORGANIC MATTER IN THE SURFACE OCEAN .....6

3. CONCLUSION.....39

LIST OF TABLES

Page

Table 2.1: Substrates added to medium after pre-growth to late exponential phase on glucose in the low substrate addition RNAseq experiment.....23

Table 2.2: Substrate binding proteins of interest based on transporters upregulated by *R. pomeroyi* DSS-3 when growing in co-culture with a marine diatom. .....24

Table 2.3: Tara Ocean samples (Sunagawa *et al.*, 2015) used for SBP analysis.....25

Table 2.4: Percent of cells with orthologs to the indicated substrate binding proteins at each station.....26

LIST OF FIGURES

Page

Figure 2.1: Heat map of targeted transporter operons response to acetate addition .....28

Figure 2.2: Reactions of the ethylmalonyl-CoA pathway .....29

Figure 2.3: Heat map of targeted transporter operons response to taurine addition .....30

Figure 2.4: Select sulfonate degradation pathways found in *R. pomeroyi* .....31

Figure 2.5: Map showing Tara Ocean stations analyzed .....32

# CHAPTER 1

# INTRODUCTION AND LITERATURE REVIEW

*Dissolved organic matter and microbial interactions*- Marine microbes comprise the largest part of the oceanic biomass and are responsible for a majority of the nutrient and chemical fluxes that occur there (Pomeroy *et al.*, 2007). One of these important fluxes is the transformation of dissolved organic matter (DOM), including the metabolites produced by phytoplankton and released into seawater (Moran *et al*., 2016). Marine phytoplankton do not exist naturally without a cadre of associated heterotrophic bacteria (Ramanan, *et al.*, 2016). It is well established that heterotrophic bacteria are often found inhabiting the nutrient rich boundary surrounding phytoplankton referred to as the phycosphere (Amin, *et al.*, 2012; Luo, *et al.*, 2014). Due to their relatively large size and silicate frustules, diatoms are major players in the biological pump (i.e., the downward flux of organic carbon into the ocean depths; Durkin, *et al.*, 2016) and the cycling of iron, nitrogen, and carbon in the surface ocean. Diatoms are an important source of metabolites found in the marine DOM pool that are consumed by members of the ocean microbiome (Amin, *et al.*, 2012).

The processing of phytoplankton-derived metabolites has profound impacts on the food webs that support ocean life (Azam & Malfatti, 2007) and the cycling of all major elements on Earth (Moran *et al.*, 2016). In fact, it is estimated that half of all carbon fixed by phytoplankton in the surface ocean passes through the DOM pool and is used within hours to days by heterotrophic bacteria (Hansell, 2013). Yet, the compounds involved in this exchange are difficult to identify because they are present in very low concentrations and are not easily distinguished from the background organic matter (Moran *et al.*, 2016). Traditional chemical

approaches to the characterization of DOM has resulted in the identification of the high molecular-weight (HMW) compounds, primarily because this is the fraction of DOM that can be physically separated from the salt matrix, yet much of this is not accessible for degradation by marine microbes. Microbiologists have worked to characterize the smaller and more labile compounds in DOM using Michaelis-Menten kinetic rate calculations and turnover rates of single organic compounds (such as dissolved adenosine triphosphate or dissolved free amino acids). These measure bacterial DOM transformations of individual compounds that are thought to be representative of the compounds supporting bacterial heterotrophy (Hollibaugh & Azam, 1983; Hodson *et al.,* 1981; Ferguson & Sunda, 1984). Although chemical methods for identifying labile molecules from seawater are still laborious, methods involved in the analysis of the environmental microbiome have made enormous strides of the past several decades and hold promise for attacking this challenging question.

*Marine environmental 'omics* - The development of 16S rRNA sequencing allowed scientists to examine a large percentage of the microbiome that was not accessible through traditional laboratory methods. Whole-genome shotgun sequencing of the Sargasso Sea took culture independent methods one step further by inventorying the genetic content of entire communities in complex environmental samples (Venter *et al.,* 2004). Environmental 'omics gradually emerged from asking questions about the identity of the microbes in an environment to addressing more complex questions such as community gene expression and the proteins microbes utilize when transforming dissolved organic compounds (Venter *et al.,* 2004; Moran *et al.,* 2016).

Larger sampling expeditions such as the Global Ocean Sampling project and the Tara Oceans expedition collected comprehensive environmental 'omic datasets on microbial

communities in the ocean, sampling surface and deep seawater from the major ocean basins (Rusch, *et al.,* 2007; Sunagawa, *et al.,* 2015). The publicly available portion of the Tara Oceans dataset is composed of metagenomic samples and environmental data collected from 210 stations across the global ocean and at multiple depths, containing a genetic survey of microbial capabilities for metabolite uptake from the DOM pool (Sunagawa, *et al.,* 2015). Understanding these communities and their metabolic capabilities provides an opportunity for better characterization of the rapidly metabolized compounds in the DOM pool, provided the genes that mediate their uptake can be correctly annotated and identified in the metagenomic datasets. However, metagenomes alone do not necessarily deliver functional information, and can only provide hints about the biogeochemical cycles linked to the community members.

*Bacterial transporters* – Both ABC (Zheng *et al.,* 2013) and TRAP (Mulligan, Fischer, & Thomas, 2011) transporters are ubiquitous in marine bacterial genomes. These multi-protein transporter systems utilize substrate binding proteins (SBP), proteins that physically bind the molecule and transfer it to the machinery responsible for translocation across the cell membrane (Mulligan, *et al.,* 2007). The specificity of these SBPs may be broad (capable of binding more than one substrate) or narrow. They can provide critical hints about the metabolites that are rapidly assimilated by bacteria (Maqbool, *et al.,* 2015; Mulligan, *et al.,* 2011). The experimental verification of the substrates targeted by microbial transporters is critical to the understanding the rapid flux of carbon through the marine microbial food web.

## References


Amin SA, Parker MS, Armbrust EV. 2012. Interactions between diatoms and bacteria. *Microbiol Mol Biol Rev* 76: 667-84

Azam F, Malfatti F. 2007. Microbial structuring of marine ecosystems. *Nat Rev Microbiol* 5: 782-91

Durkin CA, Van Mooy BAS, Dyhrman ST, Buesseler KO. 2016. Sinking phytoplankton associated with carbon flux in the Atlantic Ocean. *Limnol Oceanogr* 61: 1172-87

Ferguson RL, Sunda WG. 1984. Utilization of amino acids by planktonic marine bacteria: Importance of clean technique and low substrate additions 1,2. *Limnol Oceanogr* 29: 258-74

Hansell DA. 2013. Recalcitrant dissolved organic carbon fractions. *Ann Rev Mar Sci* 5: 421-45

Hodson RE, Azam F, Carlucci AF, Fuhrman JA, Karl DM, Holm-Hansen O. 1981. Microbial uptake of dissolved organic matter in McMurdo Sound, Antarctica. *Mar Biol* 61: 89-94

Hollibaugh JT, Azam F. 1983. Microbial degredation of dissolved proteins in seawater. *Limnol Oceanogr* 28: 11104-1116

Maqbool A, Horler RS, Muller A, Wilkinson AJ, Wilson KS, Thomas GH. 2015. The substrate-binding protein in bacterial ABC transporters: dissecting roles in the evolution of substrate specificity. *Biochem Soc Trans* 43: 1011-7

Moran MA, Kujawinski EB, Stubbins A, Fatland R, Aluwihare LI, et al. 2016. Deciphering ocean carbon in a changing world. *Proc Natl Acad Sci USA* 113: 3143-51

Mulligan C, Fischer M, Thomas GH. 2011. Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. *FEMS Microbiol Rev* 35: 68-86

Mulligan C, Kelly DJ, Thomas GH. 2007. Tripartite ATP-independent periplasmic transporters: application of a relational database for genome-wide analysis of transporter gene frequency and organization. *J Mol Microbiol Biotechnol* 12: 218-26

Pomeroy LR, Williams PJ, Azam F, Hobbie JE. 2007. The microbial loop. *Oceanography* 20: 28-33

Ramanan R, Kim BH, Cho DH, Oh HM, Kim HS. 2016. Algae-bacteria interactions: Evolution, ecology and emerging applications. *Biotechnol Adv* 34: 14-29

Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, et al. 2007. The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol* 5: e77

Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, et al. 2015. Structure and function of the global ocean microbiome. *Science* 348: 1261359-59

Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch DB, et al. 2004. Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science* 304: 66-74

Zheng WH, Vastermark A, Shlykov MA, Reddy V, Sun EI, Saier MH, Jr. 2013. Evolutionary relationships of ATP-Binding Cassette (ABC) uptake porters. *BMC Microbiol* 13: 98

# CHAPTER 2

# BACTERIAL TRANSPORTERS AS BIOSENSORS OF LABILE DISSOLVED ORGANIC MATTER IN THE SURFACE OCEAN [1]

[1]Thomas, C. M., Sharma, S., Smith, C.B., and Moran, M.A.
To be submitted to *Environmental Microbiology*.

**Abstract**

Marine dissolved organic matter (DOM) is a mixture of thousands of molecules that impact ocean life. Characterization of the most rapidly-cycling subset of DOM (metabolites produced by marine microbes) has lagged behind that of other chemical compounds of this marine organic matter reservoir. However, it is important for understanding of the flux of recently-fixed carbon and revealing interactions occurring between members of the ocean microbiome. Previous research identified transporter operons in the bacterium *Ruegeria pomeroyi* with enriched expression levels when co-cultured with the diatom *Thalassiosira pseudonana*. The research presented here aims to identify the substrates that trigger expression of these operons and to quantify the abundance of these genes in the global ocean. Two of the targeted transporter operons increased in expression following the addition of acetate and taurine. Further, several of the operons were associated with diverse groups of bacterial families in the surface ocean.

