A multi-omics perspective on microbial mediated methane oxidation in the Saanich Inlet water column

by

Mónica Torres Beltrán

B.Sc. Oceanography, Universidad Autónoma de Baja California, 2007

M.Sc. Coastal Oceanography, Universidad Autónoma de Baja California, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Microbiology and Immunology)

The University of British Columbia

(Vancouver)

November 2018

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

A multi-omics perspective on microbial mediated methane oxidation in the Saanich Inlet water column

submitted by **Mónica Torres Beltrán** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** in **Microbiology and Immunology**.

Examining Committee:

Steven J. Hallam, Microbiology and Immunology
Supervisor

Maria T. Maldonado
University Examiner

Kristin Orians
University Examiner

Samantha B. Joye
External Examiner

Additional Supervisory Committee Members:

Thomas J. Beatty, Microbiology and Immunology
Supervisory Committee Member

William W. Mohn,Microbiology and Immunology
Supervisory Committee Member

Martin Hirst, Microbiology and Immunology
Supervisory Committee Member

Abstract

Microbial communities play an integral role in the biogeochemical cycling of carbon, nitrogen and sulfur throughout the biosphere. These communities interact, forming metabolic networks that change and adapt in response to availability of electron donors and acceptors. Oxygen minimum zones (OMZs) are regions of the ocean where oxygen (O_2) is naturally depleted (<20 μM). In OMZs microbial communities use alternative terminal electron acceptors such as nitrate, sulfate and carbon dioxide, resulting in fixed nitrogen loss and production of greenhouse gases including methane (CH_4). In this thesis, I explored microbial community structure, dynamics and metabolic interactions as they relate to CH_4 cycling in Saanich Inlet, a seasonally anoxic fjord on the coast of British Columbia Canada that serves as a model ecosystem for studying microbial processes in OMZs. Leveraging decadal time series observations in Saanich Inlet, I developed a geochemical dataset consisting of nutrient and gas measurements, coupled with multi-omic (DNA, RNA and protein) sequence information to chart microbial community structure and dynamics along defined redox gradients. I conducted methods optimization comparing *in situ* and on-ship sampling paradigms and used correlation analysis to infer putative microbial interaction networks in relation to water column CH_4 oxidation. Methanotrophic bacteria in Saanich Inlet were identified associated with three uncultivated Gammaproteobacteria clades termed OPU1, OPU3 and symbiont-related that partitioned in the water column during periods of prolonged stratification. Water column distribution of the OPU3 clade was found to correlate with nitrite (NO_2^-). Based on these results, I conducted incubations with labelled CH_4 and NO_2- to test this correlation and constrain potential metabolic interactions between methanotrophs and other one-carbon utilizing microorganisms under low O_2 conditions. Using multi-omic information derived from these incubations I confirmed the role of OPU3 in coupling CH_4 oxidation to NO_2^- reduction and uncovered potential metabolic interactions between OPU3 and other co-occurring microorganisms including Methylophilales, Planctomycetes and Bacteroidetes. Evidence for a communal function in CH_4 oxidation expands the role of OPU3 in the global carbon budget and provides a conceptual foundation for the development of numerical models to predict CH_4 flux from OMZs as they expand throughout the global ocean.

Lay Summary

In nature, microbes do not grow in isolation, they form communities of interacting cells, resulting in distributed networks that drive nutrient and energy cycling. Oxygen plays a crucial role in the development of these networks. Marine oxygen minimum zones (OMZs) are currently expanding and intensifying due to climate change. In OMZs, the use of alternative electron acceptors by microbes results in the production of greenhouse gases such as methane (CH_4). In this thesis, I explored the structure and function of microbial communities in relation to CH_4 cycling processes in a model ecosystem, Saanich Inlet, to study microbial processes in coastal and open ocean OMZs. Coupling multi-molecular (DNA, RNA and protein) sequence information and geochemical parameter information, I identified microbial networks driving CH_4 consumption under low oxygen conditions. Understanding how this greenhouse gas is cycled in expanding OMZs will help us to better predict climate change effects at a global scale.

Preface

The work presented in this thesis was made possible by support of collaborators, former and current member of the Hallam laboratory, and contractors as described in the following paragraphs.

Chapter 2: The Saanich Inlet geochemical datasets are the results of the tireless effort of many post-doctoral fellows, students, technicians, and volunteers who think big about the microbial world. In chapter 2, I presented a compendium of time-series observations encompassing historical oxygen (O_2) measurements and more recent monthly (2006-2014) monitoring efforts, representing over 100 independent sampling expeditions. This chemical compendium partners with a compendium of multi-omic sequence information from the Saanich Inlet water column detailing time-series microbial multi-omic datasets. Combined, these compendiums provide a community-driven framework for observing and predicting microbial community responses to changing levels of O_2 deficiency extensible to open ocean OMZs.

Dr. Steven Hallam designed and supervised the Saanich Inlet time-series project. Environmental chemical and multi-omic data collection was carried out aboard the *RSV Strickland*. Sample collection was conducted with the valuable help and support of Captain Ken Brown and his crew. Sample collection for the generation of chemical, physical and multi-omic datasets was possible thanks to the extensive logistical support and planning of several Chief Scientists. David Walsh, Elena Zaikova, Olena Shevchuk, Craig Mewis, Alyse K. Hawley, and I held the Chief Scientist position. Sea going technicians Chris Payne and Larysa Pakhomova were keystones for sample collection, and CTD operation and calibration. Alyse K. Hawley and David Capelle carried out dissolved gas measurements and quality control under the supervision of Dr. Philippe Tortell.

Multi-omic datasets were the result of many hours of bench work carried out by many collaborators, technicians, students, and volunteers. Alyse K. Hawley, Melanie Scofield, Sam Kheirandish, Andreas Mueller, Payal Sipahimalani, Olena Shevchuk, and I generated metagenomic and small subunit tag datasets. Sequencing was carried out at the Joint Genome Institute (JGI) and Génome Québec Innovation Centre at the McGill University. Metatranscriptomic datasets were generated following the protocol designed by Alyse K. Hawley. Alyse K. Hawley and I carried out RNA extractions. Sequencing was carried out at the JGI. Alyse K. Hawley and Heather Brewer at the Environmental Molecular Science Laboratory (EMSL) and the Pacific Northwest National Laboratory (PNNL) conducted Metaproteomic data including extractions and protocol design.

Portions of text, protocols and figures in chapter 2 were published in Nature Scientific Data as: **Torres-Beltrán, M.**, Hawley, A.K. *et al*. 2017. A compendium of geochemical information from the Saanich Inlet water column. Sci.Data. 4. doi:10.1038/sdata.2017.159 Additional manuscript fully detailing multi-omic

sequence data protocols was published in Nature Scientific Data as: Hawley, A.K., **Torres-Beltrán, M.** *et al*. 2017. A compendium of multi-omic sequence information from the Saanich Inlet water column. Sci.Data. 4. doi:10.1038/sdata.2017.160.

Chapter 3: I used small subunit ribosomal RNA (SSU rRNA) gene 454 sequencing data to compare and cross-calibrate *in situ* sampling devices such as McLane PPS with conventional bottle sampling methods by testing bottle effects on microbial community composition, and potential activity when using different filter combinations and filtration methods. I used SSU rDNA and rRNA 454 pyrosequencing data generated during the SCOR Working Group "Microbial Community Responses to Ocean Deoxygenation" workshop held in Vacnouver, B.C in July 2014. Collection of samples was carried out aboard the *RSV Strickland* operated by Captain Ken Brown. Chris Payne and Larysa Pakhomova conducted CTD deployment and data processing. Dr. Virginia Edgcomb, Dr. Maria Pachiadaki, Andreas Mueller, Melanie Scofield, and I collected and processed DNA and RNA samples on ship. Dr. Craig Taylor supervised sample collection from *in situ* filtering systems. Andreas Mueller, Melanie Scofield and I extracted DNA and RNA samples. Kateryna Tyshchenko and I constructed cDNA libraries and performed PCR reactions of samples to be submitted for 454 sequencing at Génome Québec Innovation Centre at the McGill University. I conducted the downstream analysis of sequences using QIIME, and all statistical analyses using R scripts I wrote. I generated all figures, and wrote the chapter under the supervision of Dr. Steven Hallam.A version of this chapter will be submitted to a peer reviewed journal.

Chapter 4: I used geochemical information and SSU rRNA gene 454 sequencing data to address microbial eukaryote community dynamics over a 12-month period, uncovering significant correlations between parasitic dinoflagellates within Syndiniales and other eukaryotic taxa during months of peak water column stratification. I used methods in this chapter to gain insight into the use of analytical and statistical tools that later I used to asses the methanotrophic community composition, dynamics, and metabolism as detailed in chapters 5 and 6. I used 454 sequencing data of the SSU rRNA gene generated from 159 DNA samples collected in the Saanich Inlet between May 2008 and April 2009. Collection of samples was carried out aboard the *RSV Strickland* operated by Captain Ken Brown. Chris Payne and Larysa Pakhomova conducted CTD deployment and data processing. Past members of the Hallam laboratory extracted DNA samples until April 2009. I extracted DNA and RNA samples from July 2014. With assistance of Melanie Scofield and Andreas Mueller, and Kateryna Tyshchenko, I performed PCR reactions of samples to be submitted for 454 sequencing at Génome Québec Innovation Centre at the McGill University. I conducted the downstream analysis of sequences using the QIIME software, and correlation analyses using CoNet. Taylor Sehein from the Edgcomb laboratory at Woods Hole Oceanographic Institution and Dr. Maria Pachiadaki from the Bigelow Laboratory contributed greatly on data interpretation key for the chapter development. Taylor Sehein and I generated all figures. Dr. Virginia Edgcomb from Woods Hole Oceanographic Institution closely supervised and provided constructive feedback on data analysis and manuscript production. Dr. Steven Hallam provided constructive feedback on chapter and manuscript production and presentation.

Text, protocols, and figures presented in this chapter were published as: **Torres-Beltrán, M**., Sehein, T. *et al*. 2018. Protistan parasites along oxygen gradients in a seasonally anoxic fjord: a network approach to assessing potential host-parasite interactions. Deep Sea Res. II. doi.org/10.1016/j.dsr2.2017.12.026

Chapter 5: I used SSU rRNA gene 454 sequencing data generated from 288 DNA samples collected in the Saanich Inlet between May 2008 and July 2010. I surveyed the microbial community along seasonal redox gradients and focused on identifying microbial agents driving methane (CH_4) oxidation. Collection of samples was carried out aboard the *RSV Strickland* operated by Captain Ken Brown. Chris Payne and Larysa Pakhomova conducted CTD deployment and data processing. Past members of the Hallam laboratory extracted DNA samples collected until April 2009. I extracted DNA samples from May 2009 to July 2010. With assistance of Melanie Scofield and Andreas Mueller, I performed PCR reactions of samples to be submitted for 454 sequencing at Génome Québec Innovation Centre at the McGill University. I conducted the downstream analysis of sequences using the QIIME software, and conducted further phylogenetic assignment of methanotrophic taxa. Claire Stilwell constructed the particulate methane monooxygenase subunit β (*pmoA*) gene libraries. Alyse K. Hawley extracted the DNA and RNA samples to generate the metagenomic and metatranscriptomic datasets to be sequenced at JGI on the Illumna MiSeq platform. Maya Bhatia processed the multi-omics datasets through MetaPathways, designed and built by Niels Hanson and Kishori Konwar. David Capelle carried out dissolved CH_4 measurements and quality control. Evan Durno and I conducted the correlation analyses using CoNet and R scripts (written by Evan Durno). For the published article I generated all figures and wrote the manuscript under the supervision of Dr. Steven Hallam.

Text, protocols, and figures presented in chapter 5 were published as: **Torres-Beltrán, M.** *et al*. 2016. Methanotrophic community dynamics in a seasonally anoxic fjord: Saanich Inlet, British Columbia. Front. Mar.Sci. doi.org/10.3389/fmars.2016.00268.

Chapter 6: I used multi-omic sequence information (DNA, RNA and protein) derived from CH_4 incubation experiments on Saanich Inlet O_2 deficient waters to identify community-level interactions related to CH_4 oxidation. David Capelle and I collected DNA, RNA and protein samples aboard the *RSV Strickland* operated by Captain Ken Brown. Chris Payne and Larysa Pakhomova conducted CTD deployment and data processing. I extracted DNA samples and performed PCR reactions of samples to be submitted for 454 sequencing at Génome Québec Innovation Centre at the McGill University. I extracted RNA samples at EMSL under the supervision of Lye Meng Markillie. Lye Meng Markillie and Dr. Hugh Mitchell conducted the RNA sequencing on the Ion Torrent platform and the initial quality control and processing of sequences at EMSL. Kateryna Tyshchenko and I constructed cDNA libraries on remaining RNA and performed PCR reactions of samples to be submitted for 454 sequencing at Génome Québec Innovation Centre at the McGill University. I extracted protein samples at EMSL with the help and supervision of Heather Brewer. Generation of peptide spectra was done by Heather Brewer, and spectra mapped to protein database by Angela Norbeck and Samuel Purvine at EMSL. Dr. John Cliff and I conducted nanoSIMS quantification, and Dr. Jim Moran conducted bulk isotopic incorporation. Connor Morgan-Lang assembled the metagenomic datasets. I conducted the downstream analysis of sequences using the QIIME software and metagenomic analysis using MetaPathways, designed by Niels Hanson and Kishori Konwar. I conducted statistical analyses using R scripts that I wrote. I generated all figures, and wrote the chapter under the supervision of Dr. Steven Hallam. A version of this chapter will be submitted to a peer reviewed journal.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AOM	Anaerobic oxidation of methane
BLAST	Basic local alignment search tool
CTD	Conductivity, temperature and depth measuring instrument
CO_2	Carbon dioxide
CH_4	Methane
CH_3OH	Methanol
CH_2O	Formaldehyde
$CHOO^-$	Formate
DNA	Deoxyribonucleic acid
cDNA	Complementary DNA
rDNA	Ribosomal DNA
EMSL	Environmental Molecular Sciences Laboratory
ETNP	Eastern Tropical North Pacific
ETSP	Eastern Tropical South Pacific
H_2S	Hydrogen sulfide
H_4SiO_4	Silicic acid
ISA	Indicator species analysis
JGI	Joint Genome Institute
$MgCl_2$	Magnesium chloride
MALV	Marine alveolate
MMO	Methane monooxygenase
MPP	Masterflex peristaltic pump
N_2	Dinitrogen

n-damo	Nitrite dependent anaerobic methane oxidation
NH_4^+	Ammonium
NO_3^-	Nitrate
NO_2^-	Nitrite
N_2O	Nitrous oxide
O_2	Oxygen
OMZs	Oxygen minimum zones
OPU	Particulate methane monooxygenase encoding phylogenetic group
OTU	Operational taxonomic unit
PAR	Photosynthetically active radiation
PCR	Polymerase chain reaction
PNNL	Pacific Northwest National Laboratory
PMO	Particulate methane monooxygenase
PO_4^{3-}	Orthophosphate
PPS	Phytoplankton sampler
QIIME	Quantitative Insights Into Microbial Ecology
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
RuMP	Ribulose monophosphate
SCOR	Scientific Committee on Oceanographic Research
sMMO	Soluble methane monooxygenase
SSU	Ribosomal small subunit
TEA	Terminal electron acceptor

Acknowledgments

First of all I would like to thank my supervisor Dr. Steven Hallam for his support and for always pushing me to do better. Thank you Dr. Hallam for your patience and generosity on life matters. I would like to thank my committee members, Dr. Thomas J. Beatty, Dr. William W. Mohn and Dr. Martin Hirst, for all their support and input. I would like to thank past and present members of the Hallam lab for all the smiles. Thank you Alyse K. Hawley for your friendship and your continuous support both emotional and logistical. Thank you David Capelle,Maya Bhatia, Esther Gies, Elena Zaikova, Evan Durno, Aria Hahn, Ashley Arnold, Kateryna Levdokymenko for all your support and company over these years. Thank you Andreas Mueller and Melanie Scofield for being the best team in the world and remind me everything is awesome!. I would like to thank Chris Payne and Larysa Pakhomova for their help and support make possible our Saanich cruise. Thank you Captain Ken Brown and his crew for their engaged effort on every cruise aboard the RSV Strickland. Thank you to members of the Tortell, Crowe and Suttle Laboratories at UBC and members of the Varela Lab at UVic for logistical support.

I would like to thank to my collaborators at the Joint Genome Institute for sequencing so many samples, and at Environmental Sciences Laboratory at Pacific Northwest National Labs for your support in trasncriptomics, proteomics, and stable isotope analysis none of this work would have been possible without you. Thank you to Consejo Nacional de Ciencia y Tecnología (CONACyT) and the Tula Foundation for the funding that made possible my degree.

Muchas gracias mamita linda querida, gracias por todo tu esfuerzo y apoyo durante toda mi vida, gracias por ayudarme a siempre levantarme y seguir adelante. Este trabajo también es tuyo. Te amo.

Gracias a mis suegros por su incondicional apoyo en estos años.

Mi querida Citlali Márquez, infinitas gracias amiga mía por todo tu apoyo, por tu compañia y por las largas platicas que hicieron más fácil y llevaderos los años del doctorado. Gracias por hacerme sentir en familia. Gracias mi querida Natalie Millán por tu compañia sin importar la distancia, te quiero mucho amiga mia!.Gracias a mi querida familia, Ricardo Mendoza y Marta Reta, por estar presentes en los momentos más difíciles.

Mi amado Luis Malpica Cruz, mi Boo, mi papoy, sin ti esto simplemente no se hubiera logrado, literalmente. Gracias por haber sido mis manos en muchos instantes, gracias por tu paciencia y apoyo en estos años, gracias por regalarme perspectiva en momentos difíciles, gracias por darme la mano. Esta tesis, es tan mía como lo es tuya. Gracias por vivir esta etapa a mi lado y por haberla sobrevivido. Te amo con todo mi corazón. Eres mi mejor amigo y el amor de mi vida.

Dedication

Para Luis y Marina

En memoria de mi abuelo

Chapter 1

Introduction

1.1 Global ocean deoxygenation

The ocean is changing. In addition to the tons of plastic polluting the ocean and the continued decline of fish stocks worldwide, unseen beneath the waves along with increasing acidification, another problem arises: the loss of oxygen (O_2) from ocean waters. This O_2 loss, or deoxygenation, is one of the most important changes occurring in an ocean increasingly modified by human activities (Breitburg *et al*., 2018).

Greenhouse gas-driven global warming is likely the ultimate cause of this ongoing deoxygenation in many parts of the open ocean (Bopp *et al*., 2013). Decreased O_2 solubility due to a warming ocean is estimated to account for $\sim$15% of the total current global O_2 loss and $>$50% of the O_2 loss in the upper 1000 m of the ocean (Helm *et al*., 2011; Schmidtko *et al*., 2017; Breitburg *et al*., 2018). Higher temperatures also raise metabolic rates, thus accelerating O_2 consumption (Breitburg *et al*., 2018) and resulting in a spatial redistribution of available O_2 (Brewer *et al*., 2017). For instance, as O_2 levels decline, aerobic organisms shelter in more oxygenated waters, resulting in habitat compression accompanied by the diversion of energy into microbial metabolism (Diaz and Rosenberg, 2008; Wright *et al*., 2012). Intensified stratification may account for the remaining 85% of global ocean O_2 loss by reducing O_2 transport into the ocean interior and affecting the nutrient supply controlling the production of organic matter and its subsequent sinking from the ocean surface (Breitburg *et al*., 2018). Sinking organic particles, or marine snow, provide a nucleation point for otherwise suboxic or anoxic processes adding to energy diversion towards microbial metabolism in oxygenated waters (Wright *et al*., 2012).

As of 1960, global dissolved O_2 concentration observations show ongoing regional changes in oceanic O_2 (Whitney *et al*., 2007; Bograd *et al*., 2008; Stramma *et al*., 2008; Keeling *et al*., 2010; Stramma *et al*., 2010; Helm *et al*., 2011; Schmidtko *et al*., 2017), with an estimated 2% O_2 loss (77 billion metric tons) over the past 50 years (Schmidtko *et al*., 2017). As a consequence, open-ocean O_2 minimum zones (OMZs) have expanded in area (4.5 million Km^2, based on water with O_2 concentrations $<$70 μM at 200 m depth) (Bopp *et al*., 2002; Keeling *et al*., 2010; Keller *et al*., 2016; Schmidtko *et al*., 2017), and the volume with water completely devoid of O_2 (anoxic) has more than quadrupled over the same period (Schmidtko *et al*., 2017).

1.1.1 Marine oxygen minimum zones (OMZs)

Marine oxygen minimum zones (OMZs) are widespread, naturally occurring water column features that arise from the respiration of organic matter in subsurface waters with restricted circulation. Operationally defined by O_2 concentrations between 0 to 20 μM, the differential accumulation of nitrite (NO_2^-), and reduced sulphur compounds, OMZs currently constitute 1-7% of the global ocean volume (Fuenzalida *et*

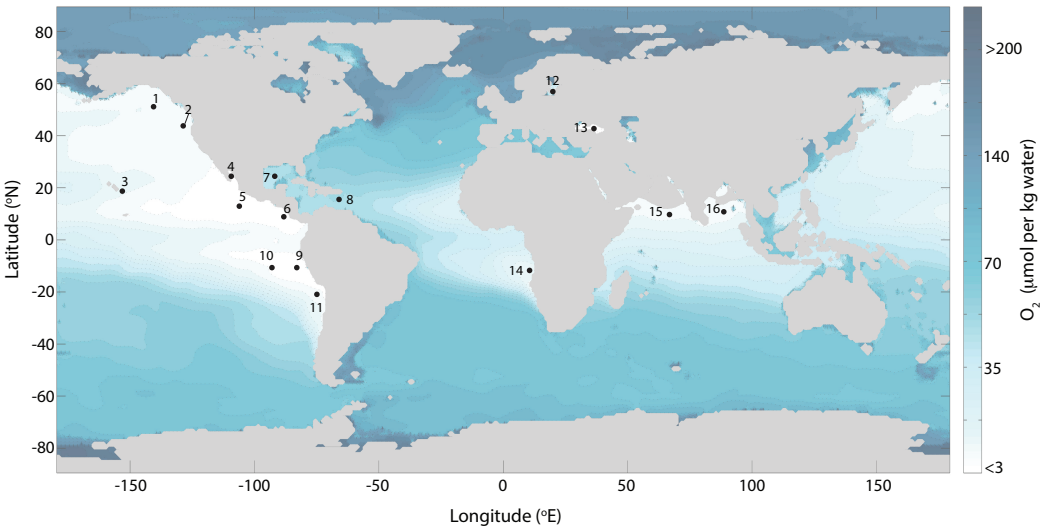

Figure 1.1: Global OMZ distribution. The global distribution of OMZs includes: (1) the Northeastern Subarctic Pacific Ocean, (2) the Saanich Inlet, (3) the Hawaii Ocean Time-series, (4) the Guaymas Basin, (5) the Eastern Tropical North Pacific, (6) Costa Rica, (7) the Gulf of Mexico, (8) the Cariaco Basin, (9) Peruvian, (10) Eastern Tropical South Pacific , (11) Chilean, (12) the Baltic and (13) Black Seas, (14) the Namibian upwelling system, (15) Arabian Sea and (16) Bay of Bengal. Contour plot depicts the minimum oxygen (O_2) concentrations (μmol per kg water) for different ocean regions. Figure was modified from Wright *et al*., 2012.

al., 2009; Paulmier and Ruiz-Pino, 2009; Ulloa and Pantoja, 2009; Codispoti, 2010; Lam and Kuypers, 2011; Ulloa *et al*., 2012; Wright *et al*., 2012). (Figure 1.1).

As O_2 levels decline, nutrients and energy are increasingly diverted away from higher trophic levels into microbial community metabolism (Diaz and Rosenberg, 2008; Wright *et al*., 2012), increasing nutrient and energy cycling through the use of alternative terminal electron acceptors (TEAs) (Diaz and Rosenberg, 2008). As a result, OMZs are hotspots for the biogeochemical cycling of carbon, nitrogen and sulphur with resulting feedback on nitrogen loss processes and greenhouse gases production including nitrous oxide (N_2O) and methane (CH_4) (Lam *et al*., 2009; Ward *et al*., 2009; Canfield *et al*., 2010; Lam and Kuypers, 2011; Naqvi *et al*., 2010), which influence global warming (Lam *et al*., 2009; Ward *et al*., 2009). Both of these gases contribute to warming by increasing the amount of solar energy that is absorbed by the planet measured as radiative forcing in Watts per square meter. However, the global warming potential (GWP) of N_2O and CH_4 varies substantially from the most common greenhouse gas, carbon dioxide (CO_2), by 300- and 30-fold, respectively (based on one-hundred year atmospheric residence times) (IPCC, 2013). Current research efforts are defining the interaction networks underlying microbial metabolism in OMZs and generating new insights into coupled biogeochemical processes in the ocean driving the cycling of climate-active trace gases (Canfield *et al*., 2010; Hawlet *et al*., 2014; Louca *et al*., 2016; Tsementzi *et al*., 2016; Hawley *et al*., 2017a).

1.2 Methane in the ocean

Methane is the most abundant hydrocarbon in the atmosphere and the second most important climate active trace gas after CO_2 (Solomon *et al*., 2007). In the ocean, CH_4 is primarily produced in anoxic marine

sediments by methanogenic archaea (Kiene, 1991). However, the majority of the CH_4 produced in marine sediments is oxidized by anaerobic methanotrophic (ANME) archaea. Microbial anaerobic oxidation of CH_4 (AOM) plays a role in CH_4 regulation and reduces methane release from marine systems (Knittel and Boetius, 2009). Anaerobic oxidation of CH_4 is well known to be coupled with sulfate reduction (sulfate-dependent AOM) in marine ecosystems (Knittel and Boetius, 2009, Milucka *et al*., 2012, McGlynn *et al*., 2015, Wegener *et al*., 2015), and recently it was demonstrated that AOM may be coupled with the dissimilatory reduction of metals, including Fe(III) and Mn(IV) (metal-dependent AOM, metal-AOM) (Ettwig *et al*., 2016, Fu *et al*., 2016, Scheller *et al*., 2016). As with sediment sources, new CH_4 production in the water column has been described. For example, CH_4 concentrations in oxic surface waters are 5-75% supersaturated with respect to the atmosphere resulting in a "CH_4 paradox" (Kiene, 1991; Karl *et al*., 2008). Ocean particles, i.e. marine snow and fecal pellets, contain anoxic microenvironments and have been considered a potential mechanism for water column CH_4 production and transport (Karl and Tilbrook, 1994). Furthermore, recent molecular studies support the hypothesis that surface ocean CH_4 is in part derived from the breakdown of methylphosphonates produced by ammonia-oxidizing archaea (Metcalf *et al*., 2012). Pelagic aerobic or anaerobic methanotrophic bacteria may still oxidize any CH_4 leaking into the water column. For instance, aerobic CH_4 oxidation has been estimated to consume $>$50% of CH_4 in the water column (Fung *et al*., 1991; Reeburgh *et al*., 1991) forming a final barrier and limiting its escape to the atmosphere (Reeburgh *et al*., 1991; Blumenberg *et al*., 2007; Kessler *et al*., 2011; Heintz *et al*., 2012) (Fig. 1.2). Pelagic CH_4 oxidation in marine environments is rarely quantified, but along the margins of an OMZ, where CH_4 intersects traces of O_2, it could be a significant process (Mau *et al*., 2013) that likely has the greatest influence on the CH_4 budget before its emission to the atmosphere (Reeburgh, 2007).

Marine OMZs encompass large reservoirs of CH_4 (Zhang *et al*., 2011; Pack *et al*., 2015). For instance, the Eastern Tropical North Pacific (ETNP) OMZ is both the largest OMZ (Paulmier and Ruiz-Pino, 2009) and the largest reservoir of oceanic CH_4 in the world (Sansone *et al*., 2001; Reeburgh, 2007; Naqvi *et al*., 2010), potentially releasing $\sim$1 Tg CH_4 yr^{-1} (Naqvi *et al*., 2010). In OMZs, O_2 is consumed faster than it is resupplied, resulting in a layer of hypoxic waters surrounding a functionally anoxic core (Thamdrup *et al*., 2012) where CH_4 accumulates (Wright *et al*., 2012). Thickening OMZs will likely move these large CH_4 pools closer to the ocean surface (Stramma *et al*., 2008; Keeling *et al*., 2010; Helm *et al*., 2011). Characterizing CH_4 -consuming microbial populations in OMZs is critical for understanding greenhouse gas and nutrient budgets under conditions of global warming and ocean deoxygenation.

1.3 Methanotrophs in OMZs

In the water column, aerobic methanotrophic bacteria or methanotrophs utilize CH_4 as the primary source of carbon and energy (Hanson and Hanson, 1996; Park *et al*., 2002; Tol *et al*., 2003; Murrell, 2010). Methanotrophs harbor genes encoding the enzyme CH_4 monooxygenase (MMO), which oxidizes CH_4 to methanol (CH_3OH). Methanol is subsequently converted to formaldehyde (CH_2O) and ultimately to CO_2 (CH_2O is oxidized to formate ($CHOO^-$) and $CHOO^-$ is oxidized to CO_2), generating reducing power for biosynthesis. Carbon is assimilated into biomass at the oxidation level of formaldehyde via the Ribulose monophosphate (RuMP) or Serine cycles (Murrell, 2010). There are two structurally and biochemically distinct forms of MMO, particulate (pMMO) and soluble CH_4 monooxygenase (sMMO). Differential expression of the two MMOs is controlled by the concentration of copper in the growth medium, with sMMO only produced at low copper concentrations ($<$1 M) (Takeda and Tanaka, 1980; Hanson and Hanson, 1996). The pMMO is universal to all methanotrophs, while only a small subset harbors both sMMO and pMMO (Heyer *et al*., 2002; Murrell, 2010; Semrau *et al*., 2010). Therefore, functional gene probes targeting pMMO i.e. the particulate CH_4 monooxygenase subunit β (*pmoA*) gene has been used to classify and identify novel methanotrophs in different environments (McDonald and Murrell, 1997; Dunfield *et al*., 2007; Ettwig *et al*., 2010; Swan *et al*., 2011).

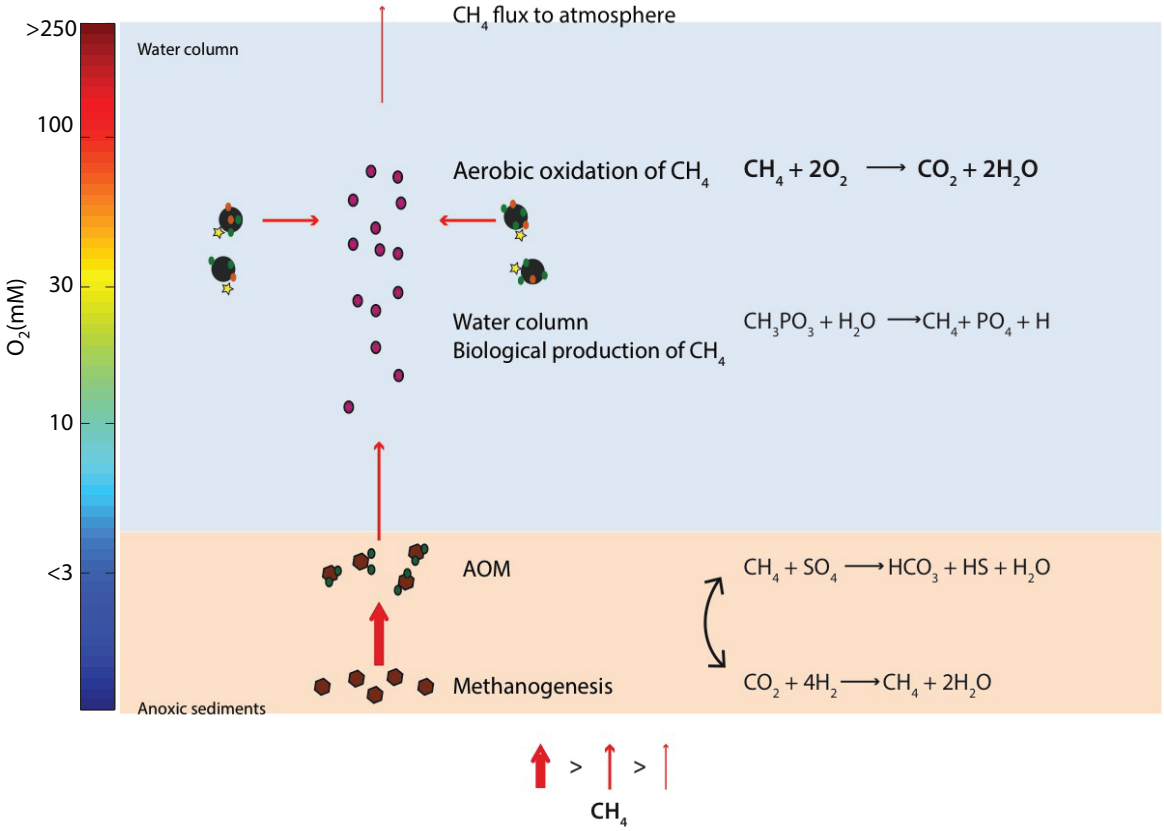

Figure 1.2: Methane cycling processes in the ocean. Methane (CH_4) is produced in deep sediments by methanogenic archaea (brown diamonds). It diffuses upwards and is consumed by the anaerobic oxidation of CH_4 (AMO), i.e. sulfate-dependent AOM, where sulphate reducers (green rods) use the hydrogen generated maintaining conditions that allow CH_4 oxidation to proceed. Anoxic microenvironments in ocean particles, i.e. marine snow and fecal pellets (gray circles) and the breakdown of methylphosphonates (yellow stars) produced by ammonia-oxidizing archaea (orange and light green rods), have been considered a mechanisms for CH_4 production and transport into the water column. Methane released to the water column is available for aerobic oxidation by methanotrophs (pink rods).

Although OMZs contribute mainly to global ocean CH_4 cycling, information is scarce regarding the composition and activity of the microbes associated with CH_4 cycling in these O_2 deficient environments. Previous surveys have quantified CH_4 oxidation rates in oxic (>90 μmol O_2 kg^{-1}) and anoxic (<1μmol O_2 kg^{-1}) OMZ waters as well as in anoxic sediments (Ward *et al.*, 1989; Ward and Kilpatrick, 1990; 1993). Results have indicated that sediment AOM occurred at a much slower rate than aerobic CH_4 oxidation in the water column (0.72^{-1}.42 nmol $L^{-1}h^{-1}$ and 0.4-2 nmol $L^{-1}h^{-1}$, respectively) providing evidence of the importance of aerobic CH_4 oxidation as a second biological filter after sediment AOM (Ward *et al.*, 1989; Ward and Kilpatrick, 1990; 1993). Subsequent small subunit ribosomal RNA (SSU rRNA) gene surveys from diverse OMZs, such as the Eastern Tropical South Pacific, the Namibian Upwelling and the Black Sea (Woebken *et al.*, 2007; Stevens and Ulloa, 2008; Glaubitz *et al.*, 2010), showed that canonical methanotrophs within the Alpha and Gammaproteobacteria, are present albeit rare members of the microbial community, i.e. they are rare biosphere constituents (<0.01% of the total community).

Parallel efforts to describe the distribution, abundance and potential metabolic activity of the functional gene *pmoA* identified pMMO-encoding phylogenetic groups (OPUs) OPU1 to OPU4 affiliated with Methylococcales in OMZs waters (Hayashi *et al*., 2007). For instance, OPU1 and OPU3 are commonly recovered in molecular gene surveys in the open ocean and coastal OMZs (Hayashi *et al*., 2007; Tavormina *et al*., 2010; Tavormina *et al*., 2013; Knief, 2015; Torres-Beltrán *et al*., 2016; Padilla *et al*., 2017) exhibiting differential abundance and distribution patterns associated with water column O_2 concentrations, i.e. OPU3 was observed to be more abundant under low O_2 concentrations ($<20\ \mu$M) (Tavormina *et al*., 2013). In addition to canonical methanotrophs, more recent studies have identified a number of non-canonical microbial groups with the potential to mediate CH_4 cycling in the O_2-deficient water column including bacteria affiliated with Verrucomicrobia (Dunfield *et al*., 2007), SAR324 within the Deltaproteobacteria (Swan *et al*., 2011) and the NC10 candidate division (Ettwig *et al*., 2010) expanding the phylogenetic range of potential CH_4-oxidizing phenotypes.

1.3.1 Methanotrophy under water column O_2 deficiency

Nitrite dependent anaerobic CH_4 oxidation in NC10

Studies of freshwater habitats have further linked AOM to the reduction of oxidized nitrogen compounds, including nitrate (NO_3^-) and NO_2^- (Raghoebarsing *et al*., 2006; Haroon *et al*., 2013). Recently discovered NO_2^--reducing bacteria of the NC10 phylum couple CH_4 oxidation to dinitrogen (N_2) production through a unique NO_2^--dependent anaerobic CH_4 oxidation (n-damo) pathway (Raghoebarsing *et al*., 2006; Ettwig *et al*., 2009). Characterized in the bacterium Candidatus Methylomirabilis oxyfera from freshwater sediments (Ettwig *et al*., 2010), the n-damo pathway reduces NO_2^- to nitric oxide (NO), which is then putatively dismutated into N_2 and O_2 gas, with O_2 serving as the oxidant for intra-aerobic methanotrophy.

Oxygen minimum zones provide potential niches for NO_2^--dependent anaerobic CH_4 oxidation (n-damo) carried out by diverse bacteria linking CH_4 oxidation to pathways of nitrogen loss under anoxic conditions. A recent study confirmed that OMZs harbor transcriptionally active bacteria affiliated with the candidate division NC10. Observations suggest a niche for NC10 in nitrogen and CH_4 cycling in OMZs (Padilla *et al*., 2016).

Alternative n-damo in canonical methanotrophs

Recent taxonomic and functional screening studies provide insight into methanotrophic community structure and activity. These studies suggest that some of the classic aerobic methanotroph species play an important role in nitrogen loss in OMZs as they thrive in O_2 deficient environments by directly using NO_3^- or NO_2^- as terminal oxidants in CH_4 oxidation pathways (Costa *et al*., 2000; Modin *et al*., 2007; Stein and Klotz, 2011; Beck *et al*., 2013; Hernandez *et al*., 2015). Evidence for the use of NO_3^- and NO_2^- in type I methanotrophs affiliated with Methylococcales, has been found in both culture-dependent and -independent studies of diverse low O_2 environments (Kalyuzhnaya *et al*., 2013; Chistoserdova, 2015; Kits *et al*., 2015a,b; Danilova *et al*., 2016). Furthermore, environmental expression of *pmoCAB* for group OPU3 was first demonstrated in a metatranscriptome from the Guaymas Basin OMZ (Lesniewski *et al*., 2012), and recently alternative modes of CH_4 oxidation by OPU3 haven been observed in the Costa Rica OMZ (Padilla *et al*., 2017). In addition, genes mediating dissimilatory NO_3^- and NO_2^- reduction were identified in the OPU3 binned genome, and were found to be transcribed in conjunction with key enzymes catalyzing formaldehyde assimilation, suggesting partial denitrification linked to CH_4 oxidation (Padilla *et al*., 2017).

1.3.2 Methanotrophic interactions

Incubation studies enriched with CH_4 revealed a marked increase of SSU rRNA gene sequences, which suggest cooperative metabolism between methanotrophic bacteria and potential microorganisms that utilize single-carbon compounds (Sauter *et al*., 2012). For instance, results derived from incubation experiments using sediment samples from Lake Washington showed a simultaneous response between Bacteroidetes, Methylophilales and canonical methanotrophs to CH_4 addition over a range of O_2 concentrations (15-75 μM) (Beck *et al*., 2013; Hernandez *et al*., 2015). Although phylogenetic-based observations alone cannot explain the underlying mechanisms of metabolite exchange, co-occurrence observations may shed light on community-level interactions that support the metabolic requirements of methanotrophic agents in OMZs.

Co-occurrence patterns and networks

The application of next-generation sequencing technologies revolutionized the field of microbial ecology (MacLean *et al*., 2009). It is now possible to study hundreds of samples of microbial communities simultaneously and with great sequencing coverage (Hamady *et al*., 2008). This has allowed for co-occurrence analysis resulting in microbial association networks consisting of nodes and edges from which ecological relationships between different microorganisms (Fig. 1.3), and between microorganisms and their environment, can be inferred (Faust and Raes, 2012; Faust *et al*., 2015; Fuhrman *et al*., 2015). Co-occurrence analysis has been used to study the ecological interactions of microbes in lakes (Eiler *et al*., 2013; Peura *et al*., 2015), soils (Barbern *et al*., 2012), streams (Widder *et al*., 2014), the human microbiome (Faust *et al*., 2012), and the marine environment (Gilbert *et al*., 2011; Steele *et al*., 2011; Chow *et al*., 2014; Cram *et al*., 2015).

Co-occurrence analysis has been increasingly explored with network inference techniques. A number of methods are available to construct taxon co-occurrence networks from cross-sectional data (Faust and Raes, 2012), ranging from correlation combined with permutation tests (Barbern *et al*., 2012) and similarity assessments with hypergeometric distribution (Chaffron *et al*., 2010) to approaches dealing with compositionality (Faust *et al*., 2012; Friedman and Alm, 2012), indirect edges (van den Bergh *et al*., 2012) and multiple factors influencing taxon abundances (Faust *et al*., 2012; Trosvik *et al*., 2015). These network inference techniques can be applied to construct dynamic models that are needed for a more comprehensive understanding of the consequences of short- and long-term perturbations such as ocean deoxygenation on microbial metabolism, i.e. CH_4 oxidation.

As the activity and dynamics of CH_4 oxidizing microbes in OMZs remain unconstrained, long-term monitoring surveys that integrate methanotrophic activity and dynamics under changing levels of water column O_2 deficiency provide a promising environmental context to assess what constrains microbial controls of CH_4 cycling. Understanding these constraints is an essential step in determining how biotic and abiotic factors merge to generate biological filters that reduce the flux of climate-active trace gases to the atmosphere.

1.3.3 Saanich Inlet is a model ecosystem

Saanich Inlet is a seasonally anoxic fjord on the coast of Vancouver Island, British Columbia, Canada (Carter, 1932; 1934; Herlinveaux, 1962; Anderson and Devol, 1973) (Fig. 1.4). Saanich Inlet water circulation is characteristic of an inverse estuary where a glacial sill at the mouth restricts exchange between the deep basin and external waters for most of the year. Freshwater is supplied at the inlet mouth predominantly by the Cowichan and Fraser Rivers, producing horizontal density differences that result in an inward flow in the surface layer and outward flow at depth (Herlinveaux, 1962; Gargett *et al*., 2003). During the spring and summer months, high levels of primary productivity in surface waters and limited vertical mixing of basin waters below the sill result in anoxia and the accumulation of CH_4, NH_4^+ and H_2S (Lilley *et al*., 1982;

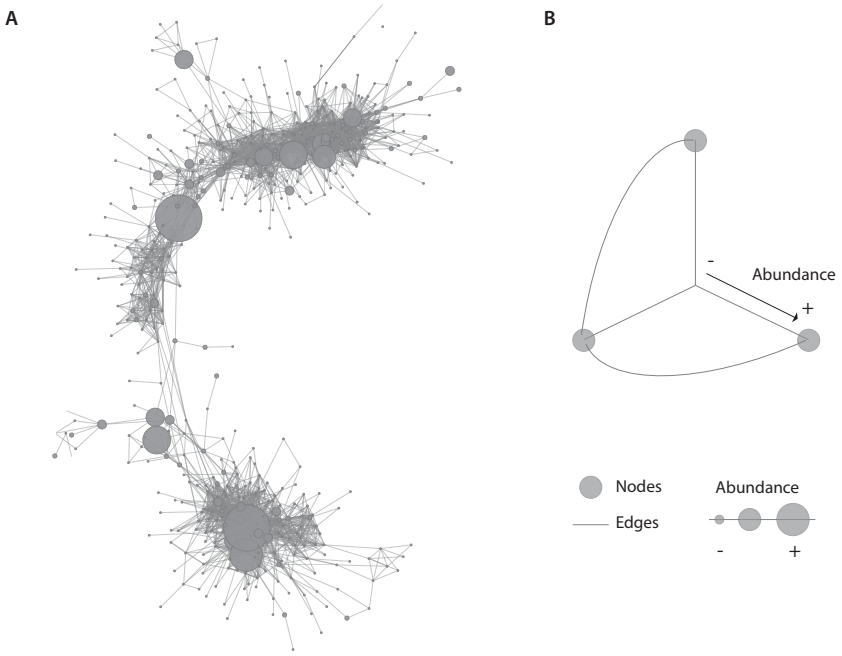

Figure 1.3: Microbial co-occurrence networks. A) Model microbial co-occurrence network based on Bray-Curtis correlations among OTUs. Nodes are depicted as gray dots whose size is relative to their abundance and edges representing significant correlations ($p < 0.001$) are depicted as lines connecting nodes. B) Model hive panel network based on Bray-Curtis correlations among OTUs. Nodes are depicted as gray dots whose position on the axis is relative to their abundance and edges representing significant correlations ($p < 0.001$) are depicted as lines connecting nodes.

Ward *et al*., 1989; Ward and Kilpatrick, 1990). In late summer and fall, neap tidal flows produce an influx of denser water from the northeastern subarctic Pacific Ocean (NESAP) that cascade over the sill, resulting in vertical mixing and the re-supplying of deep basin waters with O_2 and nutrients (Herlinveaux, 1962; Gargett *et al*., 2003). The recurring seasonal development of water column anoxia followed by deep-water renewal makes the Saanich Inlet a model ecosystem for monitoring biogeochemical responses to changing levels of water-column O_2 deficiency (Walsh *et al*., 2009; Zaikova *et al*., 2010; Walsh and Hallam, 2011; Wright *et al*., 2012).

Saanich Inlet geochemical data time-series

For over four decades, Saanich Inlet has been the site of a number of important studies evaluating water-column chemistry, and microbial community structure, function and dynamics in relation to changing levels of water-column O_2 deficiency extensible to coastal and open ocean OMZs (Walsh *et al*., 2009; Zaikova *et al*., 2010; Walsh and Hallam, 2011; Wright *et al*., 2012), including CH_4 oxidation (Ward *et al*., 1989; Ward and Kilpatrick, 1993) and ongoing metagenomic and environmental monitoring surveys (Walsh *et al*., 2009; Zaikova *et al*., 2010). Resulting data from decadal sampling efforts is available as compendiums of time-series observations encompassing historical O_2 measurements (Herlinveaux, 1962; Lee *et al*., 1999) and more recent monthly geochemical data (detailed in chapter 2) and multi-omic (DNA, RNA and protein) sequence information (Hawley *et al*., 2017b; Torres-Beltrán *et al*., 2017). In combination, these compendi-

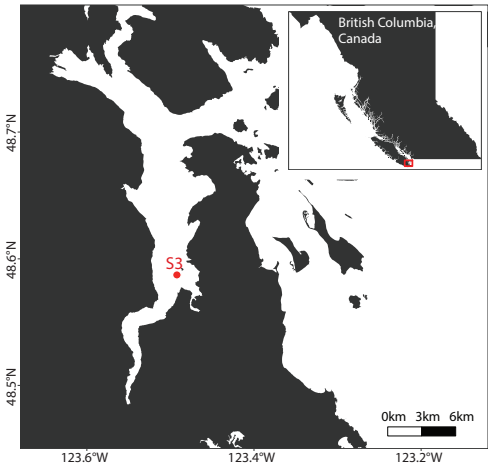

Figure 1.4: Saanich Inlet. Saanich Inlet, on the east coast of Vancouver Island, indicating sampling station S3.

ums provide a community-driven framework for observing and predicting microbial community responses to changing levels of O_2 deficiency extensible to open ocean OMZs.

Methane oxidation in Saanich Inlet

Methane oxidation rate measurements, microbial community structure surveys, *pmoA* libraries and CH_4 incubation experiments have been previously carried out in Saanich Inlet. Results from CH_4 oxidation rate measurements showed that during peak stratification, when subsurface (60-100 m) and deep (165-200 m) CH_4 maximums are observed, methanotrophs inhabiting the oxic and dysoxic compartments of the water column are expected to exhibit high CH_4 oxidation rates (2 nmol L^{-1} h^{-1}) (Ward *et al.*, 1989).

Interestingly, sequences affiliated with canonical methanotrophs such as Methylococcales and Methylomonas within the Gammaproteobacteria were recovered as rare biosphere components from SSU rRNA gene library datasets even during peak stratification (Zaikova *et al.*, 2010). This low abundance of canonical methanotrophs has also been reported for open ocean OMZs (Woebken *et al.*, 2007; Stevens and Ulloa, 2008; Glaubitz *et al.*, 2010). Moreover, *pmoA* libraries were dominated by a phylotype that clusters outside of known canonical methanotrophs within the Alpha and Gammaproteobacteria (Stilwell, 2007). More recently a microcosm study enriched with CH_4 revealed a marked increase of bacterial SSU rRNA gene sequences affiliated with Methylophilales, Methylococcales, Methylophaga, Thiotrichales and Planctomycetes (Sauter *et al.*, 2012).

Previous surveys provide insight into microbial community members potentially associated with CH_4 oxidation in the Saanich Inlet water column. However, these surveys did not couple methanotrophic bacteria taxonomic affiliation, distribution and abundance throughout water-column O_2 compartments. In addition, no linkages where made between specific community members to CH_4 oxidation gene abundance, distribution or expression patterns, constituting the "Saanich Inlet CH_4 oxidation conundrum". Moreover, no analysis was conducted comparing Saanich Inlet methanotrophic community composition and potential activity to observations reported in other coastal or open ocean OMZs.

1.3.4 Thesis objectives and overview

In this thesis, I addressed the "Saanich Inlet CH_4 oxidation conundrum" detailed above using the geochemical and sequence information available for Saanich Inlet with additional and novel experimental approaches. Results from this thesis provide a more coherent understanding of microbial community structure, function and dynamics associated with CH_4 oxidation in Saanich Inlet with important implications for understanding the mechanisms controlling CH_4 cycling throughout the global ocean.

Chapter 2: A compendium of geochemical information from the Saanich Inlet water column.

In Chapter 2, I detail the chemical and physical datasets used throughout this thesis that pair with SSU rDNA tag, metagenomic, metatranscriptomic and metaproteomic information. Chapter 2 aims to reinforce the need for model systems and to provide more in-depth information on the Saanich Inlet time-series, while supporting the using model systems to understand coupled biogeochemical processes and ecological interactions further.

Chapter 3: Sampling and processing methods impact microbial community structure and function.

In Chapter 3, I examine the effect of water sample collection and filtering methods on microbial community structure and function using coupled SSU rDNA and rRNA 454 tag sequencing data. Chapter 3 aims to highlight the need for standards of practice in model systems for extensible results and the interpretation of microbial community composition and biogeochemical cycling metabolism.

Chapter 4: Protistan parasites along water column oxygen gradients: a network approach to assessing potential host-parasite interactions.

In Chapter 4, I present correlation analyses on SSU rDNA and rRNA 454 tag sequencing information examining the O_2 effects on ecological interactions and energy flow in marine ecosystems. Chapter 4 aims to determine host-parasite interactions along water-column O_2 gradients and their potential effects on nutrient cycling in the Saanich Inlet water column.

Chapter 5: Methanotrophic community dynamics in Saanich Inlet.

In Chapter 5, I conduct correlation analyses on SSU rDNA tag sequencing information coupled with geochemical parameters to identify the significant distribution patterns of methanotrophs and interactions along water-column O_2 gradients. Chapter 5 aims to determine microbial community indicator groups and co-occurrence patterns associated with O_2, CH_4 and alternative terminal electron acceptors to gain insight into community-level interactions for CH_4 oxidation under water column O_2 deficiency.

Chapter 6: Community-level interactions support CH_4 oxidation in the Saanich Inlet O_2 -deficient water column.

In Chapter 6, I show CH_4 incubation experiments with O_2-deficient waters to generate multi-omic sequencing information and identify community-level metabolic interactions for CH_4 oxidation. Chapter 6 aims to evaluate how gene expression and metabolism related to CH_4 oxidation change under different water-column O_2 and CH_4 conditions, and to link metabolic potential to microbial community networks carrying out CH_4 oxidation under under water-column O_2 deficiency.

Chapter 2

A compendium of geochemical information from the Saanich Inlet water column[1]

This chapter details the methodologies and workflows for generating geochemical information including physical (temperature, salinity, density, irradiance, and fluorescence), chemical (PO_4^{3-}, SiO_2, NO_3^-, NO_2^-, NH_4^+, and H_2S), dissolved gas (O_2, CO_2, N_2, N_2O, CH_4), and biological (cell counts) parameter data. Moreover, this chapter reinforces the need for model systems and provide more in-depth information on the time series ending with a call to use these systems to better understand coupled biogeochemical processes and ecological interactions.

2.1 Introduction

Marine oxygen (O_2) minimum zones (OMZs) are hotspots for the biogeochemical cycling of carbon, nitrogen and sulphur with resulting feedback on nitrogen loss processes and climate active trace gas production including nitrous oxide (N_2O) and methane (CH_4) (Lam *et al*., 2009; Ward *et al*., 2009; Canfield *et al*., 2010; Lam and Kuypers, 2011; Naqvi *et al*., 2010). The effects of climate change, including increased stratification and reduced O_2 solubility in warming waters are resulting in OMZ expansion and intensification (Arrigo, 2005; Whitney *et al*., 2007; Diaz and Rosenberg, 2008; Stramma *et al*., 2008; Paulmier and Ruiz-Pino, 2009; Keeling *et al*., 2010; Schmidtko *et al*., 2017) reinforcing the need to monitor changes in water column geochemistry in O_2-deficient waters. Saanich Inlet is a seasonally anoxic fjord on the coast of Vancouver Island, British Columbia, Canada (Carter, 1932; 1934; Herlinveaux, 1962; Anderson and Devol, 1973) where the recurring seasonal development of water column anoxia followed by deep water renewal makes it a model ecosystem for monitoring biogeochemical responses to changing levels of water column O_2-deficiency (Walsh *et al*., 2009; Zaikova *et al*., 2010; Walsh and Hallam, 2011; Wright *et al*., 2012).

Oceanographic surveys in OMZ waters rely on a standard suite of measurements including temperature, salinity, density and conductivity. Additional parameters including irradiance, used to measure water column light penetration, fluorescence used to monitor chlorophyll concentrations and dissolved gases including

[1] A version of this chapter appears as **Torres-Beltrán, M.**, Hawley, A.K. *et al*. 2017. A compendium of geochemical information from the Saanich Inlet water column. Sci.Data. 4. doi:10.1038/sdata.2017.159

O_2 and carbon dioxide (CO_2) provide information on primary production (Lewis *et al*., 1985; Kolber and Falkowski, 1993; Wright *et al*., 2012). Chemical measurements of phosphate (PO_4^{3-}), silicic acid (SiO_2), and nitrate (NO_3^-) are measured as essential nutrients supporting growth and cell division (Arrigo, 2005). Nitrite (NO_2^--) and ammonium (NH_4^+) are also measured to better constrain nitrogen cycling processes (Lam *et al*., 2009; Ward *et al*., 2009; Lam and Kuypers, 2011; Wright *et al*., 2012). Because some OMZs can become completely anoxic, hydrogen sulfide (H_2S) concentrations can be used as an indicator for sulphate reduction driving chemoautotrophic metabolism (Canfield *et al*., 2010; Ulloa *et al*., 2012). Measurements of N_2O and CH_4 can also be used to monitor potential climatological impacts of OMZ expansion (Monteiro *et al*., 2006; Lam *et al*., 2009; Ward *et al*., 2009; Canfield *et al*., 2010; Lam and Kuypers, 2011; Naqvi *et al*., 2010). Collectively, these measurements define geochemical gradients in OMZ water columns that shape the conditions for coupled biogeochemical cycling.

Here I present a compendium of time-series observations encompassing historical O_2 measurements (Herlinveaux, 1962; Lee *et al*., 1999) (Fig. 2.1A) and more recent monthly monitoring efforts in Saanich Inlet from 2006 through 2014, representing over 100 independent sampling expeditions (Fig. 2B). This compendium contains physical (temperature, salinity, density, irradiance, and fluorescence), chemical (PO_4^{3-}, SiO_2, NO_3^-, NO_2^-, NH_4^+, and H_2S), dissolved gas (O_2, CO_2, N_2, N_2O, CH_4), and biological (cell counts) parameter data (Fig. 2.1B and C) useful in comparing to other oceanographic time-series from the northwest Atlantic to Eastern Tropical Pacific through the Global Ocean Sampling expeditions (Sunagawa *et al*., 2015), the Hawaii and Tara Oceans (Karl and Church, 2014; Pesant *et al*., 2015) and Bermuda Atlantic Time-series (Steinberg *et al*., 2001) and in the development of biogeochemical models. In addition, this chemical compendium partners with a compendium of multi-omic sequence information from the Saanich Inlet water column detailing time-series microbial multi-omic datasets (Hawley *et al*., 2017b). Combined, these compendiums provide a community-driven framework for observing and predicting microbial community repsonses to changing levels of oxygen deficiency extensible to open ocean OMZs.

2.2 Methods

Time-series monitoring in Saanich Inlet was conducted on a monthly basis aboard the *MSV John Strickland* at station S3 (48^o 35.500 N, 123^o 30.300 W) as previously described (Zaikova *et al*., 2010). Water samples from 16 high-resolution (HR) depths at station S3 (10, 20, 40, 60, 75, 80, 90, 97, 100, 110, 120, 135, 150, 165, 185 and 200 m) spanning oxic (¿90 μmol O_2 kg^{-1}), dysoxic (90-20 μmolO_2 kg^{-1}), suboxic (20^{-1} μmol O_2 kg^{-1}) anoxic (¡1 μmol O_2 kg^{-1}) and sulfidic water column compartments (Wright *et al*., 2012) were collected using Niskin or Go-Flow bottles for dissolved gasses: O_2, CO_2, CH_4, Nitrogen gas (N_2), N_2O; nutrients: NO_3^-, NO_2^--, NH_4^+, SiO_2, PO_4^{3-}, H_2S; and cell counts. Sampling methods for HR samples and additional six large-volume depths (10, 100, 120, 135, 150 and 200 m) collected for time-series multi-omic sequence information analyses are published in an accompanying compendium (Hawley *et al*., 2017b).

2.2.1 Environmental sampling

Historical dissolved O_2 concentrations were obtained from station S3 by sampling with Niskin bottles at discrete depths and subsequently analyzing water samples using various modifications of the Winkler method

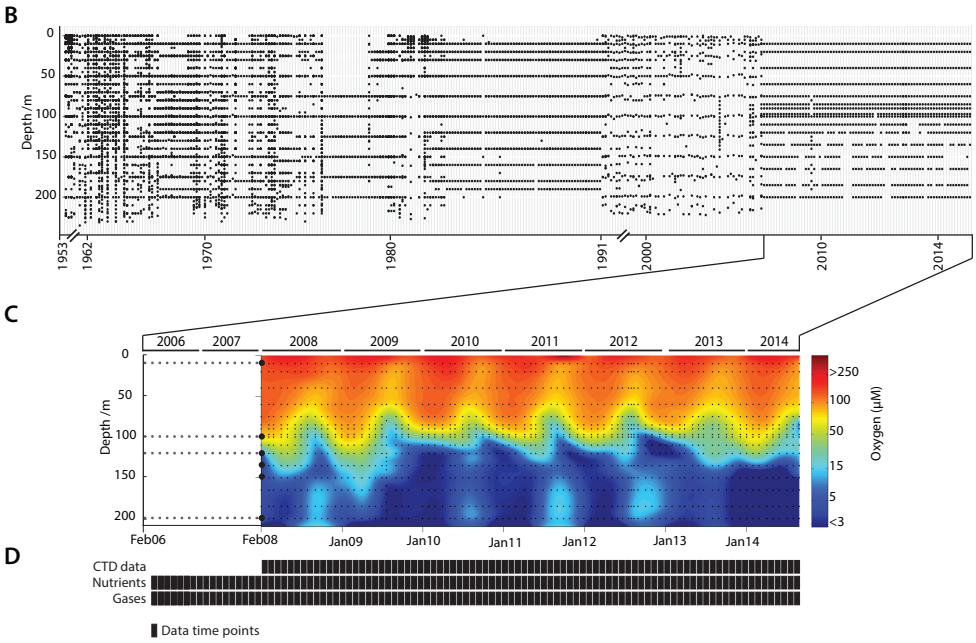

Figure 2.1: Geochemical data time series in the Saanich Inlet. A) Sampling station S3 in the Saanich Inlet B) Historical sampling effort in Saanich Inlet depicted as O_2 sampling points from 1953 to 2014. C) Oxygen concentration contour for CTD data (February 2008 onward), and points for 16 sampling depths for nutrients and gases. D) Sample inventory from February 2006 to October 2014 showing historical, CTD and nutrient datasets (solid black).

(Herlinveaux, 1962; Carpenter, 1965; Lee *et al*., 1999) (Data Citation 1). Historical water column profiles can also be accessed at the Ocean Sciences Data Inventory website hosted by the Institute of Ocean Sciences and Fisheries and Oceans Canada (http://www.pac.dfo-mpo.gc.ca/science/oceans/data-donnees/search-recherche/profiles-eng.asp). Samples collected from February 2006 to February 2008 were processed and analysed for dissolved gases and nutrients as first reported in Zaikova *et al* (Zaikova *et al*., 2010) (Fig. 2.2). Beginning on February 2008, a Sea-Bird SBE 25 CTD (conductivity, temperature and depth), with Sea-Bird SBE 43 dissolved O_2 and Biospherical Instruments PAR sensors attached was used to measure conductivity, temperature, dissolved O_2, PAR/Irradiance and fluorescence (Data Citation 1). To minimize the effects of off-gassing, waters were collected in the following order; dissolved O_2 for Winkler titration (from select depths for CTD calibration), dissolved gases (N_2O and CH_4), NH_4^+, H_2S, nutrients, cell counts (Data Citation 1) and salinity (from selected depths for CTD calibration). A detailed seawater sampling video protocol can be found online (http://www.jove.com/video/1159/seawater-sampling-and-collection).

2.2.2 Chemical Data

CTD data analysis

CTD data were downloaded, converted and pre-processed in the laboratory using the SeaBirdSeasoft software. Downcast data of the deepest cast (200 m) was extracted and converted from ASCII format

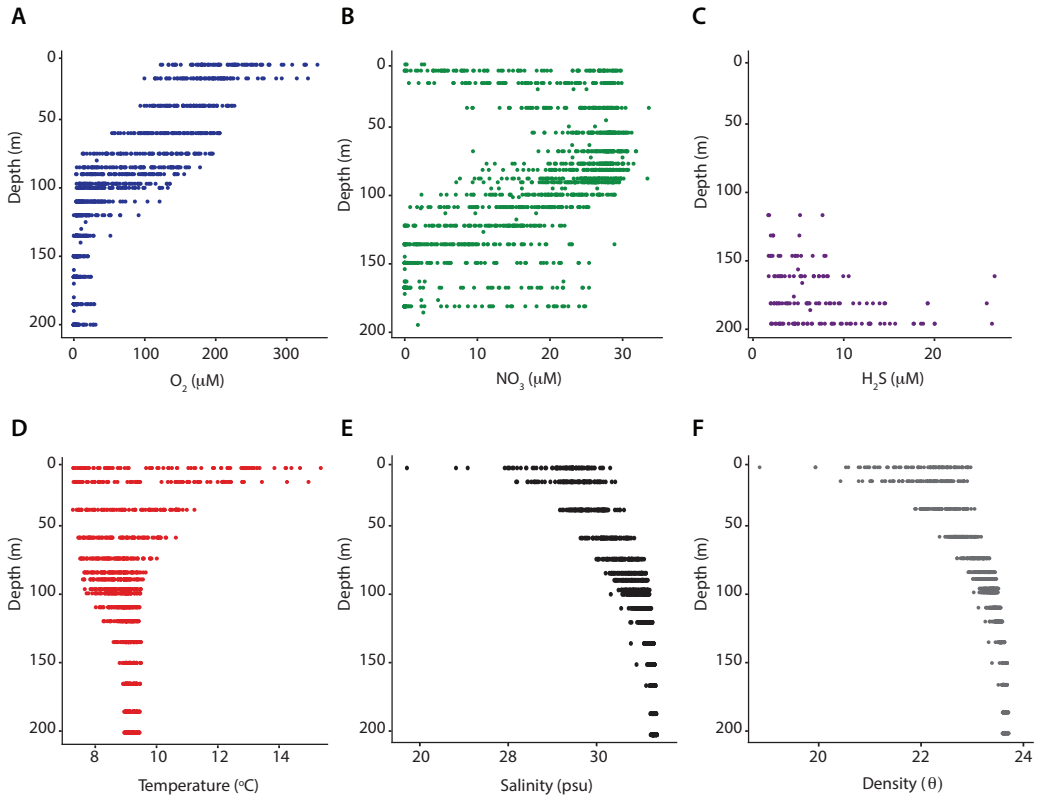

Figure 2.2: Time series environmental parameters water column profiles. Panel showing dot plots for oxygen (O_2 ; blue), nitrate (NO_3^- ; green), hydrogen sulphide (H_2S ; purple), temperature (^oC ; red), salinity (psu ; black) and density (θ; gray) measurements along the depth profile for samples taken from February 2008 to October 2014 at Station S3 in Saanich Inlet.

into a .cnv file for manual curation. Salinity and density were calculated using the Derive module with the corrected conductivity measurements. Temperature and salinity were exported using an ITS-90 scale. Oxygen sensor measurements collected in millilitre per litre (ml/L) were converted to micromolar (μM) units (Data Citation 1). Discrete Winkler analyses from water samples spanning LV depths were used to calibrate the CTD O_2 measurements (Data Citation 1).

Nitrate, Phosphate and Silicic acid

For each depth, sample water was filtered through a 0.2 μm acrodisc (Millipore) and used to rinse a 15 ml tube three times before filling with 14 ml. Samples were stored on ice and later in the lab at -20^o C for up to four months prior to analysis. A Bran Luebbe AutoAnalyser 3 using air-segmented continuous-flow and standard colorimetric methods was used for analysis. In brief, nitrate (NO_3^-) was reduced to nitrite (NO_2^-) by a copper-cadmium reduction column. Nitrite was then quantified by a modified colorimetric assay (Armstrong *et al.*, 1967), reading sample absorbance at 550 nm. Orthophosphate (PO_4^{3-}) was quantified based on the colorimetric method for reduced phospho-molybdenum complex, reading samples absorbance at 880 nm (Murphy and Riley, 1962). Silicic acid (H_4SiO_4) was quantified by reduction to a molybdenum blue

complex, reading sample absorbance at 820 nm. Oxalic acid was added to remove phosphate interference (Armstrong *et al*., 1967) (Data Citation 1).

Ammonium

A fluorometric measurement protocol for ammonium (NH_4^+) analysis was carried out as previously described in Holmes *et al* for marine samples (Holmes *et al*., 1999). For each depth, glass amber scintillation bottles were rinsed three times, then filled to overflowing and capped immediately to minimize off-gassing of NH_4^+ and stored on ice for 1-3 hours before processing. A total of 5 ml of sample water was transferred to vials with 7.5 ml o-phthaldialdehyde (OPA; Sigma) in triplicate. Simultaneously, 7.5 ml of OPA was added to prepared NH_4^+ standard curve (0.025 10.0 μM NH_4Cl) and stored at room temperature for up to 4 hours. Fluorescence at 380ex/420emm was read using a Turner Designs TD-700 fluorometer (2006-2009) or Varioskan plate reader (2009-2014) in triplicate with 300 μl of sample or standard in a 96-well round bottom plate (Corning) (Fig. 2.3) (Data Citation 1).

Nitrite

The protocol for NO_2^- analysis was carried out as previously described in Armstrong *et al* modified for marine samples (Armstrong *et al*., 1967). For each depth, sample water was filtered through a 0.2 μm acrodisc (Millipore) and used to rinse a 15 ml tube three times before filling with 14 ml filtered sample water and stored on ice for 1-3 hours before processing. A total of 2 ml of sample water was transferred to 4 ml plastic cuvettes in triplicate, and 100 μl sulphanilamide (Sigma) and 100 μl nicotinamide adenine dinucleotide (NAD; Sigma) were added. Simultaneously, reagents were added to prepared standards (0.025 5.0 μM $NaNO_2^-$). Cuvettes were inverted and stored on ice for up to 4 hrs. Absorbance at 542 nm was read using a Cary60 spectrometer (Fig. 2.3) (Data Citation 1).

Hydrogen sulfide

The protocol for hydrogen sulfide (H_2S) was carried out as previously described in Cline (Cline, 1969) modified for marine samples. For each depth, 10 ml sample water was collected directly into a 15 ml tube containing 200 μL 20% Zinc Acetate (Sigma) and stored on ice for 4-24 hours before processing. Samples were mixed prior transferring a total of 300 μL of sample into triplicate wells of a 96-well round or flat-bottom plate (Corning), and 6 μL Hach Reagent (Hach) 1 and 2 for sulphide assay were added to each well. After 5 min incubation, absorbance at 670 nm was read using a spectrophotometer (2008-2009) or Varioskan plate reader (2009-2014) (Data Citation 1).

Cell counts

For each depth, 10 ml sample water was collected directly into a sterile 15 ml tube containing 1.1 ml of 37% formaldehyde and stored on ice. Back at the lab, samples were stored at 4^o C for up to two days prior to cell counting using a BD LSR II flow cytometer (2008- 2012) or MACS Quant Analyzer (2012 2014) based on the following protocols. For BD LSR II, a dye mixture was prepared by diluting 3 μL of the SYBR Green I (Invitrogen) dye in 1830 μL of sterile water. Six drops (Alignflow) alignment beads were then added to this mixture. In a round-bottom polystyrene tubes, 25 μL of the dye mix was added to 475 μL of the water sample (in triplicates). The cells and beads were then counted using BD LSR II flow cytometer. For MAXSQuant, a dye mixture was prepared by diluting 240 μL of seawater sample with 10 μL of SYBR Green I (Invitrogen) dye mix which contains 6 μm flow cytometry blue laser alignment beads (Alignflow),

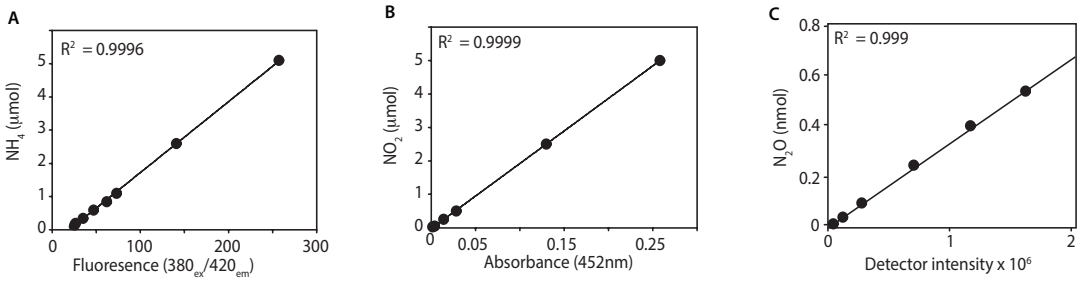

Figure 2.3: Validation for environmental parameters. A-B) Typical standard curves for chemical parameters ammonium (NH_4^+) and nitrite (NO_2^-), and C) gas concentration nitrous oxide (N_2O).

for calibration purposes. SYBR Green mix was prepared by diluting 4 μL of the dye in 1570 μL of sterile water following an addition of 30 μL beads. Samples are prepared in triplicates in a 96-well flat bottom black plate (Corning) and run on MACSQuant Analyzer (MiltenyiBiotec) (Data Citation 1).

Dissolved Gases

For each depth, sample water was collected through silicon tubing ($\sim$15 cm long and 1/4“ thick, pre-flushed for a few seconds with sample water) into a 30 or 60 ml borosilicate glass serum vial, overflowing three times the volume and taking care to remove air bubbles from the tubing and vial during filling. The vials were spiked with 50 μL saturated mercuric-chloride solution, then crimp-sealed with a butyl-rubber stopper and aluminium cap. Samples were stored in the dark at 4^o C until processing. Dissolved gases were analysed using either headspace for CO_2, CH_4, N_2 and N_2O (2006-2009, samples stored for up to 2 years) or automated purge-and-trap for CH_4 and N_2O only (2009-2014, samples stored for $<$3 months) coupled with gas chromatography-mass spectrometry (GC-MS) (Capelle *et al.*, 2015) (Data Citation 1). Samples with $>$20% standard deviation between replicates were excluded to discard any long storage effects.

2.2.3 Data Records

Data record 1 The Saanich Inlet O_2 historical data (1953-2000) is accessible in comma-separated-value format file *Historical_O_2_DATA.csv* on the Dryad Digital (Data Citation 1).

Data record 2 The Saanich Inlet time-series CTD data is accessible in comma-separated-value format file *Saanich_TimeSeries_CTD_DATA.csv* on the Dryad Digital Repository (Data Citation 1).

Data record 3 The Saanich Inlet time-series chemical data is accessible in comma-separated-value format file *Saanich_TimeSeries_Chemical_DATA.csv* on the Dryad Digital Repository (Data Citation 1).

Data record 4 The Saanich Inlet time-series Winkler O_2 data is accessible in comma-separated-value format file *Saanich_TimeSeries_Winkler_DATA.csv* on the Dryad Digital Repository (Data Citation 1).

2.3 Technical Validation

Data quality control

Data in the Saanich Inlet time series was collected and processed by experienced scientists with extensive

training in the sampling methods and data processing steps described above. People interested in becoming part of the scientific crew were invited to participate in training sessions with experienced scientists in the field and laboratory to gain practical experience. Once in the field, trainees were carefully supervised during sample collection for a minimum of 3 months for quality assurance. Following each cruise, the acting chief scientist compiled all chemical and physical data collected and conducted initial quality controls, checking for outliers and verifying standard curves. Data were then entered into an in-house database along with field notes and precise records of volumes of water filtered informing downstream analyses.

CTD and chemical data validation

The SeaBird 43 dissolved O_2 sensor was calibrated by Winkler O_2 measurements (Winkler, 1888). Samples from selected depths were collected into Winkler glass Erlenmeyer flasks using latex tubing, overflowing three times to ensure no air contamination. Oxygen concentration was determined using a Brinkman autotitrator, routinely calibrated with a potassium iodide standard. Stability of CTD O_2 measurements was determined by comparing the high values with Winkler measurements, and low values with sulfidic profiles where the sensors levels off. Where H_2S is detected I consider O_2 measurements to be 0 μM based on spontaneous auto-oxidation reaction of H_2S with O_2. I have estimated our limit of detection for the automated Winkler method at $\sim$0.007 ml/L or $\sim$0.3 μM. The SeaBird conductivity sensor was calibrated using salinity samples collected at selected depths. Salinity glass bottles were rinsed 4 times and filled with water sample, stored at room temperature and analyzed within 4 months on a Guildline Portasal salinometer.

For each cruise, standard curves for NH_4^+ and NO_2^- were prepared. Stock solutions and reagents for both assays were freshly made every three months and stored in the dark at 4^o C and were tested prior to being used for analysis. Stock solution quality and assay validation was carried out using linear regression and calculating the r squared value (r^2 0.90) on the absorbance data (Fig. 2.3). Standard curve stock solutions and reagents for H_2S assay were evaluated every three months based on manufacturers instructions. I have estimated our limit of detection for these assays to be 0.001 μM NH_4^+, 0.0006 μM NO_2^-, and 1.7 μM H_2S. Samples for NO_3^-, PO_4^{3-} and H_4SiO_4 were run in single measurements. Autoanalyzer estimated limit of detection for these measurements are 0.020 μM NO_3^-, 0.012 μM PO_4^{3-} and 0.100 μM H_4SiO_4.

Flow cytometry validation

Concentration of flow cytometry (FL) alignment beads was determined by microscopy using a hemocytometer. Bead counts for each FL run were then used to calculate the volume of sample measured. Two blanks were included in each FL run, and consisted of sterile water bead/dye solution with sterile water in place of sample water, to ensure instrument cleanliness and optics function. Size gates were set to include beads and bacterial and archaeal cell sizes and to reduce noise of any small particulate debris.

Gas analysis validation

A thorough review of the Purge and Trap GCMS (PT-GCMS) method validation has been previously described (Capelle *et al.*, 2015). Standard curves were run at the start of each batch of 25 samples by injecting precisely measured quantities of a standard gas mixture (CH_4, N_2O, CO_2 and N_2) calibrated against National Ocean and Atmospheric Administration (NOAA) certified reference gas mixture. Single standards were also measured every $\sim$2 hours (5 - 6 sample per run) to monitor instrument drift. The precision of CH_4 and N_2O measurements based on replicate measurements of air-equilibrated water samples was $<$4%. Accuracy

was confirmed by measuring dissolved N_2O and CH_4 in carefully prepared air-equilibrated, temperature-controlled Milli-Q water and comparing this to expected concentrations based on gas-solubility equations (Wiesenburg and Guinasso, 1979; Weiss and Price, 1980). Detection limits depend on the volume of sample being purged, and were 0.8 nM for CH_4 and 0.5 nM for N_2O for the samples analyzed in this time-series (2009-2014) (Fig 2.3). Samples were run in duplicate or triplicate to ensure reproducible readings. The relative standard deviation between replicate samples was calculated and included in the output data. The output data are also carefully inspected to ensure optimal instrument performance during sample analysis before being submitted to the database.

2.4 Data Citation

1. M. Torres-Beltran, A.K. Hawley *et al*. Dryad Digital Repository. http://dx.doi.org/10.561/dryad.nh035 (2017).

2.5 Conclusion and application

The combined use of geochemical and multi-omic sequence information have led to new insights into coupled biogeochemical cycling of carbon, nitrogen and sulfur between key microbial players and the development of a predictive ecosystem model describing the flow of multi-omic sequence information and process rates along O_2 gradients (Hawley *et al*., 2014; Louca *et al*., 2016; Hawley *et al*., 2017a). Thus, time-series data from Saanich Inlet provides a community-driven framework for observing and predicting microbial community responses to ocean deoxygenation across multiple scales of biological organization.

In this thesis, I used time-series geochemical data to correlate O_2, CH_4, NO_3^- and NO_2^- with methanotrophic community members and predict community-level interactions for CH_4 oxidation throughout O_2 gradients in the Saanich Inlet water column.

Chapter 3

Sampling and processing methods impact microbial community structure and function

This chapter reinforces the need for standards of practice in model systems for extensible results and focuses on the sample collection problem for multi-omics information. In addition, the methods used here form the basis of microbial community composition and structure analyses carried out throughout chapters 4 to 6 in the investigation of community-level interactions along water column O_2 gradients.

3.1 Introduction

Current research efforts are defining the interaction networks underlying microbial metabolism in oxygen minimum zones (OMZs) and generating new insights into coupled biogeochemical processes driving nutrient and energy flow among and between trophic levels on ecosystem scales (Hawley *et al*., 2014;Cram *et al*., 2015;Louca *et al*., 2016;Torres-Beltrán *et al*., 2016a). However, marine microbial responses at the individual, population and community levels to OMZ expansion, and the concomitant impact of these responses on global-scale nutrient and energy cycling remain poorly constrained due in part to inconsistent, and perhaps inadequate sampling methods that limit cross-scale comparisons between locations and may cloud our view of *in situ* microbial processes.

Over the past twenty years, oceanographic researchers have increasingly used multi-omic (DNA, RNA, protein and metabolites) methods to determine microbial community structure, function and activity in relation to physical, chemical and biological oceanographic processes. A wide range of sample collection and processing methods have been used to generate these data sets without standardization (Stein *et al*., 1996;Fuhrman and Davis, 1997;Massana *et al*., 1997;Acinas *et al*., 1999;Crump *et al*., 1999;Murray *et al*., 1999;James *et al*., 2000;Moeseneder *et al*., 2001;LaMontagne and Holden, 2003;Venter *et al*., 2004;DeLong *et al*., 2006;Hewson and Fuhrman, 2006;Rusch *et al*., 2007;Waidner and Kirchman, 2007;Brown *et al*., 2009;Ghiglione *et al*., 2009;Walsh *et al*., 2009;Zaikova *et al*., 2009;Canfield *et al*., 2010;Gilbert *et al*., 2010;Hurwitz *et al*., 2013;Rodriguez-Mora *et al*., 2013;Smith *et al*., 2013;D'Ambrosio *et al*., 2014;Ganesh

et al., 2014;Parris *et al*., 2014;Brum *et al*., 2015;Ganesh *et al*., 2015;Orsi *et al*., 2015;Padilla *et al*., 2015;Pesant *et al*., 2015;Suzuki and Shimodaira, 2015). While large-scale patterns appear to be consistent between studies with respect to major taxonomic groups and water column compartments, our ecological perspective is blurred by inconsistencies in marker gene selection and coverage.

Recent reports have begun to evaluate potential biases in OMZ microbial community structure, function and activity with emphasis on sample collection and filtration methods. Water column sampling typically involves the use of collection bottles (Niskin or GO-FLO) and on-ship filtration to concentrate microbial biomass into two primary size fractions, a larger particle associated (>1-30 μm) and smaller free-living (<1-0.2 μm) fraction (Padilla *et al*., 2015). Size fractionation surveys conducted in the Eastern Tropical South Pacific (ETSP) and Eastern Tropical North Pacific (ETNP) showed differential microbial community structure and nitrogen cycling functional gene distribution and expression across size fractions (Ganesh *et al*., 2014;Ganesh *et al*., 2015). In addition to filter fractionation, Padilla and colleagues working in the Manzanillo Mexico OMZ observed variation in microbial community structure based on filtered water volume (Padilla *et al*., 2015). Most recently, a study in the Cariaco Basin observed that particles sinking on timescales relevant to sample collection and filtration can influence microbial community structure in Niskin or GO-FLO bottles (Suter *et al*., 2016). A comparison of metatranscriptome data obtained from bathypelagic Mediterranean Sea samples collected using Niskin bottles followed by shipboard filtration vs. filtration and fixation in situ found significant shifts in gene expression for particular groups of microorganisms (Edgcomb *et al*., 2016). These findings reinforce the need for continued evaluation of the methods used for sample collection and processing in order to establish standards of practice that reduce collection bias and enable more robust cross-scale comparisons. This is particularly relevant when conducting process rate measurements in which community structure variation due to bottle effects can result in potential rates that do not reflect of *in situ* microbial activity (Stewart *et al*., 2012a).

The Scientific Committee on Oceanographic Research (SCOR) initiated Working Group 144 Microbial Community Responses to Ocean Deoxygenation to investigate and recommend community standards of practice for compatible multi-omic and process rate measurements in OMZs and other oxygen deficient waters in order to facilitate and promote future cross-scale comparisons that more accurately reflect *in situ* microbial community structure, function and activity (http://omz.microbiology.ubc.ca/page4/index.html). The inaugural workshop of SCOR Working Group 144 was held in British Columbia Canada during the week of July 14, 2014. During the workshop attendees participated in practical sampling and experimental activities in Saanich Inlet. During spring and summer months, restricted circulation and high levels of primary production lead to progressive deoxygenation and the accumulation of methane (CH_4), ammonium (NH_4^+) and hydrogen sulfide (H_2S) in deep basin waters. In late summer and fall, oxygenated nutrient rich waters flow into the inlet from the Haro Strait renewing deep basin waters (Carter, 1932;1934;Herlinveaux, 1962;Anderson and Devol, 1973;Zaikova *et al*., 2010;Walsh and Hallam, 2011;Torres-Beltrán *et al*., 2016b). The seasonal pattern of water column anoxia and renewal makes the inlet a model ecosystem for evaluating changes in microbial community structure, function and activity in response to changing levels of water column deoxygenation. Saanich Inlet is thus a tractable environment to test different water sample collection and processing methods relevant to OMZs.

Table 3.1: Biomass collection scheme for DNA and RNA *in situ* and on-ship samples. Table shows the filtering methods used for a given sampling condition (*in situ* vs. on-ship), including pre-filter size cutoff, biomass collection filter, filtered volume, and molecular sample obtained.

Filtration type	Pre-filter size (μm)	Collection filter size (μm)	Volume (L)	Molecular sample
in situ		0.4	2	DNA, RNA
On ship	0.4	0.22	2.5 1.5 0.5 0.25	DNA, RNA*
	2.7	0.22	2.5 1.5 0.5 0.25	DNA

Experiments carried out during the workshop were designed to compare and cross-calibrate *in situ* sampling with conventional bottle sampling methods including the use of different filter combinations and sample volumes. Here, I describe the effect of these parameters on microbial community structure and potential activity and discuss community standards development in relation to the mandate of the SCOR Working Group 144.

3.2 Methods

3.2.1 Environmental sampling

Environmental parameter data and molecular (DNA and RNA) collection methods used during the SCOR workshop were similar to those previously described (Zaikova *et al*., 2010; Walsh and Hallam, 2011; Torres-Beltrán *et al*., 2016b). In brief, water sampling was conducted on aboard the *MSV John Strickland* at station SI03 (48^{o} 35.500 N, 123^{o} 30.300 W) on July 16, 2014. Samples were collected using 12 L GoFLo bottles on a winch system for dissolved gases, nutrients and CTD with a PAR and O_2 sensor attached was used to measure temperature, salinity, PAR/Irradiance, fluorescence, conductivity, density, and dissolved O_2 at 165 and 185 meter depth intervals spanning anoxic ($<1\mu$mol O_2 kg^{-1}) and sulfidic water column compartments (Fig. 3.1).

3.2.2 Workshop microbial biomass collection

For comparison to Niskin bottle sampling, water samples were collected and preserved *in situ* using a McLane Phytoplankton Sampler (PPS) system deployed at 165 and 185 m depth intervals spanning anoxic ($<1\mu$mol O_2 kg^{-1}) and sulfidic water column compartments (Fig.3.1B). Sample volumes of 2 L were filtered onto 0.4 μm GFF membrane filters (47 mm diameter). Filter biomass was directly frozen and stored at -80 ^{o}C for downstream DNA and RNA extraction. On-ship samples were collected using Niskin bottles

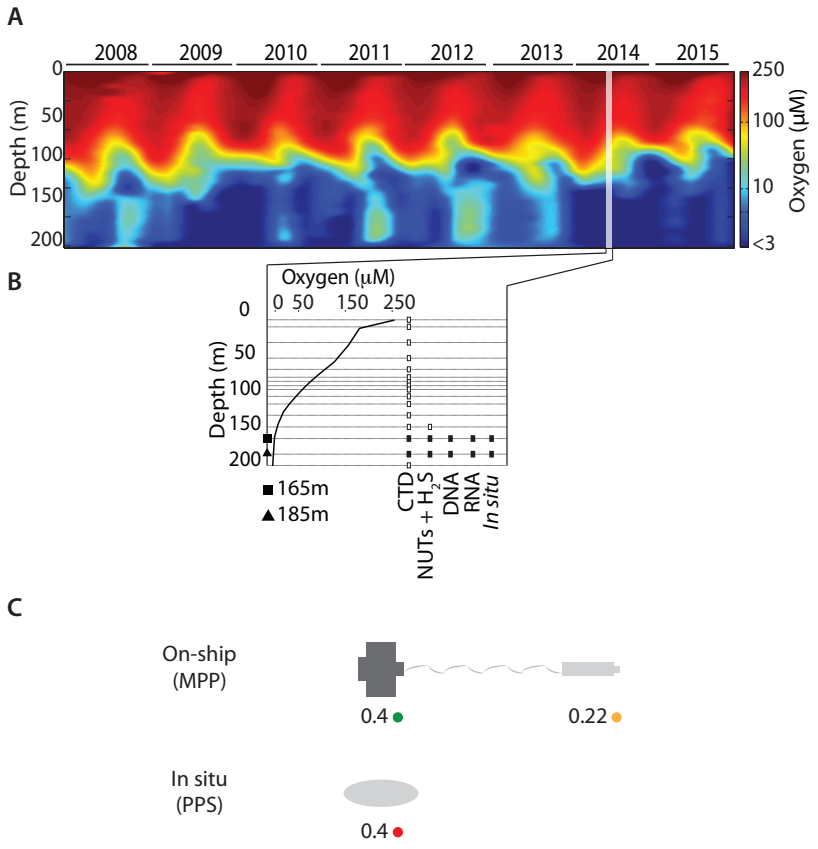

Figure 3.1: SCOR workshop sampling schemeA) Saanich Inlet time-series CTD oxygen (O_2) concentration contour (2008 -2015) showing seasonal water column stratification and deep-water renewal events. B) Sample inventory for SCOR workshop held in July 2014 includes CTD, nutrients (NO_3^-, NO_2^-, NH_4^+, PO_4^{-3} and SiO_2), hydrogen sulfide (H_2S), and molecular on-ship and *in situ* data. Data shown in this survey is depicted as solid symbols (*in situ* and on ship DNA and RNA (rectangle) collected at 165 (square) and 185 (triangle) meters. C) Schematic model for filter combinations and filtration methods showing on-ship (MPP) filtration onto 0.4 μm pre-filters (green), filtration onto 0.22 μm filters with in-line pre-filtration using the 0.4 μm pre-filters (yellow), and *in situ* filtration (PPS) onto 0.4 μm filters without pre-filtration (red).

from 165 and 185 m as described above (Fig. 3.1B) and concentrated for DNA and RNA extractions with a MasterFlex peristaltic pump (MPP) ($\sim$60 mL min^{-1}) using different filter combinations (0.4 μm polycarbonate or 2.7 μm GF/D pre-filters in-line with a 0.22 μm Sterivex polycarbonate filter cartridge) (Fig. 3.1C) and water volumes (250 ml, 500 ml, 1 L and 2.5 L) (Table 3.1). Filtered biomass on Sterivex filters was preserved in 1.8 ml of sucrose lysis buffer (DNA analyses) or RNA later (RNA analyses) prior to storage at -80 ^{o}C. Pre-filters (0.4 μm only) were also preserved in 1.8 ml of lysis buffer (DNA) or RNA later (RNA) prior to storage at -80 ^{o}C.

3.2.3 Time-series microbial biomass collection

Time-series samples were collected as previously described (Walsh *et al*., 2009; Zaikova *et al*., 2010; Hawley *et al*., 2017b; Torres-Beltrán *et al*., 2016a). Briefly, large volume (10 L) samples were collected from February 2006 to February 2011 at six depths (10, 100, 120, 135, 150 and 200 m) and filtered with an in-line 2.7 μm GDF glass fiber pre-filter onto a 0.22 μm Sterivex polycarbonate cartridge filter. Hi-resolution (2 L) samples were collected from May 2008 to July 2010 from 16 depths (10 to 200 m) and filtered directly onto a 0.22 μm Sterivex polycarbonate cartridge filter. All time-series samples were preserved in 1.8 ml of sucrose lysis buffer (DNA) prior to storage at -80 ^{o}C.

3.2.4 Nucleic acid extraction

Genomic DNA was extracted from Sterivex filters as previously described (Zaikova *et al*., 2010; Hawley *et al*., 2017b). Briefly, after defrosting Sterivex on ice, 100 μl lysozyme (0.125 mg ml^{-1}; Sigma) and 20 μl of RNAse (1 μl ml^{-1}; ThermoFisher) were added and incubated at 37^{o}C for 1 h with rotation followed by addition of 50 μl Proteinase K (Sigma) and 100 μl 20% SDS and incubated at 55 ^{o}C for 2 h with rotation. Lysate was removed by pushing through with a syringe into 15 mL falcon tube (Corning) and with an additional rinse of 1 mL of lysis buffer. Filtrate was subject to chloroform extraction (Sigma) and the aqueous layer was collected and loaded onto a 10K 15 ml Amicon filter cartridge (Millipore), washed three times with TE buffer (pH 8.0) and concentrated to a final volume of between 150-400 μl. Total DNA concentration was determined by PicoGreen assay (Life Technologies) and genomic DNA quality determined by visualization on 0.8% agarose gel (overnight at 16V). Genomic DNA was extracted from 0.4 and 2.7 μm pre-filters as follows. The filter was cut in half using sterile scissors. One half was minced into in smaller pieces and used for DNA extraction while the remaining half was stored at -80 ^{o}C. Filter pieces were transferred to a 15mL falcon tube followed by addition of 1.8 mL lysis buffer and 150 μL 20% SDS. In order to ensure biomass removal form filter, 3 and 2 mm zirconium beads were added for beat beading using a vortex mixer at maximum speed. Filters were shaken for 4 minutes in 2 minutes laps then subjected to chloroform extraction and processed in the same way as described above for Sterivex filters.

Total RNA was extracted from Sterivex filters using the mirVana Isolation kit (Ambion) (Shi *et al*., 2009; Stewart *et al*., 2010) protocol modified for sterivex filters (Hawley *et al*., 2017b). Briefly, after thawing the filter cartridge on ice RNA later was removed by pushing through with a 3 ml syringe followed by rinsing with an additional 1.8 mL of Ringer's solution and incubated at room temperature for 20min with rotation. Ringer's solution was evacuated with a 3 ml syringe followed by addition of 100 μl of 0.125 mg ml^{-1} lysozyme and incubated at 37 ^{o}C for 30 min with rotation. Lysate was removed from the filter cartridge and subjected to organic extraction following the mirVana kit protocol. DNA removal and clean up and purification of total RNA were conducted following the TURBO DNA-free kit (ThermoFisher) and the RNeasy MinElute Cleanup kit (Qiagen) protocols respectively. Total RNA concentration was determined by RiboGreen analysis (Life Technologies) prior to synthesize first strand cDNA using the SuperScript III First-Strand Synthesis System for RT-qPCR (Invitrogen) according to manufacturer instructions. Total RNA was extracted from 0.4 and 2.7 μm filters as follows. The filter was cut in half using sterile scissors. One half was minced into in smaller pieces and used for RNA extraction while the remaining half was stored at

-80 ^{o}C. Filter pieces were transferred to a 15 mL falcon tube followed by addition 1.8 ml MirVana Lysis Buffer and 100 μl of 0.125 mg ml^{-1} lysozyme. In order to ensure biomass removal form filter, 3 and 2 mm zirconium beads were added for beat beading using a vortex mixer at maximum speed. Filters were shaken for 4 minutes in 2 minutes laps followed by an incubation at 37 ^{o}C for 30 min with rotation, then processed in the same way as described above for Sterivex filters.

3.2.5 Small subunit ribosomal RNA sequencing

Extracted DNA and cDNA from 165 and 185 meter depth intervals (2.7, 0.4 and 0.22 μm) was used to generate SSU rDNA and rRNA pyrotags with three domain resolution. PCR amplification procedures were carried out as previously described (Hawley *et al*., 2017b). In brief, pyrotag libraries were generated by PCR amplification using multi-domain primers targeting the V6-V8 region of the SSU rRNA gene (Allers *et al*., 2013): 926F (5'-cct atc ccc tgt gtg cct tgg cag tct cag AAA CTY AAA KGA ATT GRC GG-3') and 1392R (5'-cca tct cat ccc tgc gtg tct ccg act cag-<XXXXX>-ACG GGC GGT GTG TRC-3'). Primer sequences were modified by the addition of 454 A or B adapter sequences (lower case). In addition, the reverse primer included a 5 bp barcode designated <XXXXX> for multiplexing of samples during sequencing Twenty-five microliter PCR reactions were performed in triplicate and pooled to minimize PCR bias. Each reaction contained between 1 and 10 ng of target DNA, 0.5 μl Taq DNA polymerase (Bioshop inc. Canada), 2.5 μL Bioshop 10x buffer, 1.5 μl 25 mM Bioshop $MgCl_2$, 2.5 μL 10 mM dNTPs (Agilent Technologies) and 0.5 μL 10 mM of each primer. The thermal cycler protocol started with an initial denaturation at 95 ^{o}C for 3 minutes and then 25 cycles of 30 s at 95 ^{o}C, 45 s at 55 ^{o}C, 90 s at 72 ^{o}C and 45 s at 55 ^{o}C. Final extension at 72 ^{o}C for 10 min. PCR products were purified using the QiaQuick PCR purification kit (Qiagen), eluted elution buffer (25 μL), and quantified using PicoGreen assay (Life Technologies). SSU rDNA and rRNA amplicons were pooled at 100 ng for each sample. Emulsion PCR and sequencing of the PCR amplicons were sequenced on Roche 454 GS FLX Titanium at the Department of Energy Joint Genome Institute (DOE-JGI), or the McGill University and Génome Québec Innovation Center.

A total of 1,027,601 small subunit ribosomal rDNA and rRNA pyrotags were processed together using the Quantitative Insights Into Microbial Ecology (QIIME) software package (Caporaso *et al*., 2010). Reads with length shorter than 200 bases, ambiguous bases, and homopolymer sequences were removed prior to chimera detection. Chimeras were detected and removed using chimera slayer provided in the QIIME software package. Sequences were then clustered into operational taxonomic units (OTUs) at 97% identity using identity using UCLUST with average linkage algorithm. Prior to taxonomic assignment, singleton OTUs (OTUs represented by one read) were omitted, leaving 29,589 OTUs. Representative sequences from each non-singleton OTU were queried against the SILVA database release 111 using the BLAST algorithm (Altschul *et al*., 1990).

3.2.6 Statistical analysis and data visualization

Statistical analyses were conducted using the R software package (RCoreTeam, 2013). Pyrotag datasets were normalized to the total number of reads per sample. Hierarchical cluster analysis (HCA) and nonmetric multidimensional scaling (NMDS) were conducted to identify community compositional profiles associated

with water column compartments using the pvclust (Suzuki and Shimodaira, 2015) and MASS (Venables and Ripley, 2002) packages with Manhattan Distance measures, and statistical significance to the resulting clusters as bootstrap score distributions with 1,000 iterations and NMDS stress value <0.05.Diversity indexes (Shannon and alpha diversity) were calculated to identify changes in community structure based on filtration parameters using the vegan (Oksanen *et al.*, 2015) package. Non-parametric Friedman test block design was conducted to test the significance of the volume variation on the community composition using the stats (RCoreTeam, 2013) package. In addition, one-way ANOVA was conducted to test the significance of filtration combinations on taxa relative abundance using the ggpbur (Kassambara, 2017) package.

Multi-level indicator species analysis (ISA) was conducted to identify OTUs specifically associated with different filtering parameters based on groups resolved in HCA using the indicspecies package (De Caceres and Legendre, 2009). The ISA/multi-level pattern analysis calculates p values with Monte Carlo simulations and returns indicator values (IV) and *p*-values with α <0.05. The IVs range between 0 and 1, where indicator OTUs considered in the present chapter for further community analysis shown an IV >0.7 and *p*-value <0.001. ISA groups abundance was visualized as dot plots using the *bubble.pl* pearl script (http://hallam.microbiology.ubc.ca/LabResources/Software.html). Taxonomic distribution of identified OTUs was visualized using the ggplot2 (Wickham, 2009) package. The total SSU rRNA: rDNA ratios were calculated for the subset of matching samples (165 m 250 mL and 2.5 L on-ship 0.22 μm filters with in-line 0.4 μm pre-filtration, and 185 m 500 mL and 2.5 L on-ship 0.22 μm filters with in-line 0.4 μm pre-filtration) to account for variation in taxon abundance in the DNA pool (Frias-Lopez *et al.*, 2008; Stewart *et al.*, 2012b) and compared for a subset of microbial groups to explore how filtration parameters influence recovery of potentially active OTUs. I then selected OTUs based on ISA results and their shifts in abundance among filtering conditions.

3.2.7 Data deposition

SSU rDNA and rRNA pyrotag sequences reported in this chapter have been submitted to the The National Center for Biotechnology Information (NCBI) under BioSample numbers: SAMN05392373 - SAMN05392466.

3.3 Results

3.3.1 Water column conditions

Samples were collected during a stratification period characteristic of summer months (June-August) in Saanich Inlet (Carter, 1932; 1934; Herlinveaux, 1962; Zaikova *et al.*, 2010). Below 150 m, water column CTD O_2 concentrations were below <3 μM dissolved O_2 consistent with water column anoxia during peak stratification (Fig 3.1A; Table A.1). In addition, increasing levels of H_2S (13.95 μM) and NH_4^+ (6.1 μM) at 185 m were also observed indicating anoxic and sulfidic conditions in deep basin waters. NO_3^- and NO_2^- concentrations peaked at 150 m reaching 12 μM and 0.6 μM, respectively. Phosphate concentrations ranged between 4.5 and 5.8 μM from 150 to 185 m, and SiO_2 concentration peaked at 185 m reaching 110 μM (Table A.1).

3.3.2 Benchmarking workshop and Saanich Inlet time-series results

I evaluated microbial community structure using 521 time-series samples traversing the Saanich Inlet water column (Fig. A.1) and 29 samples collected during the SCOR workshop using rDNA pyrotag sequences to compare and cross-calibrate *in situ* sampling with the McLane PPS system and bottle sampling methods. Non-metric multidimensional scaling indicated workshop samples clustered together primarily with high-resolution suboxic and anoxic samples from 165 to 200 m depth intervals (0.22 μm Sterivex filters without pre-filtration) collected during summer months (Fig. 3.2A). Similarities between time-series and workshop samples provided an internal check on experimental design and a rationale for examining more granular differences between community structure and potential activity resulting from different filtration parameters.

Workshop samples collected at 165 and 185 m depth intervals formed three groups in NMDS analyses associated with on-ship 0.4 μm filters (group I), on-ship 0.22 μm filters with either 0.4 or 2.7 μm in-line pre-filtration (group II), and *in situ* 0.4 μm filters (group III) (Fig. 3.2B) were resolved based on NMDS analysis. For the most part, samples within groups partitioned by depth with small variation between sample volumes from the same depth. Consistent with NMDS, HCA resolved three groups (AU $>$ 70, 1000 iterations) associated with on-ship 0.4 μm filters (group I), on-ship 0.22 μm filters with in-line 0.4 μm pre-filters (group II), and *in situ* filtration 0.4 μm filters (group III) (Fig. 3.3). Within each group, samples partitioned by depth and filtration volume. Friedman block test results indicated that differences in community structure driven by sample volume were not significant.

3.3.3 Size-fractionation effects on community structure

Based on NMDS and HCA results, I focused on changes in OTU relative abundance and taxon identity between groups. Microbial community structure was primarily comprised of OTUs (>0.1% relative abundance) affiliated with ubiquitous OMZ taxa including Marine Group A, SAR11, SAR324, SUP05 (Field *et al*., 1997; Fuhrman and Davis, 1997; Brown and Donachie, 2007; Tripp *et al*., 2008; Lam *et al*., 2009; Walsh *et al*., 2009; Zaikova *et al*., 2010; Walsh and Hallam, 2011; Wright *et al*., 2012) as well as Bacteroidetes, Desulfobacterales, and Euryarchaeota (Fig. 3.3B). Interestingly, several of these groups were not detected in on-ship 0.4 μm pre-filter samples but were recovered from in-line 0.22 μm filters (group II). These included Marine Benthic Group E and Halobacteriales, SAR406 within the Marine Group A, Methylophilales within the Betaproteobacteria, Desulfuromonadales and Desulfarculales within the Deltaproteobacteria and SUP05 within the Gammaproteobacteria. Conversely, Acidimicrobiales within the Actinobacteria, Cyanobacteria, Rhodobacterales within the Alphaproteobacteria, OM190 within the Planctomycetes, and eukaryotic phyla including Cnidaria and Arthropoda within the Opisthokonta and Diatoms within the Stramenopiles were detected in on-ship 0.4 μm filter samples but not recovered on in-line 0.22 μm filters or *in situ* 0.4 μm filters. Eukaryotic phyla affiliated with Alveolata were recovered on 0.22 μm filter samples with in-line 0.4 μm pre-filters (group I) and Phycisphaerales within the Planctomycetes were detected in *in situ* 0.4 μm filter samples (group III), respectively (Fig. 3.3B).

Filtration methods, including the use of different pre-filters and volumes, resulted in a significant source of variation (p <0.001) for the relative abundance of several bacterial phyla including Bacteroidetes, Deferribacteres, Alpha-, Delta- and Gammaproteobacteria, Planctomycetes, archaeal phyla including Thaumar-

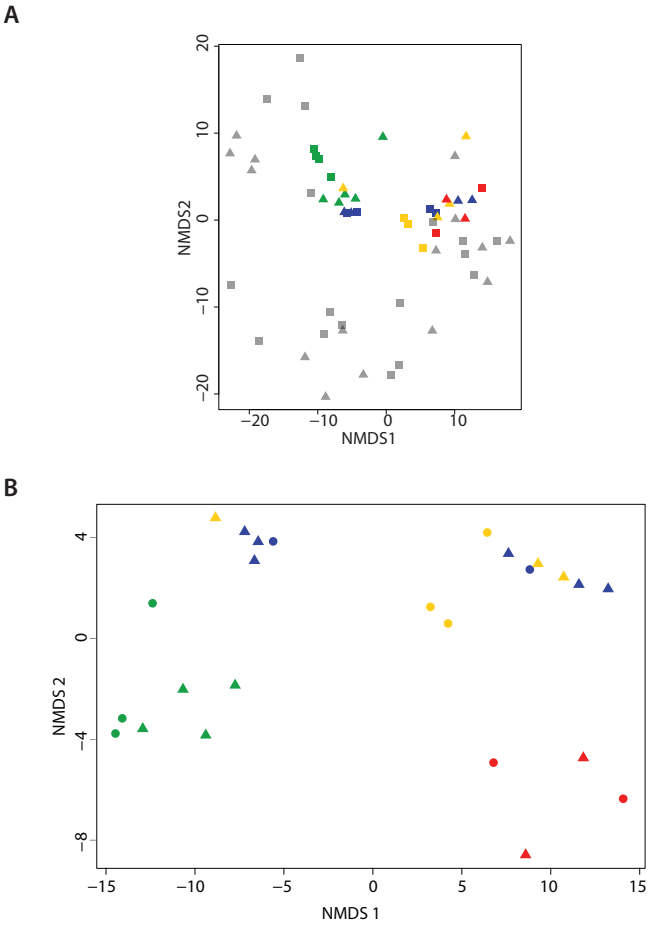

Figure 3.2: Cluster analysis of Saanich Inlet time-series and SCOR workshop tag data. Cluster analysis of small subunit ribosomal RNA (SSU rDNA) gene pyrotag data for the Saanich Inlet time-series (2006-2011) and SCOR workshop samples collected at 165 and 185 m. A) NMDS based on Manhattan distance (1000 iterations) of time-series and SCOR pyrotag data showing microbial community partitioning based on oxygen gradients spanning suboxic (20-1 μmolO$_2$ kg-1) and anoxic-sulfidic ($<$1 μmol O$_2$ kg^{-1}). B) NMDS based on Manhattan distance (1000 iterations) of SCOR pyrotag data showing microbial community partitioning based on filtering conditions. In both biplots samples are depicted by depth (165m = square and 185m = triangle) and filter type (PPS *in situ* 0.4 μm (red), MPP on-ship 0.4 μm (green), 0.22 μm pre-filtered onto 0.4 μm filters(yellow), 0.22 μm pre-filtered onto 2.7 μm filters (blue), and time-series 0.22 μm (black)).