Keywords: Metabolites, Transporter, Microbial Interactions

## Introduction

Marine microbes are responsible for a majority of the nutrient and chemical fluxes that occur in the global ocean, including the metabolites produced and released by phytoplankton into the oceanic pool of dissolved organic matter (DOM) (Pomeroy *et al.,* 2007; Moran *et al*., 2016). Previous work identified several transporter operons that were upregulated when the model heterotrophic marine bacterium *Ruegeria pomeroyi* DSS-3 was grown in co-culture with the marine diatom *Thalassiosira pseudonana* (Durham *et al.,* 2017). Since most gene annotations are based on sequence similarity to phylogenetically distant organisms and are not experimentally verified in each bacterial species, the specific metabolites that triggered the transcriptional response by *R. pomeroyi* were not known. The research presented here works towards identification of the substrates targeted by the upregulated transporters. The broad goal of this work is to increase our understanding of the labile metabolites passed from autotrophs to heterotrophs in the surface ocean, and to examine the global distribution of these transporters through bioinformatic analysis of marine metagenomic data.

The marine Roseobacter clade is nested within the Rhodobacteraceae family and the class Proteobacteria, subclass Alphaproteobacteria. It is abundant in the ocean, having successfully invaded a variety of ecological niches, likely due to the group's diverse metabolic capabilities (Wagner-Dobler & Biebl, 2006; Buchan *et al.,* 2005). In addition to being ecologically important, many roseobacters are easily maintained in laboratory culture (Buchan *et al.,* 2014). The strain used for these studies, *Ruegeria pomeroyi* DSS-3, was the first heterotrophic marine bacterium to have its genome completely sequenced (Moran *et al.,* 2004; Buchan *et al.,* 2005; Moran *et al.,* 2007; Luo & Moran, 2014). *R. pomeroyi* DSS-3 was initially isolated from seawater collected from the southeastern coast of the United States and has since played a major

role in the study of microbially processed DOM (Gonzalez *et al.*, 2003; Rivers *et al.*, 2014). Like most roseobacters, *R. pomeroyi* has a relatively high number of tripartite ATP-independent periplasmic (TRAP, 28 operons) and ATP-binding cassette (ABC, 119 operons) transporters (Newton *et al.*, 2010; Moran *et al.*, 2007; Rivers *et al.*, 2014; http://www.membranetransport.org/transportDB2/index.html; http://roseobase.org/index.html). These characteristics as well as the extensive body of previous research associated with this model organism made it an ideal candidate for this study.

ABC transporters target a diverse set of substrates such as amino acids, sugars, vitamins, peptides, and lipids (Davidson *et al.*, 2008; Denger *et al.*, 2011). For the transporters encoded in the *R. pomeroyi* genome, current annotations indicate uptake of amino acids (polar and branched chain), C4-dicarboxylates, carbohydrates, choline, glycine betaine, proline, glutamate/glutamine, aspartate/asparagine, glutathione, peptides, polyamines, sulfonates, taurine, *N*-acetyltaurine, xylose, monosaccharides, threonine, and homoserine (http://roseobase.org/index.html). Although ABC transporters were discovered in the 1970s, there is still limited understanding of the substrates targeted by these transporter complexes, including the diversity and specificity of the substrate binding proteins (SBP) that initiate the transport process (Maqbool, *et al.*, 2015).

The TRAP transporter family is thought to translocate a narrower range of substrates. They were identified first as transporters specializing in dicarboxylic acids, although this has expanded to include glutamate, sialic acid, taurine, lactate, and gluconate (Mulligan, et al., 2011). In *R. pomeroyi*, TRAP transporter annotations are currently limited to dihydroxypropanesulfonate (DHPS), isethionate, ectoine/hydroxyectoine, mannitol, and C4-dicarboxylates (Mulligan, *et al.*, 2007; Mayer, *et al.*, 2010; http://roseobase.org/index.html;

Rivers, *et al*., 2014). The transporters for DHPS, isethionate, and ectoine/hydroectoine have been experimentally verified (Mayer *et al.*, 2010; Weinitschke *et al.*, 2010; Schulz *et al*., 2017).

The function of many genes carried by ocean microbes remains unclear. This is indeed the case for transporters, whose annotations are usually based on similarity to a few experimentally characterized genes (Moran *et al.*, 2016). In a co-culture study of *R. pomeroyi* and centric diatom *T. pseudonana* in which the diatom provided the only source of organic carbon, Durham and colleagues (2017) found that several operons predicted to transport organic compounds were upregulated by the bacterium. Two of the upregulated transporter operons had been experimentally verified to translocate sulfonate compounds DHPS (via a TRAP transporter; Mayer *et al.*, 2010) and *N*-acetyltaurine (via an ABC transporter; Denger *et al.*, 2011). Additionally, a deletion mutant missing the three-gene *R. pomeroyi* transporter system for DHPS, named *hpsKLM*, showed deficient growth in co-culture with *T. pseudonana* compared to the wild type, with cell numbers decreased by as much as 6-fold (Landa *et al.*, In press). While the DHPS transporter was the most highly upregulated transporter system in the Durham *et al.* (2015) study, eight additional systems showed an increase in expression as well (Durham *et al.*, 2017). With the exception of the DHPS and *N*-acetyltaurine transporters, the remaining seven upregulated transporter operons have not been experimentally verified, and their characterization offers an opportunity for further identification of labile metabolites in the marine DOM pool.

Following the work previously described by Durham *et al.*, (2015, 2017), Landa *et al.* (In press) showed that 13 transporters annotated for uptake of organic substrates were upregulated in *R. pomeroyi* when metabolites were provided by the diatom *T. pseudonana* but not when metabolites were provided by the dinoflagellate *Alexandrium tamarense*. Landa *et al.* (In press) found that of the nine transporter systems targeted in this research, four were enriched in the

diatom dominated culture when compared to the dinoflagellate dominated culture (represented by substrate-binding proteins SPO0591, SPO1112, SPO2658, and SPO3665), three were enriched in the dinoflagellate culture (SPO0660, SPO2433, and SPO3705), and two did not display a differential expression between the communities (SPO1021 and SPO2441). These examples provide distinct indications that marine heterotrophic bacteria, and specifically *R. pomeroyi,* are capable of responding transcriptionally to metabolites released by phytoplankton taxa. More broadly, they offer evidence that bacterial-phytoplankton associations are important links in the microbial food web, that the compounds that mediate these links are diverse, and that the links influence the flux of elements such as carbon, nitrogen, and sulfur in ocean surface waters.

In this study, my research goal was to use transcriptomic approaches to test a variety of potential substrates that might trigger gene expression of the *R. pomeroyi* transporter systems that were upregulated during co-growth with the diatom. This provides an alternative approach to identifying metabolites by traditional chemical analysis. Further, I used a transporter identification pipeline to mine the existing Tara Ocean metagenomics data for information on the abundance and distribution of this gene set and address its broader ecological relevance in the ocean microbiome.

## Materials and Methods

*RNASeq experiments* - *Ruegeria pomeroyi* DSS-3was grown on ½ YTSS agar plates at 30°C and a single colony transferred to liquid ½ YTSS medium and shaken at 200 rpm. After 48 h the culture was transferred into marine minimal media (MBM) containing 2 mM glucose at a final OD ($A_{600nm}$) of ~1 x $10^{-3}$. Growth was monitored on a plate reader for 28-30 h and after

indications of a decrease in exponential growth rate, a substrate was added at a concentration of either 100 µM (acetate, proline, valine, leucine, isoleucine, taurine, galacturonic acid, glutamate, and aspartate) or 200 µM (DHPS and glycine betaine). Following the substrate addition, the culture was incubated at 30°C and 200 rpm for 45 min. All treatments were carried out in duplicate.

*RNA extraction*- Samples were mechanically disrupted for 10 min using 0.5 mL of 0.1 mm zirconia beads, and centrifuged for 1 min at 5,000 rpm. The supernatant was transferred into a 1.5 mL collection tube, and 100% ethanol was added and mixed with a 21-gauge needle. RNA was extracted using a modified RNeasy mini kit (Qiagen) protocol according to the manufacturer's protocol. DNA was removed from the samples using a modified (double DNase enzyme treatment) Ambion® TURBO DNA-free™ (Invitrogen) kit.

Samples were concentrated using the RNA Clean & Concentrator™-5 (Zymo, Irvine, CA). Ribosomal RNA was removed using the Ribo-Zero rRNA Removal Kit protocol (Illumina, San Diego, CA). A second cleanup was completed, and samples were quantified with a Qubit 2.0 Fluorometer at the Georgia Genomics Facility (Athens, GA). Samples were checked for degradation and rRNA depletion using a High Sensitivity RNA ScreenTape System (Agilent, Waldbronn, Germany).

*Library preparation and sequencing*- cDNA stranded libraries were prepared using a modified protocol with the KAPA Stranded RNA-Seq kit (Kapa Biosystems, Wilmington, MA). The RNASeq low substrate addition experiment used indexes from NEBNext Multiplex Oligos Primer Set 1 and 2 (New England Biolabs). The RNASeq high substrate concentration libraries were prepared with the KAPA Single-Indexed Adapter Set A + B kit (Kapa Biosystems, Wilmington, MA). All libraries were sequenced at HudsonAlpha Institute for Biotechnology

(Huntsville, AL) on a HiSeq Illumina 2500. Reads were processed for quality control using the FASTX toolkit, imposing a minimum quality score of 20 on 80% of the read length. Reads aligning to an in-house rRNA database were removed (blastn, score cutoff ≥50). Remaining reads were mapped to the *R. pomeroyi* DSS-3 genome using Bowtie 2.2.1 (Langmead and Salzberg, 2012). Reads that aligned exactly one time to the coding strand were counted using HTSeq-count 0.6.1 (Anders *et al.*, 2015). These reads were analyzed for differential gene expression using the R package DESeq2 (Love *et al.*, 2014).