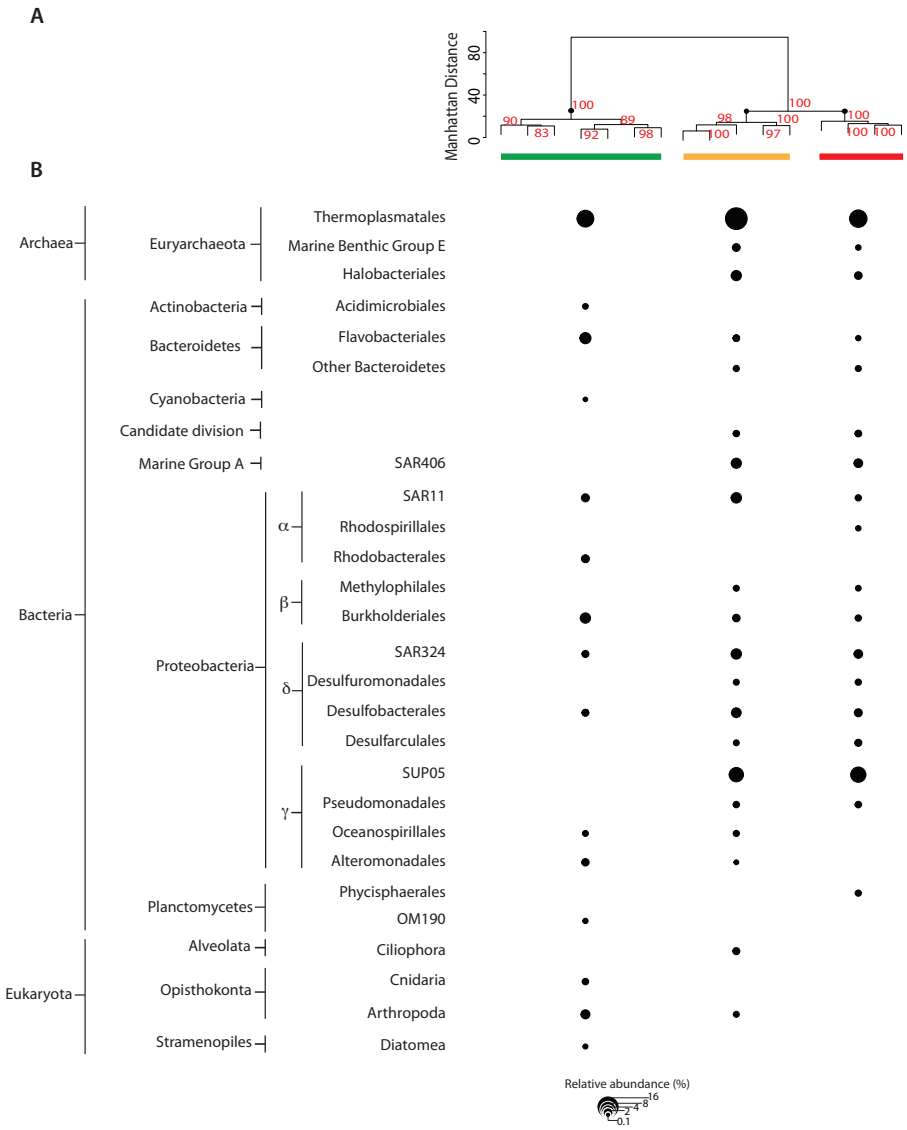

Figure 3.3: Microbial community partitioning based on filtration methods Top: Cluster analysis (AU>70, 1000 iterations) for 0.4 μm filter fractions. Samples are depicted with a colored bar by filter type (PPS *in situ* 0.4 μm (red), and MPP on-ship 0.4 μm (green) and 0.22 μm pre-filtered onto 0.4 (yellow). Bottom: Abundant taxa (>0.1% relative abundance from total reads in sample) observed among filter combinations for PPS *in situ* 0.4 μm (red), and MPP on-ship 0.4 μm (green) and 0.22 μm pre-filtered onto 0.4 μm filters (yellow). The size of dots depicts the relative abundance for each taxa as indicated in key.

Table 3.2: Microbial taxa with significant abundance differences among tested filtering conditions. Significant *p*-values ($p < 0.001$) obtained from one-way ANOVA testing for filtration methods effect on microbial relative abundance.

Taxa	*p*-value
Thaumarchaeota	0.0000033
Bacteroidetes	0.000051
Deferribacteres	0.001
α-proteobacteria	0.01
δ-proteobacteria	$2.1e^{-8}$
γ-proteobacteria	0.0004
Planctomycetes	$4.1e^{-8}$
Opisthokonta	0.01
Rhizaria	0.01

chaeota, and eukaryotic phyla including Opisthokonta and Rhizaria (Table 3.2 and Fig. 3.4). For example, the relative abundance of Bacteroidetes (Flavobacteriales), Alphaproteobacteria (SAR11, Rhodobacterales and Rhodospirillales) and Opisthokonta (Maxillopoda) associated with on-ship 0.4 μm filters increased 5-fold compared to 0.22 μm filters with in-line 0.4 μm pre-filters and *in situ* 0.4 μm filters, while the relative abundance of Deferribacteres, Deltaproteobacteria (SAR324, Desulfobacterales and Desulfarculales) and Gammaproteobacteria (Oceanospirillales mainly affiliated with SUP05, Pseudomonadales and Alteromonadales) associated with on-ship 0.4 μm filters decreased 5-fold compared to 0.22 μm filters with in-line 0.4 μm pre-filters and in situ 0.4 μm filters (Fig. 3.4). Conversely, the relative abundance of Planctomycetes (Phycisphaerales and OM190) associated with *in situ* 0.4 μm filters increased 5-fold compared to on-ship 0.22 μm filters with in-line 0.4 μm pre-filters and 0.4 μm filters (Fig. 3.4).

Together these results indicate that filter selection and in-line positioning can introduce bias into microbial community structure data and reinforce the idea that filtration methods should be taken into consideration more carefully when interpreting microbial count data.

3.3.4 Size fractionation effects on indicator OTUs (DNA analyses)

To identify OTUs associated with specific filtration methods I conducted multi-level indicator species analysis (ISA) on HCA groups I-III. As expected, resulting indicator OTUs varied with respect to filtration methods used (Fig. 3.5). The largest differences with respect to indicators were detected between *in situ* and on-ship 0.4 μm filter samples. Indicator OTUs detected in on-ship 0.4 μm filter samples were mostly affiliated with bacterial phyla including Actinobacteria, Bacteroidetes, Cyanobacteria, Deferribacteres, Proteobacteria, and Verrucomicrobia, archaeal phyla including Euryarchaetoa, and eukaryotic phyla including Alveolata, Opisthokonta, Rhizaria and Stramenopiles (Fig. 3.5). Indicator OTUs detected in *in situ* 0.4 μm filter samples were mostly affiliated with bacterial phyla including Candidate divisions (WS3, OD1 and BRC1), Chloroflexi, Deferribacteres, Firmicutes, Lentisphareae, Nitrospirae, Alpha-, Beta-, Delta- and Gammaproteobacteria, and Planctomycetes, archaeal phyla including Euryarchaeota, and eukaryotic phyla including Alveolata, Excavata and Opisthokonta (Fig. 3.5). Indicator OTUs detected in 0.22 μm filter sam-

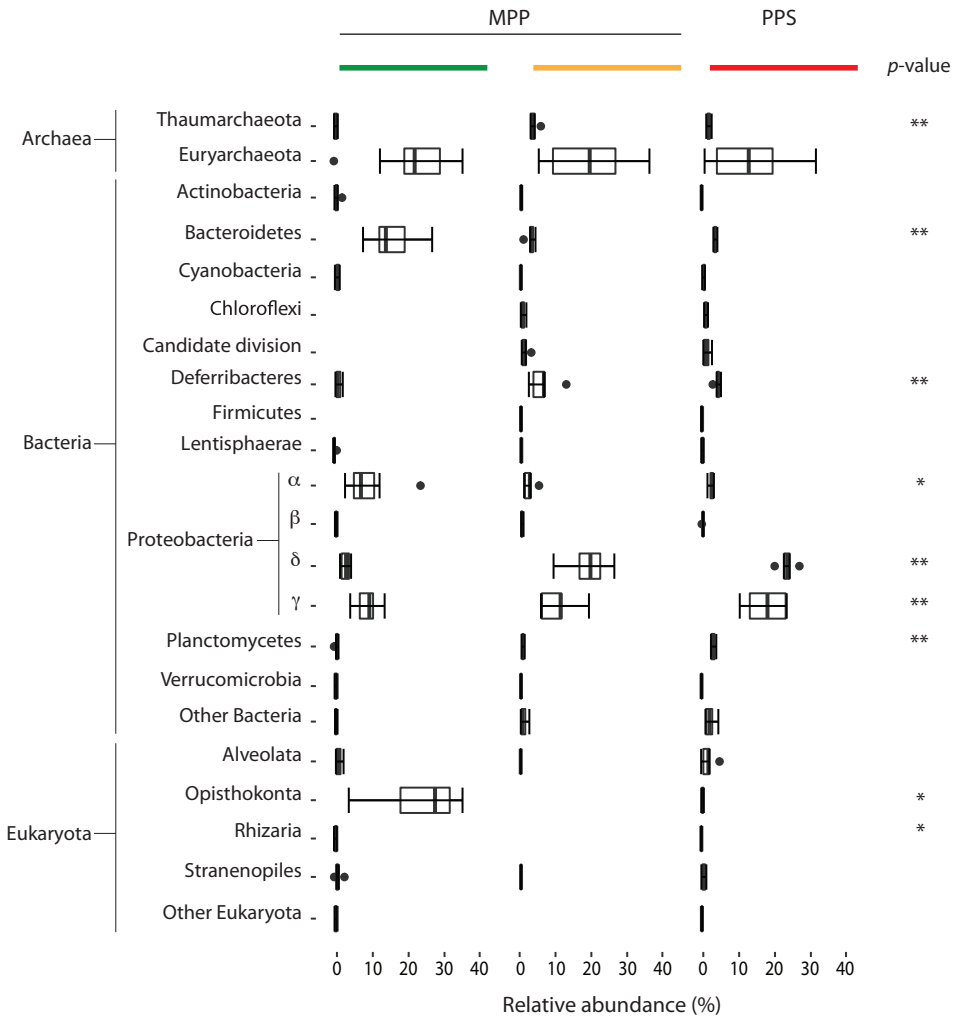

Figure 3.4: Microbial community abundance shifts in relation to filtration methods. Community shifts occurred among and within filter combinations for PPS *in situ* 0.4 μm (red), and MPP on-ship 0.4 μm (green) and 0.22 μm pre-filtered onto 0.4 μm filters (yellow). The size of each box represents the average of the percentage of relative abundance throughout the water column over this period. For both plots, extended dashed lines (whiskers) represent at the base the lower and upper quartiles (25% and 75%) and at the end the minimum and maximum values encountered. The middle line represents the median. p-values for significant relative abundance shifts among filter groups are depicted as $p < 0.01$ (*) and $p < 0.001$(**) aside corresponding taxa.

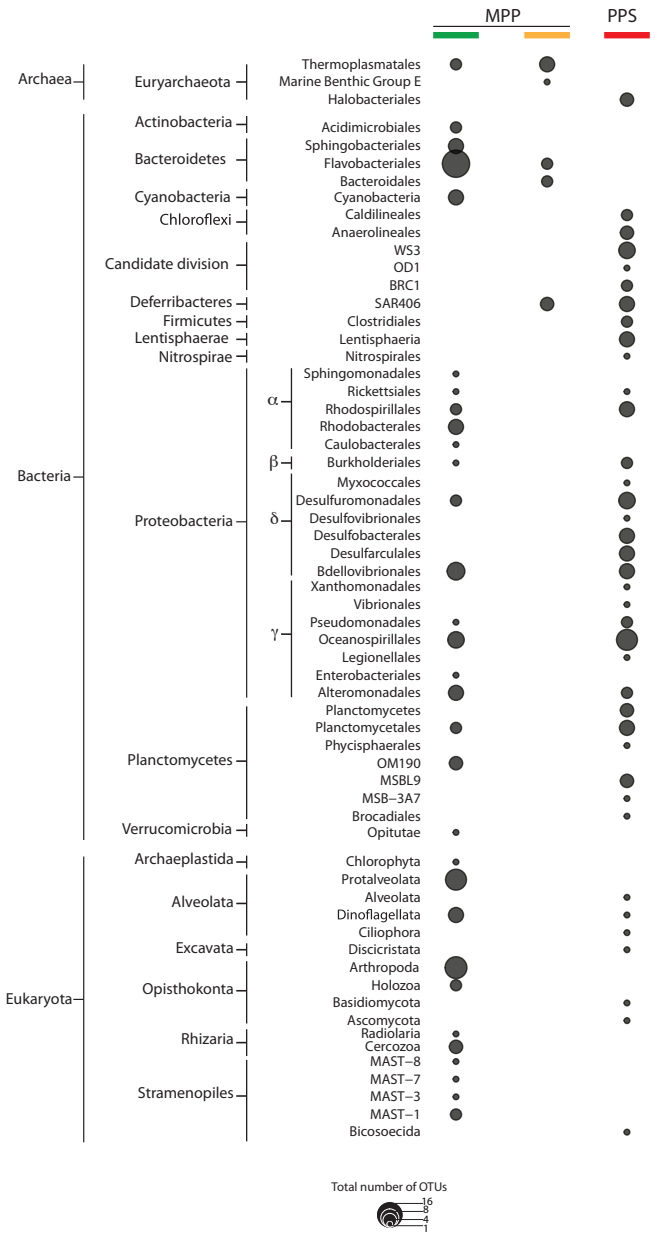

Figure 3.5: Indicator OTUs associated with specific filtration methods. Indicator OTUs for filter groups PPS *in situ* 0.4 μm (red), and MPP on-ship 0.4 μm (green) and 0.22 μm pre-filtered onto 0.4 μm filters (yellow). The size of dots depicts the total number of indicator OTUs affiliated to specific taxa.

ples with in-line 0.4 μm pre-filter were mostly affiliated with bacterial phyla including Deferribacteres and Bacteroidetes, and archaeal phyla including Euryarchaeota (Fig. 3.5).

The differences observed between in situ and on-ship indicator OTUs reinforce the effect of size fractionation on microbial community structure and raise important questions about metabolic reconstruction efforts based solely on on-ship filtration methods.

3.3.5 Size-fractionation effects on expressed OTUs within specific populations (rRNA analyses)

To further evaluate the impact of size fractionation on detection of active microbial groups I compared SSU rRNA:rDNA ratios of OTUs between 0.4 μm filters collected and preserved *in situ* vs. on-ship 0.22 μm filters with in-line 0.4 μm pre-filters. I focused on OTUs exhibiting ratios >1 as a proxy for cellular activity (Blazewicz *et al*., 2013). Ratios for Candidate divisions, Desulfobacterales, SUP05, Phycisphaerae, and Halobacteria were highest in 0.4 μm *in situ* filter samples while SAR11, Rhodospirillales, Methylophilales and Burkholderiales within Betaproteobacteria, and Verrucomicrobia and eukaryotic phyla affiliated with Alveolata, Opisthokonta, Rhizaria and Stramenopiles were highest in on-ship 0.22 μm filter samples with in-line 0.4 μm pre-filtration (Fig. A.3).

I detected OTUs affiliated with SUP05, Marine Group A, SAR11 and SAR324 that showed ratios >1, with differential expression between *in situ* and on-ship filters. For instance, a total of 4 SUP05 OTUs with ratios ranging from 1-2, and 3 SAR324 OTUs with ratios equal to 3 were exclusively detected in 0.4 μm *in situ* filter samples (Fig. 3.6). I also observed 6 Marine Group A OTUs with ratios equal to 2 in on-ship 0.22 μm filter samples with in-line 0.4 μm pre-filtration and 1 exclusively active *in situ* OTU (Fig. 3.6).

Interestingly, I observed 8 SAR11 OTUs with ratios ranging from 1-3 exclusively in on-ship 0.22 μm filter samples with in-line 0.4 μm pre-filtration (Fig. 3.6). Candidate divisions BCR1 and WS3, Delta- and Gammaproteobacteria, and Planctomycetes OTUs also manifested higher ratios in 0.4 μm *in situ* filter samples than in on-ship samples (Fig. 3.7). These differences showed some depth specificity. For example, I observed BCR1 and WS3 OTUs with high ratio values at 165 m while Desulfovibrionales and Desulfarculales within the Deltaproteobacteria had the highest ratio values at 185 m (Fig. 3.7). Similarly, most OTUs affiliated with Planctomycetes (Phycisphaerae, OM190, Brocadiales and Plactomycetales) had the highest ratio values at 185 m (Fig. 3.7). In contrast, Flavobacteriales within Bacteroidetes and Alphaproteobacteria had higher ratio values in on-ship 0.22 μm filter samples with in-line 0.4 μm pre-filtration than 0.4 μm *in situ* sampled at both 165 and 185 m (Fig. 3.7).

3.4 Discussion

In this chapter I used SSU rDNA and rRNA count data generated during the SCOR Working Group 144 Microbial Community Responses to Ocean Deoxygenation workshop to determine the effects of collection and filtration methods on microbial community structure and potential activity in the anoxic water column of Saanich Inlet. Observed differences in microbial community structure and potential activity associated with *in situ* versus on-ship size fractionation suggest potential sources of error when linking field processes to microbial agents based on genomic sequence information in isolation. In particular, *in situ* results detected several microbial groups implicated in the sulfur-cycle that are underrepresented in publicly available amplicon and shotgun sequencing data sets (Hawley *et al*., 2017). Overall, results from this study provide useful information on how different sampling methods can contribute to bias in experimental outcomes and reinforce the need for more integrated studies based on standardized sampling protocols that increasingly incorporate *in situ* measurements.

Microbial community shifts associated with size fractionation

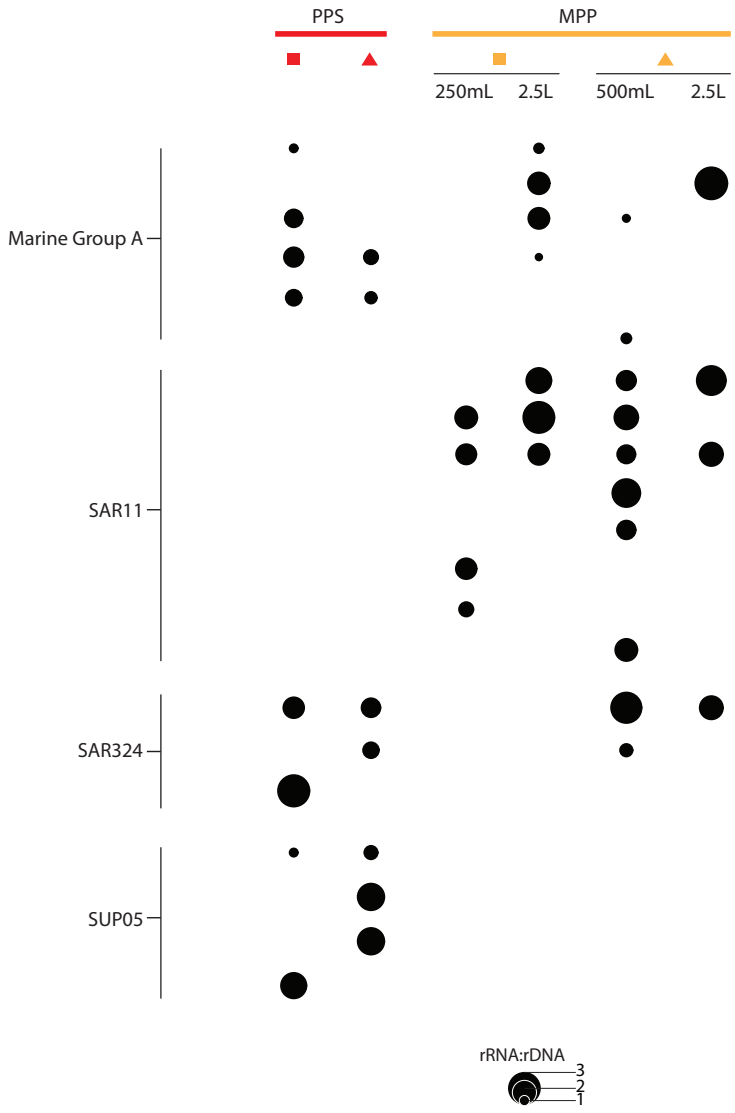

Figure 3.6: Activity differences for abundant OTUs in relation to filtration methods. Small subunit (SSU) rRNA: rDNA ratio for abundant and ubiquitous taxa in OMZs i.e SUP05, SAR406, SAR11 and SAR324, observed at PPS *in situ* 0.4 μm (red) and MPP on-ship 0.22 μm pre-filtered onto 0.4 μm filters (yellow). The size of dots depicts ratio values for individual OTUs as indicated on plot size key.

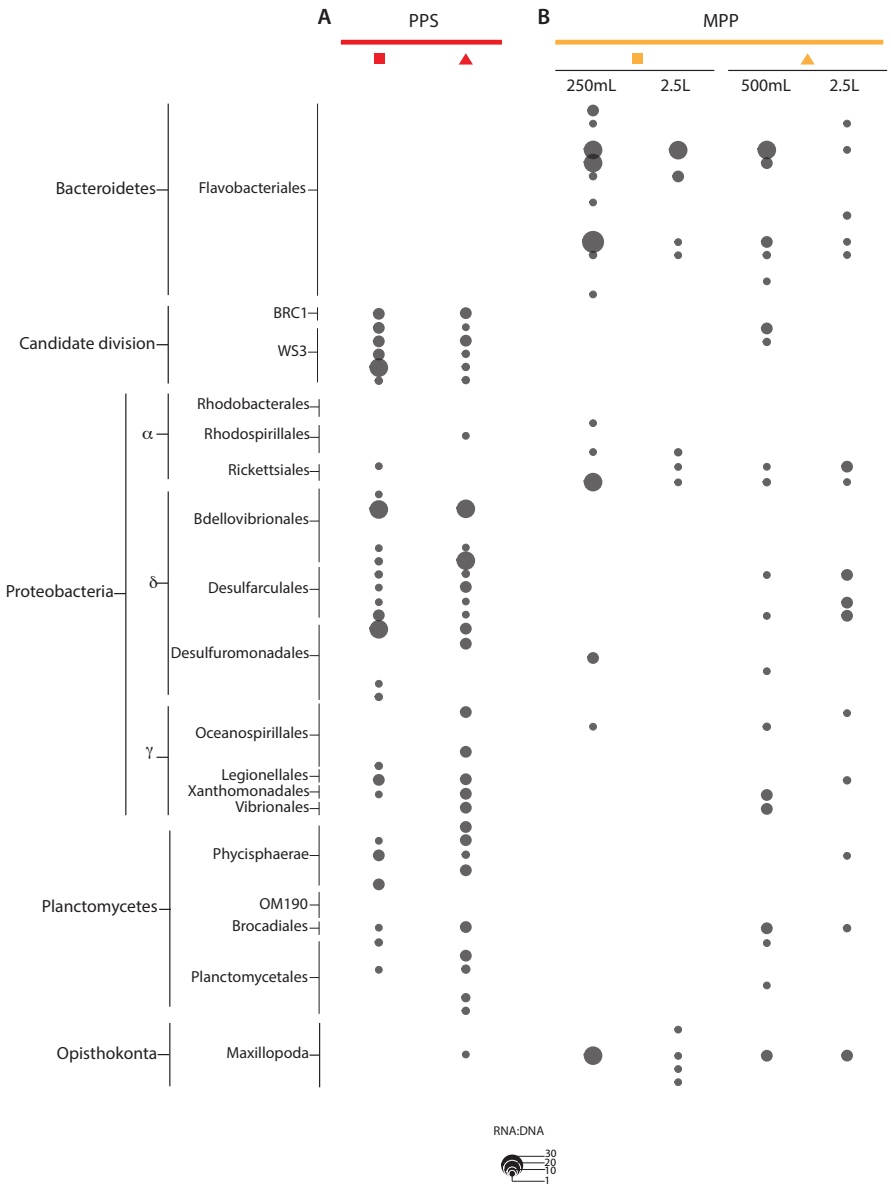

Figure 3.7: Activity differences for indicator OTUs in relation to filtration methods. Small subunit (SSU) rRNA: rDNA ratio for indicator OTUs observed at PPS *in situ* 0.4 μm (red) and MPP on-ship 0.22 μm pre-filtered onto 0.4 μm filters (yellow). The size of dots depicts ratio values for individual OTUs as indicated on plot size key.

Understanding how microorganisms interact within the ocean at different scales is integral to linking microbial food webs to nutrient and energy flow processes (Azam and Malfatti, 2007). Particles play a salient role in structuring microbial community interactions and the interplay between "particle-associated" and "free-living" microbiota creates a dynamic metabolic network driving biogeochemical transformations (Smith *et al.*,*et al.*, 1992;DeLong *et al.*, 1993;Crump *et al.*, 1999;Simon *et al.*, 2002;Grossart, 2010;Ganesh *et al.*, 2014). Previous observations from OMZ waters implicate particle maxima as hotspots for metabolic coupling (Garfield *et al.*, 1983;Naqvi *et al.*, 1993;Whitmire *et al.*, 2009;Ganesh *et al.*, 2014). However, the definition of "particle-associated" versus "free-living" can sometimes seem arbitrary and the degree to which microorganisms alternate between these two fractions in different water compartments is not firmly established. Typically, anything >0.4 μm has been considered particle associated (Azam and Malfatti, 2007) although most studies use a 0.2-1.6 or 2.7 μm cut-off to concentrate microbial biomass. In the current chapter I compared 0.4 μm *in situ* filtration without a pre-filtration step to on-ship 0.22 μm filtration with in-line 0.4 μm pre-filtration in order to evaluate "particle-associated" versus "free-living" fractions.

The core microbial community detected *in situ* versus on-ship was similar to time-series observations in suboxic-anoxic water column compartments (165-185 m) during summer months. However, at a more granular OTU level important differences in community structure and potential activity for a number of microbial groups were resolved that reinforce and expand on previous size fractionation studies in open ocean OMZs. For example, Padilla and colleagues have shown that mode and magnitude of sampling bias depends on filter type and pore size, particle load, and community complexity (Padilla *et al.*, 2015). In the present chapter, sample volume had a non-significant effect (p >0.05) on microbial diversity although wire time and filtration duration likely impacted particle size, as did potential bottle effects due to settling when processing on-ship samples (Fig. A.2). Several studies have shown that particles can settle in sampling bottles on timescales relevant to on-ship processing (Gardner, 1977;Suter *et al.*, 2016).

With respect to size fractionation, community structure differences were driven by shifts in abundance and activity of many known microorganisms. For example, as observed in the ETNP OMZ, OTUs affiliated with Deferribacteres were enriched in the smaller size fraction (<0.4 μm) consistent with an autotrophic lifestyle (Ganesh *et al.*, 2014). Similarly, indicator OTUs affiliated with Bacteroidetes (Bacteroidales and Flavobacteriales), Lentisphareae, Deltaproteobacteria (Myxococcales and Desulfobacterales), Planctomycetes, and Verrucomicrobia were enriched in the larger size fraction (>0.4 μm) as observed in both ETNP and ETSP OMZs, consistent with attachment to sinking aggregates or zooplankton (Crump *et al.*, 1999;Simon *et al.*, 2002;Eloe *et al.*, 2011;Allen *et al.*, 2012;Fuchsman *et al.*, 2012;Ganesh *et al.*, 2014;Padilla *et al.*, 2015). These similarities transcended domain boundaries with eukaryotic phyla including Dinoflagellata (Alveolata), Radiolaria (Rhizaria) and Syndiniales (Stramenopiles) enriched in the 0.4 μm filter samples (Guillou *et al.*, 2008;Duret *et al.*, 2015).

With respect to *in situ* versus bottle collection methods, previous studies have identified changes in community gene expression profiles (Feike *et al.*, 2012;Stewart *et al.*, 2012a) and process rate measurements (Taylor and Doherty, 1990;Stewart *et al.*, 2012a;Taylor *et al.*, 2015;Edgcomb *et al.*, 2016). Here I identified changes in microbial community structure that potentially explain variance in gene expression or process rates. For example, indicator OTUs detected in *in situ* samples indicate a previously unrecognized

role for sulfate-reducers and candidate divisions WS3, OD1 and BRC1 in the Saanich Inlet water column. Previous studies have implicated WS3 and OD1 metabolism in sulfur cycling and methanogen provisioning (Kirkpatrick *et al*., 2006;Wrighton *et al*., 2012). Similarly, Desulfobacterales and Desulfovibrionales, are commonly underestimated in abundance in OMZs waters (Suter *et al*., 2016) as they are more prevalent on particles that likely settle during on-ship processing.

Working in the Cariaco Basin OMZ, Suter and colleagues provide a compelling description of bottle settling rates (approximately 12 min for a 1mm particle to sink into below-spout space of an 8 L Niskin bottle or 18 min for a 12 L bottle) that can result in sampling bias (Suter *et al*., 2016). On-ship processing during the SCOR workshop took between 20-30 minutes. Shipboard processing times can often be even longer than this. In addition, wire-time and turbulence associated with sampling moment and vibration (Suter *et al*., 2016), and filtration across the membrane (Duret *et al*., 2015) can impact particle size and stability in bottles e.g. production of smaller particles derived from larger aggregates. Based on sinking rates estimated in the Cariaco Basin it is possible that observed community structure differences between *in situ* and on-ship samples in Saanich Inlet could be explained by a combination of particle settling and turbulence associated with sample collection and filtration. This could affect our perception on the microbial metabolic network with respect to transient spatial interactions that are altered or disrupted using on-ship methods.

Implications of size fractionation for inferring microbial activity in OMZs

In contrast to only examining rDNA sequences, combining those with analysis of rRNA sequences can provide a robust proxy for past, and present or emerging cellular activities (Blazewicz *et al*., 2013) that can inform hypotheses related to life strategies and metabolic interactions within microbial communities (Lepp and Schmidt, 1998;Barnard *et al*., 2013). Here, I considered rRNA:rDNA ratio values >1 as an indicator for potentially active microbial community members. While, using ratios to infer activity at higher taxonomic levels e.g. Phylum, Class, Order, can promote inconsistent results (Blazewicz *et al*., 2013;Ganesh *et al*., 2015), focusing on the OTU level can identify ecologically relevant populations with the potential to play integral roles in nutrient and energy cycling within the ecosystem under study. For example, SUP05 has been determined to be an abundant member of the Saanich Inlet microbial community comprising between 20-30% of total bacteria at 165 and 185 m, respectively (Walsh *et al*., 2009; Zaikova *et al*., 2010; Walsh and Hallam, 2011). I detected 85 OTUs affiliated with SUP05 based on rDNA sequences. However, only 4 had rRNA:rDNA ratio values >1 indicating population level variation in potential activity.

Consistent with previous observations in the ETSP using metatranscriptomic data (Padilla *et al*., 2015), examination of the taxonomic affiliation of indicator OTUs produced using *in situ* versus on-ship methods identified differences between the active microbial community in samples. Some candidate divisions recovered in *in situ* samples have not been previously well described in the Saanich Inlet water column based on rDNA sequences due to their low abundance ($<0.1\%$). Interestingly, the rRNA:rDNA ratios observed for indicator WS3 and BCR1 OTUs (ratios equal to 7 and 2, respectively) were greater than those observed for OTUs affiliated with ubiquitous and abundant taxa, including SUP05. Similar observations were made for OTUs affiliated with Deltaproteobacteria, Chloroflexi, Firmicutes, Lentisphaera, Nitrospina and Marine Group A, reinforcing the idea that multi-omic sequences and process rate measurements sets sourced from on-ship samples have the potential to underestimate the contribution of some active microbial groups present

in the water column. Such groups may be sensitive to settling, turbulence or other factors including oxygen exposure, necessitating in situ sampling to reveal their contributions to the metabolic network.

Hawley and colleagues used metaproteomics to develop a conceptual model of coupled carbon, nitrogen and sulfur cycling (Hawley *et al*., 2014). Louca and colleagues incorporated these ideas into a numerical model integrating multi-omic sequence and geochemical information to predict metabolic fluxes and growth yields under near steady-state conditions (Louca *et al*., 2016). Both conceptual and numerical models were based on interactions between Thaurmarchaeota, SAR11, SUP05 and Planctomycetes (Walsh *et al*., 2009; Zaikova *et al*., 2010; Walsh and Hallam, 2011; Wright *et al*., 2012). Although, the metabolic potential of WS3, OD1 and sulfate-reducing Deltaproteobacteria in the Saanich Inlet water column remains to be determined, the potential role of these groups at the nexus of sulfur cycling and methanogenesis (Kirkpatrick *et al*., 2006) presents an opportunity for new hypothesis development and testing to integrate these groups into prevailing conceptual and numerical models for coupled biogeochemical cycling.

3.4.1 Conclusion

As the research community transitions away from descriptive studies of marine microorganisms to more quantitative comparisons at ecosystem scales integrating multi-omic information with process rates and modeling, the need for standards of practice that reduce sampling bias becomes increasingly important. Moreover, a general lack of consensus related to assignments to particle-associate versus free-living size fractions further confounds robust cross-scale comparisons. Results from this chapter suggest that *in situ* sampling approaches have the potential to limit many biases by providing a more authentic representation of microbial activity than on-ship sampling methods. At the same time, consistent on-ship methods need to be established that limit bottle effects and harmonize filtration practices for effective cross-scale comparisons. Several promising devices such as the PPS (Edgcomb *et al*., 2016), Environmental Sample Processor (ESP) (Jones *et al*., 2008;Preston *et al*., 2009;Ottesen *et al*., 2011;Robidart *et al*., 2014), Automatic Flow Injection Sampler (AFIS) (Feike *et al*., 2012), and Clio (Jakuba *et al*., 2014) have been developed with the potential to support *in situ* sampling under a variety of operational scenarios. For example, recent studies with the ESP have enabled dynamic intermittent sampling during light dark cycles in surface waters revealing conserved patters of gene expression on ocean basin scales (Ottesen *et al*., 2014;Aylward *et al*., 2015). Although community adoption of these new technologies remains in early stages due in part to accessibility, price point, and operating constraints, these devices and their "descendents" likely reflect the future of microbial sampling in the ocean given their autonomous and programmable designs extensible to time series or event response monitoring. Based on the information provided, I recommend a replicated study of different *in situ* sampling technologies that incorporates multi-omic sequencing and measurements of process rates focused on coupled carbon, nitrogen and sulfur cycling in coastal and open ocean OMZs.

Chapter 4

Protistan parasites along water column oxygen gradients: a network approach to assessing potential host-parasite interactions[1]

This chapter introduce oxygen (O_2) effects on ecological interactions and energy flow in aquatic ecosystems. Using methodologies from Chapter 3 for small subunit ribosomal (SSU) DNA and RNA tags I established eukaryotic community structure and interactions highlighting the potential impact parasitic interactions may have on population dynamics, extensible to nutrient cycling process, during seasonal water column stratification in Saanich Inlet. In addition, I explored correlation and network analyses, and visualization tools using time-series tag data that form the basis for community-level interactions analyses carried out in Chapters 5 and 6. This chapter is the basis for interpreting microbial community structure, composition and interactions that serve as framework for constructing microbial community networks with potential metabolic implications for nutrient cycling throughout water column O_2 compartments.

4.1 Introduction

As oxygen (O_2) levels decline, energy in oxygen minimum zones (OMZs) is increasingly diverted into microbial community metabolism with resulting feedback on metazoan organisms through habitat compression. This has a disproportionate impact on large multicellular eukaryotes that cannot permanently occupy anoxic OMZ cores (Parris *et al*., 2014). Despite this tendency, some vertically-migrating crustacea, zooplankton, chaetognaths and fish are known to seek temporary refuge from predation in OMZ waters (Wishner *et al*., 1998; Escribano *et al*., 2009; Wishner *et al*., 2013). Moreover, a diverse community of microeukaryotes and zooplankton can operate in OMZ waters where they help shape biogeochemistry and

[1] A version of this chapter appears as **Torres-Beltrán, M**., Sehein, T. et al. 2018. Protistan parasites along oxygen gradients in a seasonally anoxic fjord: a network approach to assessing potential host-parasite interactions. Deep Sea Res. II. doi.org/10.1016/j.dsr2.2017.12.026

ecology. Protists in particular, participate in wide-ranging interactions that couple metabolic processes and behaviour to nutrient and energy flow patterns in the environment (Arstegui *et al*., 2009).

Throughout marine water columns and sediments, protists shape pools of bioavailable carbon and other nutrients through production, grazing and symbiotic associations. Phagotrophic protists and viruses are considered as the main sources of mortality for marine organisms (Suttle, 2005; Arstegui *et al*., 2009). Organic carbon in the dissolved fraction is transferred to higher trophic levels (to Metazoa) in a heterotrophic food chain involving bacteria, small and large nanoflagellates (2-20 μm), and larger flagellates and ciliates, termed the "microbial loop" (Azam *et al*., 1983; Taylor GT, 1986). The relative contributions of protist grazing and viral lysis as top-down controls on marine bacterioplankton are still debated (e.g. (Pedrs-Ali *et al*., 2000; Cuevas and Morales, 2006; Chow *et al*., 2014), and undoubtedly vary depending on site-specific physico-chemical conditions and the physiological state of individual populations. Diverse communities of phagotrophic and parasitic protists have been described under low oxygen conditions, many of which exhibit putative symbiotic relationships with prokaryotes, and protist grazing has been shown to shape specific prey populations along the redoxcline (Lin *et al*., 2007)(Edgcomb *et al*., 2011; Orsi *et al*., 2011, 2012; Wright *et al*., 2012; Parris *et al*. 2014; Jing *et al*. 2015).

Parasitic protists are taxonomically diverse and a major source of mortality for other microbial eukaryotes as well as Metazoa (Chambouvet *et al*., 2008), likely making them important shapers of food web structure (Jephcott *et al*., 2016). The Alveolata comprise one of the largest eukaryotic lineages, which includes four protist groups, the Ciliophora, Dinoflagellata, Protalveolata, and Apicomplexa, the last two of which are comprised primarily of parasites (Edgcomb, 2016). Exclusive marine parasitoid protists consistently recovered in high-throughput molecular studies were initially assigned to novel marine alveolate (MALV) Groups I and II within the Syndiniales (Lopez-Garcia *et al*., 2001; Moon-van der Staay *et al*., 2001; Massana *et al*., 2004b; Romari and Vaulot, 2004; Not *et al*., 2007; Edgcomb *et al*., 2011). Group I consists of numerous undescribed species while Group II contains sequences affiliated with the genus *Amoebophrya*, including species influencing "red tide" blooms (Chambouvet *et al*., 2008). Over time the original two groups have expanded to include five (Guillou *et al*., 2008) or eight (Richards and Bass, 2005) deeply diverging and abundant clades. For example, in euphotic zone samples from the Tara Oceans expedition 36% of protistan operational taxonomic units (OTUs) were attributed to parasites, almost half of which were affiliated with a single MALV Group I clade (de Vargas *et al*., 2015)(de Vargas *et al*. 2015). Molecular signatures of Syndiniales are almost universally retrieved from picoplankton ($<$3 μm) samples, although some clades within these groups appear to be restricted to specific habitats (Guillou *et al*., 2008). Variations in sequence abundances in ribosomal RNA- vs. DNA-based libraries, due to PCR amplification biases by taxon-specific rDNA copy number, may indicate variability in the impact of parasitism in different habitats (Not *et al*., 2009).

Infection of host cells by members of Syndiniales typically leads to host mortality. Known hosts include other Syndiniales, photosynthetic dinoflagellates, ciliates, rhizarians and metazoans (Guillou *et al*., 2008; Brte *et al*., 2012; Massana *et al*., 2014). The most studied syndinian parasites belong to the species complex, *Amoebophrya*. *Amoebophrya* infect new hosts by forming motile spores or weakly motile spores (dinospores) that penetrate the host cell membrane and travel to the host nucleus or cytoplasm to replicate

(Chambouvet *et al*., 2008). Over the course of several days, the trophont grows and ultimately ruptures the host cell membrane, releasing the ephemeral, multinucleate, (often) multiflagellate parasite vermiform into the water column, which subsequently fragments into individual dinospores (Chambouvet *et al*., 2008). In the case of some syndinian parasites (e.g. infection of tintinnid ciliates by *Eudubosquella* species), infection is thought to be a passive process, whereby the host ingests but does not digest the spores (Dolan, 2013). Other syndinian trophonts, such as those formed by *Eudubosquella* species emerge through a more elaborate process involving creation of a food vacuole encompassing the remains of the host cell, digestion of host remains, followed by serial nuclear and cytoplasmic divisions to form spores (Coats *et al*., 2012; Dolan, 2013). Increased syndinian parasite abundance during the decline of harmful algal blooms (i.e. *Alexandruim catanella*) suggest potential top-down control of phytoplankton species during bloom events (Bai *et al*., 2007; Velo-Surez *et al*., 2013; Choi *et al*., 2017). Studies of parasitism of other taxa suggest syndinian parasites infect their hosts throughout the year, but studies of ciliates indicate that infections of ciliate hosts are most common in summer months when phytoplankton biomass is apparently sufficient to support elevated grazer densities (Cachon, 1964; Coats, 1989; Coats *et al*., 1994). Thus, Syndiniales may have the potential to contribute significantly to releases of organic carbon through cell lysis during bloom termination events and may influence protist community diversity by exerting top-down control on select species within a population.

Fluorescence *in situ* hybridization (FISH) analyses has shown that Syndiniales alternate between two life stages: the free-swimming, infective dinospore that can survive without a host for up to three days, and the multinuclear trophont phase within a host (Velo-Surez *et al*., 2013). Most infections lead to lysis of the host cell; however, one study of the Syndiniales Group II taxon *Amoebophrya*, revealed dinospores could lie dormant within the resting stage of the dinoflagellate, Scrippsiella trochoidea and re-emerging simultaneously when the host entered the vegetative state, suggesting potentially complex life histories that can impact protist populations at multiple life stages (Chambouvet *et al*., 2011). In molecular surveys of protistan diversity, Syndiniales are grouped taxonomically with the novel marine alveolate (MALV) lineages I and II, and can form the majority of sequences from diverse marine habitats including coastal waters, open oceans, and stratified systems (Massana *et al*., 2004a; Not *et al*., 2007; Guillou *et al*., 2008). These parasites are thought to exhibit top-down pressures on host populations within the eukaryotic community, such as, radiolarians, dinoflagellates, ciliates, and metazoa; however, there is a paucity of data on the range of host species for different Syndiniales groups or on net ecosystem effects of parasitism by these taxa.

Saanich Inlet's seasonally stratified water column serves as a well-characterized model ecosystem for examining how deoxygenation shapes microbial community population dynamics and interactions along defined redox gradients in the ocean. Stratification and deep-water renewal cycles are known to shape microbial community composition and activity, and to lead to seasonal blooms of not only various phytoplankton (some of which are toxic species) but also metazoan species that serve as prey for commercially important fisheries in the area. A previous study indicated that protist populations occupy different niches along the redoxcline over a 12-month cycle including abundant Dinoflagellata OTUs (Orsi *et al.* 2012). In oxic waters at 10m depth, Syndiniales OTUs represented $\sim$90% of all Dinoflagellata-affiliated sequences in samples from all seasons (Orsi *et al.* 2012). At 200m depth an increase in Dinoflagellata OTUs (80%

affiliated with Syndiniales) was observed during deep water renewal periods (Orsi *et al.* 2012). While this may reflect active infection of a deep-water bloom of non-phototrophic hosts, this may alternatively reflect dinospores released from sinking infected and lysed phototrophic hosts originating from the upper water column as a responses to seasonal blooms.

Here I explore eukaryotic small subunit ribosomal (SSU) RNA- and DNA-based pyrotag datasets generated from monthly filtered samples collected at depths throughout the Saanich Inlet water column over a 12-month period to elucidate changes in microbial eukaryote populations and eukaryote-eukaryote interactions that have the potential to influence OMZ biogeochemistry and ecology. I focus on population dynamics of Syndiniales OTUs to better constrain their impact on organic matter release and bloom termination during seasonal stratification and renewal.

4.2 Methods

4.2.1 Environmental sampling

Environmental monitoring and sample collection were carried out monthly aboard the MSV John Strickland at station SI03 (48^{o} 35.500 N, 123^{o} 30.300 W) in Saanich Inlet, B.C. From May 2008 to April 2009 and July 2014 a total of 175 water samples were collected at 16 high-resolution depths (10, 20, 40, 60, 75, 80, 90, 97, 100, 110, 120, 135, 150, 165, 185 and 200 m) spanning oxic (>90 μmol O_2 kg^{-1}), dysoxic (90-20μmol O_2 kg^{-1}), suboxic (20^{-1} μmol O_2 kg^{-1}) anoxic (<1 μmol O_2 kg^{-1}) and sulfidic water column conditions (Wright *et al.*, 2012). Samples were processed and analyzed as previously described (Zaikova *et al.*, 2010; Hawley *et al.*, 2017b; Torres-Beltrán *et al.*, 2017). Briefly, water samples were collected from Niskin or Go-Flow bottles for DNA and RNA (Hawley *et al.*, 2017b). Conductivity, temperature, and depth were measured using a Seabird SBE19 CTD-device (Sea-Bird Electronics Inc., Bellevue USA), with a PAR and O_2 sensor attached. An O_2 probe attached to the CTD was also used to measure dissolved O_2 throughout the water column as previously described (Torres-Beltrán *et al.*, 2017).

4.2.2 Nucleic acid sampling and extraction

Biomass to generate SSU rDNA pyrotag datasets for microbial community composition profiling (2 L) was filtered directly onto a 0.22 μm Sterivex polycarbonate cartridge filter from high-resolution depths collected from May 2008 to April 2009, and July 2014. Biomass to generate SSU rRNA pyrotag datasets for potentially active microbial community composition profiling analysis (2 L) was collected on July 2014 at high-resolution depths and filtered directly onto a 0.22 μm Sterivex polycarbonate cartridge filter within 20 minutes of shipboard sample collection.

Genomic DNA was extracted from the Sterivex filters as previously described (Zaikova *et al.*, 2010; Hawley *et al.*, 2017b). Briefly, after defrosting Sterivex on ice, 100 μl lysozyme (0.125 mg ml^{-1}; Sigma) and 20 μl of RNAse (1 μl ml^{-1}; ThermoFisher) were added and incubated at 37^{o}C for 1 h with rotation followed by addition of 50 μl Proteinase K (Sigma) and 100 μl 20% SDS and incubated at 55 ^{o}C for 2 h with rotation. Lysate was removed by pushing through with a syringe into 15 mL falcon tube (Corning) and

with an additional rinse of 1 mL of lysis buffer. Filtrate was subject to chloroform extraction (Sigma) and the aqueous layer was collected and loaded onto a 10K 15 ml Amicon filter cartridge (Millipore), washed three times with TE buffer (pH 8.0) and concentrated to a final volume of between 150-400 μl. Total DNA concentration was determined by PicoGreen assay (Life Technologies) and genomic DNA quality determined by visualization on 0.8% agarose gel (overnight at 16V).

Total RNA was extracted from Sterivex filters using the mirVana Isolation kit (Ambion) (Shi *et al*., 2009; Stewart *et al*., 2010) protocol modified for sterivex filters (Hawley *et al*., 2017b). Briefly, after thawing the filter cartridge on ice RNA later was removed by pushing through with a 3 ml syringe followed by rinsing with an additional 1.8 mL of Ringer's solution and incubated at room temperature for 20min with rotation. Ringer's solution was evacuated with a 3 ml syringe followed by addition of 100 μl of 0.125 mg ml^{-1} lysozyme and incubated at 37 ^{o}C for 30 min with rotation. Lysate was removed from the filter cartridge and subjected to organic extraction following the mirVana kit protocol. DNA removal and clean up and purification of total RNA were conducted following the TURBO DNA-free kit (ThermoFisher) and the RNeasy MinElute Cleanup kit (Qiagen) protocols respectively. Total RNA concentration was determined by RiboGreen analysis (Life Technologies) prior to synthesize first strand cDNA using the SuperScript III First-Strand Synthesis System for RT-qPCR (Invitrogen) according to manufacturer instructions.

4.2.3 Small subunit ribosomal RNA and RNA gene sequencing and analysis

All pyrotag libraries were generated by PCR amplification using multi-domain primers targeting the V6-V8 region of the SSU rRNA gene (Allers *et al*., 2013): 926F (5'-cct atc ccc tgt gtg cct tgg cag tct cag AAA CTY AAA KGA ATT GRC GG-3') and 1392R (5'-cca tct cat ccc tgc gtg tct ccg act cag-<XXXXX>-ACG GGC GGT GTG TRC-3'). Primer sequences were modified by the addition of 454 A or B adapter sequences (lower case). In addition, the reverse primer included a 5 bp barcode designated <XXXXX> for multiplexing of samples during sequencing. Twenty-five microliter PCR reactions were performed in triplicate and pooled to minimize PCR bias. Each reaction contained between 1 and 10 ng of target DNA, 0.5 μl Taq DNA polymerase (Bioshop inc. Canada), 2.5 μL Bioshop 10 x buffer, 1.5 uL 25 mM Bioshop $MgCl_2$, 2.5 μL 10 mM dNTPs (Agilent Technologies) and 0.5 μL 10 mM of each primer. The thermal cycler protocol started with an initial denaturation at 95 ^{o}C for 3 minutes, followed by 25 cycles of 30 s at 95 ^{o}C, 45 s at 55 ^{o}C, 90 s at 72 ^{o}C and 45 s at 55 ^{o}C, and a final extension at 72 ^{o}C for 10 min. PCR products were purified using the QiaQuick PCR purification kit (Qiagen), eluted elution buffer (25 μL), and quantified using PicoGreen assay (Life Technologies). SSU rRNA and SSU rRNA gene (rDNA) amplicons were pooled at 100 ng for each sample. Emulsion PCR and sequencing of the PCR amplicons were sequenced on Roche 454 GS FLX Titanium at the McGill University and Génome Québec Innovation Center.

Pyrotag sequences from May 2008 to April 2009 were processed using the Quantitative Insights Into Microbial Ecology (QIIME) software package (Caporaso *et al*., 2010). To minimize the removal of false positive reads all 2,501,489 pyrotag sequences generated from the 159 samples were clustered together. Reads with a length shorter than 200 bases, ambiguous bases, and homopolymer sequences were removed prior to chimera detection. Chimeras were detected and removed using chimera slayer provided in the

QIIME software package. Sequences were then clustered into operational taxonomic units (OTUs) at 97% identity using uclust with average linkage algorithm. Prior to taxonomic assignment singleton OTUs (OTUs represented by one read) were omitted, leaving 251,337 protistan OTUs. Representative sequences from each non-singleton OTU were queried against the SILVA database (Pruesse *et al*., 2007) using BLAST (Altschul *et al*., 1990). Pyrotag sequences from July 2014 were recruited to the OTUs sequences from the May 2008 to April 2009 dataset using BLAST (Altschul *et al*., 1990) with an identity threshold higher than 98%. Only protistan OTUs showing the highest match (>98% identity) across the reference dataset were selected to support their potential activity under defined water column conditions.

4.2.4 Statistical analyses

Unless otherwise indicated all statistical analyses were performed using the R software (RCoreTeam, 2013). Pyrotag datasets were Hellinger transformed (Legendre and Gallagher, 2001) using the vegan package to the square root of observed total number of reads per sample. Hierarchical cluster analysis (HCA) was conducted to identify groups associated with discrete water column conditions using the pvclust (Suzuki and Shimodaira, 2015) package with Manhattan Distance measures, and statistical significance to the resulting clusters was computed as bootstrap score distributions with 1,000 iterations.

Multi-level indicator species analysis (ISA) using the indicspecies package (De Caceres and Legendre, 2009) was performed to identify OTUs specifically associated with different water column conditions defined by HCA that may be known by their parasite-host interaction. The ISA/multi-level pattern analysis calculates *p*-values with Monte Carlo simulations and returns indicator values (IV) and *p*-values with α <0.05. The IVs range between 0 and 1, where indicator OTUs considered in the present chapter for further community analysis show an IV >0.6 and *p*-value <0.001.

To generate a robust matrix of significant protistan co-occurrences between prevalent OTUs at water column conditions that may represent parasite-host interactions, correlation coefficients from May 2008 to April 2009 pyrotag samples were calculated using the Bray-Curtis and Spearmans rank correlations in the CoNet software (Faust *et al*., 2012) with OTU abundance as count data as previously used for pyrotag data (Torres-Beltrán *et al*., 2016a). Prior analysis in CoNet, the OTU matrix derived from pyrotag taxonomic analysis was transformed into presence-absence data to remove OTUs with less than 1/3 zero counts, leaving 556 OTUs for all samples that were selected in the OTU abundance matrix for correlation analysis. To construct ensemble networks, thresholds for Bray-Curtis and Spearmans rank correlations measures were set to 0.6 correlation value as a pre-filter for OTU pairs, followed by computing edge scores only between highly correlated OTU pairs. To assign statistical significance to the resulting scores, edge and correlation measure-specific permutation and bootstrap score distributions with 1,000 iterations each were computed. *p*-values were tail-adjusted so that low *p*-values correspond to co-presence and high *p*-values to exclusion, *p*-values on each final edge were corrected to *q*-values (cut-off of 0.05). Finally, only edges with at least two supporting pieces of evidence, i.e. high correlation value on both correlation measures and *q*-value below threshold were retained. The final edges and nodes matrices were exported using Cytoscape 2.8.3 (Shannon *et al*., 2003). Nodes corresponded to individual OTUs and edges were defined by computed correlations between corresponding OTU pairs. Edges were selected and exported based on specific paired OTUs such

as those affiliated with parasite-host taxonomic groups.

To gain insight into potential parasite-host activity, SSU rRNA:rDNA ratios were calculated as previously described for pyrosequencing data (Frias-Lopez *et al*., 2008; Stewart *et al*., 2012b). I selected OTUs from the July 2014 dataset based on their taxonomic affiliation (BLAST-based >98% similarity) with taxa of interest. In addition, only OTUs affiliated to indicators showing significant correlations were selected.

4.2.5 Data deposition

The SSU rDNA and rRNA pyrotag sequences reported in this chapter have been submitted to the The National Center for Biotechnology Information (NCBI) under BioSample numbers: SAMN03387532 - SAMN03387915 and SAMN05392441 - SAMN05392453.

4.3 Results

4.3.1 Water column conditions

Water column properties were monitored through the progression of stratification and deep-water renewal over a one-year period (May 2008 to April 2009). Throughout the year the water column surface temperature ranged between 7 to 16 ^{o}C during winter and summer months respectively, reaching an average of 9 ^{o}C at bottom waters. Salinity between surface and bottom waters ranged between 28 to 31 ppm throughout the year, showing the lowest values at the surface during spring months. In addition, as previously described for OMZs (Wright *et al*., 2012), in this chapter I define water column conditions on the basis of O_2 concentration ranges: oxic (>90 μM O_2), dysoxic (90-20 μM O_2), suboxic (20^{-1} μM O_2) and anoxic (<1 μM) (Wright *et al*., 2012). Between May and August 2008, as water column stratification peaked, suboxic conditions intensified corresponding with the development of deep-water anoxia. The concentration of NO_3^- between the surface and 100m ranged between 5 to 20 μM, decreasing rapidly between 100 and 135m before reaching a minimum of <1 μM in anoxic bottom waters (Fig. 4.1). The beginning of 2008 deep-water renewal occurred toward the end of September and continued through November. Initially, dissolved O_2 was observed throughout the water column, although upwards shoaling of O_2 and NO_3^- depleted bottom waters produced an intermediate suboxic layer between 100 and 135 m. By October, dissolved O_2 concentrations between 150 and 200 m ranged between 14 and 28 μM consistent with complete oxygenation of the water column. Over the same time interval, NO_3^- concentrations between surface and 100m ranged between 4 to 29 μM, decreasing rapidly below 100 m before reaching a second maximum ranging between 21 to 24 μM at 200 m (Fig. 4.1). In November, O_2 and NO_3^- concentrations continued to increase above 100 and below 135 m with intervening depth intervals experiencing increased oxygen decline. In December water column O_2 deficiency intensified below 100 m as the water column become increasingly stratified through April 2009.

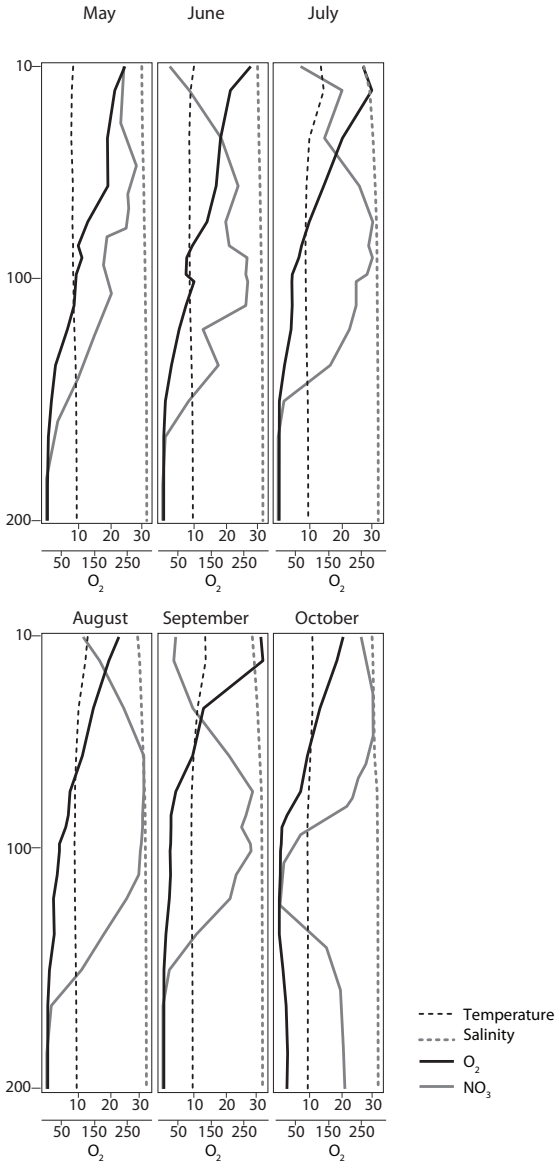

Figure 4.1: Water column chemical parameters. CTD and chemical data shown as monthly panels for temperature (^{o}C), Salinity (ppm), Oxygen (μM) and nitrate (μM) along the depth profile for samples taken from May to September 2008 at Station S3 in Saanich Inlet.

4.3.2 Eukaryotic community structure

To explore eukaryotic community structure in the Saanich Inlet water column I analyzed SSU rDNA gene pyrotag sequences from 159 samples collected at 16 depths ranging between 10 and 200 m over the time interval between May 2008 and April 2009. Overall, the opisthokont community was dominated by OTUs affiliated with metazoan sequences within the phylum Arthropoda (34% total community) (Fig 4.2A). Protistan community composition was dominated by OTUs affiliated with taxonomic groups previously identified in the Saanich Inlet water column, using different identification approaches (Orsi *et al*., 2012) including the Alveolata (41%), Stramenopiles (10%), Hacrobia (3%) and Rhizaria (2%)(Fig 2A). Nineteen percent of Alveolata OTUs were affiliated with Dinoflagellata sequences within Syndiniales groups (I, II, III and V) (Fig 4.2B). Groups I (27%) and II (66%) dominated throughout the year showing peaks of abundance during the stratification period (May- August) in the inlet (Fig 4.3B) reaching up to 19% and 43% of relative abundance, respectively, from the total protist sequences.

To describe protistan community partitioning along water column O_2 gradients I conducted hierarchical cluster analysis using rDNA pyrotag profiles from each sample. Results revealed four major groups or clusters (AU $>$70, 1,000 iterations) associated with oxic (group I), dysoxic-suboxic (group II and III), and anoxic (group IV) water column conditions that were previously observed to occur with seasonal stratification and deep water renewal events (Orsi *et al*., 2012) (Fig 4.3A). To further resolve OTUs occurring under specific water column O_2 conditions and to reveal protistan co-occurrence patterns between taxa i.e. Syndiniales and potential hosts, multi-level indicator species analysis (ISA) was conducted based on HCA groups (Table B.1-B.4). Indicator OTUs affiliated with Arthropoda, Cnidaria, Dinoflagellata and Stramenopiles were characteristic of the oxic water column compartment while indicator OTUs affiliated with Choanoflagellatea, Colpodea, Cryptophyceae, Dinoflagellata, Picozoa, Prymnesiophyceae, Stramenopiles and Telonemia were characteristic of the dysoxic-suboxic water column conditions (Fig 4.3B). Syndiniales OTUs were the most abundant indicator OTUs found in these water column conditions during peak stratification in summer (56% of total indicator OTUs) (Fig 4.3C). In addition, OTUs affiliated with potential hosts such as the ciliate subclass Choreotrichia and the metazoan class Maxillopoda, were also found as indicators for the dysoxic-suboxic water column conditions along with the Syndiniales OTUs. In comparison, indicator OTUs affiliated with Dinoflagellata, Stramenopiles and Arthropoda were identified for the anoxic water column compartment during the stratification period.

4.3.3 Exploring protistan co-occurrence patterns

To explore community co-occurrence patterns focusing primarily on potential parasitic interactions between Syndiniales and protists throughout water column O_2 gradients, I conducted a co-occurrence analysis based on Bray-Curtis and Spearman correlations among OTUs. Resulting significant (Correlation Value $>$0.6, p $<$0.001) pairs were used to determine potential interactions i.e. parasitism dynamics over water column stratification period. Correlation analysis resulted in a total of 6,273 significant pairs corresponding to co-occurrences between prevalent OTUs throughout water column conditions over time. Significant pairs were observed among different taxonomic groups suggesting a broad range of potential interactions resulting from 325 unique source OTUs affiliated with Alveolata (3,765), Apusozoa (3), Archeaplastida (139), Excavata

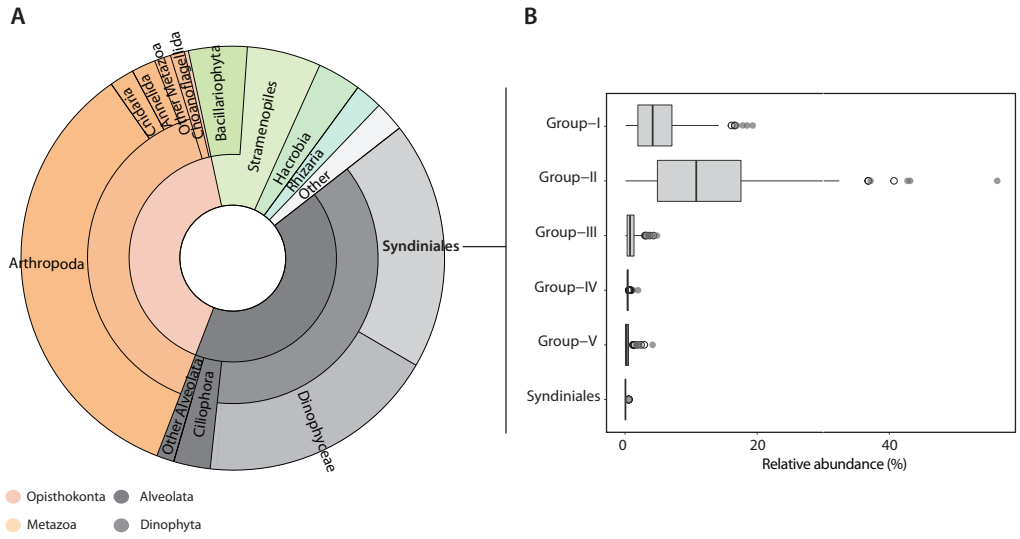

Figure 4.2: Taxonomic breakdown of eukaryotic OTUs. A) Krona chart showing the taxonomic composition of eukaryotic OTUs found in the rDNA pyrotag datasets from May 2008 to April 2009. Layers represent hierarchical taxonomy from the upper (centre) to the lowest (outer) taxonomic level. B) Relative abundance for Syndiniales OTUs divided by taxonomic groups found in pyrotag dataset. The size of each box represents the average of relative abundance (%) calculated from the total number of eukaryotic reads throughout the water column over this period. Extended dashed lines (whiskers) represent at the base the lower and upper quartiles (25% and 75%) and at the end the minimum and maximum values encountered. The middle line represents the median.

(250), Hacrobia (483), Opisthokonta (418), Rhizaria (224), and Stramenopiles (986).

Syndiniales correlations corresponded to 63% of the total pairs observed, showing co-occurrence with OTUs affiliated with Alveolata (including Apicomplexa, Ciliophora, Dinoflagellata), Chlorophyta, Metazoa and a variety of Stramenopiles including MAST groups (Fig 4.4; Appendix B). The relative proportion of taxa correlating with Syndiniales OTUs was similar between groups; however, the number of unique OTUs associated with the different Syndiniales groups varied significantly. For instance, group II was the most taxonomically diverse assemblage, exhibited the most correlations with other protistan OTUs, while Group V exhibited the fewest (Fig 4.5A). The occurrence of OTUs affiliated with amoebae, green algae, haptophytes, and Metazoa was exclusively statistically correlated with Syndiniales groups I and II (Fig 4.5B). Interestingly, significant positive correlations between Syndiniales and OTUs affiliated with taxa described as potential hosts i.e. Choreotrichia, *Phaeocystis* sp. and Maxillopoda, were observed in different oxygen regimes. For instance, while Syndiniales (groups I-V) and Choreotrichia OTUs co-occurred throughout the water column, Syndiniales groups II, III and V, and Chroreotrichia OTUs co-occurred mostly in dysoxic-suboxic (90-150 m) and anoxic (below 150 m) water column conditions as water column stratification intensified during summer months (Fig 4.6). Syndiniales-Choreotrichia co-occurrence patterns were supported by ISA results showing these OTUs to be characteristic of dysoxic-suboxic water column conditions during

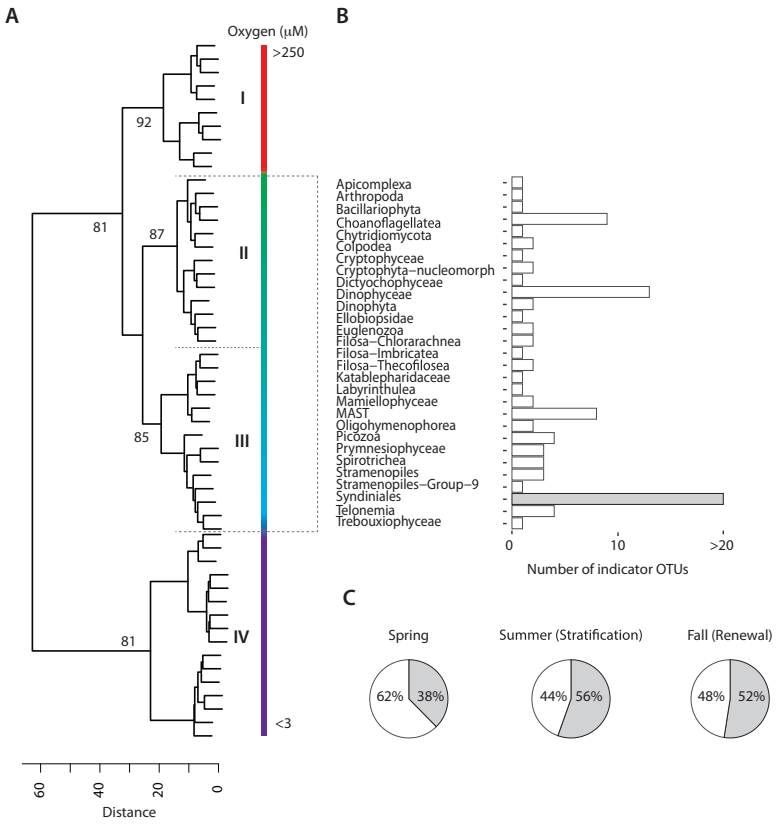

Figure 4.3: Eukaryotic community structure. A) Hierarchical clustering of eukaryotic pyrotag data during peak stratification based on Manhattan distance. Clusters are delimited by O_2 concentration range represented by number from I to IV: oxic = I (red), dysoxic- suboxic = II and III (green and blue, respectively), and anoxic =IV (purple). Bootstrap values (1000 iterations) are shown in gray. B) Indicator OTUs (Indicator value 0.6, p =0.05, $\alpha < 0.01$) for suboxic-dysoxic water column conditions during peak stratification. Bars depict the total number of OTUs for each taxonomic group observed. Total Syndiniales indicator OTUs (n=54) are shown in dark gray. C) Distribution of Syndiniales indicator OTUs over three different water column periods in the inlet (spring, summer stratification and fall deep-water renewal). Pie charts depict the percentage of Syndiniales OTUs (dark gray) out of the total number of indicator OTUs for each period.

the same time intervals (Table B.1). In addition, Syndiniales groups I and II, and *Phaeocystis* sp. OTUs co-occurred in oxic waters in May and become progressively associated with dysoxic-suboxic waters as the water column became increasingly stratified in July and August (Fig 4.7). Syndiniales groups I and II, and Maxillopoda. OTUs exhibited a similar temporal co-occurrence pattern (Fig 4.8). Syndiniales-Maxillopoda co-occurrence patterns were supported by ISA results indicating that these OTUs were characteristic of oxic and dysoxic water column conditions during the same time intervals (Table B.1). Moreover, I observed that OTUs affiliated with Syndiniales and the three target taxa were evenly distributed throughout the water column in September and October possibly reflecting vertical transport and mixing during deep water renewal.

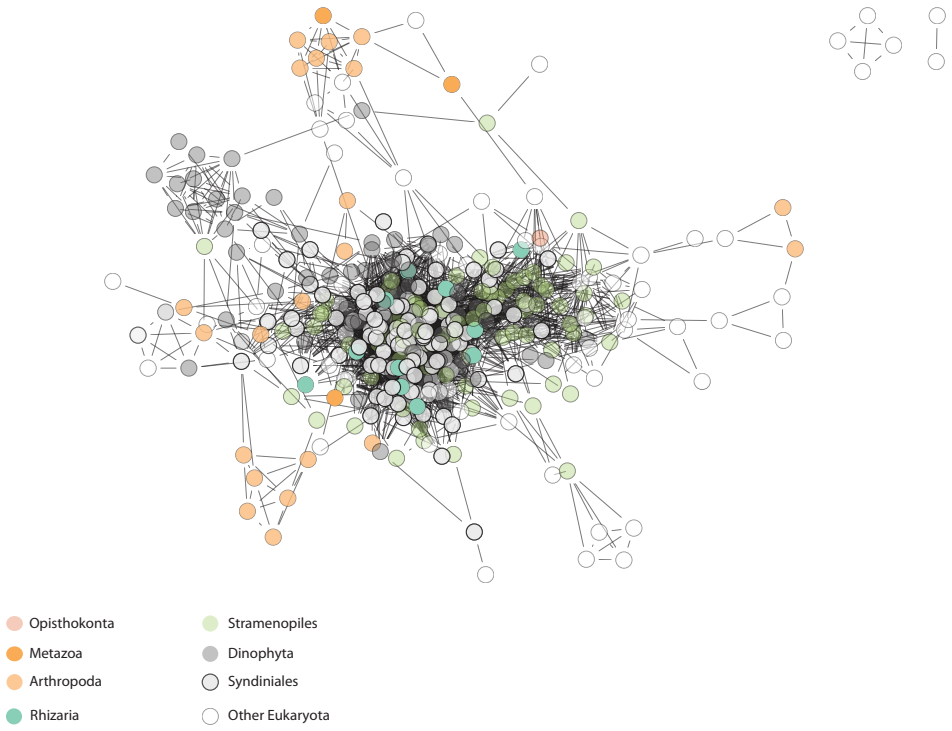

Figure 4.4: Co-occurrence network on SSU rDNA pyrotag protist data from May-August 2008. Co-occurrence network derived from Bray-Curtis and Spearman correlation measures on rDNA pyrotag protist data from May-August 2008. Nodes depict OTUs and edges co-occurrence correlations. Syndiniales OTUs are highlighted in gray and other protist OTUs coloured as indicated in color key.

4.3.4 Insight into potential Syndiniales parasitic interactions

Given the observed co-occurrence patterns between rDNA of Syndiniales and Choreotrichia, *Phaeocystis* sp. and Maxillopoda during the 2008 stratification period, I wanted to evaluate the potential activity of these groups under low oxygen conditions. I made this connection by comparing pyrotag SSU rDNA observations and SSU rRNA: rDNA ratios from samples collected at the same water column depths during peak water column stratification in July 2014.

Consistent with observed patterns in 2008, rDNA observations in 2014 showed Syndiniales group I and II OTUs were evenly distributed throughout the water column exhibiting highest relative abundance, 10 and 38% respectively, in suboxic and anoxic water column conditions (Fig 4.9). Syndiniales group III OTUs distributed from surface (4% relative abundance) to the suboxic boundary of the water column, while Syndiniales group V OTUs were primarily observed in dysoxic-suboxic water column conditions with a relative abundance $<$0.1% (Fig 4.9). Choreotrichia OTUs were also found evenly distributed throughout the water column and showed higher relative abundance values ($\sim$6%) below the dysoxic boundary of the water column and highest abundance in anoxic waters (16%) (Fig 9). Maxillopoda OTUs were found evenly

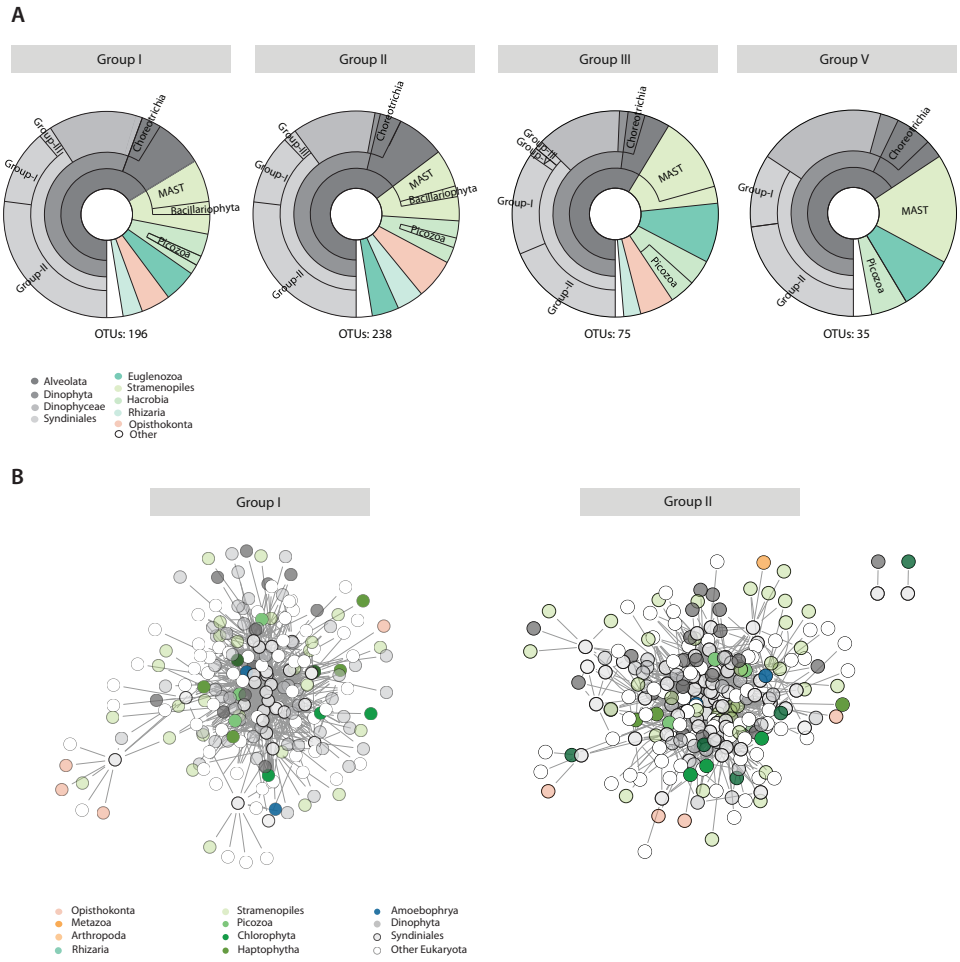

Figure 4.5: Syndiniales OTUs interactions. A) Eukaryotic taxa diversity associated with each Syndiniales group. Krona charts depict total unique OTUs with significant interactions for each Syndiniales group (I, II, III and V). Layers represent hierarchical taxonomy from the upper (centre) to the lowest (outer) taxonomic level. Total number of unique OTUs is indicated at the bottom of each chart. B) Co-occurrence network derived from Bray-Curtis and Spearman correlation measures on Sydiniales Group I and II derived from rDNA pyrotag protist data from May-August 2008. Nodes depict OTUs and edges co-occurrence correlations. Syndiniales OTUs are highlighted in gray and other protist OTUs coloured as indicated in color key.

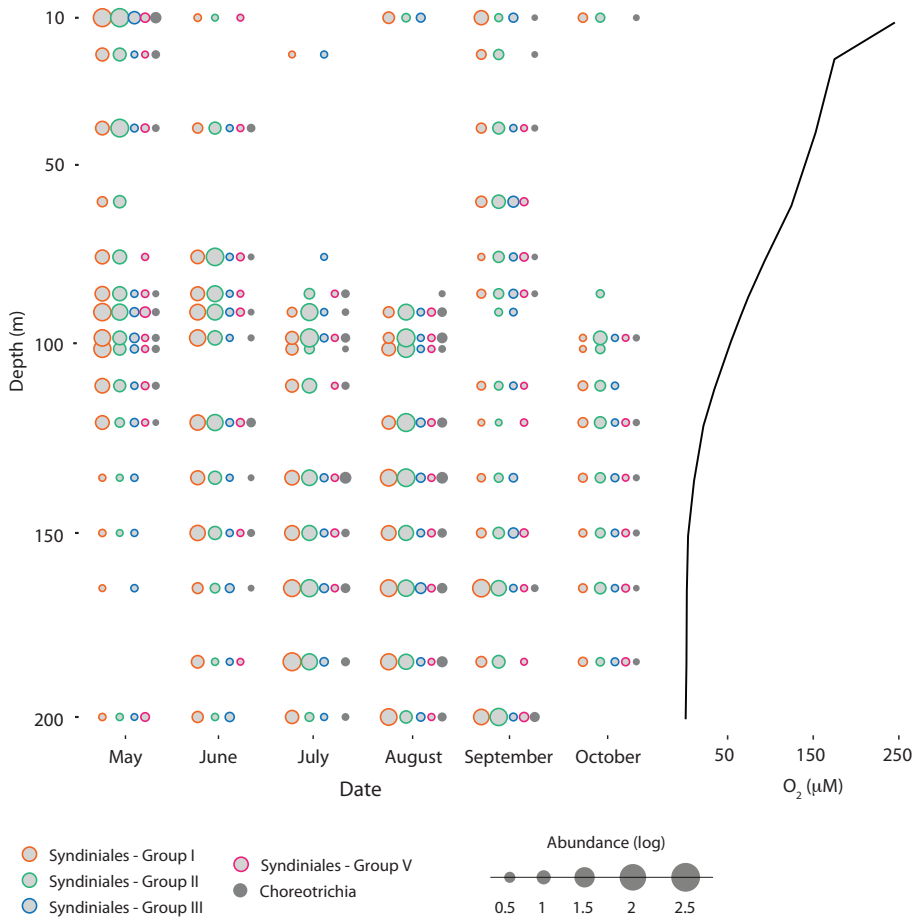

Figure 4.6: Vertical distribution and abundance of Choreotrichia and Syndiniales OTUs Choreotrichia and Syndiniales OTUs distribution through the water column from May-October 2008. Dots size depicts OTUs relative abundance in logarithmic scale.

distributed throughout the water column with a relative abundance ∼0.45%, while *Phaeocystis* sp. OTUs were mostly distributed at dysoxic-suboxic water column conditions with a relative abundance ∼0.01% (Fig 4.9).

Potential activity of these OTUs was determined by calculating SSU rRNA:rDNA ratios across depth intervals (10, 100, 120,135, 150, 165 and 185m) spanning the O_2 concentration ranges from oxic to anoxic. In contrast with rDNA observations alone, SSU rRNA observations can provide a robust proxy for past, present or emerging cellular activities (Blazewicz *et al.*, 2013). Moreover, monitoring SSU rRNA dynamics over time can inform hypotheses related to life strategies within communities (Lepp and Schmidt, 1998; Barnard *et al.*, 2013). I selected OTUs affiliated with indicator OTUs showing significant correlations, and considered values >1 as potentially occurring interactions among likely active OTUs in the same depth

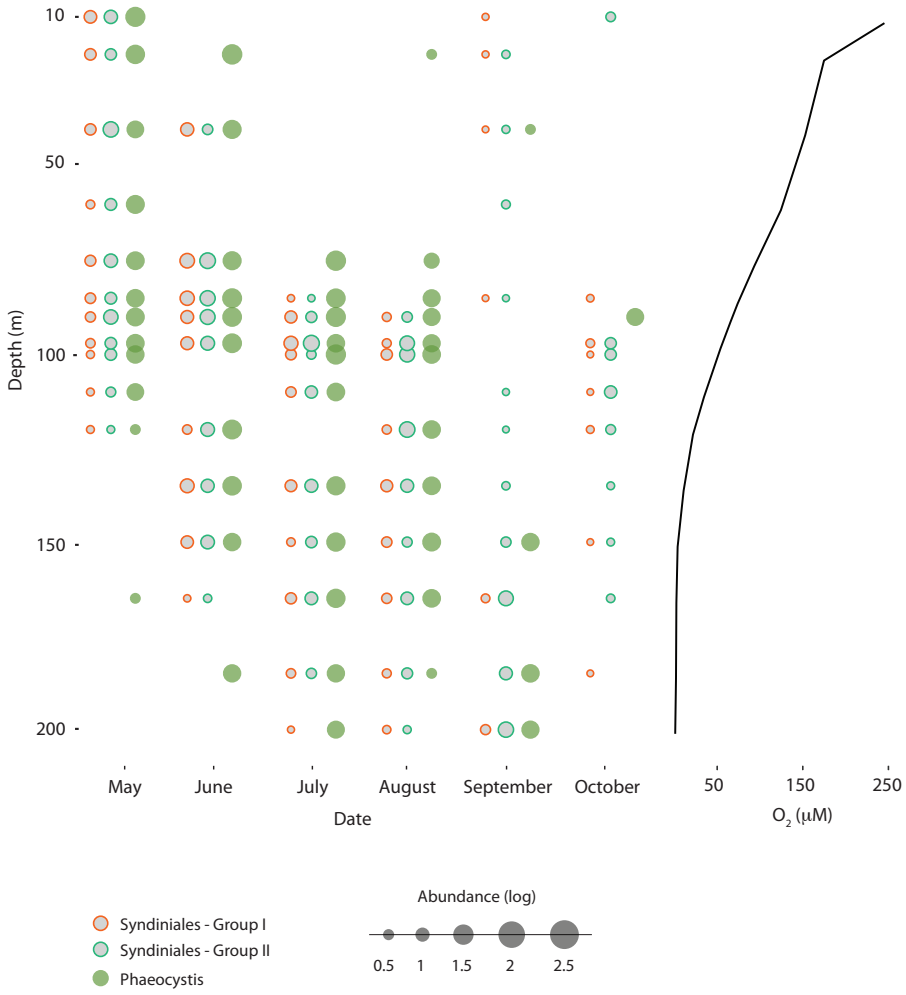

Figure 4.7: Vertical distribution and abundance of *Phaeocystis antarctica* and Syndiniales OTUs *Phaeocystis antarctica* and Syndiniales OTUs distribution through the water column from May-October 2008. Bars depict OTUs relative abundance in logarithmic scale.

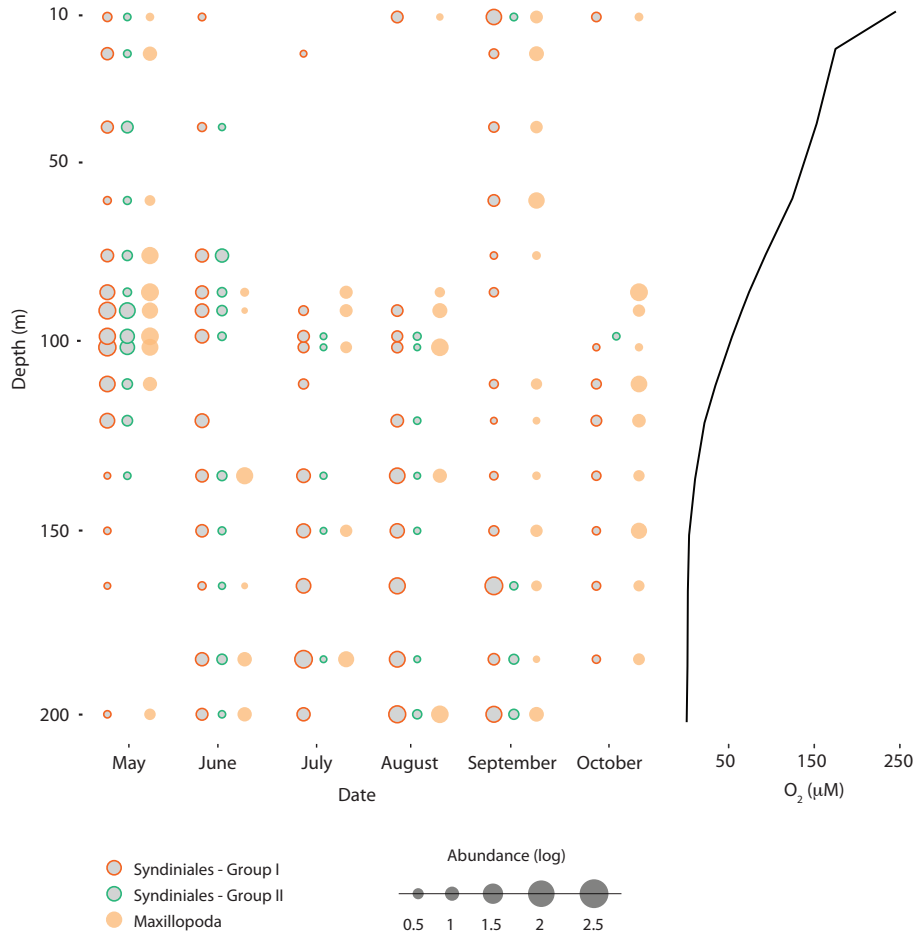

Figure 4.8: Vertical distribution and abundance of Maxillopoda and Syndiniales OTUs Maxillopoda and Syndiniales OTUs distribution through the water column from May-October 2008. Dots size depicts OTUs relative abundance in logarithmic scale.

interval. For instance, Syndiniales group II OTUs were active throughout the water column exhibiting the highest ratio (9.3) at 120m corresponding to the suboxic water column compartment (Fig 4.8). Similarly, Choreotrichia, *Phaeocystis* and Maxillopoda OTUs were found throughout the water column but exhibited the highest ratios in suboxic waters. For instance, Choreotrichia OTUs showed the highest ratio (7.1) at 135 m, while *Phaeocystis* and Maxillopoda OTUs at 150 m with ratio values equal to 14 and 8.6, respectively (Fig 4.8). Thus, it could be hypothesized that potential interactions among active OTUs could occur primarily at dysoxic-suboxic water column conditions between the 100 and 150 m.

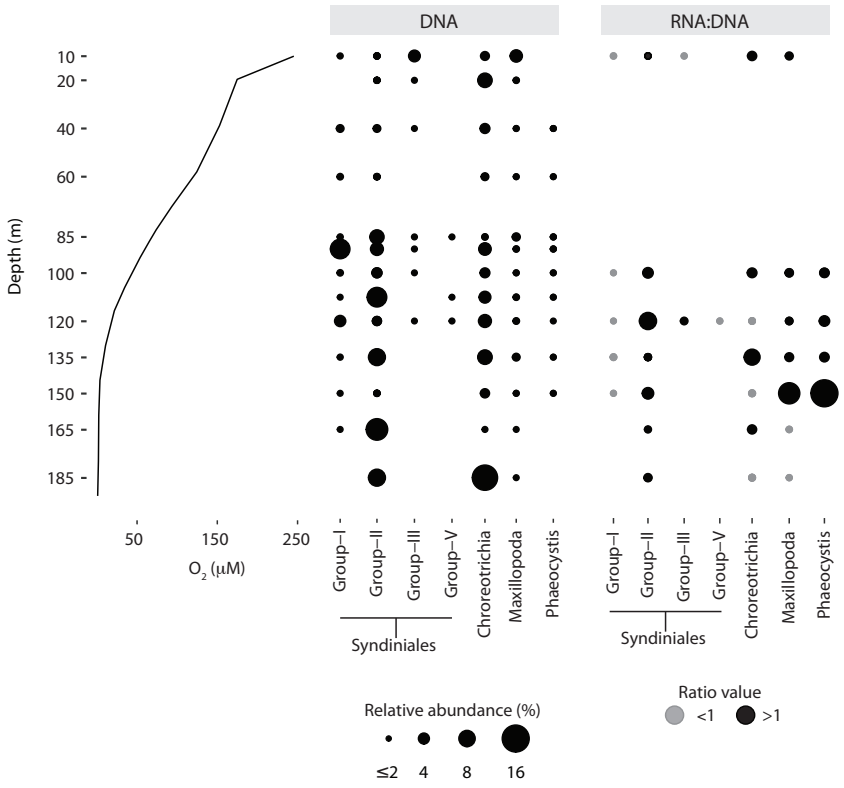

Figure 4.9: Distribution of Syndiniales and interacting Choreotrichia, *Phaeocystis* sp., and Maxillopoda OTUs throughout the water column oxygen gradient during the stratification period of July 2014. Dissolved oxygen concentration (μM) profile is shown as a black sparkline on the left figure panel. OTUs vertical distribution is shown as dots which size depicts the relative abundance (%) of the OTUs affiliated to each taxonomic group calculated out of the total number of reads in the SSU rDNA pyrotag dataset. The rRNA:rDNA ratio throughout the water column for the OTUs affiliated to each taxonomic group is shown as size and color constrained dots (gray <1 ; black >1).

4.4 Discussion

This chapter explores protistan diversity, and seasonal changes in abundance and distribution throughout the water column in Saanich Inlet, a seasonally anoxic fjord that serves as a well-characterized model ecosystem for examining how deoxygenation shapes microbial community population dynamics and interactions along defined redox gradients in the ocean. Observations primarily focused on identifying statistically significant correlations that hint to potential interactions between the protistan parasite Syndiniales and potential host taxa. Proposed interactions were based on correlational edges between OTUs obtained from Bray-Curtis and Spearmans rank correlations and reinforced by determining potential activity from SSU rRNA:rDNA ratios for these groups based on the assumption that potential parasitic interactions are most likely to occur at depths where the host is likely to be active/alive. These results provide insights into the potential for par-

asitic interactions occurring under suboxic conditions during the summer stratification period with possible implications for carbon cycling and bloom termination dynamics.

Water column seasonal stratification and deep-water renewal restructures microbial eukaryotic communities that respond to changing O_2 concentrations by altering their vertical distribution and activities (Parris *et al*., 2014;Duret *et al*., 2015;Jing *et al*., 2015). For instance, a previous study in Saanich Inlet highlighted seasonal differences in eukaryotic community structure correlated to water column O_2 conditions (Orsi *et al*. 2012). Similarly, the use of time resolved SSU rDNA pyrotag observations over this seasonal cycle allowed me to observe changes in abundance and distribution patterns of the eukaryote community along defined water column O_2 conditions. Although, the use of high-throughput sequencing data allowed detailed exploration of eukaryote diversity, I focused this analysis on Syndiniales, that comprise a diverse and abundant (43% during the summer stratification period) group of parasitic dinoflagellates that are responsive to water column stratification and deep-water renewal, and potentially play a significant role in shaping microbial biogeochemistry and ecology in diverse water column conditions. As an initial approach to select potential OTU pairs that likely showed predator-prey or parasitic interactions, I used ISA and co-occurrence correlation analyses in complementary manner. For instance, Syndiniales OTUs exhibited concurrent patterns with three likely host taxa i.e. Choreotrichia, *Phaeocystis* sp. and Maxillopoda and were selected based on statistical significance then further analyzed to provide baseline evidence for developing ecological hypotheses on predator-prey or parasitic interactions that may be relevant in stratified and O_2 deficient marine environments.

Syndiniales interactions and the significance of peak stratification

The initial co-occurrence correlation analysis considered the eukaryotic community throughout 12-month period and identified five Syndiniales OTUs that have statistically significant correlations with other cosmopolitan eukaryotic taxa including Stramenopiles from the MAST clades, picomonads, and other Syndiniales (Table B.1). These taxa are small (2-8 μm) and not previously documented in the literature as hosts of Syndiniales; however, spatial and temporal trends of these interactions reveal Syndiniales OTUs track the water column distribution of other microbial eukaryotes from surface water spring blooms to mid-waters as the inlet stratification strengthens in July (Table B.2-B.4). The abundance of Syndiniales OTUs throughout the year, yet the absence of significant correlations detected in the 12-month analysis between Syndiniales and known host taxa, suggested a much broader host range. To reveal more ephemeral interactions between Syndiniales and common bloom-forming taxa (i.e. diatoms and dinoflagellates), I repeated our co-occurrence correlation analysis by targeting the period May through August, encompassing the peak stratification period.

Co-occurrence analysis of data collected during peak stratification revealed correlations among members of a diverse protistan community. Syndiniales OTUs were at the core of complex significant correlations with other protist taxa. The number of significant correlations for this period increased 10-fold and involved more diverse taxa, including several species of diatoms, the athecate dinoflagellate Gymnodinium, the bloom-forming haptophyte *Phaeocystis*, and heterotrophs, including several species of ciliates, stramenopiles, groups of small flagellates, and Metazoa. Many eukaryotic OTUs showed significant correlations with multiple Syndiniales OTUs (Appendix B), suggesting co-occurrence might drive host plasticity

as a survival strategy that allows parasites to respond to host community shifts. Given the short-lived, free-swimming infective stage, I hypothesized co-occurrences between Syndiniales and other eukaryotes in the Saanich Inlet water column are more likely to reflect parasitism or predator-prey relationships than simple co-occurrences under specific water column conditions. Indicator analyses were compared with results from the co-occurrence correlation analysis to gain insights about taxa that were uniquely abundant during the spring and summer months and that showed significant correlation with Syndiniales over the same time period.

Based on these criteria, the distributions of three taxa were examined in detail: *Phaeocystis antarctica*, which produced a bloom in 2008 providing an opportunity to investigate if parasitic dinoflagellates were involved in the termination of the bloom; a ciliate OTU affiliated to Choreotrichia, an indicator taxon during peak stratification in micro-oxic waters; and a metazoan copepod taxon, an indicator OTU during peak stratification. Ciliates, phytoplankton, and copepod metazoa are all known to be potential hosts of Syndiniales (Guillou *et al*., 2008; Dolan, 2013; Massana *et al*., 2014). Observations of spatial and temporal distribution of OTUs affiliated to these taxa suggested that several types of parasitic Syndiniales could be potentially infecting the same host, and/or in the case of heterotrophic protistan taxa, that the heterotrophs may be potentially feeding on Syndiniales dinospores released from previously infected sinking cells. Nonetheless, I acknowledge the possibility that a fraction of these OTUs may not be involved in predator/prey or parasitic relationships, and may simply co-occur as a community under specific environmental conditions. Furthermore, single-cell and microscopic analyses could be used in the future to verify specific interactions proposed by the observations in this chapter.

Interactions with Choreotrichia

Tintinnid ciliates belonging to the order Choreotrichida are ubiquitous in the marine plankton, particularly in coastal waters where their numbers are generally higher. Tintinnids play different roles in eukaryotic communities; they are preyed upon by planktonic zooplankton (Montagnes *et al*., 2010), as well as larger protist taxa (Dolan, 2013). Also, as nanoplankton consumers, they have an important role in bloom termination and/or as vectors for transferring bloom-associated toxins to higher trophic levels (Dolan, 2013). Furthermore, some tintinnid ciliates have been observed to serve as hosts for at least two different parasites, including Syndiniales, and even single cells may be co-infected (Dolan, 2013). In this chapter I observed ciliates belonging to the subclass Choreotrichia were most abundant between 75 m and 150 m in August 2008, and these OTUs were indicator OTUs for micro-oxic mid-waters during peak stratification. Correlation analysis results showed seven indicator Choreotrichia OTUs that were significantly correlated with Syndiniales groups (I-V) OTUs (Table B.3).

However, the distribution of Syndiniales Group V OTUs most closely matched the distribution patterns of the OTUs affiliated with Choreotrichia throughout the water column between May and August (Fig 4.6). For example, in July 2008, the four Syndiniales groups and the ciliate subclass Choreotrichia were present at the surface 10m and in micro-oxic waters between 85m and 120m (Fig. 4.6). while in August, only Syndiniales Group V matched Choreotrichia abundance and distribution patterns in the mid- and deep-water samples (Fig 4.6). Analysis of SSU rRNA:rDNA ratios during peak stratification in 2014 indicated potential activity of Choreotrichia at 10 m, 97 m, 135 m, and 165 m depths in waters that ranged from

oxic to anoxic. Syndiniales Group II was also active at these depths, which may suggest either parasitism or grazing could take place under different water column O_2 conditions. Interestingly, SSU rRNA: rDNA ratios for Syndiniales Group I, III, and V indicated cells were not likely active in the upper oxic layer, which supports the hypothesis that Choreotrichia may be feeding on dinospores from these groups during peak stratification (Fig 4.9). The correlations between OTUs affiliated with Choreotrichia and the four known Syndiniales groups in Saanich Inlet indicated by the correlation analysis suggest complex interactions could occur that are associated with season water column dynamics. Further single-cell analyses are required to reveal the extent to which parasitism versus grazing is contributing to these detected potential interactions.

Interactions with Phaeocystis antarctica

Phaeocystis antarctica is a colony-forming haptophyte that contributes to primary production and dimethylsulfide formation (Kettle *et al*., 1999). The dimethylsulfide produced by bloom events can be oxidized to sulphuric acid in the atmosphere where it can function as nuclei for cloud condensation, impacting global albedo (Charlson *et al*., 1987). *P. antarctica* has a complex life cycle that includes colonial, bloom-forming aggregates, individual flagellates (microzoospores), and a recently discovered zygotic "benthic stage" (GaeblerSchwarz *et al*., 2010).

In this chapter, I observed OTUs affiliated with P. antarctica comprised 6% of the pyrotag library at 10 m depth in May 2008, consistent with a late spring surface bloom. Based on the depth of detection of *P. antarctica* pyrotags in subsequent months, cells appeared to sink through the water column and signatures of this taxon disappeared following deep water renewal in October 2008 (Fig 4.7). Four and eight OTUs affiliated with Syndiniales groups I and II, respectively, were identified as having significant correlations with this haptophyte, with OTUs from these two groups of Syndiniales tracking signatures of *P. antarctica* through the water column during peak stratification. SSU rRNA:rDNA ratio analysis indicated *P. antarctica* was not potentially active at 10m, however its highest potential activity was observed at 150 m in July 2014 as also reflected in its abundance based on rDNA observations alone. This suggests an ephemerally active bloom followed by sinking of likely infected cells and "benthic stage" zygotes (Fig 4.9). Following renewal (which occurred on October 14, 2008; Ocean Networks Canada VENUS node (http://www.oceannetworks.ca) the co-occurrence between Group I and II Syndiniales OTUs and *P. antarctica* was not detected, however those two groups of Syndiniales appeared to correlate with different putative hosts in the surface waters, i.e. OTUs affiliated with *Pseudocalanus* copepods and maxillopods.

Parasites of *P. antarctica* blooms have not been previously identified; however, based on our co-occurrence correlation analyses I hypothesize this haptophyte could host several strains of Group I and II Syndiniales, and that these two groups of parasites locate and infect new hosts when *P. antarctica* is absent from the water column. Our observation hints at a testable hypothesis for future research is that parasitism of *Phaeocystis* blooms in Saanich Inlet could play a role in controlling the duration of phytoplankton blooms that play a major role in spring primary production, and that influence dimethylsulfide synthesis.

Interactions with copepods

Common taxonomic groups of mesozooplankton in Saanich Inlet include calanoid copepods, chaetognaths, euphausiids, cnidarians, and ctenophores. These taxa are critical for trophic transfer of carbon in marine food webs (e.g. Jnasdttir *et al*., 2015) as mesozooplankton are predators of phytoplankton, het-

erotrophic protists, and bacteria, as well as prey items for larger metazoans including fish. Zooplankton observed from a time series collected off southern Vancouver Island undergo annual variations in abundance, leading to changes in overall biomass and relative species composition that can further influence the pelagic food web (Hargreaves *et al*., 1994; Beamish *et al*., 1997; Bertram *et al*., 2001). Copepods typically found in shelf and slope waters off the coast of British Columbia have regionally specific ranges, driven primarily by water temperature, in addition to inter-annual shifts in species abundance. A few occurrences of dinoflagellate syndinian parasites from marine alveolate Group II (MALV II) infecting copepods and copepod eggs have been reported in the literature (Kimmerer and McKinnon, 1990; Skovgaard *et al*., 2005; Gmez *et al*., 2009); however, the overall ecological effects of these infections are not known.

Co-occurrence correlation analysis shows that Metazoa exhibit significant correlations with Syndiniales in Saanich Inlet. While the observed correlations may imply Metazoa may be feeding directly on dinospores or more likely, infected prey, these also suggest that Syndiniales OTUs could parasitize certain Metazoa. I observed OTUs affiliated with the genera *Corycaeus* and *Oncaea* were indicator OTUs for surface waters in May when prey taxa including *Thlassiosira aestivalis*, *Chrysochromulina rotalis*, *Phaeocystis antarctica*, *Micromonas* sp., and Florenciellales diatoms were abundant, suggesting higher metazoan grazing activity during peak phytoplankton growth. Additionally, these copepods showed significant correlations with Syndiniales from Groups I and II and, like the other taxa discussed, Syndiniales OTUs appears to track the grazers throughout the water column (Fig. 4.8).

The corresponding SSU rRNA:rDNA ratios for OTUs affiliated with the crustacean class Maxillopoda (copepods) indicated potential activity of this group of metazoans in the upper oxic layer during peak stratification in July 2014 (Fig. 4.9). Syndiniales Group II SSU rRNA:rDNA ratios indicated potential activity in the water column at the same depths as the maxillopods, while Group I Syndiniales was likely inactive. The RNA-based analyses thus suggest Syndiniales Group I OTUs are likely sinking cells, while Group II could potentially have parasitic or prey interactions with members of Maxillopoda below the surface. Given the apparent significance of the interactions between metazoan groups known to play important roles in marine food webs, I consider it is important to understand the extent to which members of Syndiniales parasitize metazoans and the impact that this may have on carbon and nutrient turnover parasitize metazoans.

4.5 Conclusion

In this chapter I used molecular time series observations to detect interactions between groups of primarily uncultured, parasitic dinoflagellates within the Syndiniales and other eukaryotes along defined oxygen gradients in Saanich Inlet. Co-occurrence correlation and indicator species analyses revealed the potential for significant interactions may occur between four known Syndiniales groups and different protistan and metazoan taxa including the stramenopile MAST clades not previously known to be parasitized by Syndiniales. I focused the analysis on several eukaryotic OTUs showing significant correlations with Syndiniales during peak stratification, including *Phaeocystis antarctica*, ciliates affiliated with Choreotrichia, and copepods. These observations provide baseline understanding on the potential host range of the major parasitic Syndiniales groups that could infect key primary producing and heterotrophic populations along stratified water columns. Results presented here give insight into possible impacts these infections may have on population dynamics, extensible to nutrient cycling processes, during seasonal water column stratification in Saanich Inlet.

Chapter 5

Methanotrophic community dynamics in Saanich Inlet[1]

This chapter represents a unique correlation analysis coupling time-series geochemical and small subunit ribosomal (SSU) RNA gene observations to chart spatial and temporal patterns of methanotrophic interactions. Here the trace gas problem in OMZs is reintroduced by focusing on the CH_4 oxidation processes. Oxygen (O_2) effects on ecological interactions and energy flow were reinforced by observed redox-driven niche partitioning along changing water column redox gradients for methanotrophic bacteria and the potential use of novel metabolic strategies such as the use of alternative terminal electron acceptors for CH_4 oxidation. Correlation analyses revealed potential community-level interactions among methanotrophs and one-carbon compounds utilizing microorganisms that served as conceptual framework for Chapter 6.

5.1 Introduction

Reduced levels of dissolved O_2 ($<$20 μM kg^{-1}) enhances the use of alternative inorganic compounds as electron acceptors for anaerobic respiration by microorganisms resulting in the production of climate active trace gas such as carbon dioxide (CO_2), nitrous oxide (N_2O) and methane (CH_4) (Lam *et al*., 2009b; Ward *et al*., 2009). OMZs are the largest marine source of CH_4 flux to the atmosphere ($\sim$ 1 Tg CH_4 yr^{-1}) (Naqvi *et al*., 2010). Previous surveys have quantified CH_4 oxidation rates in oxic ($>$ 90 μmol O_2 kg^{-1}) and dysoxic-suboxic ($<$20 μmol O_2 kg^{-1}) OMZ waters as well as anoxic sediments. Although the process of anaerobic oxidation of CH_4 (AOM) can consume $\sim$75% of CH_4 in the sediment (Strous and Jetten, 2004; Knittel and Boetius, 2009), results indicated that anaerobic oxidation of CH_4 (AOM) occurs at a much slower rate than aerobic CH_4 oxidation in the water column (0.72 nmol L^{-1}h^{-1} and 2 nmol L^{-1}h^{-1} respectively) (Ward *et al*., 1989; Ward and Kilpatrick, 1990; 1993). Aerobic CH_4 oxidation has been estimated to consume $>$50% of CH_4 in the water column (Fung *et al*., 1991; Reeburgh *et al*., 1991a) likely having the largest influence on the CH_4 budget before emission to the atmosphere (Reeburgh, 2007). Thus, aerobic CH_4 oxidation provides a second biological filter following sediment AOM that reduces CH_4 flux (Ward *et al*., 1989; Ward and

[1] A version of this chapter appears as **Torres-Beltrán, M.** *et al*. 2016. Methanotrophic community dynamics in a seasonally anoxic fjord: Saanich Inlet, British Columbia. Front. Mar.Sci. doi.org/10.3389/fmars.2016.00268 .