*Tara Oceans analysis*- A custom gene database (geneDB) was constructed for identifying orthologs of selected *R. pomeroyi* transporters in public marine metagenomic data. The substrate binding proteins (SBP) from transporter systems upregulated in *R. pomeroyi* when grown with *T. pseudonana* (Durham *et al.*, 2015) (Table 1.2) were used as the target genes in the analysis. Each geneDB consisted of likely orthologs of the substrate binding proteins identified in the genome sequences of a diverse set of marine bacteria available at the Joint Genome Institute's Integrated Microbial Genomes & Microbiome Samples (IMG-M ER) website (https://img.jgi.doe.gov/cgi-bin/mer/).

A Hidden Markov Model (HMM) was generated from the proteins in each geneDB using GraftM (Boyd *et al*., 2007) and used to identify orthologs of the transporter SBPs in the Tara Ocean peptide database (consisting of protein sequences called from assembled contigs), downloaded from EMBL-EBI (http://www.ebi.ac.uk). In this bioinformatic pipeline, potential orthologs from the Tara Oceans peptide database initially identified by the HMM were retained if they had ≥40% similarity to a reference ortholog in the geneDB and a top hit to the *R. pomeroyi* ortholog in Roseoblast (http://roseobase.org/AnnotTools/blast.html). An e-value cutoff was chosen individually for each SBP based on these criteria. In the next step, the remaining

assembled contigs not initially identified in the HMM search were compared to the contigs identified by the HMM as orthologs using blastp, and any with identities of $\geq$ 90% were removed. This step eliminated any contigs that did not have a high score with the HMM but could be nonethless be orthologs; for example, partial gene assemblies. This pipeline resulted in two Tara Oceans peptide databases for each SBP; one consisted of assembled proteins identified as orthologs to the SBP proteins, and the other consisted of the rest of the assembled proteins that were not orthologs to the SBP proteins. Both datasets were used in the subsequent raw read searches.

The Tara Oceans raw read data were compared against the Tara Oceans SBP and non-SBP peptide databases to identify orthologous reads in individual samples. First, the raw reads were processed for quality control using the FASTX toolkit, imposing a minimum quality score of 20 on 80% of the read length. The quality-checked reads were then blasted against the orthologous SBP Tara Oceans peptide database generated for each SBP. This reduced data set was then blasted against the combined peptide database containing both SBP peptides and non-SBP peptides. All reads that hit to the SBP peptide database were kept for quantification.

To quantify the percent of cells in each sample that contain the targeted SBP, the previously described pipeline was used to quantify the abundance of the single copy gene *recA*. To calculate the percent of cells with the SBP, the ratio of SBP genes/*recA* genes was calculated and normalized to gene length (Howard *et al.*, 2006).

Samples used in the Tara Oceans analysis were chosen to include sites that had paired surface water and deep chlorophyll maximum samples collected on the same day and representing a size fraction of either 0.22-1.6 µm or 0.22-3.0 µm. Fourteen sites with a surface water sample and a deep chlorophyll maximum sample were analyzed (Table 2.3).

*Transporter mapping and MINE analysis* - Environmental parameters and taxonomic profiles were compared to gene abundance using the MINE program in the Maximal Information-Based Nonparametric Exploration for Variable Analysis (minerva) package in R (Filosi *et al*., 2017). Parameters used in the analysis included: depth of sample (m), temperature (ºC), salinity, oxygen (µmol/kg), nitrate (µmol/L), depth maximum $O_2$ (m), depth minimum $O_2$ (m), nitracline (m), as well as the operational taxonomic units (OTU) of prokaryotes and autotrophic eukaryotes (percent of OTUs reported). Bacterial OTUs were binned at the family level and the percent of OTUs reported was calculated.

To find representative genomes of the bacterial OTUs of interest, the 16S rRNA OTU reference sequences were downloaded from the Tara Ocean companion page (http://ocean-microbiome.embl.de/companion.html). Genomic analysis for orthologous genes from genera within bacterial families that had a maximal information coefficient (MIC) of > 0.8 in pairwise analysis with a transporter gene was completed using IMG/MER (https://img.jgi.doe.gov/cgi-bin/mer/main.cgi). Eukaryotic taxonomic profiles were downloaded from the Tara Ocean Eukaryotic Diversity companion site (http://taraoceans.sb-roscoff.fr/EukDiv). All maps depicting the percent of cells in each sample that contain the targeted SBP were generated in R using the ggplot2 package (Wickham *et al*., 2016).

**Results and Discussion**

*Transporter expression* – RNA was extracted from *Ruegeria pomeroyi* DSS-3 cultures amended with one of a suite of seven substrates (at 100 or 200 µM concentration) after pre-growth to late exponential phase on glucose (Table 2.1). Forty-five minutes after substrate amendment, RNA was extracted. Sequence libraries averaging 6.9 $\pm$ 1.5 million reads per sample (50 bp, single

end) were obtained from duplicate amendments with acetate, glutamate, DHPS, taurine, isoleucine, leucine, or valine as the added substrate, and transcript reads were mapped to the *R. pomeroyi* DSS-3 genome.

Transporters that were upregulated by *R. pomeroyi* when growing in co-culture with *Thalassiosira* as the sole source of carbon (Table 2.2) were the focus of analysis. Differential expression analysis (DESeq2) was carried out for all possible pairs of substrates, and cases in which a target transporter system was upregulated with one of the substrates relative to all others were further analyzed. Two transporter operons of interest, SPO1017-SPO1021 annotated as a branched chain amino acid ABC transporter and SPO0660-0664 annotated as an *N*-acetyltaurine ABC transporter, showed a pattern of consistent relative enrichment for one substrate compared to most of the others tested.

*ABC branched chain amino acid transporter*- The addition of 100 μM acetate to the *R. pomeroyi* culture resulted in an upregulation of the SPO1017-1021 operon compared to the other six substrate treatments (Fig. 2.1). Further, this was the most highly upregulated transporter operon under acetate amendment relative to the 147 total ABC and TRAP transporter operons in the *R. pomeroyi* genome. The substrate binding protein (SPO1021) was among the most changed in expression, with up to a 3.7-fold enrichment (DESeq2 $\log_2$ fold change, $p \leq 1E\text{-}5$) when expression in the acetate amendment was compared to expression in the other substrates. Additionally, key genes in the ethylmalonyl-CoA pathway were upregulated in the acetate treatment relative to the other six substrates (Figure 2.2). This pathway, first described in 2007 by Erb *et al.*, (2007), allows growth on acetate and other C2 substrates. The pathway generates glyoxylate and propionyl-CoA for biosynthesis and energy generation. The diagnostic gene for

the pathway, crotonyl-CoA carboxylase/reductase (Ccr) (Erb *et al.*, 2007), was significantly upregulated (up to 5-fold) in 4 of the 6 treatments (DESeq2 $\log_2$ fold change, $p \leq 0.05$).

The expression profile of key genes in the acetate amendment matched those designated as "acetate switch" genes in previous studies (Oh *et al.*, 2002; Wolfe 2005; Wolfe 2008). This phenomenon refers to a bacterium's shift from acetate excretion, when cells have an acetogenic substrate available (in this case glucose), to acetate uptake when glucose is depleted (Wolfe 2005). This response by *Escherichia coli* has been reported to involve three expression changes: upregulation of acetyl coenzyme A synthetase (*acs*), downregulation of phosphate acetyltransferase (*pta*), and downregulation of acetate kinase (*ack*) (Wolfe 2005, Oh *et al.*, 2002). In *R. pomeroyi*, the *acs* gene SPO1813 was significantly upregulated in the acetate treatment relative to all other treatments, while *pta* (SPO3560) was significantly downregulated in four of the six comparisons and *ack* (SPO0116) was downregulated in five of the six, thus matching the *E. coli* acetate switch expression profile (Oh *et al.*, 2002; Wolfe 2005). *Ruegeria pomeroyi* has an annotated acetate cation symporter (SPO1810), which resides in the same operon as the *acs* gene. This gene was significantly upregulated in acetate for only two of six treatments (DESeq2, $p \leq 0.05$) and was downregulated in acetate compared to valine. This gene may be mis-annotated or operate at higher concentrations of acetate than the 100 µM concentration used here.

The acetate switch has been linked to regulation of cell-to-cell interactions such as chemotaxis (Ramakrishnan *et al.*, 1998; Fraiber *et al.*, 2015) and quorum sensing (Studer *et al.*, 2008). Despite its frequent association with phytoplankton, *R. pomeroyi* does not have identified chemotaxis genes, yet is motile and carries two sets of *luxI/luxR* quorum sensing genes (Moran *et al.*, 2004). Other roseobacters, such as *Silicibacter* sp. Strain TM1040, exhibit chemotaxis towards phytoplankton-released metabolites (Miller *et al*., 2004). It is notable that the *luxR*-

*1/luxI-1* gene pair was upregulated in three of the six treatments up to 1.9-fold (DESeq2 $\log_2$ fold change, $p \leq 0.05$). Additionally, the *luxR-2* gene was also upregulated in 5 of the 6 comparisons (DESeq2 up to 2.6-fold, $p \leq 0.05$) and *luxI-2* was upregulated in 3 of the 6 comparisons (DESeq2 up to 2.8-fold, $p \leq 0.05$). Additional research is needed to determine whether this transporter system and the acetate switch could regulate these mechanisms in *R. pomeroyi*.

*Taurine addition*- The addition of 100 μM taurine to the *R. pomeroyi* culture resulted in an upregulation of the taurine transporter (*tauABC*), isethionate transporter (*iseKLM*), and *N*-acetyltaurine transporter (*naaABB'CC*') relative to addition of the four other substrates (Fig. 2.3). The substrate binding proteins of each of these operons again had the highest and most significant upregulation compared to the other transporter components (DESeq2, $p \leq 0.05$). Additionally, several downstream sulfonate dissimilatory genes were upregulated (Figure 2.3).

The induced expression of the taurine transporter, the taurine transcriptional regulator (*tauR*), and associated dissimilation genes (*tpa, xsc, pta*) during *R. pomeroyi* growth on taurine has been documented previously (Gorzynska *et al*., 2006; Styp von Rekowski *et al*., 2005; Krejcik *et al*., 2010). The additional upregulation of genes for the isethionate TRAP transporter, isethionate dehydrogenase, *N*-acetyltaurine ABC transporter, and *N*-acetyltaurine amidohydrolase was unexpected, however, and may result from partial sharing of the dissimilation pathways for these C2 sulfonates; indeed, all three compounds are processed by the taurine-pyruvate aminotransferase (*pta*) and sulfoacetaldehyde acetyltransferase (*xsc*) (Figure 2.4). Landa and colleagues (In press) found that *R. pomeroyi* expressed the genes for both taurine and *N*-acetyltaurine transport simultaneously when the bacterium grew in co-culture with the dinoflagellate *Alexandrium tamarense*. All three sulfonates are known metabolites of phytoplankton (Durham *et al.*, 2015), are potentially encountered together in the marine

environment, and are potentially regulated through products of the shared catabolic genes. In the *R. pomeroyi* genome, the *tauR* regulatory gene is co-located with the shared catabolic genes *xsc* and *pta,* as well as the sulfite dehydrogenase genes *soeABC*. The latter are responsible for the excretion of sulfate, a byproduct of sulfonate catabolism, and were also upregulated in taurine compared to other treatments.

*Highly expressed substrate binding proteins* - For the previously described upregulated transporter systems, the substrate binding protein (SBP) was the most highly upregulated gene in each operon. This is not the first time that the expression of the SBP subunit has been found to be the mostly highly expressed gene of a transporter operon. Marine metaproteomic studies, for example, have found SBP subunits from ABC and TRAP transporters dominating the community proteomes from the Sargasso Sea and coastal Oregon (Sowell *et al.,* 2009; Sowell *et al.,* 2011; Williams *et al.,* 2012). The increased expression of various SBPs is potentially beneficial for increasing the efficiency of substrate capture, thought to be the rate limiting step for bacterioplankton in oligotrophic environments. Paired with the large surface area-to-volume ratio of small bacterial cells, this may be an important adaptation for nutrient uptake (Sowell *et al.,* 2011).

*Other substrates* -The other treatments (DHPS, glutamate, valine, leucine, and isoleucine) did not exhibit definitive patterns in transporter expression. The SBP for DHPS (*hpsK*) was upregulated in the DHPS treatment compared to all other treatments, although not significantly. The valine amendment resulted in a significant upregulation of the *R. pomeroyi* glucose/xylose ABC transporter (SPO0861-0863) in all but one of the treatments (DESeq2, $p \leq 0.05$) with a $\log_2$ fold change of almost 7-fold. We suspect that utilization of the initial glucose in the medium lagged slightly behind in this treatment compared to the others, and the transporter was still

upregulated for glucose utilization at the time of RNA sampling. The possible allophanate TRAP transporter operon (SPO1112-1114) and the glycine/proline ABC transporter operon (SPO2441-2443) were upregulated in the isoleucine treatment, although not significantly. In the glutamate treatment, the SBP subunit for the possible galactarate TRAP transporter (SPO2433) was significantly upregulated (DESeq2, $p < 0.005$) with a $\log_2$ fold change of at least 2.8 in all but one of the treatments.

*Bioinformatic analysis of SBPs* –To quantify these specific SBPs, a bioinformatic pipeline was developed for mining the Tara Oceans metagenomic data. Fourteen stations with a paired surface (SRF) and deep chlorophyll maximum (DCM) sample (Table 2.3 and Figure 2.5) were analyzed for abundance of SBPs for 9 transporter operons and the single copy *recA* gene, with the latter used to calculate the percent of cells containing the SBP of interest (Table 2.4). Additionally, environmental parameters and taxonomic profiles were compared to gene abundance to look for significant relationships between the SBP of interest and other factors such as temperature and taxonomy.

Currently, four of the stations (4, 7, 36, and 110) have been analyzed for the abundance of the *recA* gene. Of the 9 SBP targeted (Table 2.1), 5 (*hpsk*, SPO2433, SPO0660, SPO3665, and SPO2441) have been quantified at these same stations. The five genes and eight samples (paired SRF and DCM samples from 4 stations) will be the focus of this analysis.

The percent of cells in each sample harboring the targeted SBP varied greatly (Table 2.4) from as little as 0.03% (SPO3665, Station7 DCM) to as much as 7.55% (SPO3665, Station 7 SRF). The counts of each bacterial OTU profile were binned at the family level and the percent abundance calculated. These data along with the environmental data for each of the 8 samples were analyzed using the MINE program in R, which indicated that several can be predicted by

the abundance of specific bacterial families (MIC > 0.9). The SBPs SPO0660, *hpsk*, and SPO2441 showed high MIC relationship with bacterial families in the SAR11 clade (Morris *et al.,* 2002). SAR11 families Surface 1 and 4 were previously identified in open ocean surface communities (Cram *et al.,* 2015), and the Chesapeake-Delaware Bay SAR11 family is abundant in brackish water (Ameryk *et al.,* 2014). These three families had an MIC of > 0.90 to *hpsk* and SPO2441 (SAR11 Surface 1 and Chesapeake-Delaware Bay), and SPO0660 (SAR11 Surface 4). A Pearson correlation coefficient was computed to assess the relationship between the percent of cells with these SBPs and the abundance of these OTUs at these 4 stations. The SBPs *hpsk* and SPO2441 have a positive correlation with the SAR11 Chesapeake-Delaware Bay family as well as the SAR11 Surface 1 family (all with $r > 0.76$ and $p < 0.02$). The SAR11 Surface 4 family also has a positive correlation with the *N*- acetyltaurine SBP, SPO0660 ($r = 0.75$, $p = 0.03$).

*hpsk* and SPO2441 also were significantly related to the abundance of the NS9 marine group OTU. NS9 is within the class Flavobacteria, members of which are frequently found in high abundance during phytoplankton blooms or attached to phytoplankton-produced particulate organic matter (Milici *et al*., 2017, Pinhassi *et al.,* 2004). No linear correlation was found between these SBPs and abundance of marine group NS9.

The SBP that is currently hypothesized to transport allophonate (SPO3665) was associated with the abundance of the *Rhodospirillaceae* family, another Alphaproteobacteria group. This family includes the AEGEAN-169 marine group, a sister clade to SAR11, that has been found at depths where SAR11 is less abundant (Cram *et al.,* 2015). A Pearson's test showed that the *Rhodospirillaceae* family and SPO3665 were positively correlated ($r = 0.82$, $p = 0.01$).

The gene that is currently annotated as a galactarate SBP (SPO2433) was related to the most diverse range of bacterial families, including *Phycisphaeraceae*, *Salinisphaeraceae*, and

*Comamonadaceae*. The bacterial family *Comamonadaceae* is in the class Betaproteobacteria and includes the oligotrophic marine group BAL58 (Simu & Hagstrom *et al.*, 2004; Ameryk *et al.*, 2014).

Based on the 16S rRNA OTU reference sequence for these bacterial families, genomes of representatives from each family were blasted against the SBPs to look for orthologous gene sequences. Preliminary genomic investigation of these groups failed to find orthologous genes in the genomes of families with high MIC relationships to *hpsk*, SPO2441, and SPO0660. However potential orthologs for SPO3665 and SPO2433 were found in members of the *Rhodospirillaceae* and *Comamonadaceae* families.

These preliminary results indicate that the Tara Ocean SBP pipeline is capturing a diversity of orthologous proteins associated with a wide range of bacterial groups, some of which are distantly related to *R. pomeroyi*. The complete analysis of all 14 stations and two depths will provide a more robust link between these SBPs and the abundance of bacterial taxa.

**Tables and Figures**

*Table 2.1* Substrates added to medium after pre-growth to late exponential phase on glucose in the RNAseq experiment. RNA was extracted 45 min after substrate addition to an early stationary phase culture.

| Substrate | Concentration |
|---|---|
| Acetate* | 100 µM |
| Valine* | 100 µM |
| Galacturonic Acid | 100 µM |
| Aspartate | 100 µM |
| DHPS* | 200 µM |
| Glycine Betaine | 200 µM |
| Proline | 100 µM |
| Glutamate* | 100 µM |
| Isoleucine* | 100 µM |
| Leucine* | 100 µM |
| Taurine * | 100 µM |

*Indicates that the RNA was successfully sequenced from this substrate addition.

*Table 2.2* Substrate binding proteins of interest based on transporters upregulated by *R. pomeroyi* DSS-3 when growing in co-culture with a marine diatom.

| **Locus tag** | **Possible substrate** | **Transporter type** | **Experimentally verified?** |
|---|---|---|---|
| SPO0591 (*hspK*) | DHPS | TRAP | Yes (Mayer *et al.*, 2010) |
| SPO0660 (*naaA*) | *N*-acetyltaurine | ABC | Yes (Denger *et al.*, 2011) |
| SPO1021 | branched chain amino acid | ABC | No |
| SPO1112 | Unknown | TRAP | No |
| SPO2433 | galactarate | TRAP | No |
| SPO2441 | glycine betaine/proline | ABC | No |
| SPO2658 | glutamate/aspartate | ABC | No |
| SPO3665 | allophanate | TRAP | No |
| SPO3705 | branched chain amino acid | ABC | No |