Kilpatrick, 1990; 1993).

Efforts to understand microbial agents driving methane oxidation based on small subunit ribosomal RNA (SSU rRNA) gene surveys from diverse OMZs such as the Eastern Tropical South Pacific, the Namibian Upwelling and the Black Sea (Stevens and Ulloa, 2008; Glaubitz *et al*., 2010), indicate that canonical methanotrophs within the alpha and gammaproteobacteria, are rare microbial community members i.e. $<0.01\%$ of the total microbial community. Parallel efforts to describe the distribution, abundance and potential metabolic activity of the functional gene particulate methane monooxygenase (pMMO) subunit β (*pmoA*), identified pMMO- encoding phylogenetic groups (OPUs), OPU1 to OPU4, affiliated with canonical methanotrophic groups in OMZs waters (Hayashi *et al*., 2007; Tavormina *et al*., 2013). Phylogenetically affiliated with Methylococcales OPU1 and OPU3 groups were initially observed in the Eastern Pacific Ocean OMZ (Hayashi *et al*., 2007), and exhibit differential abundance and distribution patterns. OPU3 was more abundant under low O_2 concentrations in the Costa Rica OMZ water column (Tavormina *et al*., 2013), and expression of pmoCAB for group OPU3 has recently been demonstrated in a metatranscriptome from the Guaymas Basin (Lesniewski *et al*., 2012). In addition to canonical methanotrophs, more recent studies have identified a number of non-canonical microbial groups (e.g. bacteria affiliated with the Verrucomicrobia (Dunfield *et al*., 2007), SAR324 within the deltaproteobacteria (Swan *et al*., 2011) and the NC10 candidate division (Ettwig *et al*., 2010)) with the potential to mediate CH_4 cycling in OMZs. In both known and novel cases there is limited information on the dynamics and interspecific interactions of CH_4 cycling microbes needed to constrain their biological filtering capacity.

Saanich Inlet is a seasonally anoxic fjord on the east coast of Vancouver Island British Columbia. During spring and summer months, restricted circulation and high levels of primary production lead to progressive deoxygenation and the accumulation of CH_4, ammonium (NH_4^+) and hydrogen sulfide (H_2S) in the deep waters of the Saanich Inlet basin. In late summer and fall, upwelling oxygenated nutrient rich ocean waters cascade into the inlet shoaling anoxic bottom waters upward and transforming the redox chemistry of the water column. The recurring seasonal development of water column anoxia followed by deep water renewal makes Saanich Inlet a model ecosystem for evaluating microbial community structure, function and dynamics in relation to changing levels of water column O_2 deficiency extensible to coastal and open ocean OMZs (Walsh *et al*., 2009; Zaikova *et al*., 2010; Walsh and Hallam, 2011; Wright *et al*., 2012).

Process rates and molecular surveys focused on CH_4 oxidation have been previously conducted in Saanich Inlet during peak summer stratification. Process measurements indicated that CH_4 oxidation rates were highest near the oxic-anoxic interface ($\sim$2 nmol L^{-1} d^{-1}) (Ward *et al*., 1989). Subsequently, sequences affiliated with canonical methanotrophs such as Methylococcales within the gammaproteobacteria were recovered as rare biosphere components based on full-length SSU rRNA gene sequences (Zaikova *et al*., 2010) and particulate methane monoxygenase subunit β (*pmoA*) libraries were dominated by a non-canonical phylotype (Stilwell, 2007). In the same study, anaerobic methane oxidizing archaea (ANME) were undetectable in the water column. The low abundance of canonical methanotrophs combined with measured CH_4 oxidation rates in both oxycline and deep basin waters presents a "CH_4 oxidation conundrum" (Ward *et al*., 1989; Ward and Kilpatrick, 1993; Zaikova *et al*., 2010). Here I attempt to constrain this conundrum using taxonomic survey information over two years of seasonal stratification and renewal supplemented with

functional gene information for *pmoA*.

5.2 Methods

5.2.1 Environmental sampling

Environmental monitoring and sample collection were carried out monthly aboard the MSV John Strickland at station SI03 (48^{o} 35.500 N, 123^{o} 30.300 W) in Saanich Inlet, B.C. From February 2008 to July 2010 a total of 288 water samples were collected at 16 high-resolution depths (10, 20, 40, 60, 75, 80, 90, 97, 100, 110, 120, 135, 150, 165, 185 and 200 meters) spanning oxic (>90 μmol O_2 kg^{-1}), dysoxic (90-20 μmol O_2 kg^{-1}), suboxic (20^{-1} μmol O_2 kg^{-1}) anoxic (<1 μmol O_2 kg1) and sulfidic water column compartments (Wright *et al*., 2012). Samples were processed and analyzed as previously reported for the Saanich Inlet time-series (Zaikova *et al*., 2010; Capelle *et al*., 2015; Torres-Beltrán *et al*., 2016b). Briefly, water samples were collected from Niskin or Go-Flow bottles for dissolved O_2 and CH_4, nutrients (Nitrate (NO_3^-), Nitrite (NO_2^-), Ammonium (NH4), Silicon dioxide (SiO_2), Phosphate (PO4), Hydrogen sulfide (H_2S)), DNA, and RNA. In addition, conductivity, temperature, and depth were measured using a Seabird SBE19 CTD-device (Sea-Bird Electronics Inc., Bellevue USA), with a PAR and O_2 sensor attached. CTD was also used to measure dissolved O_2 throughout the water column. Dissolved gases and nutrient measurement protocols have been previously reported for the Saanich Inlet time-series (Zaikova *et al*., 2010; Capelle *et al*., 2015; Torres-Beltrán *et al*., 2016b), and data is available through Dryad Digital Repository (www.dryad.org).

5.2.2 Nucleic acid sampling and extraction

Biomass to generate full-length small subunit (SSU) rRNA gene clone library sequences and *pmoA* gene library sequences (Stilwell, 2007), and metagenomic datasets was collected on February 2006 and February 2010 respectively, at six depths (10, 100, 120, 135, 150 and 200 m) and filtered with an in-line 2.7 μm GDF glass fibre pre-filter onto a 0.22 μm Sterivex polycarbonate cartridge filter. Biomass to generate SSU rRNA pyrotag datsets for microbial community composition profiling (2 L) and filtered directly onto a 0.22 μm Sterivex polycarbonate cartridge filter from high-resolution depths collected between May 2008 to July 2010. Biomass for RNA analysis (2 L) was collected on February 2010 at six depths (10, 100, 120, 135, 150 and 200 m), and filtered with an in-line 2.7 μm GDF glass fibre pre-filter onto a 0.22 μm Sterivex filter within 20 min of shipboard sample collection.

Environmental DNA was extracted from Sterivex filters as previously described (Wright *et al*., 2009; Zaikova *et al*., 2010). Briefly, after defrosting Sterivex on ice, 100 μl lysozyme (0.125 mg ml^{-1}; Sigma) and 20 μl of RNAse (1 μl ml^{-1}; ThermoFisher) were added and incubated at 37 ^{o}C for 1 h with rotation followed by addition of 50 μl Proteinase K (Sigma) and 100 μl 20% SDS and incubated at 55 ^{o}C for 2 h with rotation. Lysate was removed by pushing through with a syringe into 15 mL falcon tube (Corning) and with an additional rinse of 1 mL of lysis buffer. Filtrate was subject to chloroform extraction (Sigma) and the aqueous layer was collected and loaded onto a 10K 15 ml Amicon filter cartridge (Millipore), washed three times with TE buffer (pH 8.0) and concentrated to a final volume of between 150-400 μl. Total

DNA concentration was determined by PicoGreen assay (Life Technologies) and genomic DNA quality determined by visualization on 0.8% agarose gel (overnight at 16V).

Total RNA was extracted from Sterivex filters using the mirVana Isolation kit (Ambion) (Shi *et al*., 2009; Stewart *et al*., 2010) protocol modified for sterivex filters (Hawley *et al*., 2017b). Briefly, after thawing the filter cartridge on ice RNA later was removed by pushing through with a 3 ml syringe followed by rinsing with an additional 1.8 mL of Ringers solution and incubated at room temperature for 20 min with rotation. Ringers solution was evacuated with a 3 ml syringe followed by addition of 100 μl of 0.125 mg ml^{-1} lysozyme and incubated at 37 ^{o}C for 30 min with rotation. Lysate was removed from the filter cartridge and subjected to organic extraction following the mirVana kit protocol. DNA removal and clean up and purification of total RNA were conducted following the TURBO DNA-free kit (ThermoFisher) and the RNeasy MinElute Cleanup kit (Qiagen) protocols respectively.

5.2.3 Small subunit ribosomal RNA gene sequencing and analysis

To initially assess microbial community diversity in the Saanich Inlet water column full-length bacterial SSU rRNA gene sequences datasets were generated from the February 2006 LV samples as previously described (Zaikova *et al*., 2010). Briefly, DNA extracts from 10, 100, 120 and 200 m samples were amplified using SSU rRNA primers targeting the bacterial domain: B27F (5'- AGAGTTTGATCCTGGCTCAG) and U1492R (5'-GGTTAC CTTAGTTACGACTT) under the following PCR conditions: 3 min at 94 ^{o}C followed by 35 cycles of 94 ^{o}C for 40 s, 55 ^{o}C for 1.5 min, 72 ^{o}C for 2 min and a final extension of 10 min at 72 ^{o}C. Each 50 ml reaction contained 1 μl of template DNA, 1 μl each 10 mM forward and reverse primer, 2.5 U Taq (Qiagen), 5 ml 10 mM deoxynucleotides, and 41.5 μl 1X Qiagen PCR Buffer. Clone library construction and screening. The SSU rRNA gene amplicons were visualized on 1% agarose gels in 1X TAE and purified using the MinElute PCR Purification Kit (Qiagen) according to the manufacturer's instructions. Approximately 4 μl of each purified SSU rRNA gene product was cloned into a pCR4-TOPO vector using a TOPO TA cloning kit for sequencing (Invitrogen) and transformed by chemical transformation into Mach-1-T1R cells according to the manufacturer's instructions. Transformants were transferred to 96-well plates containing 180 ml LBkan50 and 10% glycerol and grown overnight at 37 ^{o}C prior to storage at -80 ^{o}C. Cloned inserts were amplified directly from glycerol stocks with M13F (5'-GTAAAACGACGGCCAG) and M13R (5'- CAGGAAACAGCTATGAC) primers using the SSU rRNA gene PCR protocol. Bidirectional end sequencing was performed on a Sanger platform at the Department of Energy Joint Genome Institute (DOE-JGI; Walnut Creek, CA).

To survey microbial community structure and dynamics throughout the complete water column profile over time, DNA extracts from the samples from May 2008 to July 2010 were used to generate SSU rRNA gene pyrotag datasets. Pyrotag libraries were generated by PCR amplification using multi-domain primers targeting the V6-V8 region of the SSU rRNA gene (Allers *et al*., 2013): 926F (5'-cct atc ccc tgt gtg cct tgg cag tct cag AAA CTY AAA KGA ATT GRC GG-3') and 1392R (5'-cca tct cat ccc tgc gtg tct ccg act cag-<XXXXX>-ACG GGC GGT GTG TRC-3'). Primer sequences were modified by the addition of 454 A or B adapter sequences (lower case). In addition, the reverse primer included a 5 bp barcode designated <XXXXX> for multiplexing of samples during sequencing. Twenty-five microliter PCR reactions were

performed in triplicate and pooled to minimize PCR bias. Each reaction contained between 1 and 10 ng of target DNA, 0.5 μl Taq DNA polymerase (Bioshop inc. Canada), 2.5 μL Bioshop 10 x buffer, 1.5 uL 25 mM Bioshop $MgCl_2$, 2.5 μL 10 mM dNTPs (Agilent Technologies) and 0.5 μL 10 mM of each primer. The thermal cycler protocol started with an initial denaturation at 95 ^{o}C for 3 minutes and then 25 cycles of 30 s at 95 ^{o}C, 45 s at 55 ^{o}C, 90 s at 72^{o}C and 45 s at 55 ^{o}C. Final extension at 72 ^{o}C for 10 min. PCR products were purified using the QiaQuick PCR purification kit (Qiagen), eluted elution buffer (25 μL), and quantified using PicoGreen assay (Life Technologies). SSU rRNA amplicons were pooled at 100 ng DNA for each sample. Emulsion PCR and sequencing of the PCR amplicons were sequenced on Roche 454 GS FLX Titanium at the DOE-JGI, or the McGill University and Génome Québec Innovation Center.

Pyrotag sequences were processed using the Quantitative Insights Into Microbial Ecology (QIIME) software package (Caporaso *et al*., 2010). To minimize the removal of false positive reads all 3,985,489 pyrotag sequences generated from the 288 samples were clustered together. Reads with a length shorter than 200 bases, ambiguous bases, and homopolymer sequences were removed prior to chimera detection. Chimeras were detected and removed using chimera slayer provided in the QIIME software package. Sequences were then clustered into operational taxonomic units (OTUs) at 97% identity using uclust with average linkage algorithm. Prior to taxonomic assignment singleton OTUs (OTUs represented by one read) were omitted, leaving 69,051 OTUs. Representative sequences from each non-singleton OTU were queried against the SILVA database (Pruesse *et al*., 2007) and the Greengenes database (DeSantis *et al*., 2006) using BLAST (Altschul *et al*., 1990).

5.2.4 pmoA gene libraries, metagenomic and metatranscriptomic sequencing and analysis

Extracted total DNA from the February 2006 samples were used to identify bacterial particulate methane monooxygenase (pMMO) subunit β (*pmoA*) presence in the Saanich Inlet water column. Briefly, *pmoA* sequences were PCR amplified (Stilwell, 2007) using gene-specific forward and reverse primers A189F (5'- GGNGACTGGGACTTCTGG) (Holmes et al., 1995) and mb661R (5'- CCGGMGCAACGTCYTTACC) (Costello and Lidstrom, 1999) and the following PCR profile: 30 cycles of 96 ^{o}C for 25 s, 54 ^{o}C for 45 s, and 72 ^{o}C for 50 s. Each 50 μl reaction contained l μl of template DNA, 2 μl each 10 μM forward and reverse primer, 2.5U Taq (Qiagen), 4 μl 10 mM deoxynucleotides, and 40.5 μl 1x Buffer (Qiagen-TaqPolymeraseKit). pmoA amplicons were purified using the MinElute PCRPurificationKit(Qiagen, CA), cloned into a pCR4-TOPO vector using a TOPO TA cloning kit for sequencing (Invitrogen, Carlsbad CA), and transformed by chemical transformation into Mach-1-T1R cells according to the manufacturer's instructions. Cloned inserts were amplified directly from glycerol stocks for fingerprint screening using the common 4-base cutter Rsa I (Invitrogen, CA). Restriction patterns were visually inspected and unique patterns selected for Sanger sequencing through the McGill University and Génome Québec Innovation Centre (Montreal, Quebec, Canada) (Stilwell, 2007). Four libraries were constructed from 10, 100, 120 and 200 m, from which a total of 33 representative sequences were identified.

Extracted total DNA (6 samples) and RNA (6 samples) from February 2010, corresponding to the 10, 100, 120, 135, 150 and 200 m depth intervals, were used to generate metagenomic and metatranscriptomic datasets at the DOE-JGI following the protocols for library production and sequencing and assembly previ-

ously described for the Saanich Inlet time-series (Hawley et al., 2016).

A total of 6 assembled metagenomes and 6 assembled metatranscriptomes were analysed using MetaPathways V2.5.1, an open source pipeline for predicting reactions and pathways using default settings (Konwar et al., 2013) (https://github.com/hallamlab/metapathways2/wiki). For each gene, reads per kilobase per million mapped (RPKM) was calculated as a proportion of the number of reads mapped to a sequence section, normalized for sequencing depth and ORF length (Konwar et al., 2015). RPKM values and relative abundance to the total number of reads were used to describe the abundance of *pmoA* genes and transcripts.

5.2.5 Phylogenetic inference

To generate a reference phylogenetic tree for methanotrophic bacteria, full-length SSU rRNA gene sequences from Saanich Inlet, and diverse environmental reference sequences affiliated with methanotrophic bacteria were first aligned and compared. Full-length SSU rRNA gene sequences from Saanich Inlet were edited manually using Sequencher software V4.1.2 (Gene Codes Corporation) and imported into the full-length SSU SILVA database (http://www.arb-silva.de) and aligned to the closest relative. Sequences affiliated with known methanotrophs were extracted from the dataset. Methanotroph reference sequences from Saanich Inlet in addition to 53 reference sequences for methanotrophic bacteria including cultured type I and II methanotrophs, OMZ representatives, mussel symbionts and envioronmental clones were clustered at 97% identity using mothur v.1.19.0 (Schloss *et al*., 2009). A total of 88 sequences affiliated with Methylococcales were recovered representing 1.3% from the total Saanich Inlet full-length SSU rRNA sequences generated. Sequences clustered at 97% identity resolved into 11 distinct clusters, 5 of which contained most of the identified sequences (93%). Representative sequences for the most abundant clusters revealed 4 subgroups with phylogenetic similarity to environmental representatives of type I methanotrophs: Methylococcaceae (Mou *et al*., 2008), putative methanotrophic group OPU3 (Hayashi *et al*., 2007; Tavormina *et al*., 2010; Tavormina *et al*., 2013), environmental seafloor clones (Santelli *et al*., 2008), and methanotrophic symbionts (Streams *et al*., 1997; Dubilier *et al*., 2008; Petersen and Dubilier, 2009). Representative sequences for the most abundant (represented by more than one sequence) clusters were identified using the get.oturep command in mothur and were included in the phylogenetic tree. Representative sequences were aligned using the SSU SILVA database and imported into the ARB software (Ludwig *et al*., 2004) for tree distance matrix and alignment generation using the ARB parsimony tool. ARB sequences were exported to Mesquite (V.2.0) and edited manually. A maximum likelihood phylogenetic tree was inferred by PHYML (Guindon *et al*., 2005) using an GTR model of nucleotide evolution where the parameter of the gamma distribution, the proportion of invariable sites and the transition/transversion ratio were estimated for each data set. The confidence of each node was determined by assembling a consensus tree of 1,000 bootstrap replicates.

To further resolve diversity of methanotrophic bacterial OTUs, I recruited pyrotag OTU sequences to full-length SSU rRNA gene sequences described above. Pyrotag sequences affiliated with methanotrophs based on BLAST-comparison in QIIME were re-clustered with SSU rRNA gene tree reference sequences using a 97% identity cut-off in mothur (Schloss *et al*., 2009). In addition, blastn was used to query representative pyrotag sequences from clusters against full-length SSU rRNA reference tree sequences. Only hits

with a perfect match across the full length of a query sequence were retrieved, and the number of pyrotags mapping to all sequences in each cluster was summed. Clusters represented by one pyrotag sequence were not used in downstream analyses. Representative pyrotag sequences were aligned in ARB to reference tree sequences, imported to Mesquite for manual edition, and finally, included in the phylogenetic tree inferred by PHYML using the parameters detailed above.

To generate a reference phylogenetic tree for PmoA, conceptually translated and annotated ORFs from the metagenomic and metatranscriptomic datasets were manually extracted from the functional annotation table <*ORF_annotation_table.txt*> in the <*results/annotation table*> output directory. Sequences were aligned and compared to diverse environmental and reference PmoA sequences. A total of 60 PmoA sequences, retrieved from metagenomic (34 sequences) and metatranscriptomic (26 sequences) datasets, were clustered over a range of identity thresholds using the UClust algorithm (USEARCH V6.0) with 52 reference sequences including Saanich Inlet *pmoA* gene libraries, cultured and environmental sequences affiliated with Type I and II methanotrophs, and novel *pmoA* phylotypes found within the SAR324 clade, Verrucomicrobia and Candidate *Methylomirabilis oxyfera* NC10. Reference sequences also included ammonium monooxygenase subunit α (AmoA). The 97% identity threshold was selected based on resolution of the OPUs and symbiont groups. Cluster representative sequences were aligned using the Multiple Sequence Comparison by Log- Expectation (MUSCLE) method (EMBL-EBI), and was manually curated in Mesquite. A maximum likelihood phylogenetic tree was inferred by PHYML (Guindon *et al*., 2005) using an WAG model of amino acid evolution where the parameter of the gamma distribution, the proportion of invariable sites and the transition/transversion ratio were estimated. The confidence of each node was determined by assembling a consensus tree of 1000 bootstrap replicates.

5.2.6 Statistical analyses

Pyrotag datasets were normalized to the total number of reads per sample, and environmental parameter data were transformed to the same order of magnitude so that each variable had equal weight. Hierarchical cluster analysis (HCA) was conducted to identify groups associated with discrete water column compartments. In addition, OTUs were correlated using nonmetric multidimensional scaling (NMDS) with environmental parameters. Hierarchical cluster and NMDS analyses of microbial community compositional profiles were done using the pvclust (Suzuki and Shimodaira, 2015) and MASS (Venables and Ripley, 2002) packages in the R software (RCoreTeam, 2013) with Manhattan Distance measures, and statistical significance to the resulting clusters was computed as bootstrap score distributions with 1,000 iterations and NMDS stress value $<$0.05.

Multi-level indicator species analysis (ISA) using the indicspecies package (De Caceres and Legendre, 2009) in the R software (RCoreTeam, 2013) was performed to identify OTUs specifically associated with different water column compartments defined by HCA. The ISA/multi-level pattern analysis calculates p values with Monte Carlo simulations and returns indicator values (IV) and p-values with $\alpha < 0.05$. The IVs range between 0 and 1, where indicator OTUs considered in the present chapter for further community analysis shown an IV > 0.7 and p-value < 0.001.

A multivariate regression analysis (Fox and Weisberg, 2011) was conducted on time-series pyrotag data

to infer significant correlations between OTUs affiliated with methanotrophic bacteria while controlling for the effect of depth. Analysis was conducted in the R environment (RCoreTeam, 2013). Parameter estimates were calculated with least squares fit between relative abundances and depth measurements. Only representative OTUs affiliated with methanotrophic bacteria were regressed. Statistical significance of correlations was determined with bootstrapping (1,000 iterations), and using a Type-I error rate of 5%. Univariate regression analyses were conducted on time-series pyrotag data to infer correlations between OTUs affiliated with methanotrophic bacteria counts and environmental variables (O_2, CH_4, NO_3^-, NO_2^-, and H_2S). Choice of model per OTU was determined by AIC testing (Akaike, 1998), using a combination of negative binomial regression (Venables and Ripley, 2002) and zero-inflated negative binomial regression (Zeileis *et al*., 2008) employed due to many zeroed observations (average: 54%). Parameter estimates and statistical significance were calculated in the R environment (RCoreTeam, 2013), using a Type-I error rate of 5%.

Co-occurrence networks To generate a robust network emphasizing co-occurrences between prevalent OTUs in water column compartments defined by HCA rather than individual depth intervals, the Bray-Curtis and Spearman's rank correlations were used. Correlation coefficients were calculated using CoNet (Faust *et al*., 2012) with OTU abundance as count data. First, an OTU matrix derived from the pyrotag taxonomic analysis was transformed into presence-absence data to remove OTUs with less than 1/3 zero counts, leaving a matrix of 780 OTUs for all samples. Next, to construct ensemble networks, measure-specific thresholds set to 0.6 were used as a pre-filter and edge scores were computed only between clade pairs. To assign statistical significance to the resulting scores, edge and measure-specific permutation and bootstrap score distributions with 1,000 iterations each were computed. p-values were tail-adjusted so that low p-values correspond to co-presence and high p-values to exclusion. After merging, p-values on each final edge were corrected to q-values (cut-off of 0.05). The positivity or negativity of each relationship was determined by consensus voting over all integrated data sources. Finally, only edges with at least two supporting pieces of evidence were retained.

The final edge matrix was visualized as a force directed network using Cytoscape 2.8.3 (Shannon *et al*., 2003). Network properties were calculated with the "Network Analysis" Plug-In. Nodes in the co-occurrence network corresponded to individual OTUs and edges were defined by computed correlations between corresponding OTU pairs. The layout revealed distinct modules, which persisted after lowering the correlation coefficient cut-off for edge creation to 0.90 reinforcing the robustness of the network. Edges from modules were selected and visualized as sub-networks using the tool Hive Panel Explorer (https://github.com/hallamlab/HivePanelExplorer/wiki) (Perez, 2015). HivePlotter allows for edge selection based on interactions within specific OTUs such as those affiliated with methanotrophic bacteria.

5.2.7 Data deposition

The SSU rRNA gene sequences reported in this chapter have been submitted to the The National Center for Biotechnology Information (NCBI) under BioSample numbers: GQ346856, GQ349233, GQ347199, HQ163221, GQ350623, GQ349295. The SSU rRNA pyrotag sequences reported in this chapter have been submitted to the NCBI under BioSample numbers: SAMN03387532 - SAMN03387915. Metagenomes reported in this chapter have been submitted to the NCBI under BioSample numbers: SAMN05224436,

SAMN05224437, SAMN05224442, SAMN05224443, SAMN05224447, SAMN05224451. Metatranscriptomes reported in this chapter have been submitted to the NCBI under BioSample numbers: SAMN05238739, SAMN05238741, SAMN05238743, SAMN05238745, SAMN05238748, SAMN05238751.

5.3 Results

5.3.1 Water column conditions

I monitored changes in water column conditions corresponding to the progression of stratification (from winter through mid-summer) and deep water renewal (late summer into fall) events over a 2-year period (May 2008 to July 2010). As previously described for OMZs (Wright *et al.*, 2012), in this chapter I define water column compartments on the basis of O_2 concentration ranges: oxic (>90 μM O_2), dysoxic (90-20 μM O_2), suboxic (<20-1 μM O_2) and anoxic (<1μM). Between May and August 2008, as water column stratification peaked, suboxic conditions intensified. This intensification corresponded with increasing levels of CH_4 and hydrogen sulphide (H_2S) below 150 m consistent with the development of deep-water anoxia. Two CH_4 concentration peaks were observed, at the subsurface (20 100 m) ranging between the $20^{-1}20$ nM, and at in the deep basin increasing steadily below 150m to a maximum of 800 nM at 200 m in late July. Over the same time interval, the concentration of H_2S ranged between 2 to 8 μM in the anoxic bottom waters (150-200m). The concentration of NO_3^- between the surface and 100m ranged between 5 to 20 μM, decreasing rapidly between 100 and 135 m before reaching a minimum of <1 μM in anoxic bottom waters. The concentration of NO_2^- between surface and 135 fm ranged between <0.1 to 0.25 μM reaching a maximum 0.3 μM in anoxic bottom waters (150-200 m) (Fig. C.1).

The beginning of 2008 deep-water renewal occurred between the end of July and early September, continuing through November. During this time interval, oxygenated nutrient-rich waters flowed over the sill, displacing anoxic bottom waters upwards and disrupting the redox gradient established during spring and summer months. Between September and October, dissolved O_2 was observed throughout the water column, although upwards shoaling O_2 and NO_3^- depleted bottom waters produced an intermediate suboxic layer between 100 and 135 m. Concomitantly, CH_4 concentrations increased transiently in the suboxic transition zone (120-135 m) ranging between 180-700 nM while decreasing below 135 m to 0 nM in bottom waters. In November, O_2 and NO_3^- and NO_2^- concentrations continued to increase above 100 and below 135m with intensification of water column O_2 deficiency within intervening depth intervals (Fig. C.1).

Beginning in December 2008 water column O_2 deficiency intensified below 100 m consistent with stratification. However, in contrast to previous studies at Saanich Inlet, these observations indicate a significantly weak renewal occurred in fall 2009 as indicated by no measurable increase in O_2 or NO_3^- in deep basin waters. This phenomenon extended O_2 deficiency below 100 m resulting in CH_4 accumulation through the summer of 2010, and corresponded with anoxic conditions below 135 m (O_2 and NO_3^- concentrations equal to 0 μM combined with high levels of H_2S). Over this period, the concentration of CH_4 between surface and 100 m ranged between 20-500 nM increasing with depth to reach a maximum 1250 nM in anoxic bottom waters in July 2010. The concentration of H_2S ranged between 2-20 μM in the anoxic waters. Over this period, the highest NO_2^- concentration (1.8 μM) was observed at 120 m in July 2010 (Fig.C.1).

5.3.2 Microbial community structure

To identify microbial agents driving methane oxidation in the Saanich Inlet water column I analyzed SSU rRNA gene pyrotag sequences from 288 samples collected at 16 depths (10-200 m) over a two year time period between May 2008 and July 2010. Results revealed consistent microbial community partitioning as commonly observed in stratified ecosystems where O_2 deficiency is associated with redox-driven niche partitioning (Alldredge and Cohen, 1987; Shanks and Reeder, 1993; Wright *et al*., 2012). For instance, hierarchical cluster analysis (HCA) resolved three major groups or clusters (AU $>$ 70, 1,000 iterations) associated with oxic (group I), dysoxic-suboxic (group II), and anoxic (group III) water column conditions (Fig. 5.1A). Consistent with oxycline formation, O_2 (R^2 = 0.80) and NO_3^- (R^2 = 0.63) were negatively correlated with groups II and III (Fig. 1B), while positively correlated with NH4 (R^2 = 0.26) and H_2S (R^2 = 0.52) (Fig. 5.1B).

Overall, microbial community composition was dominated (relative abundance $>$1%) (Rapp and Giovannoni, 2003) by OTUs affiliated with ubiquitous and abundant taxonomic groups previously identified in marine O_2 deficient environments (Field *et al*., 1997; Fuhrman and Davis, 1997; Brown and Donachie, 2007; Tripp *et al*., 2008; Lam *et al*., 2009; Walsh *et al*., 2009; Zaikova *et al*., 2010; Walsh and Hallam, 2011; Wright *et al*., 2012) including SAR11, SAR324, Nitrospina, SUP05, Marine Group A and Methylophilales within the bacterial domain, and Thaumarchaeota within the archaeal domain (Fig. 5.2). Consistent with previous observations in OMZs (Stevens and Ulloa, 2008; Glaubitz *et al*., 2010), OTUs affiliated with Methylococcaceae and *Methylomonas*, both canonical type I methanotrophs, were identified among the rare biosphere (relative abundance $<$1%) (Sogin *et al*., 2006). Additionally, OTUs affiliated with methylotrophic bacteria such as Methylobacteriaceae and *Methylophaga*, were also identified among the rare biosphere (Fig. 5.3A).

5.3.3 Methanotroph diversity and dynamics

Methanotroph diversity in Saanich Inlet was determined based on recruitment of representative OTU sequences to full-length SSU rRNA gene reference sequences (full description of reference sequences used in Methods). Initially, I identified all OTUs with a taxonomic assignment affiliated with methanotrophic bacteria using BLAST-based comparisons conducted in QIIME queried against the Silva and Greengenes reference databases. Sequences affiliated with these OTUs shared 90% identity with Methylococcales reference sequences in Silva, and 90% identity with Methylococcaceae and *Methylomonas* reference sequences in Greengenes. Subsequently, further analysis on representative methanotrophic OTUs was conducted using full-length SSU rRNA gene sequences as fragment recruitment platforms. A total of 3,804 sequences affiliated with type I methanotrophs were recovered, representing 0.095% of the total pyrotag sequences generated. Sequences clustered at 97% identity resolved into 66 distinct OTUs, 6 of which contained 75% of total methanotroph sequences. Similar to full-length SSU sequences, pyrotag OTUs revealed 4 subgroups with phylogenetic similarity to cultured and environmental representatives of type I methanotrophs: Methylococcaceae (4%) (Mou *et al*., 2008), putative methanotrophic groups OPU1 (22.55%) and OPU3 (26.23%) (Hayashi *et al*., 2007; Tavormina *et al*., 2010; Tavormina *et al*., 2013), and methanotrophic symbionts (46.89%) (Streams *et al*., 1997; Dubilier *et al*., 2008; Petersen and Dubilier, 2009) (Fig. 3).

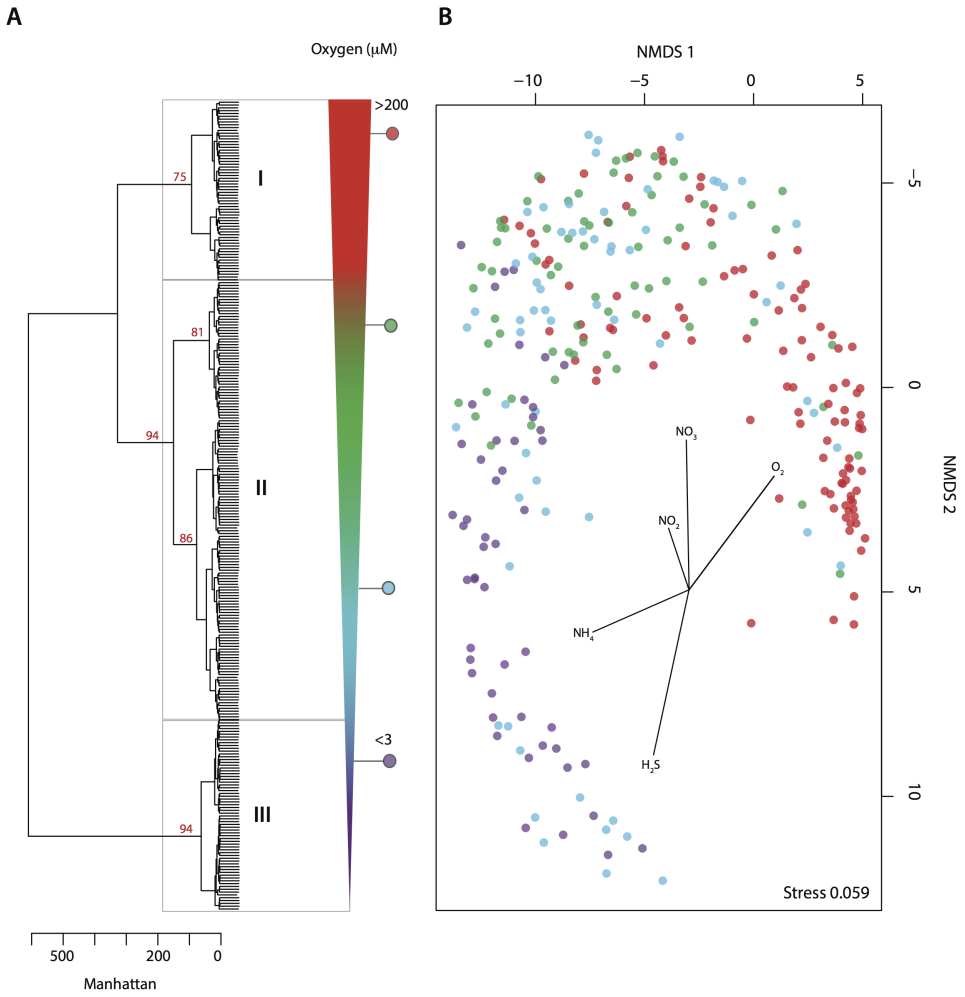

Figure 5.1: Microbial community partitioning under changing levels of water column O_2 deficiency. A) Hierarchical clustering of pyrotag data (May 2008 July 2010) based on Manhattan distance. Clusters are delimited by O_2 concentration range represented by number from I to III: oxic = I, dysoxic- suboxic = II and anoxic =III. Bootstrap values (1000 iterations) are shown in red. B) NMDS depicts microbial community partitioning along redox gradient showing correlation with environmental parameters. Samples are depicted in color dots according to O_2 concentration.

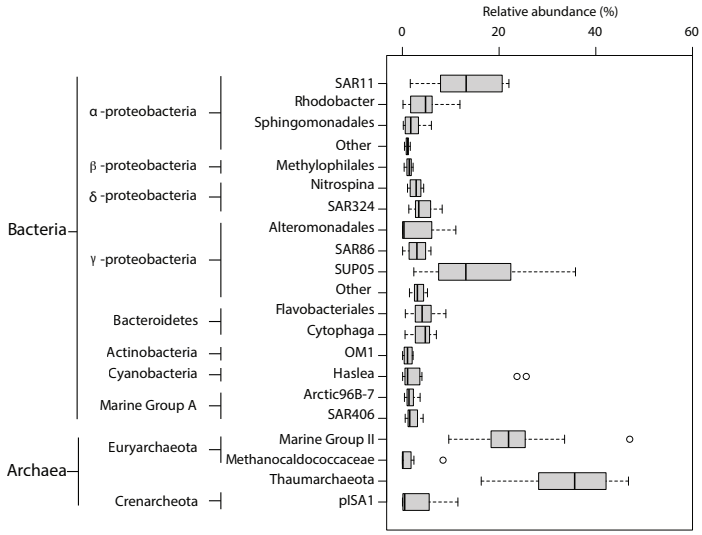

Figure 5.2: Taxonomic composition of OTUs identified in SSU rRNA gene pyrotags between 2008-2010. Abundant (>1% relative abundance) taxa found in pyrotag dataset. The size of each box represents the average of relative abundance (>0.01%) calculated from the total number of prokaryotic reads throughout the water column over this period. Extended dashed lines (whiskers) represent at the base the lower and upper quartiles (25% and 75%) and at the end the minimum and maximum values encountered. The middle line represents the median.

The most abundant OTUs were related to putative methanotrophic OPU1 (OPU_01 = OTU55333; 16.9%), OPU3 (OPU3_01 = OTU6504; 8.4%), and methanotrophic symbionts (Symbiont_01 = OTU39693; 45.5%) (Fig.5.3B), comprising ∼70% of total methanotroph sequences found in the pyrotag datasets. Based on this information I focused our analysis on these three groups.

Population dynamics of OPU1_01, OPU3_01 and Symbiont_01 were determined throughout the water column between May 2008 and July 2010. OPU1_01 and OPU3_01 were more abundant under dysoxic (< 90 μM O_2) and suboxic (<20 μM O_2) conditions while Symbiont_01 was more abundant under suboxic and anoxic (<3 μM O_2) conditions (Fig. 5.4A). Specifically, OPU1_01 showed two abundance peaks: 1) after the 2008 renewal (September November) reaching up to 0.24% relative abundance under suboxic conditions (150 and 165 m), and 2) during water column stratification (March and April 2009) reaching up to 0.15% relative abundance under oxic conditions (10 85 m). OPU3_01 peaked during stratification periods (July August 2008 and March April 2009) reaching up to 0.15% relative abundance under oxic and dysoxic conditions (10 97m), and during the extended stratification in July 2010 reaching up to 0.2% relative abundance under suboxic conditions (120-135 m) (Fig. 5.4B). Symbiont_01 also showed two abundance peaks: 1) during the 2008 stratification (May July) reaching up to 0.23% relative abundance, and 2) after the 2008 renewal (November to January 2009) reaching up to 0.6% relative abundance under oxic conditions (40-97 m), and 0.55% under suboxic anoxic (165 200 m) conditions (Fig. C.2). Cumulative abundance for methanotroph OTUs throughout the water column indicated co-occurrence and peak abundance under dysoxic-suboxic water column conditions over time, suggesting that the highest CH_4 oxidation activity is

carried out between these depth intervals in Saanich Inlet (Fig. C.2 and C.3).

Observed OTU distribution patterns for OPU1_01, OPU3_01 and Symbiont_01 became increasingly compartmentalized during the extended stratification period. For instance, OPU1_01 was observed under dysoxic conditions (90-110 m) and OPU3_01 under dysoxic-suboxic conditions (100-135 m), while Symbiont_01 under suboxic-anoxic conditions (135-200 m) (Fig. 5.4). These observations point to redox-driven niche partitioning among methanotrophic bacteria with the potential to mediate CH_4 oxidation in different water column compartments. Given these distribution patterns in relation to measured geochemical profiles i.e OPU3 and NO_2^- (Fig. 5.4B), I hypothesized the use of alternative terminal electron acceptors in the CH_4 oxidation process.

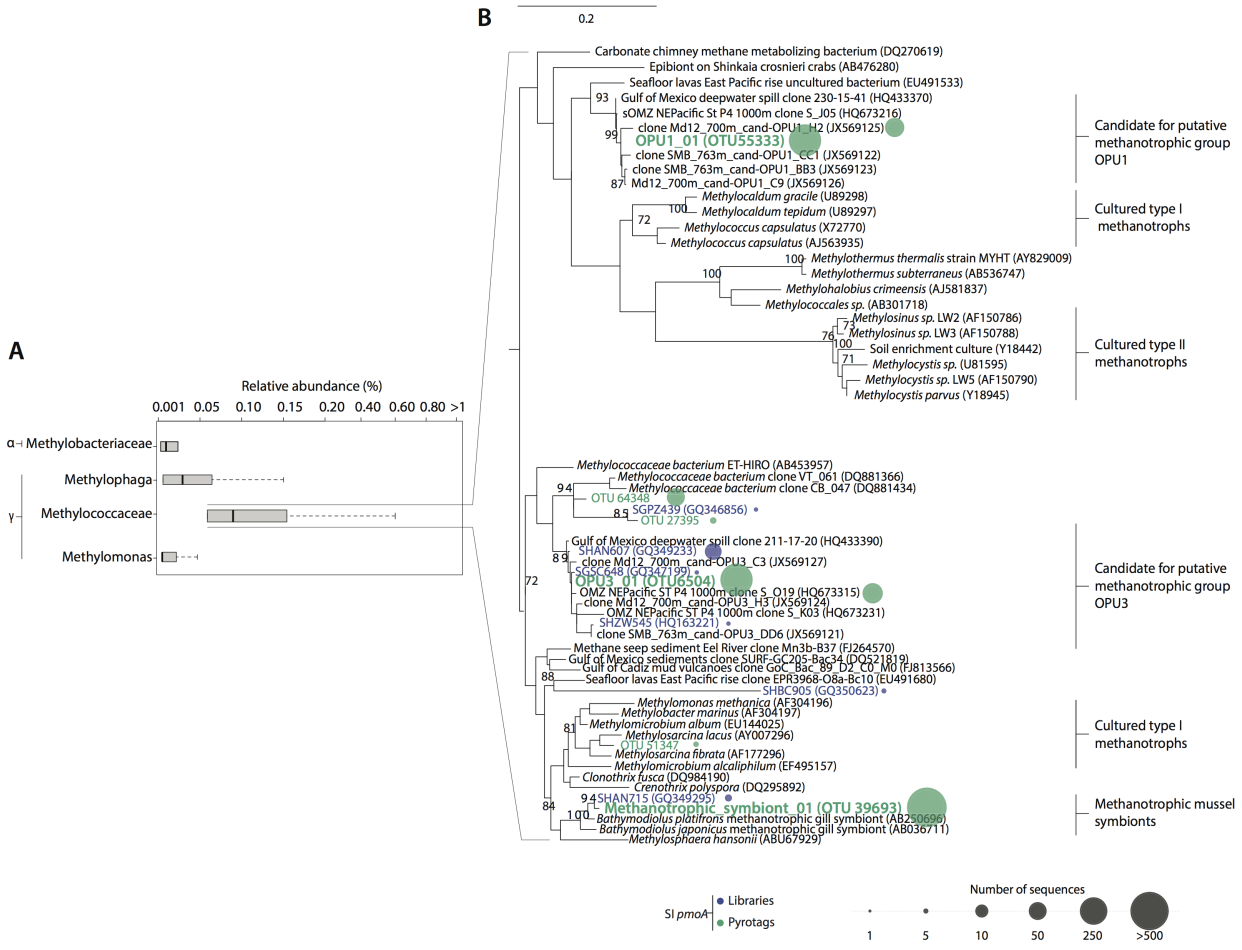

Figure 5.3: Relative abundance and phylogenetic relationships between Type I and Type II methanotroph OTUs in Saanich Inlet. A) Rare (<1% relative abundance) methanotrophic and methylotrophic taxa found in SSU rRNA gene pyrotag dataset. The size of each box represents the average of relative abundance (>0.01%) calculated from the total number of prokaryotic reads throughout the water column over this period. Extended dashed lines (whiskers) represent at the base the lower and upper quartiles (25% and 75%) and at the end the minimum and maximum values encountered. The middle line represents the median. B) Tree inferred using maximum likelihood implemented in PHYML. The percentage (>70%) of replicates in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. Reference sequences for lineages are shown in black. Representative sequences obtained from Saanich Inlet SSU rRNA gene libraries clustered at 97% similarity are shown in blue. OTUs representative sequences (97% similarity) from pyrotag datasets taxonomically identified as methanotrophs are shown in green. Sequences abundance is depicted by colored circles whose circumference indicated the total number of sequences (reads) within the cluster.

5.3.4 PmoA diversity and expression

Given the distribution of OPU1_01, OPU3_01 and Symbiont_01 OTUs in the Saanich Inlet water column I was interested in determining functional potential and activity of these groups. This was determined based on the number of conceptually translated *pmoA* genes and transcripts found throughout the water column in February 2010 and the recruitment of water column PmoA sequences to selected reference sequences (see Methods). Initially, I identified all sequences with a functional assignment affiliated with PmoA using the MetaPathways functional annotation table output. Similar to observations made for methanotroph OTU abundance, PmoA sequences recovered in the metagenomic and metatranscriptomic datasets were rare, representing <0.001% (34 sequences) and <0.018% (26 sequences) from the total number of predicted proteins in the metagenomic and metatranscriptomic datasets, respectively. Representative sequences in *pmoA* clone libraries were mostly related to OPU3 (Hayashi *et al*., 2007), and methanotrophic symbionts within the type I methanotroph clade. However, all February 2010 metagenomic, and metatrasncriptomic PmoA representative sequences clustered at 97% similarity were related to OPU3 (Hayashi *et al*., 2007) (Fig. 5.5). Corresponding metatransciptome RPKM values for *pmoA* in February 2010 showed differential expression across the redox transition zone under dysoxic-suboxic conditions. Transcript expression, depicted as dots sized based on RPKM values (Fig. 5.5), was higher at 100m (mean RPKM = 124, SD = 12.10) than at 120 (mean RPKM = 15.25, SD = 7.33) and 135m (mean RPKM = 17.05, SD = 0.58) potentially indicating the boundaries for OPU-mediated CH_4 oxidation in the Saanich Inlet under water column stratification conditions (Fig. 5.5). v

5.3.5 Methanotroph niche partitioning and co-ocurrence patterns

To better constrain niche partitioning among methanotrophic bacteria and the potential use of alternative terminal electron acceptors, i.e NO_3^- and NO_2^-, I conducted multivariate linear and beta regression analyses on the time-series data. Multivariate regression allowed me to minimize the possible linear effect of depth on OTU distribution while beta regression allowed me to use environmental and relative abundance data to model variable correlations (Kieschnick and McCullough, 2003; Ferrari and Cribari-Neto, 2004). Multivariate analysis showed significant positive correlation ($p < 0.001$) between OPU1_01 and OPU3_03 over stratification periods. However, negative correlation ($p < 0.01$) between OPU1_01 and symbiont_01 was also observed during these time periods. In addition, significant positive correlation ($p < 0.0001$) between OPU3_01 and Symbiont_01 was also observed during transitioning to stratification periods (Fig. C.2) (Table 5.1). Negative binomial regression results on the time-series data indicated OPU1_01 distribution was significantly and negatively correlated with O_2 ($p < 0.05$) and CH_4 (p= 0.001) and weakly correlated with NO_3^-, NO_2^- and H_2S. OPU3_01 distribution was significantly and negatively correlated with CH_4 (p= 0.001) and NO_2^- (p= 0.05), and weakly correlated with NO_3^- and H_2S (Table 2, Fig. 4). Symboint_01 distribution was weakly significantly and negatively correlated with CH_4, H_2S, $NO_3^-$and NO_2^- (p <0.001), and weakly correlated with O_2 (Table 5.2).

To identify characteristic OTUs occurring under specific water column oxygen conditions, multi-level indicator species analysis (ISA) was conducted based on groups resolved in HCA. A co-occurrence network was then constructed to identify potential interactions with methanotrophic OTUs. Multi-level ISA iden-

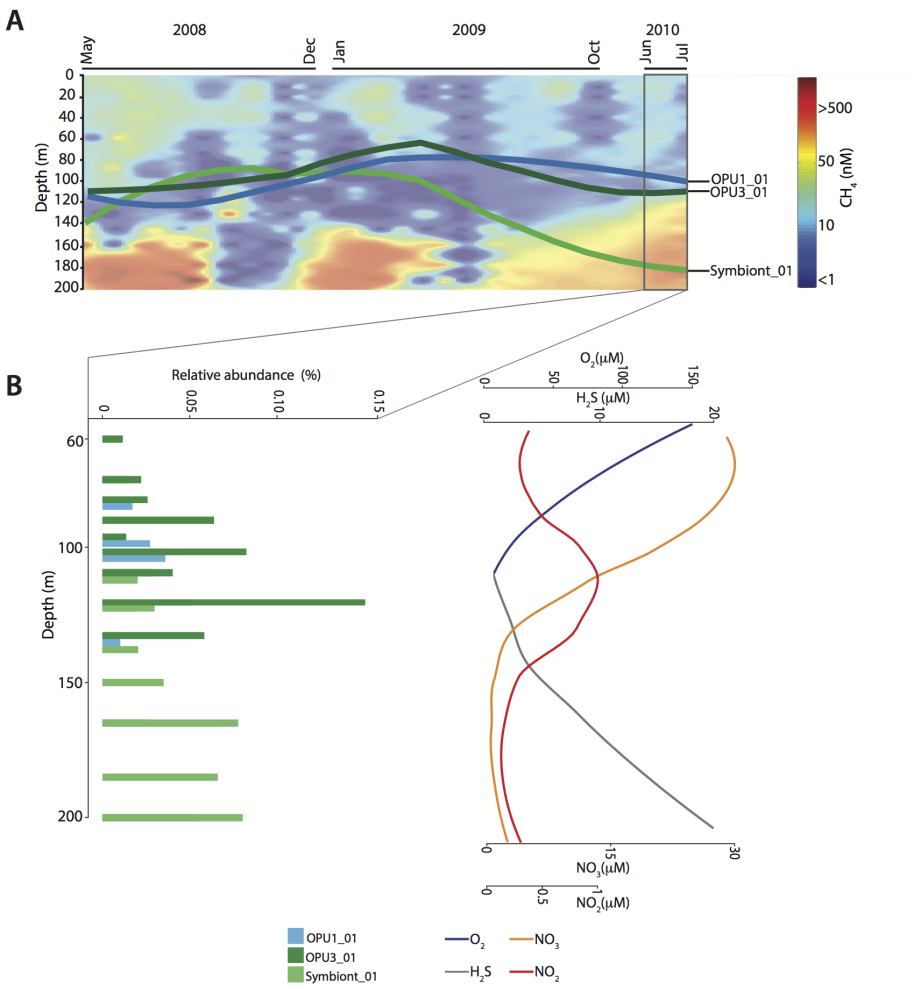

Figure 5.4: Time-series observations for methanotrophic OTUs affiliated with OPU1, OPU3, and methanotrophic symbiont groups. A) Methane contour plot for gas concentration (nM) data throughout water column from May 2008 to July 2010. Overlapped is shown the OTUs distribution mean trend throughout water column over time. B) Vertical distribution and relative abundance of methanotrophic OTUs over extended stratification period (June-July 2010). On the right, sparklines depict concentration trend for Oxygen (O_2), Sulfide (H_2S), Nitrate (NO_3^-) and Nitrite (NO_2^-).

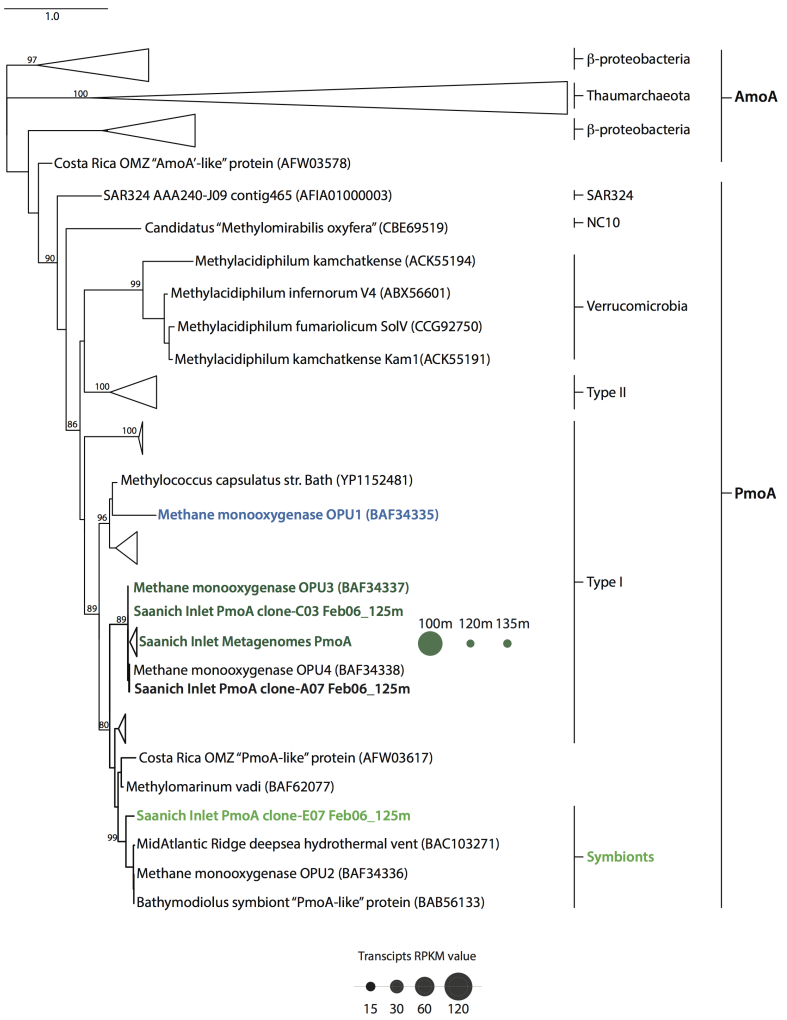

Figure 5.5: Particulate methane monooxygenase subunit β (*pmoA*) phylogenetic tree. Topography was inferred using maximum likelihood on PmoA and AmoA amino acid sequences from reference sequences including pMMO-encoding groups OPU1 and OPU3, and symbionts. Bootstrap values (%) are based on 100 replicates and are shown for branches with greater than 70% support. The scale bar represents 0.5 substitutions per site. PmoA distribution throughout water column (100- 135 m) is depicted as dots whose size represents the average RPKM value for February 2010 metatranscriptomic datasets.

Table 5.1: Multivariate regression statistics for methanotroph OTUs. Correlation values among OPU1_01, OPU3_01 and Symbiont_01 are shown with corresponding pair *p*-values.

Methanotroph OTU	Pair comparison	Correlation	*p*-value
OPU1_01	OPU3_01	0.1507	0.0012
OPU1_01	Symbiont_01	-0.0992	0.0114
OPU3_01	Symbiont_01	0.2025	0.00016

Table 5.2: Negative binomial regression statistics for methanotroph OTUs. Correlation values (estimate) for environmental parameters (variable) and standard error are shown with *z* value indicating evidence of true correlation. *p*-values for *z* (Pr(>z)) test the correlations significance.

Methanotroph OTU	Variable	Estimate	Std. Error	z value	Pr(>|z|)	Signif
OPU1_01	O_2	-0.00517	0.002331	-2.219	0.0265	<0.05
	CH_4	-0.00626	0.00189	-3.313	0.000922	<0.001
	NO_3	-0.00715	0.02036	-0.351	0.72558	
	NO_2	0.25372	0.32508	0.78	0.4351	
	H_2S	0.21858	0.12281	1.78	0.07511	
OPU3_01	O_2	0.00185	0.002623	0.707	0.4794	
	CH_4	-0.008607	0.00272	-3.165	0.00155	0.001
	NO_3	-0.04083	0.0215	-1.898	0.05766	
	NO_2	-0.60275	0.30437	-1.98	0.0476	0.05
	H_2S	0.6963	0.23799	2.926	0.0734	
Symbiont_01	O_2	0.00138	0.00225	0.616	0.5382	
	CH_4	-0.00609	0.00145	-4.209	$2.57\ e^{-5}$	< 0.001
	NO_3	-0.10017	0.02038	-4.911	$9.04\ e^{-7}$	<0.001
	NO_2	-1.3045	0.33521	-3.892	$9.95\ e^{-5}$	<0.001
	H_2S	0.30963	0.09272	3.339	0.0008	<0.001

tified OTUs affiliated with one carbon (C1) utilizing microorganisms including Methylophilales, *Methylophaga*, SAR324, Verrucomicrobia and Planctomycetes. SAR324 and Verrucomicrobia were indicators for oxic and anoxic water column conditions while Methylophilales and *Methylophaga* were indicators for suboxic and anoxic water column conditions. The indicator OTUs affiliated with Planctomycetes were evenly distributed throughout the water column (Table. 5.1 and Table C.3). OPU1_01 was identified as an indicator species for dysoxic-suboxic water column conditions, while no OTUs affiliated with OPU3_01, or Symbiont_01 were identified as indicator species for a particular water column condition. However, OPU3_01 and Symbiont_01, as well as OTU20751, affiliated with the OPU3 group, were identified as indicators for combined oxic and dysoxic-suboxic conditions (Table C.2).

Co-occurrence analysis based on Bray-Curtis and Spearman correlation values among OTUs, resulted in a microbial network (Fig. 5.6A) partitioned in modules similar to hierarchical clustered groups associated

with oxic (I), dysoxic-suboxic (II and III), and anoxic (IV) water column compartments (Appendix C). Interestingly, significant positive correlations (CV>0.6, p <0.001), shown as sub-networks (Table C.5), were observed among OPU1_01, OPU3_01 and Symbiont_01 with indicator OTUs affiliated with potential C1 utilizing microorganisms such as *Methylophaga*, Methylophilales, SAR324, Verrucomicrobia and Planctomycetes, and other ubiquitous OMZ microbes including representative taxa such as Marine Group A, Nitrospina, and SUP05 (Fig. 5.6B, Table S4). To provide further evidence for potential methanotroph interactions observed in sub-networks I compared microbial OTUs to SSU rRNA gene sequences recovered from previous CH_4 microcosm experiments using Saanich Inlet suboxic waters (Sauter *et al*., 2012) (Appendix C). Most bacterial OTUs (83% of OTUs in sub-networks) were found affiliated (>80% identity) to microcosm SSU rRNA sequences that were enriched and active after CH_4 addition, and related to Bacteroidetes, Marine Group A, Planctomycetes, Sphingomonadales, Nitrospina, Methylophilales, *Methylophaga*, SUP05 and Verrucomicrobia (Table C.5) reinforcing co-occurrence network observations.

5.4 Discussion

This chapter charts methanotroph diversity, abundance and dynamics in Saanich Inlet, a seasonally anoxic fjord that serves as a model ecosystem for understanding microbial community responses to changing levels of water column O_2 deficiency. Observations encompass an atypical extended water column stratification period in 2010 related to a relatively strong El Niño event (Blunden *et al*., 2011). This extended stratification period provided an opportunity to observe patterns of redox-driven niche partitioning among methanotrophic community members (Fig. 5.4B), hypothesize the use of alternative electron acceptors i.e NO_3^- and NO_2^- for CH_4 oxidation, and determine co-occurrence patterns between community members consistent with differential modes of metabolic coupling driving C1 metabolism along the redoxcline.

Methanotrophic community composition was primarily comprised of rare OTUs affiliated with the OPU1, OPU3, and mussel symbionts. Interestingly, no OTUs affiliated with the NC10 phylum were recovered contrasting recent observations by Padilla and colleagues in the Eastern Tropical North Pacific OMZ off the coasts of northern Mexico and the Costa Rica OMZ (Padilla *et al*., 2016). The lack of detection of NC10 could reflect differences in water column transport processes or geochemical conditions including O_2, NO_2^- or H_2S concentrations. For instance, Padilla and colleagues suggested that NC10 distribution and abundance depended on persistent anoxic conditions and high CH_4 concentrations (>1M) (Padilla *et al*., 2016). With the exception of NC10, methanotophic OTUs inhabiting Saanich Inlet were consistent with previous observations in coastal OMZs (Hayashi *et al*., 2007; Tavormina *et al*., 2013), open ocean OMZs (Stevens and Ulloa, 2008; Glaubitz *et al*., 2010) and other O_2 deficient marine environments (Elsaied *et al*., 2004; Dick and Tebo, 2010; Fuchsman *et al*., 2011; Kessler *et al*., 2011; Dick *et al*., 2013; Lke *et al*., 2016) (Fig. C.5) reinforcing the extensibility of Saanich Inlet time series as a model for understanding microbial community dynamics under water column O_2 deficiency.

The time-series observations allowed me to observe dynamic abundance and distribution patterns that changed as a function of water column redox conditions as supported by multivariate regression analysis. For instance, OPU1, OPU3, and methanotrophic symbiont OTUs tended to co-occur during periods of deep water renewal (Fig. C.2 and C.3). However, as the water column became increasingly stratified I observed

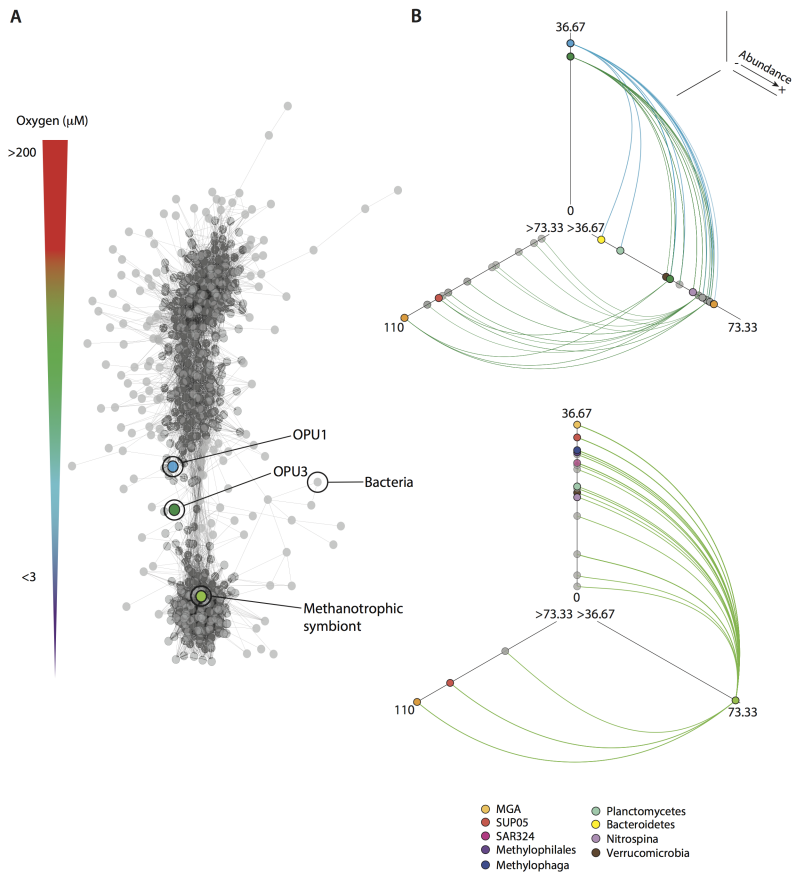

Figure 5.6: Co-occurrence patterns for methanotrophic and indicator OTUs from SSU rRNA gene pyrotag datasets.A) Network based on significant (p <0.001) Bray-Curtis and Spearman correlation values (>0.6) among OTUs that were present in at least 25% (n=72) of the total number of pyrotag samples (n=288). On top the oxygen gradient from oxic (>90 μM O_2) to anoxic (<1 μM O_2) is shown. B) Hive panels for methanotrophic OPU1 and OPU3, and symbiont OTUs showing interactions with indicator OTUs (taxa coloured as shown in key). OTUs are distributed based on abundance (log transformed) on the three axes from less abundant located closer to the center, to more abundant located towards the end of each line. Nodes are depicted according to taxonomy as indicated in network. All interactions (edges) are shown as solid lines. Direct interactions for methanotrophic OTUs to indicators are shown as solid coloured lines in hive panels.

separation of these OTUs into distinct water column compartments consistent with redox-driven niche partitioning. Our results expand on previous observations of OPU1 and OPU3 distributions in the Costa Rican OMZ water column where under suboxic conditions (O_2 >7 μM) OPU1 was more abundant than OPU3. In comparison, OTUs related to methanotrophic symbionts were more abundant under suboxic conditions where O_2 concentrations were under the detection limit (4 μM), and under anoxic sulphidic conditions, similar to observations made in other O_2 deficient waters including deep-sea hydrothermal vents (Elsaied *et al*., 2004; Dick and Tebo, 2010).

Regression analysis between OPU1, OPU3, methanotrophic symbiont OTUs, and geochemical data suggested the potential use of alternative electron acceptors including NO_3^- and NO_2^-. This statistical observation is consistent with previous enrichment and isolation studies focused on different methanotrophic groups. For example, Nitrite driven anaerobic methane oxidation has been previously reported for members of the NC10 candidate division. Although primarily identified in fresh water environments, a recent study off the coasts of northern Mexico and Costa Rica reported presence and activity of NC10 in pelagic OMZ waters (Padilla *et al*., 2016). Although no conclusive rate measurements were provided for methane oxidation by NC10, transcripts encoding nitric oxide (NO) reductase involved in NO dismutation to O_2 were detected in association with peak NO_2^- and CH_4 concentrations. The potential for OPU or methanotrophic symbionts to use NO_3^- or NO_2^- to drive CH_4 oxidation under suboxic or anoxic conditions (e.g O_2 limited methane oxidation by facultative denitrifying aerobic methanotrophs) presents important stoichiometric considerations. In order for the reaction to occur at a CH_4 concentration equal to 341 nM below 120 m, a minimum of 0.015 nmol NO_3^- and 0.042 nmol NO_2^- are required based on the CH_4: NO_3^- and CH_4: NO_2^- consumption ratio reported by Cuba and colleagues for batch reactors amended with CH_4, NO_3^- and NO_2^- (Cuba *et al*., 2011). Given the average NO_3^- and NO_2^- concentrations of 6.5 and 0.2 μM respectively in Saanich Inlet, this coupled mechanism for CH_4 oxidation and NO_3^-/NO_2^- reduction is permissive.

The use of nitrogen species NO_3^- and NO_2^- to drive CH_4 oxidation in type I methanotrophs has also been indicated under bioreactor, microcosms and isolated culture conditions (Cuba *et al*., 2011; Hernandez *et al*., 2015; Kits *et al*., 2015). Cuba and colleagues observed increased CH_4 loss in the batch reactors amended with NO_3^- (0.52 mol CH_4 g^{-1} NO_3^-) or NO_2^- (0.17 mol CH_4 g^{-1} NO_2^-) followed by community DGGE profile indicating an enrichment in Methylomonas sp. SSU rRNA sequences in amended bioreactors (Cuba *et al*., 2011). Consistent with this observation, Kits and colleagues recently reported that Methylomonas denitrificans strain FJG1T couples CH_4 oxidation to NO_3^- reduction under O_2 limiting anoxic conditions (<50 nM) resulting in N_2O production (Kits *et al*., 2015). In addition, Hernandez and colleagues reported Methylobacter sp. as a dominant methanotroph encoding respiratory nitrate and nitrite reductase genes under dysoxic or suboxic O_2 conditions (15-75 μM) in Lake Washington microcosm experiment, indicating potential use of NO_3^- and NO_2^- as alternative electron acceptors under low O_2 conditions (Hernandez *et al*., 2015). Given the absence of NC10 OTUs in our time series observations it will be of interest to determine if convergent mechanisms of NO dismutation or NO_3^- reduction are used by OPU or methanotrophic symbionts.

Using time-series measurements, Capelle and colleagues have recently identified a persistent CH_4 minimum at 110 m near the oxic-anoxic interface (Capelle *et al*., 2017) correlating with the highest CH_4

oxidation rates (2 nmol L^{-1} d^{-1}) observed in Saanich Inlet (Ward *et al*., 1989). These studies suggested methanotrophs were more abundant and/or metabolically active in the oxycline than in the upper water column. Based on cumulative methanotroph abundance (2.1-2.8% relative abundance) under suboxic conditions (100-150m) it is possible to identify OPU1, OPU3 and methanotrophic symbionts as the primary drivers of CH_4 oxidation in the Saanich Inlet water column (Fig. C.4). Interestingly, CH_4 oxidation rates measured in Saanich Inlet are very similar to those reported in the Costa Rica OMZ (2.6 nmol d^{-1}; four-fold nitrite driven anaerobic methane oxidation rate) (Padilla *et al*., 2016) where OPU1 and OPU3 were the most abundant methanotrophic groups identified. In support of this observation, I recovered *pmoA* ORFs corresponding to OPU3 during an extended stratification period in February 2010. Interestingly, only OPU3 PmoA related sequences were found for this period, probably due to the higher abundance of this group. This was reflected in the relative abundance of OPU3 SSU rRNA sequences (up to 0.15%) when compared with those from OPU1 and methanotrophic symbiont groups ($\sim$0.02% relative abundance) suggesting the potential capability for OPU3 to thrive under extended O_2 deficiency by using alternative electron acceptors. Activity of this group was inferred based on recovery of *pmoA* transcripts affiliated with OPU3 under dysoxic-suboxic water column conditions throughout 100-135 m depth intervals (Fig. 5.5). Similar observations in the Guaymas Basin indicated that *pmoA* transcripts were more abundant ($\sim$0.32% from total KEGG annotations in metatranscriptome dataset) in dysoxic deep-sea hydrothermal plume samples ($\sim$27 μM O_2) and related to sequences retrieved from the Santa Monica Basin and North Fiji hydrothermal vent field, and hydrocarbon plumes from the Deepwater Horizon oil spill in the Gulf of Mexico (Lesniewski *et al*., 2012). These sequences form the widely distributed OPU3 group of monooxygenases from uncultivated organisms, which are thought to have important roles in the oxidation of methane in O_2 deficient waters (Tavormina *et al*., 2010). Based on this information I consider the potential activity of OPU3 under dysoxic-suboxic water column conditions in Saanich Inlet. Based on the published CH_4 oxidation rate (2 nmol L^{-1} d^{-1}) and the total volume of the dysoxic-suboxic waters between 100-135m ($\sim$3.250x109 L) over 1 year in Saanich Inlet, OPU3 have the potential to consume $\sim$37.96 kg CH_4 y^{-1}. One kilogram of CH_4 has a radiative forcing equal to 0.48 W m^{-2} with an atmospheric lifetime of 12 years, and a global warming potential (GWP) of 23 (eight-fold CO_2 GWP for 100 years) (IPCC, 2013). Thus, methanotrophs affiliated to OPU3 in Saanich Inlet filter $\sim$18.2 W m^{-2} equivalent to 450 years of radiative forcing that could be released to the atmosphere each year. Although there is more to explore regarding the activity for OPU1 and methanotrophic symbionts under different water column O_2 conditions, these results highlight the potential role of OPU3 as an important sink for CH_4 along continental margins, and reinforce the extensibility of our time-series observations with implications for modeling CH_4 cycling in expanding OMZs.

In addition to methanotroph redox-driven niche partitioning, co-occurrence patterns between C1 utilizing microorganisms were observed in the Saanich Inlet water column consistent with overlapping habitat, shared niche space or preference indicating potential metabolic interactions. In particular, I observed methanotroph OTUs correlating with Bacteroidetes, Planctomycetes, and Methylophilales. These results were consistent with a previous microcosm study using Saanich Inlet waters amended with CH_4 revealing marked enrichment of SSU rRNA gene sequences affiliated with methanotrophs, Bacteroidetes, Planctomycetes, and Methylophilales (Sauter *et al*., 2012). Interestingly, cooperative metabolism between methan-

otrophic bacteria and potential C1 utilizing microorganisms affiliated with different Bacteroidetes and Methylophilales has been recently observed in incubation studies using sediment samples from Lake Washington (Beck *et al*., 2013; Hernandez *et al*., 2015). Results indicated a simultaneous response between Bacteroidetes, Methylophilales and canonical methanotrophs to CH_4 addition over a range of O_2 concentrations (15-75 μM) (Beck *et al*., 2013; Hernandez *et al*., 2015). Although phylogenetic-based observations alone cannot explain underlying mechanisms of metabolite exchange, our co-occurrence observations indicate that CH_4 oxidation in Saanich Inlet likely depend on community-level interactions that support the metabolic requirements of methanotrophic agents.

5.5 Conclusion

I used molecular time series observations in combination with geochemical information to determine the methanotrophic community composition and dynamics in Saanich Inlet revealing three rare OTUs affiliated with OPU1, OPU3, and methanotrophic symbiont groups that exhibited redox-driven niche partitioning along changing water column redox gradients. Moreover, I resolved potential novel metabolic strategies including the use of alternative terminal electron acceptors, and metabolic interactions between C1 utilizing microorganisms supporting CH_4 oxidation. Combined, these observations provide important baseline information on microbial agents that reduce the flux of climate active trace gases from ocean to atmosphere and support the potential role of OPU1, OPU3, and methanotrophic symbiont groups as a widely distributed pelagic sink for CH_4 along continental margins. Looking forward, I recommend expanded use of multi-omic sequencing in combination with process rate measurements to determine coverage of CH_4 oxidation pathways including all reactions implicated in biological CH_4 transformation and C1 transfer. In addition, isotopic labeling and incubation coupled with gene expression studies should be conducted to link CH_4 oxidation pathways and process rates to specific microbial agents along defined water column redox gradients on regional and global scales to better constrain the CH_4 filtering capacity of coastal and open ocean OMZs.

Chapter 6

Community-level interactions support CH_4 oxidation in Saanich Inlet O_2-deficient water column

This chapter is one of the first surveys that integrates multi-omic sequencing information to resolve community-level interactions for CH_4 oxidation under O_2 deficiency. Observations here support the use of NO_2^- for CH_4 oxidation by methanotrophs affiliated to OPU3, and provide evidence for coupled metabolic interactions between methanotrophs and C1 utilizing microorganisms proposed by co-occurrence patterns in Chapter 5. For instance, CH_4 oxidation coupled with NO_2^- reduction is ignited by OPU3 under low-O_2 water-column conditions, and following a fermentative metabolism OPU3 likely promotes community pathways for carbon incorporation by releasing organic compounds that are used as carbon source by co-occurring taxa. This chapter expands on available information related to community-level metabolic interactions for gas cycling that is relevant to models predicting microbial community responses to ocean deoxygenation and is extensible to biotechnological and engineering approaches using methanotrophic communities.

6.1 Introduction

Global dissolved oxygen (O_2) concentration observations since 1960 (Whitney *et al.*, 2007; Bograd *et al.*, 2008; Stramma *et al.*, 2008; Keeling *et al.*, 2010; Stramma *et al.*, 2010; Helm *et al.*, 2011; Schmidtko *et al.*, 2017) show an ongoing regional decline in oceanic O_2 concentrations, or deoxygenation, and a subsequent expansion of the mid-depth oxygen-minimum zones (OMZs) (Bopp *et al.*, 2002; Keeling *et al.*, 2010; Keller *et al.*, 2016; Schmidtko *et al.*, 2017). Ocean deoxygenation has potentially broad impacts on nutrient and greenhouse gases, i.e. CH_4 cycling in the ocean (Worm *et al.*, 2005; Diaz and Rosenberg, 2008; Vaquer-Sunyer and Duarte, 2008; Stramma *et al.*, 2011). Marine OMZs encompass large reservoirs of CH_4 (Zhang *et al.*, 2011; Pack *et al.*, 2015). For instance, The Eastern Tropical North Pacific (ETNP) OMZ is both the largest OMZ (Paulmier and Ruiz-Pino, 2009) and the largest reservoir of oceanic CH_4 (Sansone *et al.*, 2001; Reeburgh, 2007; Naqvi *et al.*, 2010) in the world. In OMZs, CH_4 accumulates in a functionally anoxic core surrounded by a layer of hypoxic waters (Thamdrup *et al.*, 2012;Wright *et al.*, 2012). Under

a warming climate, the dissolution of O_2 in seawater will decrease, whereas its consumption through respiration will likely increase (Vzquez-Domnguez *et al*., 2007) and thermal stratification could become more intense. Together, these biotic and abiotic changes will thicken OMZs, moving CH_4 pools closer to the zone of atmospheric exchange (Stramma *et al*., 2008; Keeling *et al*., 2010; Helm *et al*., 2011).

Pelagic CH_4 oxidation in marine environments is a rarely quantified process, but on the margins of an OMZ where CH_4 intersects traces of O_2, it could be a significant process (Mau *et al*., 2013). Pelagic aerobic or anaerobic CH_4 oxidation processes form a final barrier preventing iCH_4 escape to the atmosphere, i.e. aerobic CH_4 oxidation has been estimated to consume $>$50% of CH_4 in the water column (Fung *et al*., 1991; Reeburgh *et al*., 1991; Blumenberg *et al*., 2007; Kessler *et al*., 2011; Heintz *et al*., 2012). Previous taxonomic and functional screening studies provide insight into methanotrophic community structure and activity and suggest that at least some of the classic aerobic methanotroph species may be able to thrive in O_2 deficient environments by potentially utilizing alternative electron acceptors for their metabolism (Costa *et al*., 2000; Modin *et al*., 2007; Stein and Klotz, 2011; Beck *et al*., 2013; Hernandez *et al*., 2015). For instance, along continental margins two pMMO- encoding phylogenetic groups termed OPU1 and OPU3 (Hayashi *et al*., 2007) are commonly recovered in molecular gene surveys targeting aerobic methanotrophs in the OMZs (Hayashi *et al*., 2007; Tavormina *et al*., 2010; Tavormina *et al*., 2013; Knief, 2015; Torres-Beltrán *et al*., 2016; Padilla *et al*., 2017). In addition, expression of pmoCAB for the OPU3 group was first demonstrated in a metatranscriptome from the Guaymas Basin (Lesniewski *et al*., 2012) and recently, alternative modes of CH_4 oxidation by OPU3 have been observed in the Costa Rica OMZ (Padilla *et al*., 2017). Genes mediating dissimilatory nitrate (NO_3^-) and nitrite (NO_2^-) reduction were identified in the OPU3 binned genome and shown being transcribed in conjunction with key enzymes catalyzing formaldehyde assimilation, suggesting partial denitrification linked to CH_4 oxidation (Padilla *et al*., 2017). These observations provide important baseline information about methanotrophic agents that reduce the flux of CH_4 from ocean to atmosphere. However, additional functional information is necessary to determine microbial community members and reactions implicated in community-level CH_4 metabolism and derived carbon assimilation processes, under extended water-column O_2 deficiency.

Community-level CH_4 metabolism is an emerging concept from the anaerobic CH_4 oxidation processes that depend on syntrophic associations due to either energetic constraints or the necessity for toxic intermediate removal (Haroon *et al*., 2013; Skennerton *et al*., 2017). Similarly, aerobic methanotrophs tend to form communities, and it is apparent that they share carbon from CH_4 with other bacteria (Yu and Chistoserdova, 2017). It has been suggested that compounds such as methanol, formate, acetate, succinate, and possibly other organic acids released by methanotrophs, can support a broad range of microbes (Kalyuzhnaya *et al*., 2013); Modin, 2017; Tavormina *et al*., 2017). For instance, some of the species most commonly co-occurring with Gammaproteobacterial methanotrophs are non-methane-utilizing methylotrophs, i.e. Methylophilaceae family and other non-methylotrophic bacteria affiliated with Burkholderiales and Flavobacteriales (Beck *et al*., 2013; Hernandez *et al*., 2015; Oshkin *et al*., 2015; Karwautz *et al*., 2018; Kumaresan *et al*., 2018). Analyses of both natural populations (Crevecoeur *et al*., 2015; Karwautz *et al*., 2018; Kumaresan *et al*., 2018), and manipulated laboratory microcosms (Beck *et al*., 2013; Hernandez *et al*., 2015; Oshkin *et al*., 2015) suggest that these partnerships in CH_4 metabolism might not be random.

Although the communal nature of CH_4 metabolism provides a new outlook on the environmental role of the methanotrophs as essential components of food webs driven by carbon from CH_4, it is necessary to determine how resilient these interactions are to low-O_2 water-column conditions.

Saanich Inlet is a seasonally anoxic fjord on the east coast of Vancouver Island British Columbia. The seasonally stratified water column of Saanich Inlet serves as a well-characterized model ecosystem for examining how deoxygenation shapes microbial community population dynamics and interactions along defined redox gradients in the ocean (Walsh *et al*., 2009; Zaikova *et al*., 2010; Walsh and Hallam, 2011; Wright *et al*., 2012; Hawley *et al*., 2014; Louca *et al*., 2016; Torres-Beltrán *et al*., 2016). Recent observations charting methanotroph diversity, abundance and dynamics in Saanich Inlet, indicated CH_4 oxidation under low-O_2 water-column conditions likely depend on community-level interactions between OPU3 and taxa affiliated with Bacteroidetes, Planctomycetes, and Methylophilales (Torres-Beltrán *et al*., 2016). In this chapter, I aim to resolve community-level interactions controlling CH_4 oxidation occurring in the Saanich Inlet O_2-deficient water-column. I coupled incubation experiments and long-term monitoring surveys that integrate multi-omic information providing a promising environmental context to link CH_4 oxidation pathways to specific microbial agents along defined low-O_2 water-column conditions. These observations indicate methanotrophs play an important role in global metabolic cycles beyond the CH_4 cycle and provide evidence that expands our understanding of CH_4 oxidation under low-O_2 water-column conditions.

6.2 Methods

6.2.1 Incubations implementation

To explore microbial community-level interactions controlling CH_4 oxidation in the Saanich Inlet water column, I carried out an incubation experiment to test for CH_4 oxidation coupled with NO_2^- reduction under dysoxic-suboxic (<30- $\sim$3 μM O_2) water-column conditions. The rational behind this approach was based on recent observations suggesting potential novel metabolic strategies including the use of alternative terminal electron acceptors i.e. NO_2^- by methanotrophs affiliated to OPU3 under low-O_2 water-column conditions (Torres-Beltrán *et al*., 2016; Padilla *et al*., 2017). I applied a multi-omic sequencing approach (small subunit (SSU) rDNA and rRNA pyrotags, metagenomic, metatranscriptomic and metaproteomic) to infer metabolic interactions related to CH_4 oxidation under low-O_2 conditions (Fig. 6.1).

Incubations were carried out on 12 L seawater samples collected on February 15, 2015, from two depths (100 and 150 m) selected based on O_2 concentration spanning the upper and lower boundaries of the dysoxic-suboxic (<30- $\sim$3 μM O_2) water-column conditions. For each depth, sample water was collected from two 12 L Go-Flow bottles through silicon tubing ($\sim$15 cm long and 1/4" thick, pre-flushed for a few seconds with sample water) into 12 L metallic bags (pre-flushed with Helium (He)). Bags were overfilled to remove any air bubbles from the tubing during filling and sealed right after. Five bags were filled at each depth corresponding to an environmental control (background), an experimental control (no substrate addition), $^{12}CH_4$, $^{12}CH_4$ + NO_2^- and $^{13}CH_4$ + NO_2^-. Background samples for RNA (2 L) and protein (2 L) were processed on-ship immediately after sampling while the rest were stored in the dark on ice until processing in the laboratory.

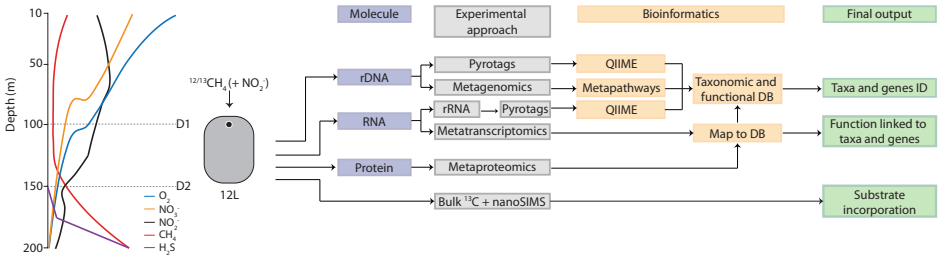

Figure 6.1: Experimental workflow for CH_4 incubation experiments carried out in February 2015. Water samples (12 L) for DNA, RNA and protein were taken at 100 (D1) and 150 m (D2) encompassing the dysoxic-suboxic water-column compartments. Growth conditions included the addition of CH_4 and $^{12,13}CH_4$, plus NO_2^-, and incubation for 72 h. A multi-omic approach (gray) including 16S pyrotag sequencing, metagenomics, metatranscriptomics, metaproteomics coupled with bulk ^{13}C measurements and nanoSIMS was carried out to obtain the final outputs (green), such as taxa, gene identification and functional information derived from substrate incorporation into cell biomass through the bioinformatics tools (orange). The environmental parameters profile is depicted as sparklines for oxygen (O_2= light blue), nitrate (NO_3^-= orange), nitrite (NO_2^-= black), methane (CH_4= red) and hydrogen sulfide (H_2S= purple).

Substrates were injected into bags through the airlock using a He-flushed glass syringe, and mixed using a plate shaker to allow substrate dissolution in the sample. Substrates were added in a 10% increase of the highest environmental concentration at each depth corresponding to 72 nM and 2 μM CH_4, and 0.03 and 0.1 μM NO_2^- for the 100 and 150 m, respectively. Water subsamples were taken before (t_0; DNA (2L)) and after 72 hour incubation (t_{72}; DNA (4L), RNA (2L) and Protein (2L)). Subsamples were filtered onto a 0.22 μm Sterivex filter as previously described for the Saanich Inlet time-series multi-omic datasets (Hawley *et al.*, 2017b).

Biomass to generate time-series metagenomic datasets was collected from June 2009 to August 2011 at six depths (10, 100, 120, 135, 150 and 200 m) and filtered with an in-line 2.7 μm GDF glass fibre pre-filter onto a 0.22 μm Sterivex polycarbonate cartridge filter (Hawley *et al.*, 2017b).

6.2.2 Carbon incorporation measurements

To first assess carbon incorporation into microbial community biomass I used $^{13}CH_4$ samples and conducted two complementary isotopic approaches. I measured the carbon isotopic composition ($\delta^{13}C$) of cell pellets (Bulk) and individual cells.

Bulk $\delta^{13}C$ on cell material pellets was measured using an elemental analyzer (EA) from Costech Analytical Technologies, Inc. (Valencia, CA) coupled to a Thermo Scientific (Bremen, Germany) Delta V Plus isotope ratio mass spectrometer (IRMS) at the Pacific Northwest National Laboratory (PNNL). Due to the small size of the cell pellets, they were suspended in an aqueous buffer to facilitate their transfer into pre-weighed tin capsules (Costech Analytical Technologies, Inc.) used for EA-IRMS analysis. I dried the samples then re-weighed the capsules to determine the total amount of biomass used for each analysis

(2^{-1}5 mg). The EA combustion reactor was loaded with cobaltic oxide and chromium oxide catalyst and maintained at 1,020 ^{o}C while the combustion reactor was loaded with copper catalyst and maintained at 650 ^{o}C. I used two in-house glutamic acid standards that were themselves calibrated against USGS 40 and USGS 41 standards (δ^{13}C = -26.39 $^{o}/_{oo}$VPBD and 37.63 $^{o}/_{oo}$VPDB respectively) and applied a two-point, slope intercept correction to the data. All δ^{13}C results are referenced to Vienna Pee Dee Belemnite and are reported in delta notation:

$$\delta^{13}C = (\frac{R_{sample}}{R_{standard}} - 1) * 1000 \tag{6.1}$$

where R_{sample} is the ^{13}C to ^{12}C ratio of the measured sample and Rstandard is the ^{13}C to ^{12}C ratio of Vienne Pee Dee Belemnite (0.0112372). Total δ^{13}C parts per mil ($^{o}/_{oo}$) were obtained for each sample. Carbon incorporation (δ^{13}C) into individual cells was measured using a 50L Cameca. Cells were fixed onto 0.2 μm precombusted polycarbonate filters (5 mL sample) and DAPI (1 μg ml^{-1} final concentration, Sigma) stained for field identification under the instrument microscope. Controls and labeled samples were measured for ^{13}C incorporation relative to the standard. δ^{13}C values of enriched cells (15-25 per field) were exported to the R software package (RCoreTeam, 2013) for results visualization.

6.2.3 Nucleic acid and protein extractions

Genomic DNA was extracted from Sterivex filters as previously described (Zaikova *et al*., 2010; Hawley *et al*., 2017b). Briefly, after thawing the filter cartridge on ice, 50 μl of 0.125 mg ml^{-1} lysozyme was added and incubated at 37 ^{o}C for 1 h with rotation followed by addition of 50 μl Proteinase K (Sigma) and 100 μl 20% SDS and incubated at 55 ^{o}C for 1 h with rotation. Lysate was removed by pushing through with a 3 ml syringe followed by rinsing with an additional 1 mL of sucrose lysis buffer. Filtrate was subject to chloroform extraction and the aqueous layer was collected onto a 10K 15 ml Amicon filter cartridge, washed three times with TE buffer (pH 8.0) and concentrated to a final volume between 150-400 μl. Total DNA concentration was determined by PicoGreen assay (Life Technologies) and genomic DNA quality determined by visualization on 0.8% agarose gel run (16h/15V).

Total RNA was extracted from Sterivex filters as previously described for the Saanich Inlet time-series (Hawley *et al*., 2017b). Briefly, after thawing the Sterivex cartridge on ice RNA later was removed by pushing through with a 3 ml syringe followed by rinsing with an additional 1.8 mL of Ringer's solution and incubated at room temperature for 20min with rotation. Ringer's solution was removed by pushing through with a 3 ml syringe followed by adding 1.8 ml MirVana Lysis Buffer and 100 μl of 0.125 mg ml^{-1} lysozyme and incubated at 37 ^{o}C for 30 min with rotation. Lysate was removed from Sterivex and subjected to organic extraction following the mirVana kit protocol. DNA removal and clean up and purification of total RNA were conducted following the TURBO DNA-free kit (Thermo Fisher Scientific) and the RNeasy MinElute Cleanup kit (Qiagen) protocols respectively. Total RNA concentration was determined by RiboGreen analysis (Life Technologies) prior to synthesize first strand cDNA using the SuperScript III First-Strand Synthesis System for RT-qPCR (Invitrogen) according to manufacturer instructions.

Total protein was extracted in the Environmental Molecular Sciences Laboratory (EMSL) at PNNL from

Sterivex as previously described for the Saanich Inlet time-series proteomic datatsets (Hawlet *et al*., 2014; Hawley *et al*., 2017b). Briefly, after thawing Sterivex filters on ice, Bugbuster (Novagen) was added and incubated at room temperature for 20-30 min with rotation. Lysate was removed by extrusion and filters were rinsed with 1 ml lysis buffer. Buffer exchange was carried out on combined lysate using Amicon Ultra 10K (Millipore) with 100 mM NH_4HCO_3 a total of three times with a final volume between 200-500 μl. Protein concentration was determined using the 2-(4-carboxyquinolin-2-yl) quinoline-4-carboxylic acid (Bicinchoninic acid or BCA) assay. Urea was added to a final concentration of 8 M and dithiothreitol added to a final concentration of 5 mM and incubated at 60 ^{o}C for 30 min, followed by 10-fold dilution with 100 mM NH_4HCO_3. Samples were then subject to trypsin digest at 37 ^{o}C for 6 h followed by C18 solid phase extraction and strong cation exchange.

6.2.4 Nucleic acid tag sequencing

Extracted DNA and cDNA was used to generate small subunit (SSU) rDNA and rRNA pyrotags with three-domain resolution. PCR amplification procedures were carried out as previously described for the Saanich Inlet time-series multi-omic data (Hawley *et al*., 2017b)., Pyrotag libraries were generated by PCR amplification using multi-domain primers targeting the V6-V8 region of the SSU rRNA gene (Allers *et al*., 2013): 926F (5'-cct atc ccc tgt gtg cct tgg cag tct cag AAA CTY AAA KGA ATT GRC GG-3') and 1392R (5'-cca tct cat ccc tgc gtg tct ccg act cag-<XXXXX>-ACG GGC GGT GTG TRC-3'). Primer sequences were modified by the addition of 454 A or B adapter sequences (lower case). In addition, the reverse primer included a 5 bp barcode designated <XXXXX> for multiplexing of samples during sequencing.

Twenty-five microliter PCR reactions were performed in triplicate and pooled to minimize PCR bias. Each reaction contained between 1 and 10 ng of target DNA, 0.5 μl Taq DNA polymerase (Bioshop inc. Canada), 2.5 μL Bioshop 10 x buffer, 1.5 μL 25 mM Bioshop MgCl2, 2.5 μL 10 mM dNTPs (Agilent Technologies) and 0.5 μL 10 mM of each primer. The thermal cycler protocol started with an initial denaturation at 95 ^{o}C for 3 minutes and then 25 cycles of 30 s at 95 ^{o}C, 45 s at 55 ^{o}C, 90 s at 72 ^{o}C and 45 s at 55 ^{o}C. Final extension at 72 ^{o}C for 10 min. PCR products were purified using the QiaQuick PCR purification kit (Qiagen), eluted elution buffer (25 μL), and quantified using PicoGreen assay (Life Technologies). SSU rDNA and rRNA amplicons were pooled at 100 ng for each sample. Emulsion PCR and sequencing of the PCR amplicons were sequenced on Roche 454 GS FLX Titanium at the McGill University and Gnome Qubec Innovation Center.

6.2.5 Meta-genomic, transcriptomic and proteomic sequencing

Extracted total DNA was used to generate Illumina paired-end metagenomic datasets at GeneWiz sequenced on the Illumina HiSeq platform. Extracted total DNA from time-series samples, were used to generate metagenomic datasets at the DOE-JGI following the protocols for library production, sequencing and assembly previously described for the Saanich Inlet time-series (Hawley *et al*., 2017b). Extracted total RNA was used to generate paired end metatranscriptomic datasets at EMSL-PNNL sequenced on the Ion PI Hi-Q platform.

Extracted total protein samples were used to generate metaproteomic datasets at EMSL-PNNL. Peptides

were analyzed by tandem MS (MS/MS), as previously described (Hawley *et al*., 2013; Hawley *et al*., 2014), using online capillary LC/MS/MS on a Thermo LTQ ion trap using data-dependent fragmentation. Detected peptides were identified from MS/MS using SEQUEST with a mass spectra generating function (MS-GF) cutoff value less than 10^{-11}, corresponding to a false discovery rate of less than 2% (Kim *et al*., 2008).

6.2.6 Tag sequence analyses

In order to relate community composition with potential activity, combined SSU rDNA and rRNA pyrotag sequences were analyzed using the Quantitative Insights Into Microbial Ecology (QIIME) software package (Caporaso *et al*., 2010). Reads with length shorter than 200 bases, ambiguous bases, and homopolymer runs were removed before to chimera detection. Chimeras were detected using the chimera slayer provided in the QIIME software package and removed before taxonomic analysis. A total of 564,694 non-chimeric sequences were clustered at 97% identity into operational taxonomic units (OTUs). Prior to taxonomic assignment, singleton OTUs (OTUs represented by one read) were omitted (Engelbrektson *et al*., 2010) leaving 19,508 OTU sequences. Representative sequences from each non-singleton OTU were queried against the SILVA database release 111 using the BLAST algorithm (Altschul *et al*., 1990). To further resolve the diversity of methanotrophic bacterial OTUs, I recruited pyrotag OTU sequences using a BLAST-comparison (>99% identity) to full-length SSU rRNA gene reference sequences including cultured and environmental Type I and II methanotrophs and pyrotag SSU rRNA gene sequences previously observed in the Saanich Inlet water column (Torres-Beltrán *et al*., 2016).

The SSU rRNA:rDNA abundance ratio was calculated to account for variation in taxon abundance in the DNA pool (Frias-Lopez *et al*., 2008; Stewart *et al*., 2012b) and compared for a subset of microbial groups to understand how incubation treatments influence changes of potentially active OTUs across treatments. Statistical analyses were conducted using the R software package (RCoreTeam, 2013). Pyrotag datasets were normalized to the total number of reads per sample. Hierarchical cluster analysis (HCA) was conducted to identify community compositional profiles associated with incubation treatments using the pvclust (Suzuki and Shimodaira, 2015) package with the Manhattan Distance measure, and statistical significance to the resulting clusters as bootstrap score distributions with 1,000 iterations. Further data curation and visualization were conducted using the dplyr (Wickham *et al*., 2015) and ggplot2 (Wickham, 2009) packages.

To further frame my observations with the previously described Saanich Inlet methanotrophic co-occurrence networks (Torres-Beltrán *et al*., 2016), I recruited incubations OTUs to time-series HR SSU rDNA pyrotag sequences (Torres-Beltrán *et al*., 2016) using BLAST-comparison (>99% identity). The time-series reference sequences included methanotrophic OTUs affiliated with the pMMO- encoding phylogenetic groups OPU1 and OPU3, and methanotrophic symbionts that showed significant correlations with alternative terminal electron acceptors, i.e. NO_2^- and taxa involved in one-carbon (C1) metabolism under suboxic water-column conditions. Only hits with a perfect match of a query sequence were selected and mapped to the reference co-occurrence network to identify potential microbial hubs carrying out methane oxidation ignited by substrate addition. Nodes were selected and colored based on taxonomy in an already existing co-occurrence network (Torres-Beltrán *et al*., 2016) using Cytoscape 2.8.3 (Shannon *et al*., 2003). Nodes size was depicted based on SSU rRNA:rDNA ratio values, showing only those OTUs with a ratio higher

than 1 that suggested potential activity under specific substrate conditions.

6.2.7 Multi-omic dataset analysis

Incubation metagenomic data were analyzed using MetaPathways V2.5.1, an open source pipeline for predicting reactions and pathways using default settings (Konwar *et al*., 2013; https://github.com/hallamlab/metapathways2/wiki), as previously described for Saanich Inlet time-series samples (Torres-Beltrán *et al*., 2016; Hawley *et al*., 2017a). Time-series metagenomic data was analysed using MetaPathways V2.5.1. For each gene, reads per kilobase per million mapped (RPKM) was calculated as a proportion of the number of reads mapped to a sequence section, normalized for sequencing depth and open reading frame (ORF) length (Konwar *et al*., 2015). RPKM values were used to describe the abundance of genes. To cross-reference incubation metagenomic datasets with the Saanich Inlet time-series metagenomic observations, conceptually translated amino acid sequences of predicted ORFs were LAST+ compared to time-series ORFs sequences. Sequence matches with higher than 70% identity were retrieved from the functional annotation table *<ORF_annotation_table.txt>* in the *<results/annotation table>* output directory and used to identify occurrence and abundance of marker CH_4 oxidation and NO_2^- reduction genes, i.e. particulate methane monooxygenase subunit β (*pmoA*) and copper-containing nitrite reductase (*nirK*) in incubation samples. Metagenomic information from incubation samples was included into the time-series reference database to be used for mapping and comparing levels of expression in metatranscriptomic and metaproteomic datasets described below.

To incorporate PmoA sequences found in incubation datasets into a reference phylogenetic tree, conceptually translated and annotated PmoA ORFs from the incubations metagenomic datasets were manually extracted from the functional annotation table *<ORF_annotation_table.txt>* in the *<results/annotation table>* output directory. Sequences were aligned and compared to diverse environmental and reference PmoA sequences (including OPU3) previously used to generate a phylogenetic reference tree for PmoA in the Saanich Inlet (Torres-Beltrán *et al*., 2016). A total of 104 PmoA sequences were clustered and representative sequences aligned as previously described to infer a maximum likelihood phylogenetic tree that included PmoA sequences observed in incubation samples. Similarly, to generate a phylogenetic reference tree for NirK, conceptually translated and annotated ORFs from the incubations metagenomic datasets were manually extracted from the functional annotation table *<ORF_annotation_table.txt>* in the *<results/annotation table>* Metapathways output directory. Sequences were aligned and compared to diverse environmental reference NirK sequences. A total of 1,818 NirK sequences were clustered over a range of identity thresholds using the UClust algorithm (USEARCH V6.0) with 28 reference sequences, including cultured and environmental sequences affiliated with Bacteroidetes, Choloflexi, Chlorobi, Nitrospinaceae, Methylococcales (including OPU3), Planctomyces, and Thaumarchaeota. The 70% identity threshold was selected based on the resolution of the taxonomic groups. Cluster representative sequences were aligned using the Multiple Sequence Comparison by the Log- Expectation (MUSCLE) method (EMBL-EBI) and were manually curated in Mesquite. A maximum likelihood phylogenetic tree was inferred using PHYML (Guindon *et al*., 2005) based on a WAG model of amino acid evolution where the parameter of the gamma distribution, the proportion of invariable sites and the transition/transversion ratio were estimated. The confidence of each

node was determined by assembling a consensus tree of 1000 bootstrap replicates.

Metatranscriptomic data was checked for read quality and sequence expansion using the htseq-qa command on the HTSeq python package (Anders *et al*., 2015). A total of 10,268,073 reads were aligned to reference metagenomic database using Torrent Mapping Alignment Program (TMAP) ((https://github.com/iontorrent/TS/tree/master/Analysis/TMAP)) using the map4 command with soft clipping enabled at both ends of the reads and the alignment output mode set to return a random alignment among those with the best score if multiple alignments receive the best score. Aligned reads were mapped to reference metagenomic database using the htseq-count command in the HTSeq python package (Anders *et al*., 2015). Annotation was converted to a suitable *.gtf* format using R costumed scripts. Expression analysis was performed using the Bioconductor package DESeq2 (Anders and Huber, 2010). Downstream enrichment analysis was performed using the modified Fisher exact test (Hosack *et al*., 2003) using R costumed scripts.

Proteomic data consisted of a total of 159,413 peptides. Peptides were searched against the metagenomic reference database. Only peptides matched to protein sequences with a peptide prophet probability (PPP) score $>$0.95 were used in further analysis. A total of 26,828 unique proteins were identified. To better characterize the metabolic processes within the microbial community for CH_4 oxidation under incubations conditions, I surveyed the expression level of a given gene transcript and later the amount of protein produced in pseudo-quantitative manner. First, I calculated a normalised spectral abundance factor (NSAF)(Hawley *et al*., 2013; Hawlet *et al*., 2014). Second, I calculated the expression ratio by dividing the NSAF value of a given protein in a treatment (t_{72}) over the control (t_{72}). Ratio values higher than 1 were considered as upregulated gene expression due to incubation conditions.

To further resolve the diversity of PmoA and NirK sequences observed in the metatranscriptomic and proteomic datasets, I retrieved transcripts and peptides that were annotated as PmoA and NirK based on ORF mapping to the metagenomic database and searched for their corresponding ORF in the gene clustering analysis conducted to build the phylogenetic trees. All PmoA and NirK transcripts and peptides found were assigned to an ORF within a cluster represented in their corresponding phylogenetic tree.

6.2.8 Pathways network and metabolic model inference

To generate a robust network emphasizing co-occurrences between prevalent pathways in the dysoxic-suboxic water-column compartment, Bray-Curtis and Spearman's rank correlations were used on time-series metagenomic data. Correlation coefficients were calculated using CoNet (Faust *et al*., 2012) with pathway abundance as RPKM data. First, a pathways matrix was constructed using the environmental Pathways/Genome Database (ePGDB) file *<SAMPLE.pwy.txt>* in the *<results/pgdb>* Metapathways output directory for each time-series metagenomic sample. The matrix was transformed into presence-absence data to remove pathways with less than 1/3 zero counts, leaving a matrix of 380 pathways for all samples. Next, to construct ensemble networks, measure-specific thresholds set to 0.6 were used as a pre-filter, and edge scores were computed only between clade pairs. To assign statistical significance to the resulting scores, edge and measure-specific permutation and bootstrap score distributions with 1,000 iterations each were computed. P-values were tail-adjusted so that low *p*-values correspond to co-presence and high p-values to exclusion. After merging, p-values on each final edge were corrected to q-values (cut-off of 0.05).

The positivity or negativity of each relationship was determined by consensus voting over all the integrated data sources. Finally, only edges with at least two supporting pieces of evidence were retained.

The final edge matrix was visualized as a force directed network using Cytoscape 2.8.3 (Shannon *et al.*, 2003). Network properties were calculated with the "Network Analysis" Plug-In. Nodes in the co-occurrence network corresponded to individual pathways and edges were defined by computed correlations between corresponding pathway pairs. The layout revealed distinct modules, which persisted after lowering the correlation coefficient cut-off for edge creation to 0.90 reinforcing the robustness of the network. Nodes and edges from modules were selected and visualized as sub-networks using the tool Hive Panel Explorer (https://github.com/hallamlab/HivePanelExplorer/wiki) (Perez, 2015). HivePlotter allows for edge selection based on interactions within specific pathways, such as those affiliated with CH_4 oxidation and C1 metabolism. ORFs within each selected pathway in a sub-network were searched for their functional and taxonomic affiliation with incubation multi-omic datasets. Pathways with ORFs showing functional representation in incubation experiments were colored in hive plots according to ORFs taxonomic affiliation.

6.2.9 Data deposition

The SSU rDNA and rRNA pyrotag sequences reported in this chapter have been deposited in the NCBI under the accession numbers: SRX2160172-SRX2160414. Metagenomes reported in this chapter have been deposited in the NCBI under the Biosample accession numbers: SAMN08283972- SAMN08283975. Metatranscriptomes reported in this chapter have been deposited in the NCBI under the Biosample accession numbers: SAMN10425508-SAMN10425517. Metaproteomes reported in this chapter have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD011287 and 10.6019/PXD011287.

6.3 Results

6.3.1 Water column conditions

Beginning in February 2015- water column O_2 deficiency intensified below 100 m and was consistent with the onset of the stratification period in the Saanich Inlet. The beginning of stratification corresponded with the intensification of suboxic conditions below 90 m and increasing levels of hydrogen sulphide (H_2S) below 150 m, which was consistent with the development of deep-water anoxia. I sampled two representative depths corresponding to the upper (100 m; 29 O_2 μM) and lower (150 m; 3.34 O_2 μM) boundaries of suboxic water-column conditions. Methane concentrations ranged from 72 nM at 100 m to 2 μM at 150 m. Hydrogen sulphide was only detected below 150 m, peaking in concentration (26 μM) at the basin. The concentration of NO_3^- ranged from 27 μM at 100 to 0.055 μM at 150 m. Similarly, the concentration of NO_2^- ranged from 0.03 μM at 100 m to 0.095 μM at 150 m.