*Table 2.3* Tara Ocean samples (Sunagawa *et al.*, 2015) used for SBP analysis. SRF=Surface Layer, DCM= Deep Chlorophyll maximum

| Station | Environmental Features | Size fraction (µm) | Ocean and sea regions |
|---|---|---|---|
| **4** | SRF | 0.22 -1.6 | North Atlantic Ocean |
| **4** | DCM | 0.22 -1.6 | North Atlantic Ocean |
| **7** | SRF | 0.22 -1.6 | Mediterranean Sea |
| **7** | DCM | 0.22 -1.6 | Mediterranean Sea |
| **36** | SRF | 0.22 -1.6 | Indian Ocean |
| **36** | DCM | 0.22 -1.6 | Indian Ocean |
| **38** | SRF | 0.22 -1.6 | Indian Ocean |
| **38** | DCM | 0.22 -1.6 | Indian Ocean |
| **52** | SRF | 0.22 -1.6 | Indian Ocean |
| **52** | DCM | 0.22 -1.6 | Indian Ocean |
| **66** | SRF | 0.22-3.0 | South Atlantic Ocean |
| **66** | DCM | 0.22-3.0 | South Atlantic Ocean |
| **68** | SRF | 0.22-3.0 | South Atlantic Ocean |
| **68** | DCM | 0.22-3.0 | South Atlantic Ocean |
| **82** | SRF | 0.22-3.0 | South Atlantic Ocean |
| **82** | DCM | 0.22-3.0 | South Atlantic Ocean |
| **98** | SRF | 0.22-3.0 | South Pacific Ocean |
| **98** | DCM | 0.22-3.0 | South Pacific Ocean |
| **109** | SRF | 0.22-3.0 | North Pacific Ocean |
| **109** | DCM | 0.22-3.0 | North Pacific Ocean |
| **110** | SRF | 0.22-3.0 | South Pacific Ocean |
| **110** | DCM | 0.22-3.0 | South Pacific Ocean |
| **137** | SRF | 0.22-3.0 | North Pacific Ocean |
| **137** | DCM | 0.22-3.0 | North Pacific Ocean |
| **142** | SRF | 0.22-3.0 | North Atlantic Ocean |
| **142** | DCM | 0.22-3.0 | North Atlantic Ocean |
| **150** | SRF | 0.22-3.0 | North Atlantic Ocean |
| **150** | DCM | 0.22-3.0 | North Atlantic Ocean |

*Table 2.4* Percent of cells with orthologs to the indicated substrate binding proteins at each station. Station number corresponds to Figure 2.5, dashes indicate that analysis at these stations is ongoing. Italicized numbers indicate that these numbers were used in the MINE analysis.

| Possible substrate | | DHPS | galactarate | *N*-acetyltaurine | Branched chain amino acid | none | allophanate | Glutamate/ aspartate | Glycine betaine/ proline | Branched chain amino acid |
|---|---|---|---|---|---|---|---|---|---|---|
| **Station** | **Depth** | **hpsK** | **SPO2433** | **SPO0660** | **SPO1021** | **SPO1112** | **SPO3665** | **SPO2658** | **SPO2441** | **SPO3705** |
| **4** | SRF | *0.14* | *0.73* | *0.81* | 0.01 | - | *5.94* | 0.00 | *0.34* | 0.00 |
| **4** | DCM | *0.11* | *0.46* | *0.90* | 0.00 | - | *5.50* | - | *0.42* | - |
| **7** | SRF | *0.40* | *0.68* | *0.68* | 0.01 | - | *7.55* | - | *1.09* | - |
| **7** | DCM | *0.31* | *0.69* | *0.55* | - | - | *0.03* | - | *0.71* | - |
| **36** | SRF | *0.72* | *0.64* | *0.57* | 0.00 | 0.00 | *5.24* | 0.01 | *0.62* | 0.00 |
| **36** | DCM | *0.58* | *0.60* | *0.50* | 0.02 | - | *5.25* | 0.00 | *0.93* | 0.00 |
| **38** | SRF | 0.24 | 0.34 | 0.19 | 0.01 | 0.00 | 4.36 | 0.00 | 0.12 | 0.00 |
| **38** | DCM | 0.09 | 0.07 | 0.37 | 0.00 | 0.00 | 4.23 | 0.00 | 0.05 | 0.00 |
| **52** | SRF | 0.14 | 0.05 | 0.56 | 0.00 | 0.00 | 5.07 | 0.00 | 0.01 | 0.01 |
| **52** | DCM | 0.07 | 0.46 | 0.43 | 0.00 | 0.00 | 5.33 | 0.00 | 0.28 | 0.09 |
| **66** | SRF | - | - | - | - | - | - | - | - | - |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **66** | DCM | - | - | - | - | - | - | - | - | - |
| **68** | SRF | - | - | - | - | - | - | - | - | - |
| **68** | DCM | - | - | - | - | - | - | - | - | - |
| **82** | SRF | - | - | - | - | - | - | - | - | - |
| **82** | DCM | - | - | - | - | - | - | - | - | - |
| **98** | SRF | - | - | - | - | - | - | - | - | - |
| **98** | DCM | - | - | - | - | - | - | - | - | - |
| **109** | SRF | - | - | - | - | - | - | - | - | - |
| **109** | DCM | - | - | - | - | - | - | - | - | - |
| **110** | SRF | *0.10* | *0.31* | *0.34* | 0.01 | 0.00 | *7.15* | 0.01 | *0.09* | 0.04 |
| **110** | DCM | *0.07* | *0.21* | *0.57* | 0.01 | 0.00 | *4.55* | 0.00 | *0.07* | 0.02 |
| **137** | SRF | - | - | - | - | - | - | - | - | - |
| **137** | DCM | - | - | - | - | - | - | - | - | - |
| **142** | SRF | - | - | - | - | - | - | - | - | - |
| **142** | DCM | - | - | - | - | - | - | - | - | - |
| **150** | SRF | - | - | - | - | - | - | - | - | - |
| **150** | DCM | 0.07 | - | 0.22 | 0.00 | - | 3.74 | - | 0.16 | 0.07 |

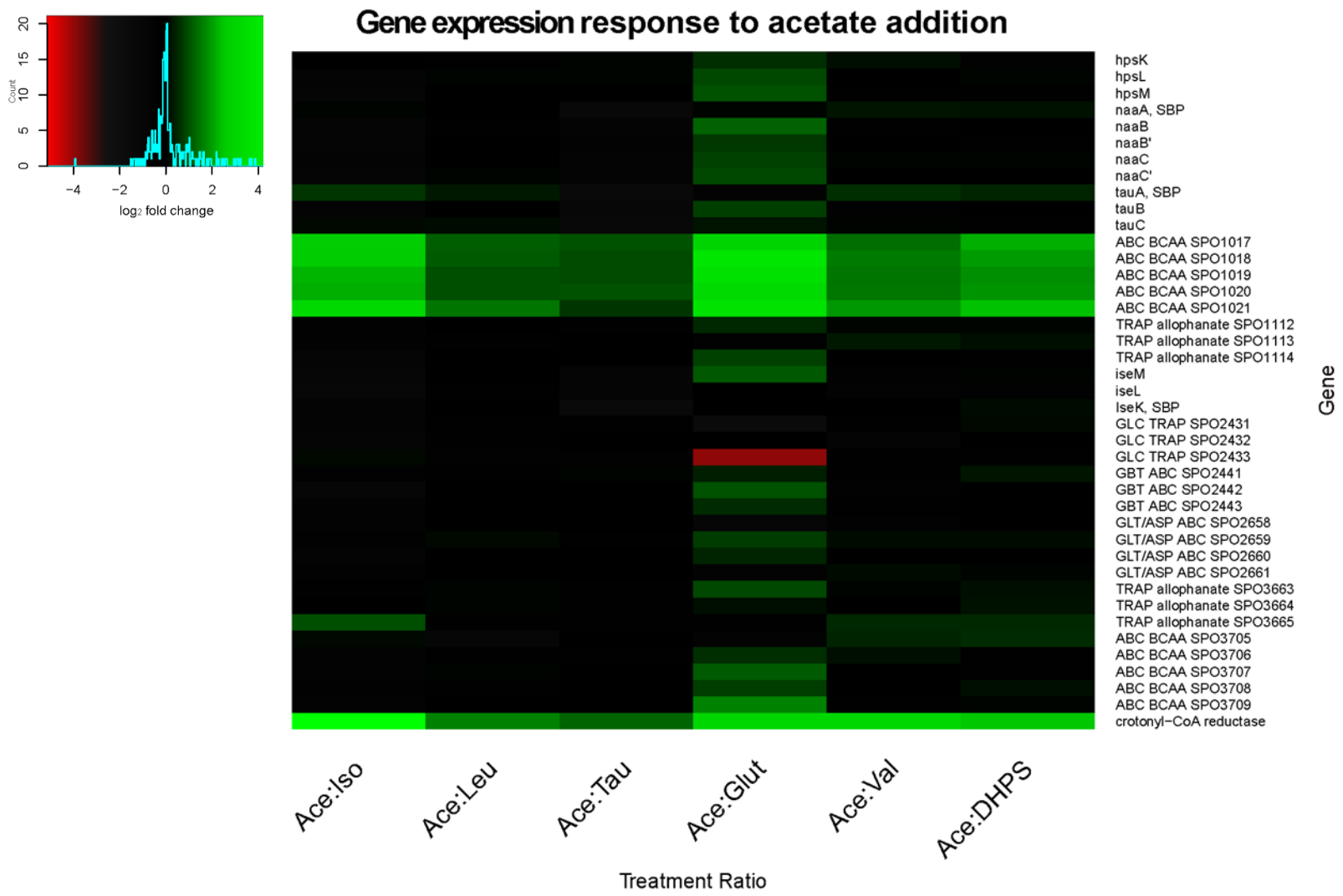


*Figure 2.1* Heat map of expression of transporter operons and crotonyl-CoA reductase (*ccrA*), the diagnostic gene of the ethylmalonyl-CoA pathway. Each row shows the ratio of expression during growth on acetate relative to one of the other six substrates.

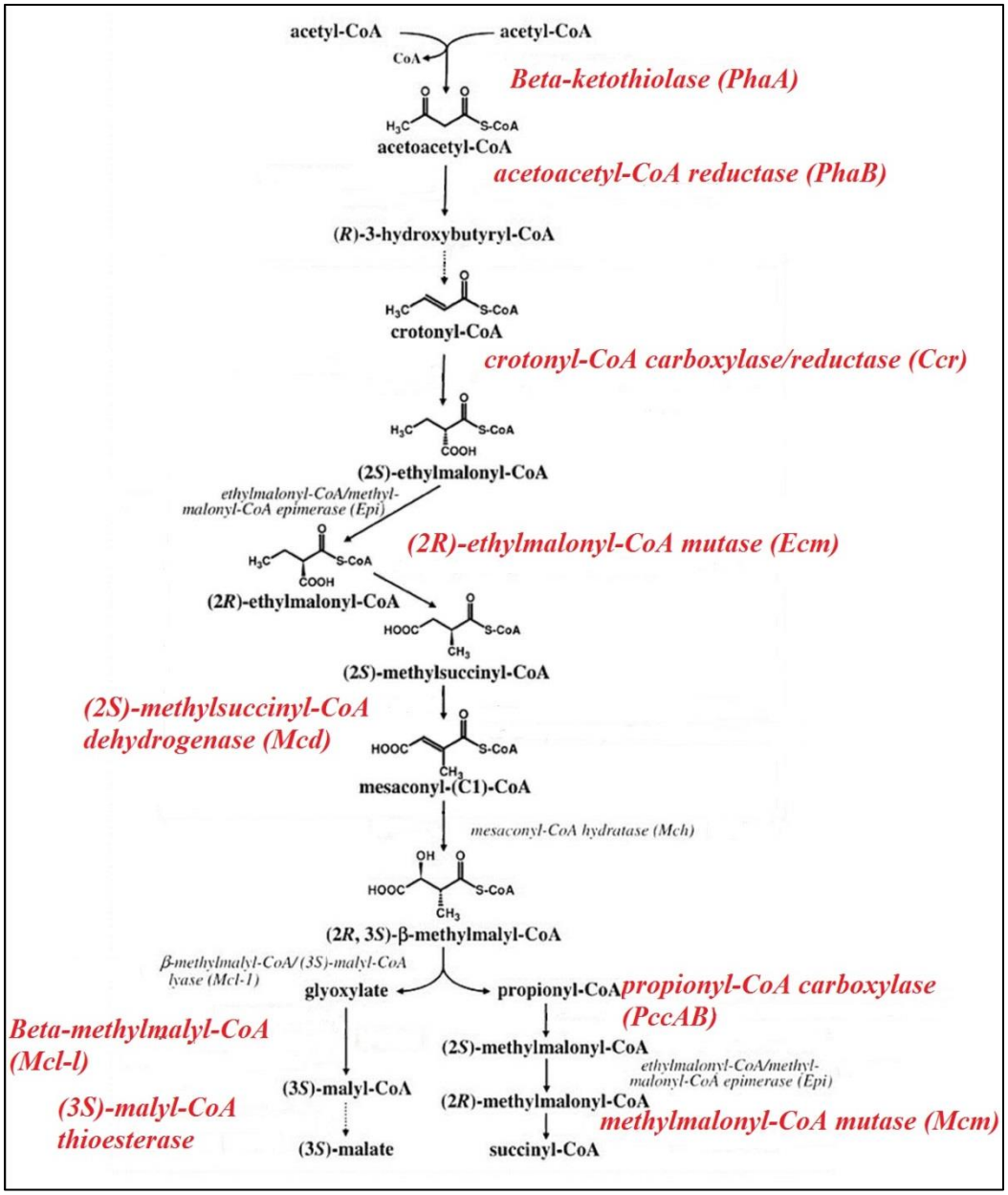


*Figure 2.2* Reactions of the ethylmalonyl-CoA pathway (adapted from Erb, *et al.* 2007).

Enzymes in red font indicate upregulation in the acetate addition in this study.

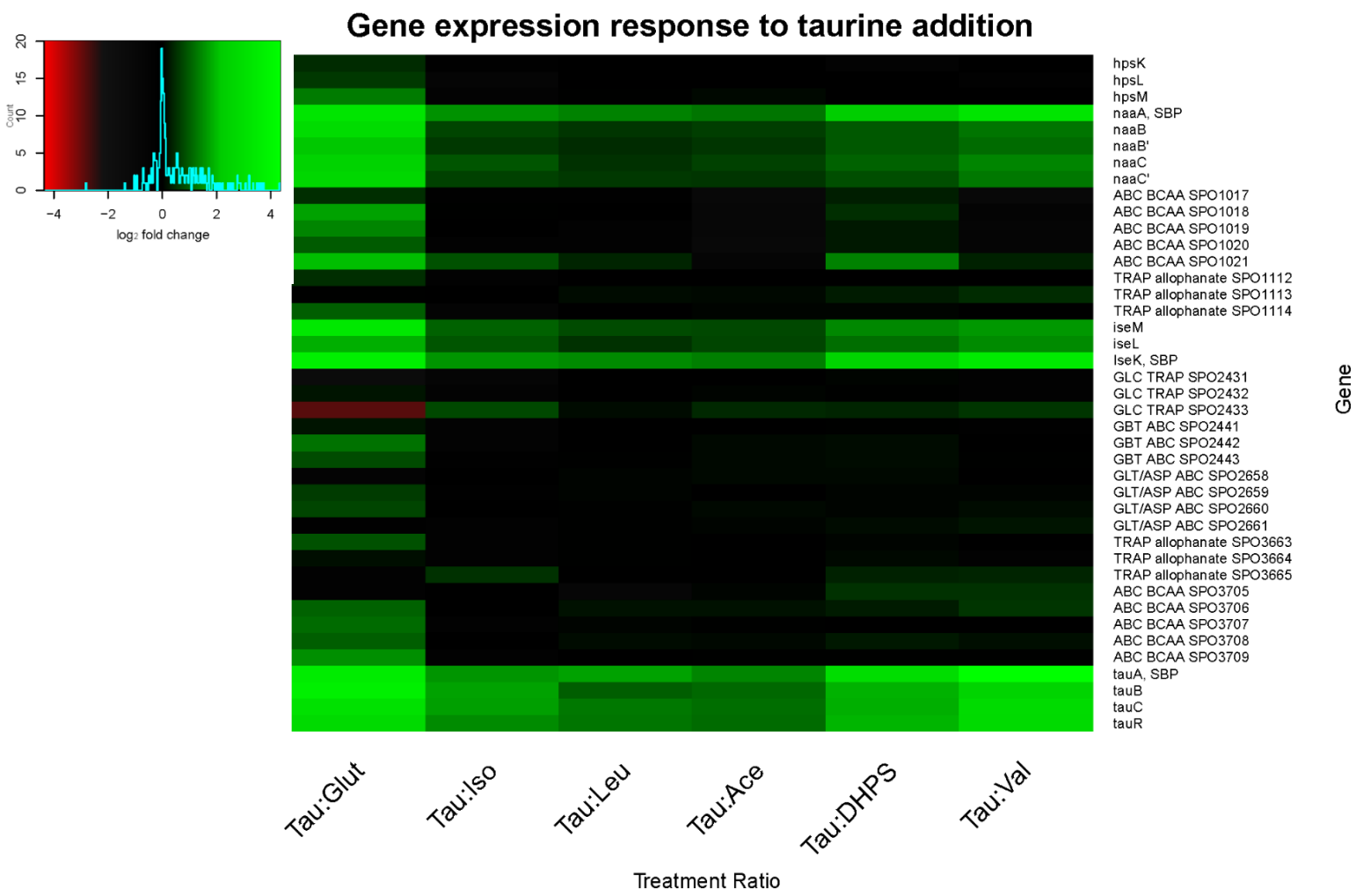


*Figure 2.3* Heat map of expression of the isethionate TRAP transporter operon (*iseKLM*), taurine ABC transporter (*tauABC*), and taurine transcriptional regulator (*tauR*). Each row shows the ratio of expression during growth on taurine relative to one of the six substrates.

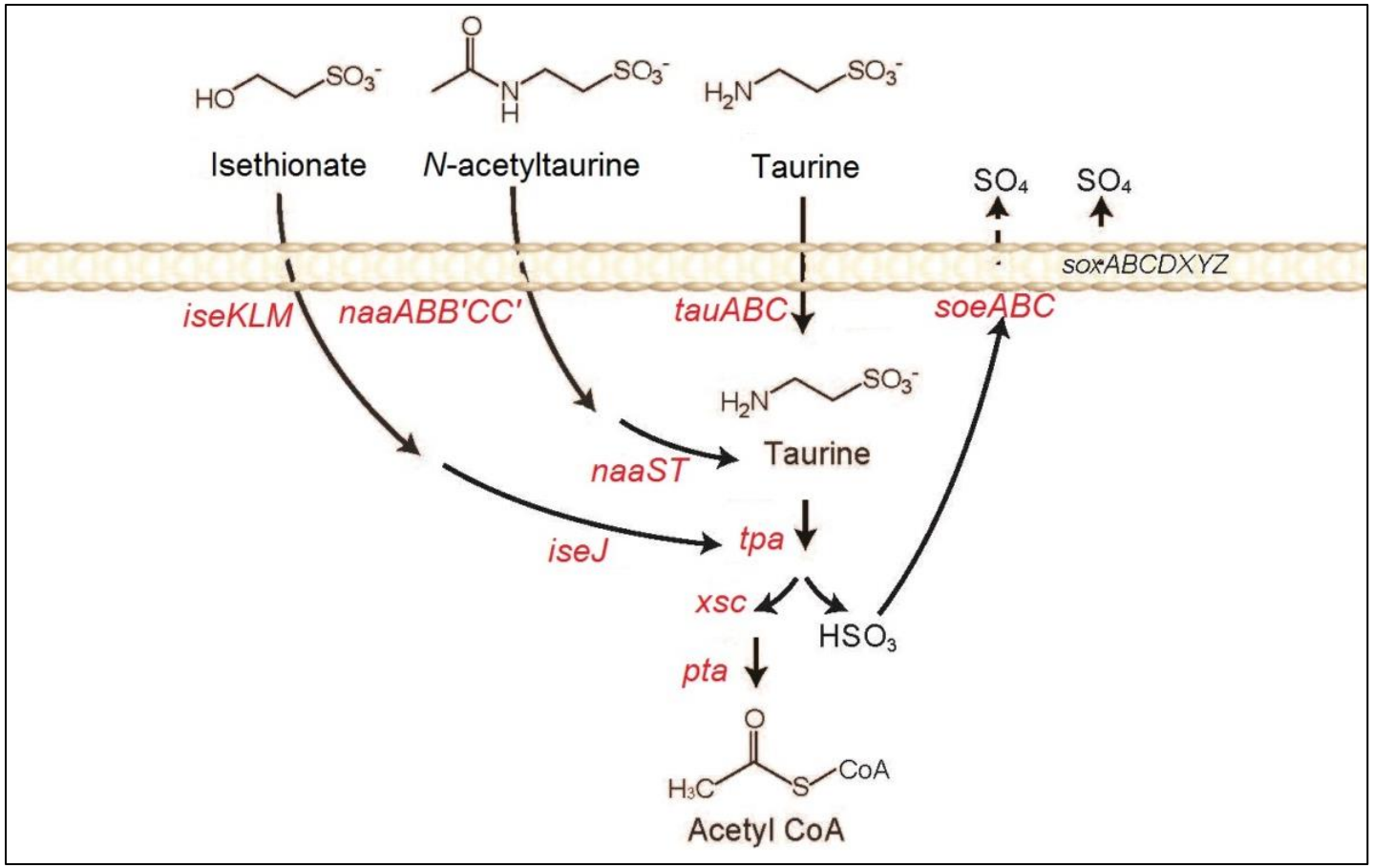


*Figure 2.4* Select sulfonate degradation pathways found in the *R. pomeroyi* genome. Genes in red indicate upregulation following taurine addition.