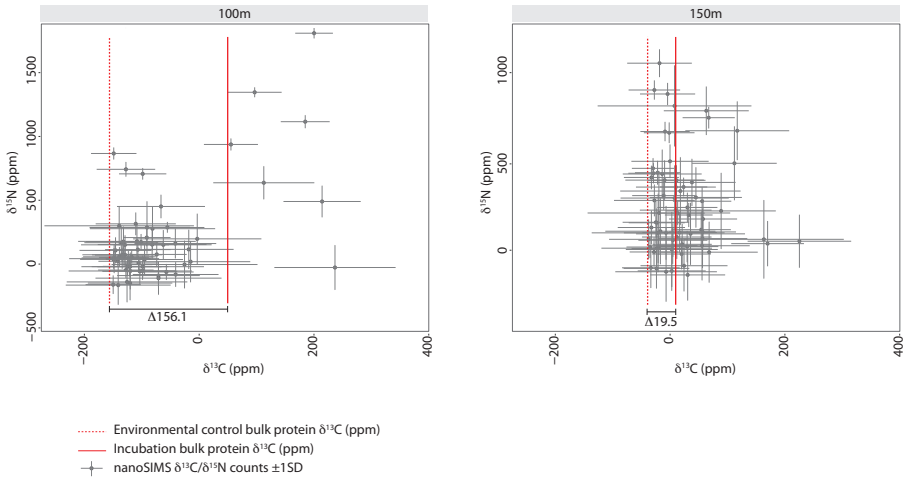

Figure 6.2: Labeled substrate incorporation into cellular biomass. Cellular nanoSIMS δ^{13}C/δ^{15}N (ppm) counts for 100 and 150 m $^{13}CH_4$ + NO_2^- incubation samples. Isotopic values are depicted as dots $\pm$1 SD. Environmental bulk protein δ^{13}C (ppm) is shown as a dotted red line while incubation sample bulk protein δ^{13}C (ppm) is shown as a solid red line. Total sample δ^{13}C enrichment (Δ) is shown between the environmental control and sample bulk protein values.

6.3.2 Carbon incorporation into biomass

To evaluate carbon incorporation into community biomass derived from CH_4 addition, I measured the isotopic carbon composition (δ^{13}C) from cell material (pellets and individual cells). Bulk δ^{13}C measurement of cell pellets showed higher carbon incorporation in 100m (181.1 $^{o}/_{oo}$) samples than in environmental control (100 m = -25 $^{o}/_{oo}$ and 150 m = -31.2 $^{o}/_{oo}$) and 150 m (11.7 $^{o}/_{oo}$) samples (Fig. 6.2). In addition, individual cell δ^{13}C values were higher in 100 m samples (1.66- 26.09 $^{o}/_{oo}$) than 150 m samples (0.02- 11.4 $^{o}/_{oo}$) (Fig. 6.2). Combined, δ^{13}C measurements suggest higher metabolic activity in 100 m samples with regard to carbon incorporation from CH_4 oxidation.

6.3.3 Microbial community composition

To initially identify potential microbial agents carrying out CH_4 oxidation coupled with NO_2^- reduction under dysoxic-suboxic water column conditions, I analyzed incubation SSU rDNA pyrotag sequences including the environmental and experimental controls, and incubation treatments ($^{12}CH_4$, $^{12}CH_4$ + NO_2^- and $^{13}CH_4$ + NO_2^-) from 100 and 150 m. Results revealed microbial community partitioning associated with depth and incubation treatment. For instance, hierarchical cluster analysis (HCA) resolved two major groups, or clusters (AU $>$ 70, 1,000 iterations), associated with 100 m (group I) and 150 m (group II), in which incubation treatment differentiate from controls (environmental and experimental) (Fig. 6.3).

Overall, microbial community composition was primarily constituted by OTUs (average relative abundance $>$1% across samples) affiliated with ubiquitous and abundant taxonomic groups previously identified

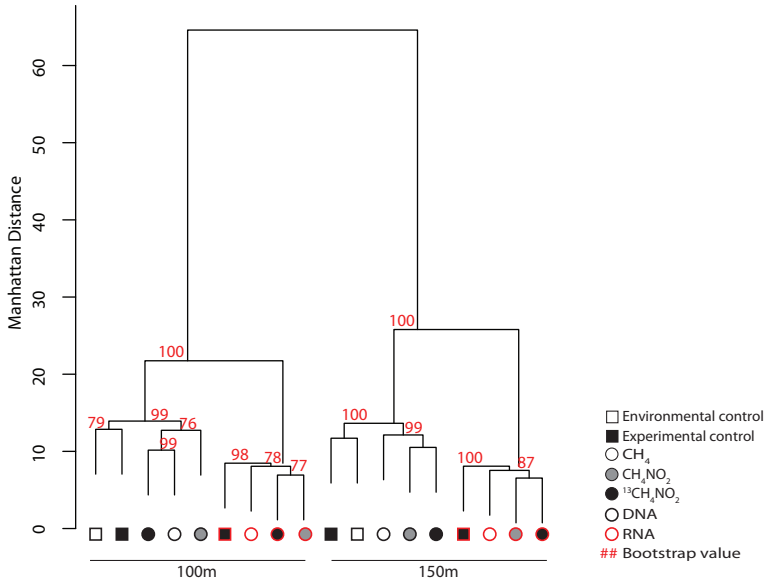

Figure 6.3: Microbial community partitioning based on depth and incubation treatments. Hierarchical clustering of pyrotag data (rDNA and rRNA) based on Manhattan distance. Clusters are delimited by depth (100 and 150 m). Samples are depicted in color dots according to incubation treatment. Bootstrap values (1000 iterations) are shown in red.

in marine surveys, including the Saanich Inlet, (Field *et al.*, 1997; Fuhrman and Davis, 1997; Brown and Donachie, 2007; Tripp *et al.*, 2008; Lam *et al.*, 2009; Walsh *et al.*, 2009; Zaikova *et al.*, 2010; Walsh and Hallam, 2011; Wright *et al.*, 2012) as Bacteroidetes, Chloroflexi, Rhodospirillaceae and Pelagibacteraceae within the Alphaproteobacteria, Methylophilales within the Betaproteobacteria, Methylococcales and SUP05 within the Gammaproteobacteria, Plactomycetes, Marine Group A, and Thaumarchaeota (Fig. 6.4).

Methanotroph diversity in incubation samples was determined based on the recruitment of representative OTU sequences to full-length and pyrotag SSU rRNA gene reference sequences as previously described (Torres-Beltrán *et al.*, 2016). I identified 339 OTU sequences affiliated with Methylococcales (2% of the total rDNA pyrotag dataset) based on BLAST comparisons against the SILVA database (>97% identity) in the incubation samples. Subsequently, I conducted further analysis on representative methanotrophic OTUs affiliated with Methylococcales using full-length SSU rRNA gene sequences as fragment recruitment platforms (≥99% identity). In all, I identified 4 subgroups with phylogenetic similarity to OPU1 (48%) and OPU3 (37%) (Hayashi *et al.*, 2007; Tavormina *et al.*, 2010; Tavormina *et al.*, 2013), methanotrophic symbionts (7%) (Streams *et al.*, 1997; Dubilier *et al.*, 2008; Petersen and Dubilier, 2009) and type I methanotrophs (8%) (Fig. D.1).

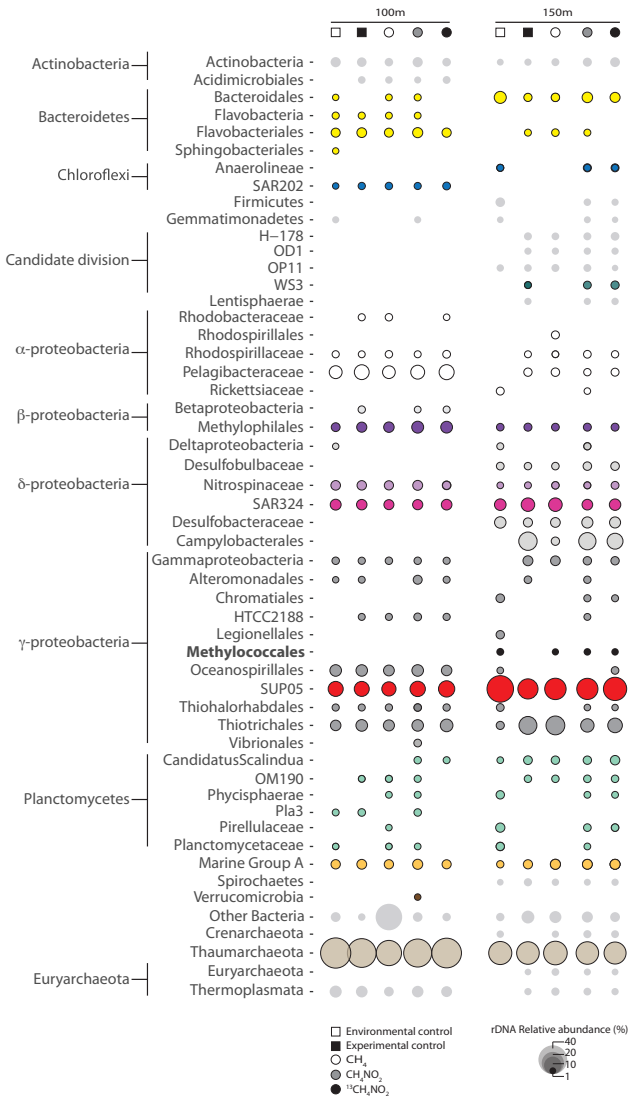

Figure 6.4: Taxonomic composition of OTUs identified in SSU rDNA pyrotags among treatments. Abundant (>1% relative abundance) taxa found in pyrotag datasets. The size of each dot represents taxa relative abundance calculated from the total number of prokaryotic reads in each treatment sample. Sample treatments are depicted as symbols: environmental control = open square, experimental control = solid square, $^{12}CH_4$ = open circle, $^{12}CH_4 + NO_2^-$ = gray circle, and $^{13}CH_4 + NO_2^-$ = solid circle.

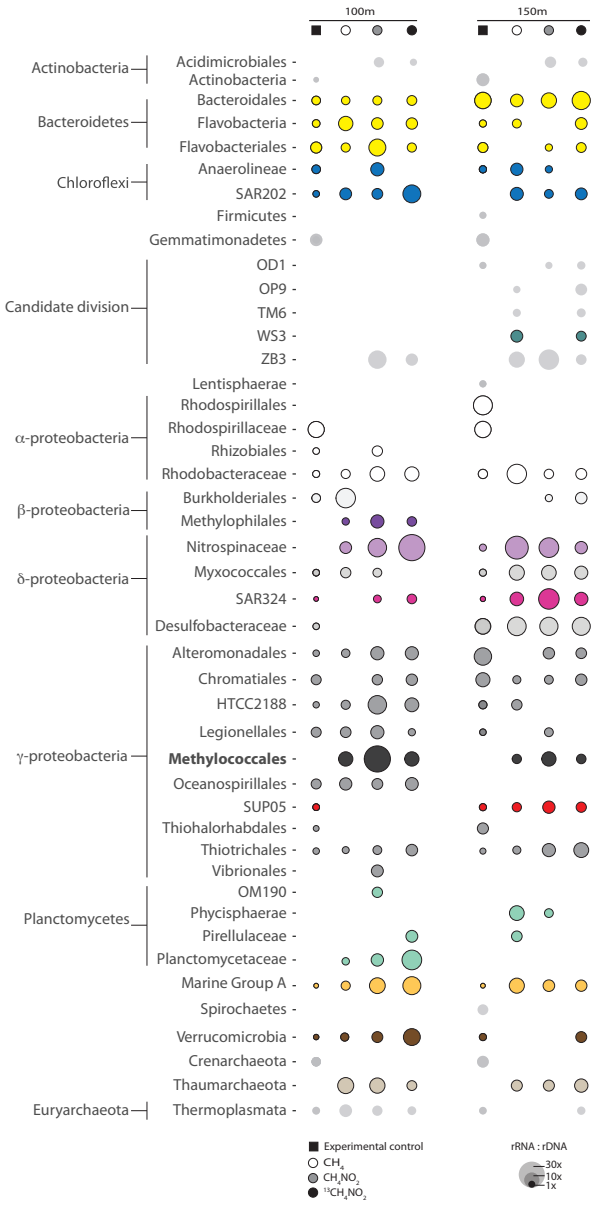

Figure 6.5: Taxonomic composition of OTUs identified as potentially active taxa under incubation treatments. Active taxa (rRNA:rDNA >1) found in SSU pyrotag datasets. The size of each dot represents the rRNA:rDNA ratio calculated from taxa relative abundance in rRNA data over taxa relative abundance in rDNA data. Sample treatments are depicted as symbols: environmental control = open square, experimental control = solid square, $^{12}CH_4$ = open circle, $^{12}CH_4 + NO_2^-$ = gray circle, and $^{13}CH_4 + NO_2^-$ = solid circle.

6.3.4 Insight into microbial community response to substrate additions

To initially identify potential changes in microbial community activity responding to CH_4 and NO_2^- addition, I analyzed SSU rRNA pyrotag sequences from the experimental control, $^{12}CH_4$, $^{12}CH_4$ + NO_2^- and $^{13}CH_4$ + NO_2^- treatments from 100 and 150 m. Overall, specific taxa dominated the SSU rRNA pyrotag datasets (average relative abundance >1%) exhibiting up to a 30-fold increase in SSU rRNA: rDNA when CH_4 and NO_2^- were added. For instance, Methylophilales (10%), Nitrospinaceae (7%), Methylococcales (17%), Planctomycetes (affiliated with OM190 (2%), Pirellulaceae (5%) and Planctomycetaceae (4%)), Marine Group A (4%), Verrucomicrobia (4%), and Thaumarchaeota (12%) OTUs showed a 10-30 fold increase in 100m CH_4 and CH_4 + NO_2^- treatments when compared to the experimental control (Fig. 6.5). In contrast, Chloroflexi affiliated with SAR202 (3%), OP9 (2%), TM6 (3%), WS3 (6%) and ZB3 (8%) Candidate divisions, SAR324 (3%), Methylococcales (6%), SUP05 (27%), Planctomycetes affiliated with Phycisphaerae (3%) and Pirellulaceae (10%), Marine Group A (3%) and Thaumarchaeota (6%) showed a 10-20 fold increase in 150 m CH_4 and CH_4 + NO_2^- treatments when compared to control (Fig. 6.5).

To further identify activity patterns among methanotroph OTUs in the SSU rRNA datasets, I compared Methylococcales OTUs in SSU rRNA pyrotag datasets to full-length and pyrotag SSU rRNA gene reference sequences as previously described (Torres-Beltrán *et al*., 2016). I identified 2 subgroups ($\geq$ 99% identity) affiliated with OPU3 (average relative abundance 13%) (Hayashi *et al*., 2007; Tavormina *et al*., 2010; Tavormina *et al*., 2013) and methanotrophic symbionts (average relative abundance 4%) (Streams *et al*., 1997; Dubilier *et al*., 2008; Petersen and Dubilier, 2009) (Fig. D.1). Based on SSU rRNA: rDNA, OPU3 and methanotrophic symbiont OTUs showed a 10-20-fold increase in 100m CH_4 and CH_4 + NO_2^- treatments when compared to the experimental control (Fig. D.1).

To provide primary evidence for potential methanotroph-community interactions occurring in incubation samples that could be representative of environmental microbial assemblies, I first BLAST compared microbial OTUs observed in incubations SSU rRNA pyrotag datasets to pyrotag sequences recovered from a previous methanotroph diversity survey carried out in Saanich Inlet waters (Torres-Beltrán *et al*., 2016). A total of 57 OTUs were found affiliated ($\geq$ 99% identity) to OTUs present in the network corresponding to 25% of total OTUs in dysoxic-suboxic and anoxic modules (Fig. 6.6A; Table D.2). These 57 OTUs were enriched and active (based on abundance and SSU rRNA:rDNA ratio values) on incubation treatments supporting our experimental community assemblies with environmental co-occurrence network observations (Fig. 6.6A; Table D.1). Sub-networks were constructed highlighting active OTUs showing significant positive correlations (CV>0.6, p<0.001) with OPU3 and methanotrophic symbiont OTUs (Fig. 6.6B). For instance, OTUs correlating with OPU3 were affiliated with Chloroflexi, Alphaproteobacteria, Nitrospina, Marine Group A, Planctomycetes, Verrucomicrobia, and Thaumarchaeota (Fig. 6.7B). In contrast, OTUs correlating with methanotrophic symbionts were affiliated with Bacteroidetes, Anaerolineae within Chloroflexi, WS3 Candidate division, Alphaproteobacteria, Methylophilales, Nitrospina, SAR324, SUP05, Marine Group A, Planctomycetes, Verrucomicrobia and Thaumarchaeota (Fig. 6.6B).

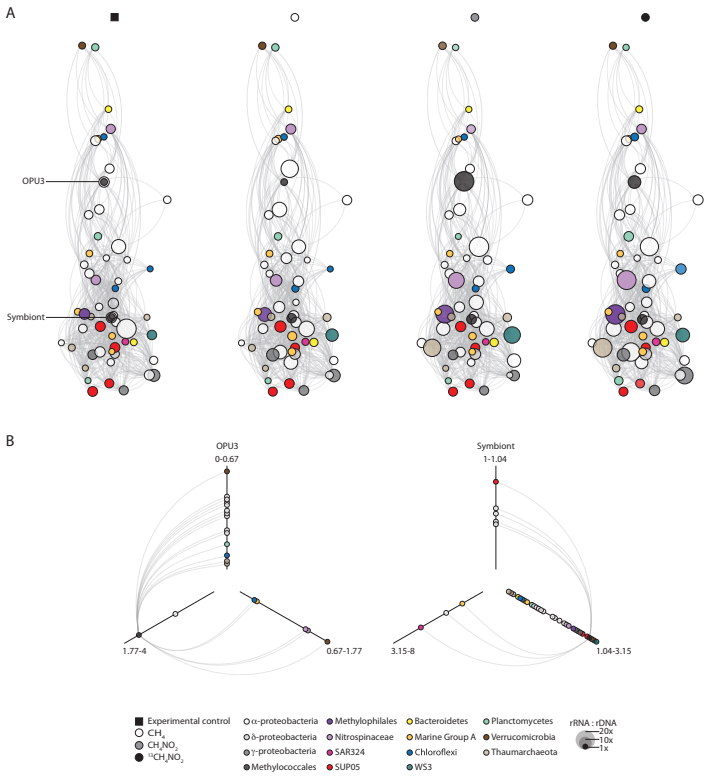

Figure 6.6: Co-occurrence patterns for methanotrophic OTUs from SSU pyrotag datasets. A) Time-series network based on significant ($p<0.001$) Bray-Curtis and Spearman correlation values (>0.6) among OTUs from SSU rRNA gene pyrotag datasets (May 2008 - July 2010). Network encompassing the sub-oxic module where nodes are depicted according to taxonomy as indicated in the legend, and all interactions (edges) are shown as solid lines. Representative networks for each treatment are shown as indicated with the symbol on top (experimental control = solid square, $^{12}CH_4$ = open circle), $^{12}CH_4 + NO_2^-$ = gray circle and $^{13}CH_4 + NO_2^-$ = solid circle). Node size represents OTUs rRNA: rDNA cumulative values for each treatment. B) Hive panels for OPU3 and symbiont OTUs showing resilient interactions throughout incubation treatments. OTUs are distributed based on average relative abundance (log transformed) on the three axes from low (located closer to the center) to high (located towards the end of each line). Nodes are depicted according to taxonomy as indicated in the network. All interactions (edges) are shown as solid lines.

6.3.5 PmoA and NirK diversity

To gain insight into the microbial community carrying out CH_4 oxidation coupled with NO_2^- reduction, I looked into the taxonomic affiliation of *pmoA* and *nirK* genes. Overall, *pmoA* sequences represented 0.002% (170 ORF sequences) of the total number of predicted ORFs in the metagenomic datasets. All *pmoA* representative sequences (clustered at 97% similarity) from incubation samples were related to OPU3 (Hayashi *et al*., 2007) (Fig. D.2). In contrast, *nirK* sequences represented 0.021% (1818 ORF sequences) of the total of predicted ORFs in the metagenomic datasets. Overall, *nirK* sequences were distributed through different taxonomic groups including Archaea (40%), Nitrospina (20%), Bacteroidetes (7%), Chlorobi (3%), and OPU3 (7%), and environmental sequences from Arabian Sea and Black Sea OMZs (3%) (Fig. D.3). Approximately 20% of *nirK* sequences fell within singleton clusters affiliated with Bacteria (clustered at 70% similarity).

6.4 Insight into microbial community response to substrates addition

To further determine the metabolically active community associated with CH_4 oxidation, CH_4-derived carbon fixation and nitrogen metabolism, I looked into expression patterns across meta-transcriptomic and proteomic datasets. Overall, transcripts and proteins related to CH_4 oxidation, CH_4-derived carbon fixation, and nitrogen metabolism found in CH_4 + NO_2^- treatments only showed higher expression with respect controls (environmental and experimental) and CH_4 addition treatments (Table 6.1).

Second, I determined the community composition in meta-transcriptomic and proteomic data by summing transcripts and proteins affiliated to a given microbial group. I observed that transcripts (2% of total transcripts) and proteins (20% total proteins) were mostly affiliated taxonomically with Flavobacteriales within Bacteroidetes, Chlorobi, Chloroflexi, Nitrospina, Rhodobacterales, Bradyrhizobiaceae and SAR 11 within the Alphaproteobacteria, Methylophilales within the Betaproteobacteria, Desulfobacterales and SAR324 within the Deltaproteobacteria, Epsilonproteobacteria, Alteromonadaceae, Methylococcaceae, SUP05, OMG and SAR86 within the Gammaproteobacteria, Planctomycetaceae within the Planctomycetes, NC10, Verrucomicrobia and Thaumarchaeota (Table 6.2), supporting our metabolic inferences based on SSU rRNA: rDNA observations.

Next, I evaluated in greater detail the abundance and taxonomic affiliation of key proteins involved in CH_4 oxidation, CH_4-derived carbon fixation and nitrogen metabolism across metatranscriptomic and metaproteomic datasets. To begin with, I observed some similarities with regard to the functional and taxonomic composition between CH_4 and CH_4 + NO_2^- treatments, such as PmoA from OPU3, formate dehydrogenase (Fdh) from Bradyrhizobiaceae and Methylobacteriaceae within the Alphaproteobacteria, and Desulfobacterales within the Deltaproteobacteria, and the serine hydroxymethyl transferase (SHMT) from OPU3, and Bacteroidales within the Bacteroidetes (Fig. 6.7). In addition, I observed transcripts for key proteins from the Calvin-Benson (CBB) cycle, such as the ribulose bisphosphate carboxylase (RuBISCO) from Gammaproteobacteria, and ribulose phosphate 3-epimerase (RPE) from Methylophilales, and Flavobacteriales within the Bacteroidetes (Fig. 6.8). Further, I observed proteins related to the TCA cycle including the phosphoenolpyruvate carboxylase (PEP) from Flavobacteriales within the Bacteroidetes, and 2-oxoglutarate ferredoxin oxidoreductase (OGOR) from Epsilonproteobacteria (Fig. 6.7). Additionally, I observed the ex-

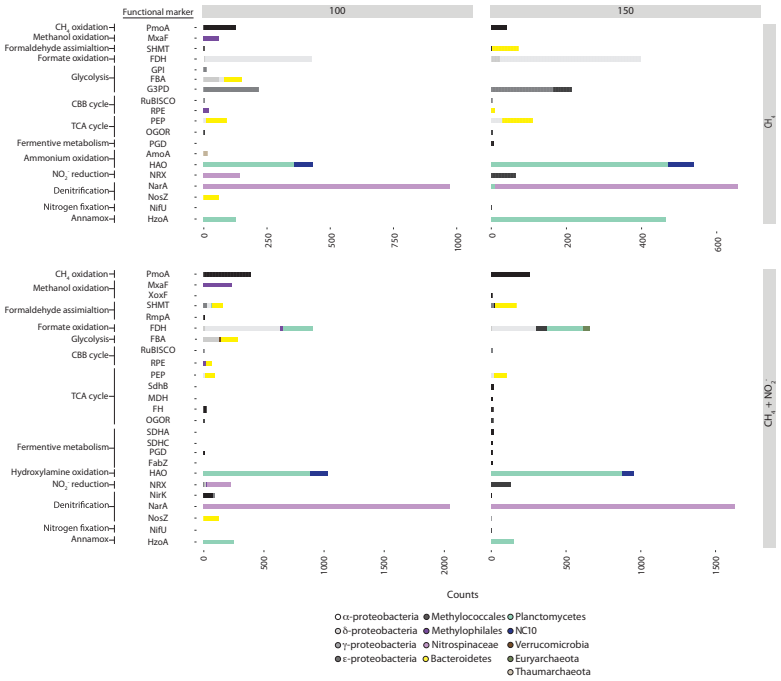

Figure 6.7: Taxonomic and functional breakdown of transcripts related to key proteins for methane and nitrogen cycling. The transcripts related to key proteins for methane and 610 nitrogen cycling are as follows: PmoA, particulate methane monooxygenase subunit β; MxaF, Ca-dependent methanol dehydrogenase; XoxF, Lanthanide-dependent methanol dehydrogenase; SHMT, serine hydroxymethyl transferase; RmpA, 3-hexulose-6-phosphate isomerase; FDH, formate dehydrogenase; GPI, glucose 6 phosphate isomerase; FBA, fructose bisphosphate aldolase; G3PD, glyceraldehyde-3-phosphate dehydrogenase; RuBISCO, ribulose bisphosphate carboxylase; RPE, ribulose phosphate 3-epimerase; PEP, phosphoenol pyruvate carboxylase; SdhB, succinate dehydrogenase iron-sulfur subunit; MDH, malate dehydrogenase; FH, fumarate hydratase; OGOR, 2-oxoglutarate ferredoxin oxidoreductase; SDHA, succinate dehydrogenase flavoprotein subunit; SDHC, succinate dehydrogenase cytochrome b556 subunit; PGD, 6-phosphogluconate dehydrogenase; FabZ, 3-hydroxyacyl ACP dehydratase; AmoA, ammonia monooxygenase subunit α; HAO, hydroxylamine reductase; NarA, nitrate reductase subunit α; NRX, nitrite oxidoreductase; NirK, nitrite reductase; NosZ, nitrous oxide reductase; NifU, nitrogen fixation protein; Hzo, hydrazine oxidoreductase. Function and taxonomy assignments were determined by the sequence identity of transcripts and peptides to metagenomic reads. The total number of reads for each protein are depicted as bars coloured according to taxonomy as indicated in the color key.

Table 6.1: Key protein ratios across meta-transcriptomic and proteomic datasets. Summed ratio values of key proteins involved in CH_4 oxidation, CH_4-derived carbon fixation, and nitrogen metabolism. Values ≥ 1 mean expression of a given transcript and protein was higher in samples than controls (environmental and experimental). Lack of value (indicated as '-') means no transcript or protein was observed in controls.

	RNA						Protein					
	Sample : Environmental Control			Sample: Experimental Control			Sample : Environmental Control			Sample: Experimental Control		
Marker protein	CH_4	$^{12}CH_4 + NO_2^-$	$^{13}CH_4 + NO_2^-$	CH_4	$^{12}CH_4 + NO_2^-$	$^{13}CH_4 + NO_2^-$	CH_4	$^{12}CH_4 + NO_2^-$	$^{13}CH_4 + NO_2^-$	CH_4	$^{12}CH_4 + NO_2^-$	$^{13}CH_4 + NO_2^-$
PmoA; particulate methane monooxygenase subunit β	4.88	6.28	7.32	12.03	15.48	18.04	1.21	7.81	6.02	1.09	3.05	3.50
MxaF; Ca-dependent methanol dehydrogenase	2.20	1.38	2.20	2.29	3.03	4.83	-	2.64	1.07	-	-	-
XoxF; Lanthanide-dependent methanol dehydrogenase	1.36	6.84	7.52	2.75	4.37	4.80	-	1.70	1.77	-	-	-
SHMT; serine hydroxymethyl transferase	2.61	4.43	4.41	1.43	3.64	3.62	-	-	-	-	-	-
RmpA; 3-hexulose-6- phosphate isomerase	7.53	9.53	11.61	2.86	7.12	8.67	-	2.50	1.53	-	-	-
FDH; formate dehydrogenase	1.39	2.95	3.43	2.58	3.26	3.50	1.75	9.70	6.55	1.18	3.36	3.00
GPI; glucose 6 phosphate isomerase	1.05	2.22	1.96	2.45	3.64	4.83	-	-	-	-	-	-
FBA; fructose bisphosphate aldolase	1.05	2.22	2.94	1.72	1.82	2.42	-	-	-	-	-	-
G3PD; glyceraldehyde 3 phosphate dehydrogenase	1.05	2.77	1.10	1.17	1.82	1.41	-	-	-	-	-	-
RuBISCO; ribulose bisphosphate carboxylase	2.09	1.48	1.47	1.00	2.73	2.42	-	-	-	-	-	-
RPE; ribulose phosphate 3-epimerase	1.46	1.01	1.47	1.00	2.73	2.42	-	-	-	-	-	-
PEP; phosphoenol pyruvate carboxylase	1.19	3.88	3.31	1.72	2.55	2.17	-	-	-	-	-	-
SdhB; succinate dehydrogenase iron−sulfur subunit	1.57	3.69	3.92	1.58	3.64	3.62	-	-	-	-	-	-
MDH; malate dehydrogenase	2.39	3.48	3.78	1.37	2.00	2.17	-	-	-	-	-	-
FH; fumarate hydratase	2.09	4.43	5.88	1.47	2.08	1.55	-	-	-	-	-	-
OGOR; 2-oxoglutarate ferredoxin oxidoreductase	1.05	2.77	2.20	1.15	1.21	1.61	-	-	-	-	-	-
SDHA; succinate dehydrogenase flavoprotein subunit	2.56	3.69	3.92	2.06	3.64	3.62	-	-	-	-	-	-
SDHC; succinate dehydrogenase cytochrome b556 subunit	2.51	1.11	1.29	2.21	1.82	1.21	-	-	-	-	-	-
PGD; 6−phosphogluconate dehydrogenase	1.25	1.77	1.76	1.72	2.43	2.42	-	-	-	-	-	-
FabZ; 3−hydroxyacyl−ACP dehydratase	2.09	11.08	14.84	2.41	4.55	1.81	-	-	-	-	-	-
AmoA; ammonia monooxygenase subunit α	4.88	3.10	3.92	2.00	2.32	2.42	-	-	-	-	-	-
HAO; hydroxylamine reductase	1.05	1.86	3.70	2.06	1.82	1.45	1.40	6.90	5.56	1.52	8.05	7.00
NarA; nitrate reductase subunit α	3.28	5.10	4.43	1.21	2.73	2.56	-	-	-	-	-	-
NRX; nitrite oxidoreductase	1.39	4.43	4.41	1.12	7.25	8.27	1.77	3.26	3.30	1.36	6.76	4.20
NirK; nitrite reductase	1.39	4.43	4.41	1.82	6.87	6.64	5.70	10.35	9.19	2.82	14.04	11.20
NosZ; nitrous oxide reductase	6.27	5.76	8.23	6.44	5.92	8.45	-	-	-	-	-	-
NifU; nitrogen fixation protein	10.46	15.51	17.64	1.72	1.59	1.81	-	-	-	-	-	-
Hzo; hydrazine oxidoreductase	1.25	1.48	1.31	1.47	2.42	2.42	1.12	5.67	4.91	1.85	5.25	5.65

pression of hydroxylamine oxidoreductase (HAO) from Planctomycetaceae and Candidate NC10, of nitrite oxidoreductase (NXR) and nitrate reductase (NarA) from Nitrospinaceae, as well as of nitrous oxide reductase (NosZ) from Chlorobi within the Bacteroidetes (only from 100 m samples) (Fig. 6.7). I also observed the expression of the nitrogen fixation protein (NifU) from Epsilonproteobacteria (only from 150 m samples), and hydrazine oxidoreductase (Hzo) from Planctomycetaceae (Fig. 6.7).

Unique expression for 100 m CH_4 samples included transcripts related to the Glycolysis pathway, such as glucose-6-phosphate isomerase (GPI) from Gammaproteobacteria (8.08 RPKM), fructose bisphosphate aldolase (FBA) from SAR11 (64.68 RPKM) within Alphaproteobacteria, SAR324 (97 RPKM) within the Deltaproteobacteria, Chlorobi (145.5 RPKM) and Flavobacteriales (16.17 RPKM) within the Bacteroidetes, and glyceraldehyde-3-phosphate dehydrogenase (G3PD) from Alteromonadaceae (41.8 RPKM), OM60

Table 6.2: Total number of transcripts and proteins affiliated with metabolically active taxa. Summed count values for all transcripts and proteins which taxonomic affiliation was related to a given microbial group across treatments. Lack of value (indicated as '-') means no protein was observed.

Taxonomy		Total number of Transcripts	Total number of Proteins
Bacteroidetes	Flavobacteriales	312	193
	Chlorobi	120	6
	Chloroflexi	360	24
	Nitrospina	7128	2728
	Rhodobacterales	2316	1673
α-proteobacteria	Bradyrhizobiaceae	960	720
	SAR11	2004	294
β-proteobacteria	Methylophilales	540	132
	Nitrosomonadales	240	-
δ-proteobacteria	Desulfobacterales	1920	370
	SAR324	288	120
ε-proteobacteria	Campylobacterales	1680	888
	Alteromonadaceae	228	180
	Methylococcaceae	5916	5904
	SUP05	2648	2258
	OMG group	216	-
	SAR86	2184	10
Planctomycetes	Planctomycetaceae	672	128
	Candidatus Methylomirabilis oxyfera NC10	600	240
	Verrucomicrobia	120	-
Archaea	Thaumarchaeota	2400	625

clade (111.61 RPKM), and SAR92 (48.51 RPKM) within Gammaproteobacteria (Fig. 6.7). Additionally, the expression of the ammonia monooxygenase subunit α (AmoA) from Thaumarchaeota (22.63 RPKM) was observed. With respect to the 150 m CH_4 sample, G3PD from Epsilonproteobacteria (168.5 RPKM) was observed particularly, as well as 6-phosphogluconate dehydrogenase (PGD) from OPU3 (5.26 RPKM). In addition, NarA from Planctomycetaceae (17.55 RPKM) was also uniquely observed in 150 m CH_4 sample (Fig. 6.7).

Differential expression patterns emerged based on O_2 concentration (related to 100 and 150 m depth intervals) for the CH_4 + NO_2^- treatments, with little variation between $^{12}CH_4$ and $^{13}CH_4$ substrates (Supplementary Information Fig. 4). Of note, was the occurrence of Ca-dependent (MxaF) and lanthanide-dependent (XoxF) methanol dehydrogenases from Methylophilales (157.2 RPKM) and OPU3 (113 RPKM), in 100 and 150 m samples, respectively (Fig.7). Also of note was the expression of 3-hexulose 6-phosphate (RmpA) from OPU3 (107 RPKM) in 150 m samples. Additionally, Fdh from Methylophilales (48.51 RPKM) and Planctomycetaceae (205.6 RPKM) were observed in 100 m samples, while Fdh from Planctomycetaceae (802 PRKM), Epsilonproteobacteria (2.27 RPKM) and Methanomicrobia (178.86 RPKM)

within the Euryarchaeota were observed in 150 m samples (Fig. 6.7). In addition, FBA from Verrucomicrobia (11.36 RPKM) was uniquely observed at 100 m. Interestingly, transcripts for the TCA cycle including succinate dehydrogenase iron-sulfur subunit (SdhB), malate dehydrogenase (MDH), and fumarate hydratase (FH) (SdhB = 8.08 RPKM, MDH = 5.68 RPKM, and FH = 5.68 RPKM), and other fermentative enzymes such as succinate dehydrogenase flavoprotein subunit (SDHA). The succinate dehydrogenase cytochrome b556 subunit (SDHC), 6-phosphogluconate dehydrogenase (PGD) and 3-hydroxyacyl ACP dehydratase (FabZ) were only observed at 150 m and were affiliated with OPU3 (SDHA=16.17 RPKM, SDHC=51.4 RPKM, PGD= 15.65 RPKM, FabZ= 8.56 RPKM) (Fig. 6.7). The expression of NirK from OPU3 was observed at 100 and 150 m (125.5 and 5.68 RPKM, respectively) (Fig. 6.7).

Protein expression was constrained to a very specific set of proteins and taxa within the pathways explored. For instance, PmoA affiliated with OPU3 was expressed across all samples, showing higher NSAF values in CH_4 + NO_2^- samples (1.64 and 0.918 NSAF at 100 and 150 m, respectively) (Fig. 6.8). MxaF and XoxF methanol dehydrogenases from Methylophilales and OPU3 were also expressed in CH_4 + NO_2^- samples (0.473 and 0.123 NSAF at 100 and 150 m, respectively) (Fig. 6.8). Formate dehydrogenase from Gammaproteobacteria was expressed across all samples, showing the highest NSAF value (0.768) in 150 m CH_4 + NO_2^- samples. Similarly, RmpA from OPU3 (0.107 NSAF) was expressed in 150 m CH_4 + NO_2^- samples (Fig. 6.8). In addition, HAO from NC10 (0.674 NSAF) was only expressed in 150m CH_4 + NO_2^- samples. Nitrite reductase (NirK) from OPU3 was only expressed (2.479 NSAF) in 100 m CH_4 + NO_2^- samples (Fig. 6.8), while NXR from Nitrospinaceae was expressed across all samples, showing the highest value (7.64 NSAF) in the 100 m CH_4 + NO_2^- samples. Hydrazine oxidoreductase, HZO, affiliated with Planctomycetaceae was expressed in 150 m samples showing the highest expression in CH_4 + NO_2^- samples (0.258 NSAF) (Fig. 6.8).

6.4.1 Elucidating microbial metabolic networks

To ultimately elucidate potential metabolic interactions between taxa, I converged taxonomic co-occurrence patterns with metabolic network information. First, a co-occurrence analysis based on Bray-Curtis and Spearman correlation values among predicted pathways at the suboxic-dysoxic water-column compartment (between 100 and 150 m) over a two-year period (June 2009 to August 2011) resulted in a microbial metabolic network (Supplementary Information S1). The co-occurrence analysis resulted in significant positive correlations (CV$>$0.6, $p<0.001$) among pathways, including CH_4 oxidation, methanol oxidation to formaldehyde, formaldehyde oxidation via the H4MPT pathway, formate oxidation, CO_2 fixation via the TCA cycle, and ammonium and NO_2^- oxidation (Fig. 6.9; Table D.2).

To provide further evidence on CH_4 oxidation as a communal function under suboxic-dysoxic water-column conditions, I BLAST-compared ($>$90% identity) incubation transcripts to conceptually translated ORFs of key proteins within correlating pathways. I observed most ORFs showed identical functional and taxonomic annotation to CH_4 + NO_2^- transcripts from 150 m, except for AmoA from Thaumarchaeota that matched transcripts from the 100 m CH_4 sample. For instance, PmoA was affiliated with OPU3, MxaF was found to be affiliated with Methylophilales, NXR was affiliated with Nitrospinaceae, Fdh was affiliated with Gammaproteobacteria and Planctomycetes, and ShdB, MDH and FH within the TCA cycle were affiliated

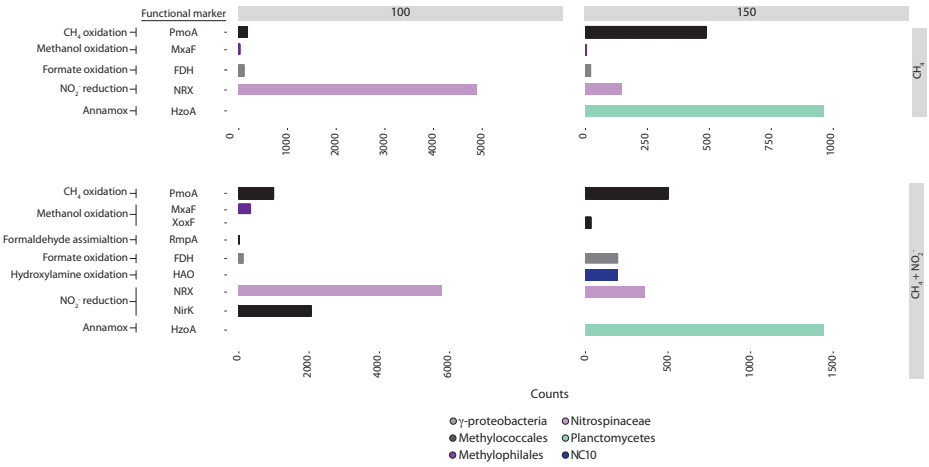

Figure 6.8: Taxonomic and functional breakdown of key proteins related to methane and nitrogen cycling. The key proteins related to methane and nitrogen cycling are as follows: PmoA, particulate methane monooxygenase subunit β; MxaF, Ca-dependent methanol dehydrogenase; XoxF, Lanthanide-dependent methanol dehydrogenase; RmpA, 3-hexulose-6- phosphate isomerase; FDH, formate dehydrogenase; HAO, hydroxylamine reductase; NRX, nitrite oxidoreductase; NirK, nitrite reductase; NosZ, nitrous oxide reductase; HzoA, hydrazine oxidoreductase. Function and taxonomy assignments were determined by the sequence identity of transcripts and peptides to metagenomic reads. The total number of reads for each protein are depicted as bars coloured according to taxonomy as indicated in the color key.

with OPU3 (Fig. 6.9).

Based on ORFs and transcript pairs I elucidated a potential mechanism from which CH_4 oxidation serves as carbon source for the microbial community under dysoxic-suboxic water column conditions. I suggest methanotrophic OPU3 goes under a fermentative state due to O_2 limitation and provides methanol to methylotrophic Methylophilales. Formate may be released by OPU3 or Methylophilales, as a byproduct of incomplete carbon assimilation, and could potentially feed the Gammaproteobacteria and Planctomycetes. Therefore, OPU3 could potentially uptake carbon via the TCA cycle. In addition, under O_2 limiting conditions, it is likely Thaumarchaeota carries out ammonium oxidation, while Nitrospinaceae carries out NO_2^- oxidation (Fig. 6.9).

6.5 Discussion

This chapter charts methanotroph community composition and metabolic responses to CH_4 and NO_2^- addition under low-O_2 conditions. The observation presented here encompass isotopic signatures demonstrating CH_4-derived carbon incorporation into microbial biomass and multi-omic information describing community shifts in abundance and metabolism related to CH_4 oxidation coupled with NO_2^- reduction. The observations indicate that OPU3 metabolic activity supplies the microbial community with organic compounds that are carbon sources. In all, the results provide evidence supporting CH_4 oxidation as community-level

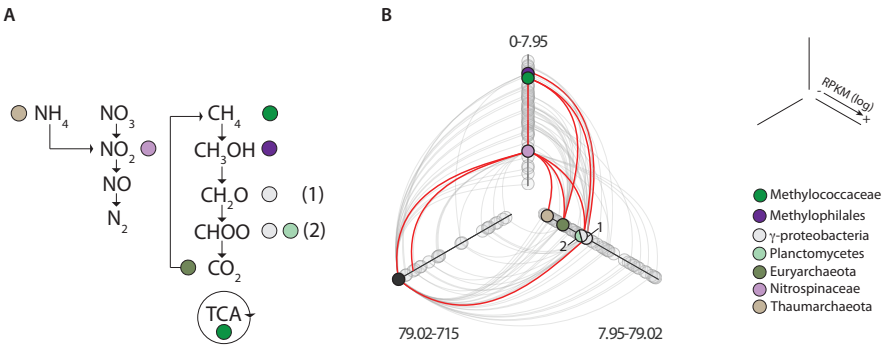

Figure 6.9: Co-occurrence patterns for dysoxic-suboxic metabolic pathways in the Saanich Inlet water column. A) Methane oxidation and carbon fixation via the TCA cycle, and Nitrogen cycling pathways reactions. Colored dots depict taxonomic affiliations for key proteins in a given reaction. B) Time-series network based on significant (p<0.001) Bray-Curtis and Spearman correlation values (>0.6) among predicted pathways from metagenomic datasets (June 2009-vAugust 2011). Hiveplot encompasses CH_4 oxidation correlating pathways where nodes are depicted according to taxonomy, as indicated in the legend, and CH_4 oxidation interactions (edges) are shown as solid red lines. Nodes are distributed along the axis according to RPKM abundance.

function knitting a functional metabolic community network around CH_4 oxidation.

Bulk ^{13}C measurements suggest microbial communities from distinct depths incorporated carbon into protein differently, underlining potential differences in their metabolism due to O_2 concentration. For instance, the community at 100 m showed higher carbon incorporation (156.1 $‰$ increase from environmental control) than the community from 150 m (19.5 $‰$ increase from the environmental control). In addition, I expected carbon incorporation to be derived from CH_4 oxidation primarily by methanotrophs. However, methanotroph abundance after 72 h incubation period (2% average relative abundance from the total number of prokaryotic reads) suggests only a small fraction of the oxidized CH_4 was converted to biomass. Recent experiments carried out in bioreactors showed that at low dissolved O_2 concentrations, 40-50% CH_4-derived carbon was extracellular acetate and formate, supporting the hypothesis that CH_4 fermentation leads to little biomass synthesis (Kalyuzhnaya *et al.*, 2013). In addition, an extended anaerobic starvation experiment on the methanotrophic isolate strain WP 12 showed that cell biomass decreased during the cultivation period and that metabolized carbon was recovered mainly as organic solutes in the starvation medium (Roslev and King, 1995).

Multi-omic information points to OPU3 as the main methanotrophic agent carrying out CH_4 oxidation coupled with NO_2^- reduction (PmoA transcripts and proteins affiliated with OPU3 showed 6-fold and 2-fold increases in CH_4 + NO_2^- treatments) under suboxic-dysoxic water column conditions in the Saanich Inlet. Recent environmental evidence based on time-series regression analyses (Torres-Beltrán *et al.*, 2016) and transcriptional data (Padilla *et al.*, 2017) from O_2-deficient marine systems support OPU3 coupling CH_4 oxidation and NO_2^- reduction as a mechanism that thrives in the O_2-deficient water columns. Furthermore, the

functional expression patterns observed for OPU3 in relation to O_2 concentrations suggest different modes of carbon fixation by the methanotroph, derived from CH_4 oxidation coupled with NO_2^- reduction. For instance, under suboxic water-column conditions ($\sim$30 μM O_2) OPU3 will likely incorporate carbon via the RuMP pathway. In contrast, under dysoxic water-column conditions ($\sim$3 μM O_2), OPU3 will likely incorporate carbon via fermentative pathways. The presence of putative fermentation genes in methanotrophs is likely widespread (Kalyuzhnaya *et al*., 2013). In fact, the expression of fermentative genes, i.e. MDH and SDH by Methylomicrobium alcaliphilum 20Z, has been recently demonstrated under low-O_2 growth conditions in a bioreactor experiment (Kalyuzhnaya *et al*., 2013), supporting the hypothesis that methanotrophs are capable of fermentation from CH_4-derived formaldehyde, leading to the formation of end products. Combined, these metabolic features and plasticity might allow OPU3 to thrive under low-O_2 water-column conditions, despite the potential bioenergetic cost related to biomass synthesis (all cellular carbon is assimilated at the oxidation level of formaldehyde via RuMP cycle, while only $\sim$ 15% via CO_2 oxidation; (Hanson and Hanson, 1996).

Co-occurrence patterns among metabolically active OTUs, identified based on SSU rRNA:rDNA ratios coupled with network analyses, suggest methanotroph partnerships with correlating community members might not be random (Sauter *et al*., 2012; Beck *et al*., 2013; Hernandez *et al*., 2015; Oshkin *et al*., 2015; Karwautz *et al*., 2018). I observed that the metabolically active community was primarily comprised of OTUs affiliated with taxa, i.e. Methylococcaceae, Methylophilales, Bacteroidetes and Planctomycetes showing a concomitant increase in transcript and protein abundance. These taxa have been described as co-occurring with methanotrophs as their abundance, based on SSU rRNA gene observations, commonly increases with CH_4 addition under laboratory conditions (Sauter *et al*., 2012; Beck *et al*., 2013; Hernandez *et al*., 2015; Oshkin *et al*., 2015; Karwautz *et al*., 2018). Observed functional patterns supported co-occurrence among metabolically active OTUs hinting to communal CH_4 metabolism based on methanotrophs feeding community members on released compounds, i.e. methanol and formate due to fermentative metabolism (Kalyuzhnaya *et al*., 2013; Yu and Chistoserdova, 2017). For instance, differential expression of OPU3 XoxF and Methylophilales MxaF indicates Methylophilales are possibly feeding on CH_4-derived methanol from OPU3 when XoxF is not transcribed/translated into protein. In fact, the methanol cross-feeding mechanism between methanotrophs and methylotrophs has been previously determined in coculture experiments using Lake Washington strains where methanol released by methanotrophic *Methylobacter* sp. supported methylotroph *Methylonera* sp. growth (Krause *et al*., 2017). In addition, concomitant with OPU3 expression patterns, the expression of FDH from Methylophilales, Planctomycetes and Methanomicrobia was observed, suggesting these taxa may feed on formate released by methanotrophs. FDH expression by Methylophilales further supports the metabolic interdependence between methanotrophs and methylotrophs when co-existing in culture conditions (Hernandez *et al*., 2015; Oshkin *et al*., 2015; Yu and Chistoserdova, 2017). Interestingly, the metabolic relationship between methanotrophs, methanogens and Planctomycetes may be based on the evolution of genes for C1 transfer reactions between the oxidation levels of formaldehyde and formate (Chistoserdova *et al*., 2004). In particular, genomic information points to the use of formate as an electron donor by anammox Planctomycetes to aid in the Wood-Ljungdahl pathway for the reduction of carbon dioxide (Strous *et al*., 2006). Another taxa commonly co-occurring with methanotrophs in mi-

crocosm experiments is Bacteroidetes; particularly, as I observed here, Flavobacteriales (Hernandez *et al*., 2015; Oshkin *et al*., 2015; Karwautz *et al*., 2018). Genes expressed by Flavobacteriales affiliated with FBA, REP and PEP, suggest that these may be feeding on polymeric compounds derived from methanotroph metabolism. Experimental results for CH_4-fed microcosm experiments suggest the specific mechanism behind the metabolic relationship between Bacteroidetes and methanotrophs relies on the former feed polymeric substances produced and released by methanotrophs (Kalyuzhnaya *et al*., 2013; Yu and Chistoserdova, 2017), i.e. biopolymers that serve as a source of carbon, energy, or reducing-power for methanotrophs in exceptional circumstances, such as under conditions of nutrient limitation (Strong *et al*., 2015). Combined functional information and environmental pathway network analyses allowed me to connect the metabolic interactions from specific taxa to the potential communal functions of CH_4 oxidation, and to ultimately pinpoint the potential metabolic mechanism by which CH_4-derived carbon could be transferred to the community in O_2-deficient water columns (Fig. 6.9).

According to the metabolic network I elucidated, $\sim$15% of the carbon incorporated by OPU3 could be synthesized into biomass (via TCA pathway under low-O_2 conditions), explaining its low relative abundance in the Saanich Inlet waters (Torres-Beltrán *et al*., 2016). The remaining $\sim$32kg C from CH_4 could potentially be distributed to co-occurring partners as polymeric compounds released by the methanotroph. This rough estimation highlights the role of methanotrophs in O_2-deficient systems beyond the CH_4 cycle and indicates that future efforts should focus on assessing the functional capacity of methanotrophs throughout suboxic-dysoxic waters within coastal and open ocean regions worldwide to identify the environmental factors regulating their metabolism and community interactions.

6.6 Conclusion

In the present chapter, I used experimental and environmental time-series multi-omic observations to determine methanotrophic community composition and activity in response to CH_4 and NO_2^- additions under low-O_2 growth conditions, as well as to understand the nature of community-level interactions related to CH_4 metabolism. Functional information revealed a communal response to CH_4 oxidation ignited by OPU3 coupling CH_4 oxidation with NO_2^- reduction. OPU3 fermentative metabolism under low-O_2 water-column conditions likely shifts the community pathways of carbon incorporation by releasing organic compounds as carbon sources for co-occurring taxa. In combination, these observations expand our understanding of the communal nature of CH_4 metabolism (reviewed in Chistoserdova and Kalyuzhnaya, 2018) and provide a new outlook on the environmental role of the methanotrophs as essential components of microbial metabolic networks beyond the CH_4 cycle. Although co-occurrence results suggested SAR324, MGA, SUP05, candidate WS3 and Verrucomicrobia may play a part in communal CH_4 oxidation, no functional information on carbon or nitrogen metabolism from these taxa was observed concomitant to methanotroph response, except for FBA from SAR324 and Verrucomicrobia. I encourage the use of tracing techniques, such as stable isotope probing coupled with RNA and protein analyses, to unveil the biochemical processes and mechanisms that could explain the co-occurring patterns of these taxa with methanotrophs. In future research, I suggest starting by addressing the biotic and abiotic factors involved in methanotrophic community co-occurrence patterns as well as the mechanisms for metabolite exchange within methanotrophic communities across sys-

tems worldwide. I recommend future efforts focus on understanding how changing levels of O_2 impact the dynamics among partnering groups and to assess the specific roles of microbes preventing CH_4 emission. This information is necessary to better inform models that evaluate and predict system responses to ocean deoxygenation. Furthermore, knowledge gained from methanotrophic community partnerships will greatly aid in overcoming synthetic community manipulation by providing the knowledge necessary to select a set of strains with characterized genomes and physiologies that can be used in ongoing biotechnological and engineering approaches aimed at methanotroph industrial applications and commercialization.