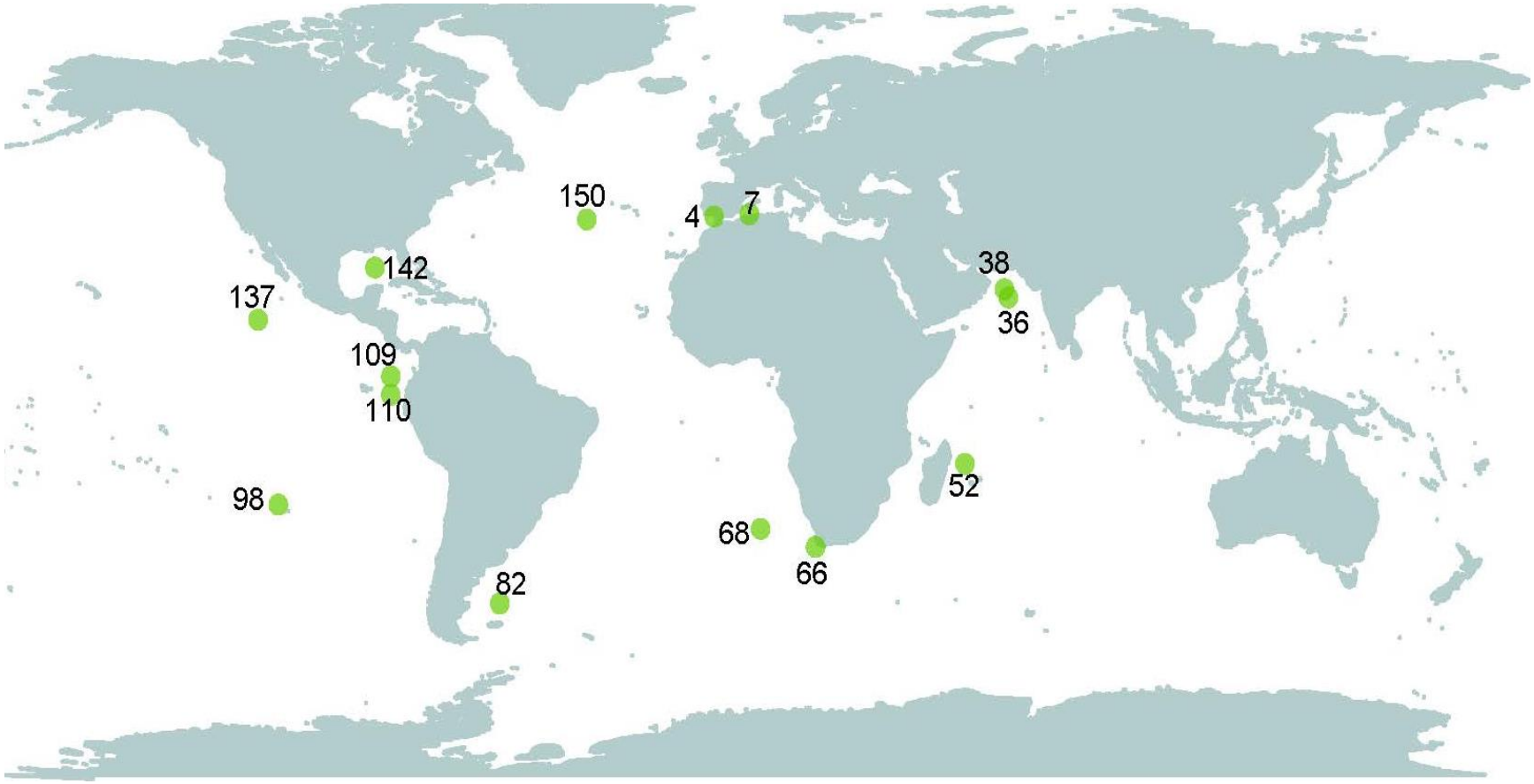


*Figure 2.5* Location of Tara Ocean stations analyzed. Station numbers are as assigned in Sunagawa *et al.*, 2015 companion website.

## References


Ameryk A, Hahnke R, Gromisz S, Kownacka J, Zalewski M, et al. 2014. Bacterial community structure influenced by *Coscinodiscus sp.* in the Vistula river plume. Oceanologia 56: 825-56

Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31: 166-9

Boyd J, Woodcroft B, Tyson G. 2015. GraftM. http://geronimp.github.io/graftM/

Buchan A, Gonzalez JM, Moran MA. 2005. Overview of the marine roseobacter lineage. *Appl Environ Microbiol* 71: 5665-77

Buchan A, LeCleir GR, Gulvik CA, Gonzalez JM. 2014. Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nat Rev Microbiol* 12: 686-98

Cram J, Chow TC, Sachdeva R, Needham DM, Parada AE, et al. 2015. Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *ISME J* 9: 563-80

Davidson AL, Dassa E, Orelle C, Chen J. 2008. Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev* 72: 317-64

Denger K, Lehmann S, Cook AM. 2011. Molecular genetics and biochemistry of *N*-acetyltaurine degradation by *Cupriavidus necator* H16. *Microbiology* 157: 2983-91

Durham BP, Dearth SP, Sharma S, Amin SA, Smith CB, et al. 2017. Recognition cascade and metabolite transfer in a marine bacteria-phytoplankton model system. *Environ Microbiol* 19: 3500-1

Durham BP, Sharma S, Luo H, Smith CB, Amin SA, et al. 2015. Cryptic carbon and sulfur cycling between surface ocean plankton. *Proc Natl Acad Sci U S A* 112: 453-7

Erb TJ, Berg IA, Brecht V, Muller M, Fuchs G, Alber BE. 2007. Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: The ethylmalonyl-CoA pathway. *Proc Natl Acad Sci* 104: 10631-36

Filosi M, Visintainer R, Albanese D, Ricadonna S, Jurman G, Furlanello C. 2017. Maximal information-based nonparametric exploration for variable analysis. ed. M Filosi

Fraiber M, Afanzar O, Cassidy CK, Gabashvili A, Schulten K, et al. 2015. CheY's acetylation sites responsible for generating clockwise flagellar rotation in Escherichia coli. *Mol Microbiol* 95: 231-44

Gonzalez JM, Covert JS, Whitman WB, Henriksen JR, Mayer F, et al. 2003. *Silicibacter pomeroyi* sp. nov. and *Roseovarius nubinhibens* sp. nov., dimethylsulfoniopropionate-demethylating bacteria from marine environments. *Int J Syst Evol Microbiol* 53: 1261-69

Gorzynska AK, Denger K, Cook AM, Smits TH. 2006. Inducible transcription of genes involved in taurine uptake and dissimilation by *Silicibacter pomeroyi* DSS-3. *Arch Microbiol* 185: 402-06

Howard EC, Henriksen JR, Buchan A, Reisch CR, Burgmann H, et al. 2006. Bacterial taxa that limit sulfur flux from the ocean. *Science* 314: 649-52

Krejcik Z, Hollemeyer K, Smits TH, Cook AM. 2010. Isethionate formation from taurine in *Chromohalobacter salexigens*: purification of sulfoacetaldehyde reductase. *Microbiology* 156: 1547-55

Landa M, Burns AS, Roth SJ, Moran MA. 2017. Bacterial transcriptome remodeling during sequential co-culture with a marine dinoflagellate and diatom. *ISME J*. In press.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9: 357-9

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15: 550

Luo H, Moran MA. 2014. Evolutionary ecology of the marine Roseobacter clade. *Microbiol Mol Biol Rev* 78: 573-87

Maqbool A, Horler RS, Muller A, Wilkinson AJ, Wilson KS, Thomas GH. 2015. The substrate-binding protein in bacterial ABC transporters: dissecting roles in the evolution of substrate specificity. *Biochem Soc Trans* 43: 1011-7

Matsen FA, Kodner RB, Armbrust EV. 2010. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* 11: 538

Mayer J, Huhn T, Habeck M, Denger K, Hollemeyer K, Cook AM. 2010. 2,3-Dihydroxypropane-1-sulfonate degraded by *Cupriavidus pinatubonensis* JMP134: purification of dihydroxypropanesulfonate 3-dehydrogenase. *Microbiology* 156: 1556-64

Milici M, Vital M, Tomasch J, Badewien TH, Giebel HA, et al. 2017. Diversity and community composition of particle-associated and free-living bacteria in mesopelagic and bathypelagic Southern Ocean water masses: Evidence of dispersal limitation in the Bransfield Strait. *Limnol Oceanogr* 62: 1080-95

Miller TR, Hnilicka K, Dziedzic A, Desplats P, Belas R. 2004. Chemotaxis of *Silicibacter* sp. strain TM1040 toward dinoflagellate products. *Appl Environ Microbiol* 70: 4692-701

Moran MA, Belas R, Schell MA, Gonzalez JM, Sun F, et al. 2007. Ecological genomics of marine Roseobacters. *Appl Environ Microbiol* 73: 4559-69

Moran MA, Buchan A, Gonzalez JM, Heidelberg JF, Whitman WB, et al. 2004. Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature* 432: 910-3

Moran MA, Kujawinski EB, Stubbins A, Fatland R, Aluwihare LI, et al. 2016. Deciphering ocean carbon in a changing world. *Proc Natl Acad Sci U S A* 113: 3143-51

Morris RM, Rappe MS, Connon SA, Vergin KL, Sieblod WA, et al. 2002. SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* 420: 806-10

Mulligan C, Fischer M, Thomas GH. 2011. Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. *FEMS Microbiol Rev* 35: 68-86

Mulligan C, Kelly DJ, Thomas GH. 2007. Tripartite ATP-independent periplasmic transporters: application of a relational database for genome-wide analysis of transporter gene frequency and organization. *J Mol Microbiol Biotechnol* 12: 218-26

Newton RJ, Griffin LE, Bowles KM, Meile C, Gifford S, et al. 2010. Genome characteristics of a generalist marine bacterial lineage. *ISME J* 4: 784-98

Oh M, Rohlin L, Kao KC, Liao JC. 2002. Global expression profiling of acetate-grown *Escherichia coli*. *J Biol Chem* 277: 13175–83

Pinhassi J, Sala MM, Havskum H, Peters F, Guadayol O, et al. 2004. Changes in bacterioplankton composition under different phytoplankton regimens. *Appl Environ Microbiol* 70: 6753-66

Pomeroy LR, Williams PJ, Azam F, Hobbie JE. 2007. The microbial loop. *Oceanography* 20: 28-33

Ramakrishnan R, Schuster M, Bourret RB. 1998. Acetylation at Lys-92 enhances signaling by the chemotaxis response regulator protein CheY. *Proc Natl Acad Sci* 95: 4918-23

Rivers AR, Smith CB, Moran MA. 2014. An updated genome annotation for the model marine bacterium *Ruegeria pomeroyi* DSS-3. *Stand Genomic Sci*

Schulz A, Stoveken N, Binzen IM, Hoffman T, Heider J, Bremer E. 2017. Feeding on compatible solutes: A substrate-induced pathway for uptake and catabolism of ectoines and its genetic control by EnuR. *Environ Microbiol* 19: 926-46

Simu K, Hagstrom A. 2004. Oligotrophic bacterioplankton with a novel single-cell life strategy. *Appl Environ Microbiol* 70: 2445-51

Sowell SM, Abraham PE, Shah M, Verberkmoes NC, Smith DP, et al. 2011. Environmental proteomics of microbial plankton in a highly productive coastal upwelling system. *ISME J* 5: 856-65

Sowell SM, Wilheml LJ, Norbeck AD, Lipton MS, Nicora CD, et al. 2009. Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J* 3: 93-105

Studer SV, Mandel MJ, Ruby EG. 2008. AinS quorum sensing regulates the *Vibrio fischeri* acetate switch. *J Bacteriol*: 5915-23

Styp von Rekowski K, Denger K, Cook AM. 2005. Isethionate as a product from taurine during nitrogen-limited growth of *Klebsiella oxytoca* TauN1. *Arch Microbiol* 183: 325-30

Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, et al. 2015. Structure and function of the global ocean microbiome. *Science* 348: 1261359-59

Wagner-Dobler I, Biebl H. 2006. Environmental biology of the marine Roseobacter lineage. *Annu Rev Microbiol* 60: 255-80

Weinitschke S, Sharma PI, Stingl U, Cook AM, Smits TH. 2010. Gene clusters involved in isethionate degradation by terrestrial and marine bacteria. *Appl Environ Microbiol* 76: 618-21

Wickham H, Chang W. 2016. Create elegant data visualisations using the grammar of graphics.

Williams TJ, Long E, Evans F, DeMaere MZ, Lauro FM, et al. 2012. A metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal surface waters. *ISME J* 6: 1883-900

Wolfe AJ. 2005. The acetate switch. *Microbiol Mol Biol Rev* 69: 12-50

Wolfe AJ. 2008. Quorum sensing "flips" the acetate switch. *Journal of Bacteriology* 190

# CHAPTER 3

# CONCLUSION

Interactions between marine bacteria and phytoplankton are fundamental to the food web that supports marine organisms. Additionally, phytoplankton and the metabolites they produce influence the flux of carbon and all other major elements in the ocean, such as nitrogen, phosphorous, and sulfur. Traditional chemical methods struggle to identify the specific metabolites being produced by phytoplankton and processed by the heterotrophic bacteria that are invariably associated with them. The goal of this research was to take an alternative approach to the identification of phytoplankton-produced labile compounds using bacterial transporters as a biosensor. To do this, the transporter expression of a model heterotrophic bacterium was studied after amendment with a potential target substrate following early stationary phase growth on glucose. Operons that were previously shown to respond to the metabolites produced by phytoplankton were analyzed. In addition to investigating the transcriptional response of a model bacterium, we mined the Tara Ocean metagenomic dataset for orthologs to the substrate binding proteins in the operons of interest. Understanding the global prevalence of the gene encoding these proteins provides a broader view of the ecological importance of the transporter systems. Environmental parameters and taxonomic profiles were analyzed with the gene data to look for significant relationships.

The transcriptomic analysis indicated a significant upregulation in expression when *R. pomeroyi* was amended with taurine and acetate, thus generating hypotheses regarding the substrates that trigger two of the seven transporter operons of interest. The acetate

addition resulted in the upregulation of an operon previously annotated as a branched chain amino acid transporter (SPO1017-1021) and of the diagnostic gene for the ethymalonyl-CoA pathway. The transcriptional profile matched a metabolic response termed the “acetate switch”, which has been linked to quorum sensing in *Vibrio fischeri* and chemotaxis in *E. coli* (Studer *et al.,* 2008; Oh *et al.,* 2002; Fraiber *et al.,* 2015). The addition of taurine to the medium resulted in the upregulation of three transporters: the taurine transporter (*tauABC*), the isethionate transporter (*iseKLM*), and the *N*-acetyltaurine transporter (*naaABB’CC*’), which are responsible for importing C2 sulfonates into the cell. These three substrates share downstream steps in sulfonate catabolism pathway in *R. pomeroyi* (Mayer *et al.,* 2010; Gorzynska *et al.,* 2006; Denger *et al.,* 2011), and the genes involved in the shared pathway were also upregulated. The other transporter operons did not exhibit a clear response to the addition of the added substrates. Cells may not have been sampled at the precise time to catch a response to the substrate addition, or the concentration of substrates (100 µM or 200 µM) may not have been high enough to elicit a detectable response. No information was gathered regarding the other five uncharacterized transporter systems.

In addition to characterizing the expressional response of *R. pomeroyi* transporter operons that were upregulated in a previous phytoplankton co-culture experiment (Durham *et al.,* 2017), we investigated the ecological significance of these genes in the global ocean. Nine SBPs were used in a transporter identification pipeline that targeted orthologous proteins at 14 stations from the Tara Oceans sampling project. Of the 4 stations currently analyzed for 5 of the SBPs, the MINE program as part of the minerva package in R identified several relationships between the percent of cells carrying a gene and taxonomic profiles. The results showed a strong relationship between *hpsk* and SPO2441 and the SAR11 families Surface 1 and Chesapeake-Delaware Bay,

as well as between the SBP for *N*-acetyltaurine (SPO0660) and SAR11 family Surface 4. A search of representative genomes in these groups, however, did not result in any potential orthologous proteins.

Two other SBPs annotated as transporters of allophanate (SPO3665) and galactarate (SPO2433) also had a high MIC scores associating them with bacterial families. The SBP for allophanate was linearly correlated with Alphaproteobacteria family *Rhodospiralles*, and orthologous genes to SPO3665 were found in representative genomes of this family. The galactarate SBP had a relationship to three bacterial families; orthologs SPO2433 were found in members of the *Comamonadaceae* family but not the other two. These results suggest that the SBPs that respond in our model bacterium to phytoplankton metabolites are associated with a diverse set of bacteria.

## References


Denger K, Lehmann S, Cook AM. 2011. Molecular genetics and biochemistry of *N*-acetyltaurine degradation by *Cupriavidus necator* H16. *Microbiology* 157: 2983-91

Durham BP, Dearth SP, Sharma S, Amin SA, Smith CB, et al. 2017. Recognition cascade and metabolite transfer in a marine bacteria-phytoplankton model system. *Environ Microbiol* 19: 3500-13

Fraiber M, Afanzar O, Cassidy CK, Gabashvili A, Schulten K, et al. 2015. CheY's acetylation sites responsible for generating clockwise flagellar rotation in *Escherichia coli*. *Mol Microbiol* 95: 231-44

Gorzynska AK, Denger K, Cook AM, Smits TH. 2006. Inducible transcription of genes involved in taurine uptake and dissimilation by *Silicibacter pomeroyi* DSS-3. *Arch Microbiol* 185: 402-06

Mayer J, Huhn T, Habeck M, Denger K, Hollemeyer K, Cook AM. 2010. 2,3-Dihydroxypropane-1-sulfonate degraded by *Cupriavidus pinatubonensis* JMP134: purification of dihydroxypropanesulfonate 3-dehydrogenase. *Microbiology* 156: 1556-64

Oh M, Rohlin L, Kao KC, Liao JC. 2002. Global expression profiling of acetate-grown *Escherichia coli*. *J Biol Chem* 277: 13175–83

Studer SV, Mandel MJ, Ruby EG. 2008. AinS quorum sensing regulates the *Vibrio fischeri* acetate switch. *J Bacteriol*: 5915-23