Chapter 7

Conclusions

Oxygen (O_2) is fundamental to biological and biogeochemical processes in the ocean (Breitburg *et al*., 2018). The effects of O_2-dependent nutrient-cycling processes are communicated by oceanic water mass circulation; as such, changes within oxygen minimum zones (OMZs) can influence microbial-driven biogeochemical processes on regional and global scales. As OMZs expand and their upper boundaries advance toward surface waters, OMZ microbial communities will shoal towards the upper water column. Many of the groups that play important roles in greenhouse gas cycling within OMZs may respond to habitat shifts by changing their abundance, diversity or composition with potential downstream ecosystem function implications. Microbial communities play significant roles driving ocean biogeochemistry (Falkowski *et al*., 2008; Beman and Carolan, 2013). Microbial community structure, in turn, is strongly influenced by the physical and chemical environment (Margalef, 1968; Tozzi *et al*., 2004). As the ecology and biogeochemistry of the oceans are tightly interconnected, the response and resiliency of microbial communities to O_2 loss may have far-reaching effects on ocean biogeochemistry that could results in feedbacks to global warming driven by greenhouse gases. Thus, in conjunction with rising ocean deoxygenation, there is the growing need to generate more information to understand how O_2 loss is altering microbial pathways and the rates of processes related to greenhouse gas cycling within the ocean water column.

7.1 Standards of practice for sample collection and data analysis

The Saanich Inlet time-series data compendiums are community-driven research frameworks for observing and predicting microbial community responses to ocean deoxygenation across multiple scales of biological organization. By generating more compatible geochemical and multi-omic sequence data, robust information may be generated across ecosystems that allows for more accurate global predictions. In combination, the geochemical and multi-omic sequence data are powerful tools with the potential to uncover tight relationships among key microbial players related to the biogeochemical cycling of carbon, nitrogen and sulfur. Elucidating the metabolic networks entwined in these relationships is crucial for the comprehensive understanding of community composition and function shifts due to O_2 loss, and thus, the generation of more accurate predictive tools.

Advances in multi-omic sequencing technology are enabling the study of microbial communities at

unprecedented scales (Hahn *et al*., 2016). Research community efforts on developing multi-omic time-series data have greatly enriched our understanding of the metabolism of key taxa that directly impact nutrient cycling in OMZs. However, time-series data analyses encounter a major flaw; namely, the incompatible collection and filtering strategies between sampling efforts make it challenging to compare observations among systems and restrain the extensiveness of observations on a global scale. I consider developing time-series datasets as reference databases and coupling *in situ* with on-ship sampling methods as a crucial strategy to improve our understanding of microbial community metabolism across marine systems.

In Chapter 3, I presented observations supporting the 2014 SCOR workshop effort to understand the effect of collection methods on microbial ecology research and the interpretation of the microbial communities in O_2-deficient water columns. Results using three-domain SSU rDNA and rRNA 454 tag sequencing data demonstrated shifts in microbial community composition, structure and function associated with collection and filtration methods. Microbial community composition and structure showed considerable abundance shifts for specific taxa that proved to be sensitive to collection time (*in situ* vs. on ship).For instance, Planctomycetes, Chloroflexi and Candidate divisions were only observed in *in situ* samples. Additionally, I provided evidence that size fractionation and particle fragmentation due to filtration had an impact on specific taxa abundance, i.e Bacteroidetes, Alphaproteobacteria and Opisthokonta increased 5-fold in abundance on large size filters (0.4 μm) compared with small size filters (0.22 μm). Coupling SSU rRNA and rDNA tag sequencing data, I demonstrated the impact of filtration methods on indicator OTUs providing insight into community function differences between *in situ* vs. on ship samples. Of note, the identification of microbial taxa within the uncultivable majority such as Candidate divisions (BCR1 and WS3), Desulfarculales, Desulfuromandales, and Phycisphaerae as these taxa are disregarded from the microbial community based on SSU rDNA studies due to their low abundance. However, these taxa may play important roles in sulfur cycling, carbon fixation and fermentation under O_2-deficient water-column conditions. Bias against these rare taxa may lead to neglecting important localized processes and may potentially compromise biogeochemical interpretations (Suter *et al*., 2016) based on on-ship observations alone.

Observations made in Chapter 3 elaborate on previous contributions provided by Ganesh *et al*., Padilla *et al*. and Suter *et al*., indicating the potential effects of sample collection methods in our understanding of microbial community structure and function in OMZs. Combined, these observations highlight the importance of considering the effects of biomass size-fractionation, filtered water volume, and collection timing in experimental design. Additionally, these findings indicate the need for establishing compatible molecular data generation techniques that facilitate cross-scale comparisons and that more accurately assess *in situ* microbial community composition and function. Furthermore, microbial community composition bias due to collection methods may impact our understanding of microbial community function and consequently affect current models estimations of biogeochemical cycling processes in OMZs worldwide that are undergoing ocean deoxygenation.

7.2 Using co-occurrence network analysis to chart ecological interactions

Correlation analyses are powerful tools that can unveil environmental co-occurrence trends that predict a variety of microbial interactions. Hypotheses generated from correlation analyses can be tested by coupling

microscopy methods and diverse emerging technologies i.e cell sorting and high-throughput co-culturing coupled with genomics to provide benchmark data for the evaluation and improvement of network inference approaches. For instance, network inference combined with co-culture experiments can deliver the growth rates and interaction strengths required for mathematical simulations of the microbial community. Faust *et al*. thoroughly reviewed the technological advances and methodological innovations recently used in the field of microbial modeling for the discovery of cooperative and competitive relationships between species and described how these techniques are opening the way towards global ecosystem network prediction and the development of ecosystem-wide dynamic models. Such technological and methodological efforts are key to developing comprehensive and realistic mathematical models required to better predict the effects of microbial communities on nutrients cycling in expanding OMZs. Nowadays, microbial ecological modeling is essential for translating the highly complex, nonlinear and evolving systems of microbial communities to fundamental knowledge and a better understanding of our changing oceans. In addition to the analysis of time-series microbial community observations, the combination of community analyses in natural environments under controlled conditions, i.e. micro-mesocosm and enrichment cultures, will allow us to understand the physiological and regulatory mechanisms at cellular level that ultimately control activity and affect the dispersal of taxa driving the important cycling of greenhouse gases such as CH_4.

In this thesis, I used correlation methods to unravel O_2-driven community-level interactions that affect nutrient and gas cycling in OMZs. First, in Chapter 4, I used correlation analysis to explore protistan parasitic interactions throughout defined O_2 gradients in the Saanich Inlet water column. Bray-Curtis and Spearman's rank correlations and indicator species analyses on SSU rDNA 454 tag sequencing data revealed potential significant interactions occurring between four known Syndiniales groups and different protistan and metazoan taxa including *Phaeocystis antarctica* (ciliates affiliated with Choreotrichia) and copepods during periods of water-column stratification. In surveying host-parasite potential interactions, I provide a baseline understanding of the potential host range of the major parasitic Syndiniales groups that could infect key primary producers and heterotrophic populations in stratified water columns. Furthermore, I give insight into the possible impacts that these infections may have on population dynamics, extensible to nutrient cycling processes, during seasonal water-column stratification in the Saanich Inlet.

In Chapter 5, I used different correlation methods to chart methanotroph dynamics over an atypical extended water-column stratification period in 2010 related to a relatively strong El Niño event. Correlation analyses on SSU rDNA 454 tag sequencing data in combination with geochemical information allowed me to resolve potential novel metabolic strategies, including the use of alternative terminal electron acceptors, i.e. NO_2^- and metabolic interactions among C1-utilizing microorganisms supporting CH_4 oxidation. Results derived from multivariate analysis allowed me to identify significant correlations among methanotrophic OTUs revealing redox-driven niche partitioning along changing water-column redox gradients among taxa, supporting the role of O_2 in shaping microbial community structure and function. Furthermore, negative binomial regressions enabled me to identify potential novel metabolic strategies, including the use of alternative terminal electron acceptors such as NO_2^- by OPU3 ($p < 0.05$). In using a co-occurrence analysis, based on Bray-Curtis and Spearman correlation values among OTUs, I further resolved potential metabolic interactions between OPU1, OPU3 and symbiont groups with Methylophaga, Methylophilales, SAR324,

Verrucomicrobia and Planctomycetes. In all, I provide important baseline information on microbial agents that reduce the flux of climate-active trace gases from ocean to atmosphere and support the potential role of OPU3 as substantial pelagic sink for CH_4 (18.2 W m^{-2}, equivalent to 450 years of radiative forcing that could be released to the atmosphere each year) along continental margins.

To date, the relevant abiotic and biotic environmental conditions shaping methanotrophic communities and influencing their activity have been studied in detail (reviewed in Chistoserdova and Kalyuzhnaya, 2018). However, the question regarding how the different factors act alone and in combination on the members of methanotrophic communities in different ecosystems remains unanswered. The correlation results I presented in Chapter 5 allowed me to provide evidence linking ecosystem function with community composition and on the dependence of environmental parameters. I also contributed novel knowledge on the otherwise unknown niche differentiation and habitat preferences among methanotrophic OTUs relevant to widely-distributed marine clusters. Finally, results in Chapter 5 represent pioneering knowledge that provides valuable insight into understanding the different responses of methanotrophs to low-O_2 and nitrogen concentrations relevant to CH_4 cycling models in O_2-deficient marine waters.

Looking forward, I recommend the use of multi-omic sequencing information coupled with process rate measurements in order to determine the coverage and efficiency of CH_4 oxidation and derived carbon fixation pathways. In addition, detailed incubation experiments using labeled substrates coupled with gene expression studies should be conducted to link CH_4 oxidation pathways and process rates to specific microbial agents. Process rates measurements along defined water column redox gradients are particularly important on regional and global scales to better constrain the CH_4 filtering capacity of coastal and open ocean OMZs.

7.3 Integrative analysis of coupled biogeochemical processes

Community-level CH_4 metabolism has been previously suggested based on environmental and laboratory microbial community surveys (reviewed in Chistoserdova and Kalyuzhnaya, 2018). The observations I presented in Chapter 6 shed light on communal CH_4 metabolism in nature and support a new outlook on the environmental role of methanotrophs as essential components of food webs driven by CH_4-derived carbon fixation. Combined, these observations support the premise that methanotrophs in nature may play important roles in global biogeochemical processes beyond the CH_4 cycle. In Chapter 6, using multi-omic (DNA, RNA and protein) data generated from an incubation experiment on suboxic-dysoxic Saanich Inlet waters. I 1) pinpointed key taxa involved in CH_4 oxidation coupled with NO_2^- reduction, 2) identified the metabolic pathways driving carbon assimilation from CH_4 under low O_2 conditions, and 3) determined microbial metabolic interactions among C1-utilizing microorganisms fueled by CH_4 oxidation.

Observations throughout multi-omic datasets supported and highlighted the potential role of OPU3 as major microbial agent carrying out CH_4 oxidation coupled with NO_2^- reduction in the Saanich Inlet suboxic-dysoxic water-column compartments. Based on the observed metabolic potential of OPU3 for thriving under O_2 deficiency, I assessed the global distribution and occurrence of *pmoA* and *nirK* genes affiliated with OPU3. I found these genes co-occur in dysoxic-suboxic O_2 conditions throughout open ocean and coastal OMZs (Fig. E.1). The global distribution of these functional marker genes throughout dysoxic-suboxic

water OMZs indicates that OPU3 may represent a significant biological sink for CH_4. However, future efforts focusing on assessing the global functional expression of these genes are required in order to better determine the environmental conditions regulating CH_4 oxidation coupled with NO_2^- reduction in order to better estimate the metabolic efficiency of OPU3 and prevent CH_4 loss to the atmosphere from currently expanding OMZs. Padilla *et al.* (2017) provided first evidence on OPU3 *pmoA* and *nirK* expression under water column oxygen deficiency in the Costa Rica OMZ. As a first approach summing to this effort, I overviewed the distribution, abundance, and functional expression of *pmoA* and *nirK* genes affiliated to OPU3 in the Saanich Inlet water column. I used metagenomic and metatranscriptomic time-series data from June 2009 to August 2011, and observed genes expression overlapped at water column stratification periods peaking at suboxic water column conditions (between the 100-150 m) (Fig. E.2). I encourage future studies to focus on determining OPU3's NO_2^--dependent CH_4 oxidation rates that could be used in mathematical modeling approaches.

Open questions remain to be answered regarding the specific nature and rate of metabolites provided by methanotrophs to the community, as well as the potential benefits that methanotrophs may gain in return. Potential metabolic interactions presented in Chapter 6 provide baseline knowledge that can be used to further investigate the mechanisms of syntrophic interaction in aerobic CH_4 oxidation under low-O_2 conditions. The synthetic community manipulation approach combined with the already existing multi-omic databases is a promising tool to efficiently achieve the experimental validation of these metabolic interactions (reviewed in Chistoserdova and Kalyuzhnaya, 2018). Obtaining new insights into the communal function of CH_4 oxidation would not only provide the necessary knowledge for predicting the activities of methylotrophs in environmental settings but would also enable an effective application of these organisms and their metabolism in industrial processes, i.e. metabolite synthesis or toxin bioremediation. The future of synthetic methylotrophy has been recently discussed and reviewed by Chistoserdova and Kalyuzhnaya (2018) and highlights the promising and emerging potential methanotrophs have as platforms for biotechnological applications.

7.3.1 Incubation experiment limitations

I initially developed the mesocosm experiment based on the theoretical concept of stable isotope probing (SIP). I tried using RNA and protein SIP techniques as these have proven to be high-resolution approaches for identifying active anaerobic and aerobic CH_4-oxidizing bacteria with considerably low levels of substrate incorporation (2% of ^{13}C incorporation) (Manefield *et al*. 2002; McDonald *et al*. 2005; reviewed in Seifert *et al*. 2012). Based on the available literature for SIP experiments and the expected relative small substrate incorporation required, I decided to limit the incubation period to 72 h in order to accurately link substrate incorporation to specific microbial agents. However, the low natural abundance and slow metabolic rates of methanotrophic bacteria in the Saanich Inlet water column did not allow for enough substrate incorporation within the incubation period, making RNA and protein SIP methodological unfeasible. This methodological flaw hindered the quantification of substrate incorporation rates. Based on these limitations, the data in Chapter 6 is restricted to an overview of related assimilation processes. I strongly suggest that future efforts develop time-series incubation experiments with a longer terminal incubation time and the use of

replicates for the different growth conditions. Furthermore, I recommend the use of the multi-omics SIP approach coupled with microscopy techniques, i.e. DNA/RNA-SIP with phylogeny-specific probes (fluoresce *in situ* hybridization; FISH) and secondary ion mass spectrometry (SIMS). The proper development of specific probes for methanotrophic bacteria (OPU3 is currently possible considering the genomic information available) will increase the resolution and accuracy of linking phylogeny and function to specific microorganisms and will provide 1) process rates linked to specific taxa, and 2) further evidence on the nature and composition of methanotrophic consortia. Finally, I strongly suggest future research efforts should focus on determining 1) the OPU3 NO_2^--dependent CH_4 oxidation mechanism in nature and 2) the specific nature and rate of metabolites provided by methanotrophs, i.e. OPU3, to the community. These research efforts will contribute to the better understanding of how microbial communities will likely thrive under ocean deoxygenation conditions and prevent climate-active gases loss to the atmosphere.

7.4 The significance of this research

This thesis encompasses multi-omic time-series observations resolving complex microbial interactions and highlights the use of correlation analysis and co-occurrence network approaches as novel tools to generate baseline knowledge with the aim of developing hypotheses to address environmental questions related to active-trace gas loss to the atmosphere in a time of climate change and ocean deoxygenation. Moreover, this thesis answers the need of the scientific community to develop standard workflows to survey the composition, structure and function of microbial communities that can be scaled to global ocean genomic data surveys. This scientific community effort will allow for better informed global models predicting microbial metabolism responses associated with global ocean deoxygenation.

Finally, this thesis constitutes an integral and comprehensive survey on methanotrophic community strategies that thrive under low-O_2 water-column conditions in a model coastal OMZ. The results presented here on methanotroph interactions, both taxonomic and metabolic, represent fundamental environmental knowledge that can be used in biotechnological and engineering approaches.

References

- Acinas, S.G., Antn, J., and Rodrguez-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, 514-522.
- Akaike, H. (1998). "A New Look at the Statistical Model Identification," in Selected Papers of Hirotugu Akaike, eds. E. Parzen, K. Tanabe G. Kitagawa. (New York, NY: Springer New York), 215-222.
- Alldredge, A.L., and Cohen, Y. (1987). Can Microscale Chemical Patches Persist in the Sea: Microelectrode Study of Marine Snow, Fecal Pellets. Science 235, 689-691.
- Allen, L.Z., Allen, E.E., Badger, J.H., Mccrow, J.P., Paulsen, I.T., Elbourne, L.D.H., Thiagarajan, M., Rusch, D.B., Nealson, K.H., Williamson, S.J., Venter, J.C., and Allen, A.E. (2012). Influence of nutrients and currents on the genomic composition of microbes across an upwelling mosaic. ISME J 6, 1403-1414.
- Allers, E., Wright, J.J., Konwar, K.M., Howes, C.G., Beneze, E., Hallam, S.J., and Sullivan, M.B. (2013). Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean. ISME J 7, 256-268.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. Journal of Molecular Biology 215, 403-410.
- Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biology 11, R106.
- Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166-169.
- Anderson, J.J., and Devol, A.H. (1973). Deep water renewal in Saanich Inlet, an intermittently anoxic basin. Estuarine and Coastal Marine Science 1, 1-10.
- Arstegui, J., Gasol, J.M., Duarte, C.M., and Herndld, G.J. (2009). Microbial oceanography of the dark ocean's pelagic realm. Limnology and Oceanography 54, 1501-1529.

- Armstrong, F.a.J., Stearns, C.R., and Strickland, J.D.H. (1967). The measurement of upwelling and subsequent biological process by means of the Technicon Autoanalyzer and associated equipment. Deep Sea Research and Oceanographic Abstracts 14, 381-389.
- Arrigo, K.R. (2005). Marine microorganisms and global nutrient cycles. Nature 437, 349-355.
- Azam, F., Fenchel, T., Field, J.G., Gray, J.S., Meyer-Reil, L.A., and Thingstad, F. (1983). The ecological role of water-column microbes in the sea. Marine ecology progress series. Oldendorf 10, 257-263.
- Azam, F., and Malfatti, F. (2007). Microbial structuring of marine ecosystems. Nat Rev Micro 5, 966-966.
- Bader, G.D., and Hogue, C.W.V. (2003). An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 4, 2-2.
- Bai, X., Adolf, J.E., Bachvaroff, T., Place, A.R., and Coats, D.W. (2007). The interplay between host toxins and parasitism by Amoebophrya. Harmful Algae 6, 670-678.
- Barbern, A., Bates, S.T., Casamayor, E.O., and Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. The ISME Journal 6, 343-351.
- Barnard, R.L., Osborne, C.A., and Firestone, M.K. (2013). Responses of soil bacterial and fungal communities to extreme desiccation and rewetting. The ISME Journal 7, 2229-2241.
- Beamish, R.J., Neville, C.M., and Cass, A.J. (1997). Production of Fraser River sockeye salmon (Oncorhynchus nerka) in relation to decadal-scale changes in the climate and the ocean. Canadian Journal of Fisheries and Aquatic Sciences 54, 543-554.
- Beck, D.a.C., Kalyuzhnaya, M.G., Malfatti, S., Tringe, Susannah g., Glavina Del Rio, T., Ivanova, N., Lidstrom, Mary e., and Chistoserdova, L. (2013). A metagenomic insight into freshwater methane-utilizing communities and evidence for cooperation between the Methylococcaceae and the Methylophilaceae. PeerJ 1, e23.
- Beman, J.M., and Carolan, M.T. (2013). Deoxygenation alters bacterial diversity and community composition in the ocean's largest oxygen minimum zone. Nature Communications 4, 2705.
- Bertram, D.F., Mackas, D.L., and Mckinnell, S.M. (2001). The seasonal cycle revisited: interannual variation and ecosystem consequences. Progress in Oceanography 49, 283-307.
- Blazewicz, S.J., Barnard, R.L., Daly, R.A., and Firestone, M.K. (2013). Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. The ISME Journal 7, 2061-2068.
- Blumenberg, M., Seifert, R., and Michaelis, W. (2007). Aerobic methanotrophy in the oxic-anoxic transition zone of the Black Sea water column. Organic Geochemistry 38, 84-91.

- Blunden, J., Arndt, D.S., and Baringer, M.O. (2011). State of the Climate in 2010. Bulletin of the American Meteorological Society 92, S1-S236.

- Bograd, S.J., Castro, C.G., Di Lorenzo, E., Palacios, D.M., Bailey, H., Gilly, W., and Chavez, F.P. (2008). Oxygen declines and the shoaling of the hypoxic boundary in the California Current. Geophysical Research Letters 35.

- Bopp, L., Le Qur, C., Heimann, M., Manning, A.C., and Monfray, P. (2002). Climate-induced oceanic oxygen fluxes: Implications for the contemporary carbon budget. Global Biogeochemical Cycles 16, 6-1-6-13.

- Brte, J., Krabberd, A.K., Dolven, J.K., Ose, R.F., Kristensen, T., Bjrklund, K.R., and Shalchian-Tabrizi, K. (2012). Radiolaria Associated with Large Diversity of Marine Alveolates. Protist 163, 767-777.

- Breitburg, D., Levin, L.A., Oschlies, A., Grgoire, M., Chavez, F.P., Conley, D.J., Garon, V., Gilbert, D., Gutirrez, D., Isensee, K., Jacinto, G.S., Limburg, K.E., Montes, I., Naqvi, S.W.A., Pitcher, G.C., Rabalais, N.N., Roman, M.R., Rose, K.A., Seibel, B.A., Telszewski, M., Yasuhara, M., and Zhang, J. (2018). Declining oxygen in the global ocean and coastal waters. Science 359.

- Brown, M.V., and Donachie, S.P. (2007). Evidence for tropical endemicity in the Deltaproteobacteria Marine Group B/SAR324 bacterioplankton clade. Aquatic Microbial Ecology 46, 107-115.

- Brown, M.V., Philip, G.K., Bunge, J.A., Smith, M.C., Bissett, A., Lauro, F.M., Fuhrman, J.A., and Donachie, S.P. (2009). Microbial community structure in the North Pacific ocean. ISME J 3, 1374-1386.

- Brum, J.R., Ignacio-Espinoza, J.C., Roux, S., Doulcier, G., Acinas, S.G., Alberti, A., Chaffron, S., Cruaud, C., De Vargas, C., Gasol, J.M., Gorsky, G., Gregory, A.C., Guidi, L., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Poulos, B.T., Schwenck, S.M., Speich, S., Dimier, C., Kandels-Lewis, S., Picheral, M., Searson, S., Bork, P., Bowler, C., Sunagawa, S., Wincker, P., Karsenti, E., and Sullivan, M.B. (2015). Patterns and ecological drivers of ocean viral communities. Science 348.

- Cachon, J. (1964). Contribution a l'etude des peridiniens parasites. . Cytologie, cycles evolutifs. Annales des Sciences Naturelles -Zoologie et Biologie Animale 6, 1-158.

- Canfield, D.E., Stewart, F.J., Thamdrup, B., De Brabandere, L., Dalsgaard, T., Delong, E.F., Revsbech, N.P., and Ulloa, O. (2010). A cryptic sulfur cycle in oxygen-minimum-zone waters off the Chilean coast. Science 330, 1375-1378.

- Capelle, D., Dacey, J., and Tortell, P.D. (2015). An automated, high-throughput method for accurate and precise measurements of dissolved nitrous-oxide and methane concentrations in natural seawaters. Limnology and Oceanography: Methods.

- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., Mcdonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., and Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7, 335-336.
- Cardoso, J.C.S., Baquero, C., and Almeida, P.S. (Year). "Probabilistic Estimation of Network Size and Diameter", in: Dependable Computing, 2009. LADC '09. Fourth Latin-American Symposium on), 33-40.
- Carpenter, J.H. (1965). The Chesapeake Bay Institute Technique for the Winkler Dissolved Oxygen Method. Limnology and Oceanography 10, 141-143.
- Carter, N.M. (1932). The oceanography of the fjords of southern British Columbia. Fish. Res. Bd. Canada Prog. Rept. Pacific Coast Sta. 12, 7-11.
- Carter, N.M. (1934). Physiography and oceanography of some British Columbia fjords. Proc. Fifth. Pacific Sci. Cong. 1, 721-733.
- Chaffron, S., Rehrauer, H., Pernthaler, J., and Von Mering, C. (2010). A global network of coexisting microbes from environmental and whole-genome sequence data. Genome Research 20, 947-959.
- Chambouvet, A., Alves-De-Souza, C., Cueff, V., Marie, D., Karpov, S., and Guillou, L. (2011). Interplay between the parasite Amoebophrya sp.(Alveolata) and the cyst formation of the red tide dinoflagellate Scrippsiella trochoidea. Protist 162, 637-649.
- Chambouvet, A., Morin, P., Marie, D., and Guillou, L. (2008). Control of toxic marine dinoflagellate blooms by serial parasitic killers. Science 322, 1254-1257.
- Charlson, R.J., Lovelock, J.E., Andreae, M.O., and Warren, S.G. (1987). Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. Nature 326, 655-661.
- Chistoserdova, L., Jenkins, C., Kalyuzhnaya, M.G., Marx, C.J., Lapidus, A., Vorholt, J.A., Staley, J.T., and Lidstrom, M.E. (2004). The Enigmatic Planctomycetes May Hold a Key to the Origins of Methanogenesis and Methylotrophy. Molecular Biology and Evolution 21, 1234-1241.
- Chistoserdova, L., and Kalyuzhnaya, M.G. (2018). Current Trends in Methylotrophy. Trends in Microbiology 26, 703-714.
- Choi, C.J., Brosnahan, M.L., Sehein, T.R., Anderson, D.M., and Erdner, D.L. (2017). Insights into the loss factors of phytoplankton blooms: The role of cell mortality in the decline of two inshore Alexandrium blooms. Limnology and Oceanography 62, 1742-1753.
- Chow, C.-E.T., Kim, D.Y., Sachdeva, R., Caron, D.A., and Fuhrman, J.A. (2014). Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. ISME J 8, 816-829.

- Cline, J.D. (1969). Spectrophotometric determination of hydrogen sulfide in natural waters. Limnology and Oceanography 14, 454-458.
- Coats, D.W., And Heisler, J.J. (1989). Spatial and temporal occurrence of the parasitic dinoflagellate Dubusquella cachoni and its tintinnine host Eutintinnus pectinis in Chesapeake Bay. Marine Biology 101, 401-409.
- Coats, D.W., Bachvaroff, T.R., and Delwiche, C.F. (2012). Revision of the family Duboscquellidae with a description of Euduboscquella crenulata n. gen., n. sp. (Dinoflagellata, Syndinea), an intracellular parasite of the cilliate Favella panamensis Kofoid Campbell, 1929. Journal of Eukaryotic Microbiology 59, 1-11.
- Coats, D.W., Bockstahler, K.R., Berg, G.M., and Sniezek, J.H. (1994). Dinoflagellate infections of Favella panamensis from two North American estuaries. Marine Biology 119, 105-113.
- Codispoti, L.A. (2010). Interesting Times for Marine N2O. Science 327, 1339.
- Costa, C., Dijkema, C., Friedrich, M., Garca-Encina, P., Fernndez-Polanco, F., and Stams, A.J.M. (2000). Denitrification with methane as electron donor in oxygen-limited bioreactors. Applied Microbiology and Biotechnology 53, 754-762.
- Costello, A.M., and Lidstrom, M.E. (1999). Molecular Characterization of Functional and Phylogenetic Genes from Natural Populations of Methanotrophs in Lake Sediments. Applied and Environmental Microbiology 65, 5066-5074.
- Cottrell, M.T., and Kirchman, D.L. (2000). Natural Assemblages of Marine Proteobacteria and Members of the Cytophaga-Flavobacter Cluster Consuming Low- and High-Molecular-Weight Dissolved Organic Matter. Applied and Environmental Microbiology 66, 1692-1697.
- Cram, J.A., Xia, L.C., Needham, D.M., Sachdeva, R., Sun, F., and Fuhrman, J.A. (2015). Cross-depth analysis of marine bacterial networks suggests downward propagation of temporal changes. ISME J 9, 2573-2586.
- Crump, B.C., Armbrust, E.V., and Baross, J.A. (1999). Phylogenetic Analysis of Particle-Attached and Free-Living Bacterial Communities in the Columbia River, Its Estuary, and the Adjacent Coastal Ocean. Applied and Environmental Microbiology 65, 3192-3204.
- Cuba, R.M.F., Duarte, I.C., Saavedra, N.K., Varesche, M.B.A., and Foresti, E. (2011). Denitrification coupled with methane anoxic oxidation and microbial community involved identification. Brazilian Archives of Biology and Technology 54, 173-182.
- Cuevas, L.A., and Morales, C.E. (2006). Nanoheterotroph grazing on bacteria and cyanobacteria in oxic and suboxic waters in coastal upwelling areas off northern Chile. Journal of Plankton Research 28, 385-397.

- D'ambrosio, L., Ziervogel, K., Macgregor, B., Teske, A., and Arnosti, C. (2014). Composition and enzymatic function of particle-associated and free-living bacteria: a coastal/offshore comparison. The ISME Journal 8, 2167-2179.

- De Caceres, M., and Legendre, P. (2009). Associations between species and groups of sites: indices and statistical inference. Ecology. http://sites.google.com/site/miqueldecaceres/.

- De Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., Lara, E., Berney, C., Bescot, N.L., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J.-M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Horak, A., Jaillon, O., Lima-Mendez, G., Lukes, J., Malviya, S., Morard, R., Mulot, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Coordinators, T.O., Acinas, S.G., Bork, P., Bowler, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J., Sieracki, M.E., Speich, S., Stemmann, L., Sunagawa, S., Weissenbach, J., Wincker, P., and Karsenti, E. (2015). Eukaryotic planktonic diversity in the sunlit ocean. Science 348.

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

- Delong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J., Frigaard, N.-U., Martinez, A., Sullivan, M.B., Edwards, R., Brito, B.R., Chisholm, S.W., and Karl, D.M. (2006). Community Genomics Among Stratified Microbial Assemblages in the Ocean 039;s Interior. Science 311, 496.

- Desantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., and Andersen, G.L. (2006). Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. Applied and Environmental Microbiology 72, 5069-5072.

- Diaz, R.J., and Rosenberg, R. (2008). Spreading dead zones and consequences for marine ecosystems. Science 321, 926-929.

- Dick, G.J., Anantharaman, K., Baker, B.J., Li, M., Reed, D.C., and Sheik, C.S. (2013). The microbiology of deep-sea hydrothermal vent plumes: ecological and biogeographic linkages to seafloor and water column habitats. Frontiers in Microbiology 4.

- Dick, G.J., and Tebo, B.M. (2010). Microbial diversity and biogeochemistry of the Guaymas Basin deep-sea hydrothermal plume. Environmental Microbiology 12, 1334-1347.

- Dolan, J.R., Montagnes, D.J.S,, Agathgam S,., Coats, D.W.,, Stoecker, D.K. Eds. (2013). The Biology and Ecology of Tintinnid Ciliates, Models for Maroine Plankton. . Wiley-Blackwell, Hoboken, NJ, USA.

- Dubilier, N., Bergin, C., and Lott, C. (2008). Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. Nat Rev Micro 6, 725-740.

- Dunfield, P.F., Yuryev, A., Senin, P., Smirnova, A.V., Stott, M.B., Hou, S., Ly, B., Saw, J.H., Zhou, Z., Ren, Y., Wang, J., Mountain, B.W., Crowe, M.A., Weatherby, T.M., Bodelier, P.L.E., Liesack, W., Feng, L., Wang, L., and Alam, M. (2007). Methane oxidation by an extremely acidophilic bacterium of the phylum Verrucomicrobia. Nature 450, 879-882.

- Dunne, J.A., Williams, R.J., and Martinez, N.D. (2002). Food-web structure and network theory: The role of connectance and size. Proceedings of the National Academy of Sciences 99, 12917-12922.

- Duret, M.T., Pachiadaki, M.G., Stewart, F.J., Sarode, N., Christaki, U., Monchy, S., Srivastava, A., and Edgcomb, V.P. (2015). Size-fractionated diversity of eukaryotic microbial communities in the Eastern Tropical North Pacific oxygen minimum zone. FEMS Microbiology Ecology 91, fiv037-fiv037.

- Edgcomb, V., Orsi, W., Bunge, J., Jeon, S., Christen, R., Leslin, C., Holder, M., Taylor, G.T., Suarez, P., Varela, R., and Epstein, S. (2011). Protistan microbial observatory in the Cariaco Basin, Caribbean. I. Pyrosequencing vs Sanger insights into species richness. The ISME Journal 5, 1344-1356.

- Edgcomb, V.P. (2016). Marine protist associations and environmental impacts across trophic levels in the twilight zone and below. Current Opinion in Microbiology 31, 169-175.

- Edgcomb, V.P., Taylor, C., Pachiadaki, M.G., Honjo, S., Engstrom, I., and Yakimov, M. (2016). Comparison of Niskin vs. in situ approaches for analysis of gene expression in deep Mediterranean Sea water samples. Deep Sea Research Part II: Topical Studies in Oceanography 129, 213-222.

- Eiler, A., Drakare, S., Bertilsson, S., Pernthaler, J., Peura, S., Rofner, C., Simek, K., Yang, Y., Znachor, P., and Lindstrm, E.S. (2013). Unveiling Distribution Patterns of Freshwater Phytoplankton by a Next Generation Sequencing Based Approach. PLOS ONE 8, e53516.

- Eloe, E.A., Shulse, C.N., Fadrosh, D.W., Williamson, S.J., Allen, E.E., and Bartlett, D.H. (2011). Compositional differences in particle-associated and free-living microbial assemblages from an extreme deep-ocean environment. Environmental Microbiology Reports 3, 449-458.

- Elsaied, H.E., Hayashi, T., and Naganuma, T. (2004). Molecular Analysis of Deep-Sea Hydrothermal Vent Aerobic Methanotrophs by Targeting Genes of 16S rRNA and Particulate Methane Monooxygenase. Marine Biotechnology 6, 503-509.

- Engelbrektson, A., Kunin, V., Wrighton, K.C., Zvenigorodsky, N., Chen, F., Ochman, H., and Hugenholtz, P. (2010). Experimental factors affecting PCR-based estimates of microbial species richness and evenness. ISME J 4, 642-647.

- Escribano, R., Hidalgo, P., and Krautz, C. (2009). Zooplankton associated with the oxygen minimum zone system in the northern upwelling region of Chile during March 2000. Deep Sea Research Part II: Topical Studies in Oceanography 56, 1083-1094.

- Ettwig, K.F., Butler, M.K., Le Paslier, D., Pelletier, E., Mangenot, S., Kuypers, M.M.M., Schreiber, F., Dutilh, B.E., Zedelius, J., De Beer, D., Gloerich, J., Wessels, H.J.C.T., Van Alen, T., Luesken, F.,

Wu, M.L., Van De Pas-Schoonen, K.T., Op Den Camp, H.J.M., Janssen-Megens, E.M., Francoijs, K.-J., Stunnenberg, H., Weissenbach, J., Jetten, M.S.M., and Strous, M. (2010). Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. Nature 464, 543-548.

- Ettwig,K.F., B. Zhu, D. Speth, J.T. Keltjens, M.S. Jetten, B. Kartal. (2016). Archaea catalyze iron-dependent anaerobic oxidation of methane. Proc. Natl. Acad. Sci. U. S. A., 113,12792-12796.

- Faust, K., Lima-Mendez, G., Lerat, J.-S., Sathirapongsasuti, J.F., Knight, R., Huttenhower, C., Lenaerts, T., and Raes, J. (2015). Cross-biome comparison of microbial association networks. Frontiers in Microbiology 6, 1200.

- Faust, K., and Raes, J. (2012). Microbial interactions: from networks to models. Nat Rev Micro 10, 538-550.

- Faust, K., Sathirapongsasuti, J.F., Izard, J., Segata, N., Gevers, D., Raes, J., and Huttenhower, C. (2012). Microbial Co-occurrence Relationships in the Human Microbiome. PLoS Computational Biology 8, e1002606.

- Feike, J., Jurgens, K., Hollibaugh, J.T., Kruger, S., Jost, G., and Labrenz, M. (2012). Measuring unbiased metatranscriptomics in suboxic waters of the central Baltic Sea using a new in situ fixation system. ISME J 6, 461-470.

- Ferrari, S., and Cribari-Neto, F. (2004). Beta Regression for Modelling Rates and Proportions. Journal of Applied Statistics 31, 799-815.

- Field, K.G., Gordon, D., Wright, T., Rapp, M., Urback, E., Vergin, K., and Giovannoni, S.J. (1997). Diversity and depth-specific distribution of SAR11 cluster rRNA genes from marine planktonic bacteria. Applied and Environmental Microbiology 63, 63-70.

- Fox, J., and Weisberg, S. (2011). Multivariate linear models in R. An R Companion to Applied Regression 2nd Ed.

- Frias-Lopez, J., Shi, Y., Tyson, G.W., Coleman, M.L., Schuster, S.C., Chisholm, S.W., and Delong, E.F. (2008). Microbial community gene expression in ocean surface waters. Proceedings of the National Academy of Sciences of the United States of America 105, 3805-3810.

- Friedman, J., and Alm, E.J. (2012). Inferring Correlation Networks from Genomic Survey Data. PLOS Computational Biology 8, e1002687.

- Fu,L., S.W. Li, Z.W. Ding, J. Ding, Y.Z. Lu, R.J. Zeng. (2016). Iron reduction in the DAMO/Shewanella oneidensis MR-1 coculture system and the fate of Fe(II). Water Res., 88,808-815.

- Fuchsman, C.A., Kirkpatrick, J.B., Brazelton, W.J., Murray, J.W., and Staley, J.T. (2011). Metabolic strategies of free-living and aggregate-associated bacterial communities inferred from biologic and chemical profiles in the Black Sea suboxic zone. FEMS Microbiology Ecology 78, 586-603.

- Fuchsman, C.A., Staley, J.T., Oakley, B.B., Kirkpatrick, J.B., and Murray, J.W. (2012). Free-living and aggregate-associated Planctomycetes in the Black Sea. FEMS Microbiology Ecology 80, 402-416.

- Fuenzalida, R., Schneider, W., Garcs-Vargas, J., Bravo, L., and Lange, C. (2009). Vertical and horizontal extension of the oxygen minimum zone in the eastern South Pacific Ocean. Deep Sea Research Part II: Topical Studies in Oceanography 56, 992-1003.

- Fuhrman, J.A., Cram, J.A., and Needham, D.M. (2015). Marine microbial community dynamics and their ecological interpretation. Nature Reviews Microbiology 13, 133.

- Fuhrman, J.A., and Davis, A.A. (1997). Widespread Archaea and novel bacteria from the deep sea as shown by 16S rRNA gene sequences. Oceanographic Literature Review 44, 1025.

- Fukunaga, Y., Kurahashi, M., Sakiyama, Y., Ohuchi, M., Yokota, A., and Harayama, S. (2009). *Phycisphaera mikurensis gen. nov., sp. nov.*, isolated from a marine alga, and proposal of *Phycisphaeraceae fam. nov.*, *Phycisphaerales ord. nov.* and *Phycisphaerae classis nov*. in the phylum *Planctomycetes*. The Journal of General and Applied Microbiology 55, 267-275.

- Fung, I., John, J., Lerner, J., Matthews, E., Prather, M., Steele, L.P., and Fraser, P.J. (1991). Three-dimensional model synthesis of the global methane cycle.

- Gaebler-Schwarz, S., Davidson, A., Assmy, P., Chen, J., Henjes, J., Nthig, E.M., Lunau, M., and Medlin, L.K. (2010). A new cell stage in the haploid-diploid life cycle of the colony forming haptophyte Phaeocystis antartica and its ecological implications. Journal of Phycology 46, 1006-1016.

- Ganesh, S., Bristow, L.A., Larsen, M., Sarode, N., Thamdrup, B., and Stewart, F.J. (2015). Size-fraction partitioning of community gene transcription and nitrogen metabolism in a marine oxygen minimum zone. ISME J 9, 2682-2696.

- Ganesh, S., Parris, D.J., Delong, E.F., and Stewart, F.J. (2014). Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. ISME J 8, 187-211.

- Gardner, W.D. (1977). Incomplete extraction of rapidly settling particles from water samplers1. Limnology and Oceanography 22, 764-768.

- Garfield, P.C., Packard, T.T., Friederich, G.E., and Codispoti, L.A. (1983). A subsurface particle maximum layer and enhanced microbial activity in the secondary nitrite maximum of the northeastern tropical Pacific Ocean. Journal of Marine Research 41, 747-768.

- Gargett, A.E., Stucchi, D., and Whitney, F. (2003). Physical processes associated with high primary production in Saanich Inlet, British Columbia. Estuarine, Coastal and Shelf Science 56, 1141-1156.

- Ghiglione, J.-F., Conan, P., and Pujo-Pay, M. (2009). Diversity of total and active free-living vs. particle-attached bacteria in the euphotic zone of the NW Mediterranean Sea. FEMS Microbiology Letters 299, 9-21.

- Gies, E.A., Konwar, K.M., Beatty, J.T., and Hallam, S.J. (2014). Illuminating Microbial Dark Matter in Meromictic Sakinaw Lake. Applied and Environmental Microbiology 80, 6807-6818.

- Gilbert, J.A., Meyer, F., Schriml, L., Joint, I.R., Mhling, M., and Field, D. (2010). Metagenomes and metatranscriptomes from the L4 long-term coastal monitoring station in the Western English Channel. Standards in Genomic Sciences 3, 183-193.

- Gilbert, J.A., Steele, J.A., Caporaso, J.G., Steinbrck, L., Reeder, J., Temperton, B., Huse, S., Mchardy, A.C., Knight, R., Joint, I., Somerfield, P., Fuhrman, J.A., and Field, D. (2011). Defining seasonal marine microbial community dynamics. The Isme Journal 6, 298.

- Glaubitz, S., Labrenz, M., Jost, G., and Jrgens, K. (2010). Diversity of active chemolithoautotrophic prokaryotes in the sulfidic zone of a Black Sea pelagic redoxcline as determined by rRNA-based stable isotope probing. FEMS Microbiology Ecology 74, 32-41.

- Gmez, F., Moreira, D., and Lpez-Garca, P. (2009). Life cycle and molecular phylogeny of the dinoflagellates Chytriodinium and Dissodinium, ectoparasites of copepod eggs. European journal of protistology 45, 260-270.

- Grossart, H.-P. (2010). Ecological consequences of bacterioplankton lifestyles: changes in concepts are needed. Environmental Microbiology Reports 2, 706-714.

- Guillou, L., Viprey, M., Chambouvet, A., Welsh, R., Kirkham, A., Massana, R., Scanlan, D.J., and Worden, A. (2008). Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). Environmental Microbiology 10, 3349-3365.

- Guindon, S., Lethiec, F., Duroux, P., and Gascuel, O. (2005). PHYML Online: a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Research 33, W557-W559.

- Hahn, A.S., Konwar, K.M., Louca, S., Hanson, N.W., and Hallam, S.J. (2016). The information science of microbial ecology. Curr Opin Microbiol 31, 209-216.

- Hamady, M., Walker, J.J., Harris, J.K., Gold, N.J., and Knight, R. (2008). Error-correcting barcoded primers allow hundreds of samples to be pyrosequenced in multiplex. Nature methods 5, 235-237.

- Hanson, R.S., and Hanson, T.E. (1996). Methanotrophic bacteria. Microbiological Reviews 60, 439-471.

- Hargreaves, N., Ware, D., and Mcfarlane, G. (1994). Return of Pacific sardine (Sardinops sagax) to the British Columbia coast in 1992. Canadian Journal of Fisheries and Aquatic Sciences 51, 460-463.

- Haroon, M.F. (2013). Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage. Nature. 500, 567570

- Hawley, A.K., Brewer, H.M., Norbeck, A.D., Pasa-Tolic, L., and Hallam, S.J. (2014). Metaproteomics reveals differential modes of metabolic coupling among ubiquitous oxygen minimum zone microbes. Proc Natl Acad Sci U S A 111, 11395-11400.

- Hawley, A.K., Nobu, M.K., Wright, J.J., Durno, W.E., Morgan-Lang, C., Sage, B., Schwientek, P., Swan, B.K., Rinke, C., Torres-Beltrán, M., Mewis, K., Liu, W.-T., Stepanauskas, R., Woyke, T., and Hallam, S.J. (2017a). Diverse Marinimicrobia bacteria may mediate coupled biogeochemical cycles along eco-thermodynamic gradients. Nature Communications 8, 1507.

- Hawley, A.K., Torres-Beltrán, M., Zaikova, E., Walsh, D.A., Mueller, A., Scofield, M., Kheirandish, S., Payne, C., Pakhomova, L., Bhatia, M., Shevchuk, O., Gies, E.A., Fairley, D., Malfatti, S.A., Norbeck, A.D., Brewer, H.M., Pasa-Tolic, L., Del Rio, T.G., Suttle, C.A., Tringe, S., and Hallam, S.J. (2017b). A compendium of multi-omic sequence information from the Saanich Inlet water column. Scientific Data 4, 170160.

- Hayashi, T., Obata, H., Gamo, T., Sano, Y., and Naganuma, T. (2007). Distribution and phylogenetic characteristics of the genes encoding enzymes relevant to methane oxidation in oxygen minimum zones of the Eastern Pacific Ocean. Res. J. Environ. Sci. 1, 275-284.

- Heintz, M.B., Mau, S., and Valentine, D.L. (2012). Physical control on methanotrophic potential in waters of the Santa Monica Basin, Southern California. Limnology and Oceanography 57, 420-432.

- Helm, K.P., Bindoff, N.L., and Church, J.A. (2011). Observed decreases in oxygen content of the global ocean. Geophysical Research Letters 38,.

- Herlinveaux, R.H. (1962). Oceanography of Saanich Inlet in Vancouver Island, British Columbia. Journal of the Fisheries Research Board of Canada 19, 1-37.

- Hernandez, M.E., Beck, D.a.C., Lidstrom, M.E., and Chistoserdova, L. (2015). Oxygen availability is a major factor in determining the composition of microbial communities involved in methane oxidation. PeerJ 3, e801.

- Hewson, I., and Fuhrman, J.A. (2006). Improved Strategy for Comparing Microbial Assemblage Fingerprints. Microbial Ecology 51, 147-153.

- Heyer, J., Galchenko, V.F., and Dunfield, P.F. (2002). Molecular phylogeny of type II methane-oxidizing bacteria isolated from various environments. Microbiology 148, 2831-2846.

- Holmes, A.J., Costello, A., Lidstrom, M.E., and Murrell, J.C. (1995). Evidence that participate methane monooxygenase and ammonia monooxygenase may be evolutionarily related. FEMS Microbiology Letters 132, 203-208.

- Holmes, R.M., Aminot, A., Krouel, R., Hooker, B.A., and Peterson, B.J. (1999). A simple and precise method for measuring ammonium in marine and freshwater ecosystems. Can J Fish Aquat Sci 56, 1801-1808.

- Hosack, D.A., Dennis, G., Sherman, B.T., Lane, H.C., and Lempicki, R.A. (2003). Identifying biological themes within lists of genes with EASE. Genome Biology 4, R70.
- Hurwitz, B.L., Deng, L., Poulos, B.T., and Sullivan, M.B. (2013). Evaluation of methods to concentrate and purify ocean virus communities through comparative, replicated metagenomics. Environmental Microbiology 15, 1428-1440.
- Ipcc (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.
- James, T.H., Patricia, S.W., and Michael, C.M. (2000). Similarity of particle-associated and free-living bacterial communities in northern San Francisco Bay, California. Aquatic Microbial Ecology 21, 103-114.
- Jephcott, T.G., Alves-De-Souza, C., Gleason, F.H., Van Ogtrop, F.F., Sime-Ngando, T., Karpov, S.A., and Guillou, L. (2016). Ecological impacts of parasitic chytrids, syndiniales and perkinsids on populations of marine photosynthetic dinoflagellates. Fungal Ecology 19, 47-58.
- Jing, H., Rocke, E., Kong, L., Xia, X., Liu, H., and Landry, M.R. (2015). Protist communities in a marine oxygen minimum zone off Costa Rica by 454 pyrosequencing. Biogeosciences Discuss. 2015, 13483-13509.
- Jnasdttir, S.H., Visser, A.W., Richardson, K., and Heath, M.R. (2015). Seasonal copepod lipid pump promotes carbon sequestration in the deep North Atlantic. Proceedings of the National Academy of Sciences 112, 12122-12126.
- Kalyuzhnaya, M.G., Yang, S., Rozova, O.N., Smalley, N.E., Clubb, J., Lamb, A., Gowda, G.a.N., Raftery, D., Fu, Y., Bringel, F., Vuilleumier, S., Beck, D.a.C., Trotsenko, Y.A., Khmelenina, V.N., and Lidstrom, M.E. (2013). Highly efficient methane biocatalysis revealed in a methanotrophic bacterium. Nature Communications 4, 2785.
- Karl, D.M., Beversdorf, L., Bjorkman, K.M., Church, M.J., Martinez, A., and Delong, E.F. (2008). Aerobic production of methane in the sea. Nature Geosci 1, 473-478.
- Karl, D.M., and Church, M.J. (2014). Microbial oceanography and the Hawaii Ocean Time-series programme. Nat Rev Microbiol 12, 699-713.
- Karl, D.M., and Tilbrook, B.D. (1994). Production and transport of methane in oceanic particulate organic matter. Nature 368, 732-734.
- Karwautz, C., Kus, G., Stckl, M., Neu, T.R., and Lueders, T. (2018). Microbial megacities fueled by methane oxidation in a mineral spring cave. The ISME Journal 12, 87-100.
- Kassambara, A. (2017). ggpubr: 'ggplot2' Based Publication Ready Plots. . R package version 0.1.5.

- Keeling, R.E., Kortzinger, A., and Gruber, N. (2010). Ocean deoxygenation in a warming world. Ann Rev Mar Sci 2, 199-229.
- Keller, D.P., Kriest, I., Koeve, W., and Oschlies, A. (2016). Southern Ocean biological impacts on global ocean oxygen. Geophysical Research Letters 43, 6469-6477.
- Kessler, J.D., Valentine, D.L., Redmond, M.C., Du, M., Chan, E.W., Mendes, S.D., Quiroz, E.W., Villanueva, C.J., Shusta, S.S., Werra, L.M., Yvon-Lewis, S.A., and Weber, T.C. (2011). A Persistent Oxygen Anomaly Reveals the Fate of Spilled Methane in the Deep Gulf of Mexico. Science 331, 312-315.
- Kettle, A., Andreae, M., Amouroux, D., Andreae, T., Bates, T., Berresheim, H., Bingemer, H., Boniforti, R., Curran, M., and Ditullio, G. (1999). A global database of sea surface dimethylsulfide (DMS) measurements and a procedure to predict sea surface DMS as a function of latitude, longitude, and month. Global Biogeochemical Cycles 13, 399-444.
- Kiene, R.P. (1991). Production and consumption of methane in aquatic systems. American Society for Microbiology.
- Kieschnick, R., and Mccullough, B.D. (2003). Regression analysis of variates observed on (0, 1): percentages, proportions and fractions. Statistical Modelling 3, 193-213.
- Kim, S., Gupta, N., and Pevzner, P.A. (2008). Spectral Probabilities and Generating Functions of Tandem Mass Spectra: A Strike against Decoy Databases. Journal of Proteome Research 7, 3354-3363.
- Kimmerer, W., and Mckinnon, A. (1990). High mortality in a copepod population caused by a parasitic dinoflagellate. Marine Biology 107, 449-452.
- Kirkpatrick, J., Oakley, B., Fuchsman, C., Srinivasan, S., Staley, J.T., and Murray, J.W. (2006). Diversity and Distribution of Planctomycetes and Related Bacteria in the Suboxic Zone of the Black Sea. Applied and Environmental Microbiology 72, 3079-3083.
- Kits, K.D., Klotz, M.G., and Stein, L.Y. (2015). Methane oxidation coupled to nitrate reduction under hypoxia by the Gammaproteobacterium Methylomonas denitrificans, sp. nov. type strain FJG1. Environmental Microbiology 17, 3219-3232.
- Knief, C. (2015). Diversity and Habitat Preferences of Cultivated and Uncultivated Aerobic Methanotrophic Bacteria Evaluated Based on pmoA as Molecular Marker. Frontiers in Microbiology 6, 1346.
- Knittel, K., and Boetius, A. (2009). Anaerobic Oxidation of Methane: Progress with an Unknown Process. Annual Review of Microbiology 63, 311-334.
- Kolber, Z., and Falkowski, P.G. (1993). Use of active fluorescence to estimate phytoplankton photosynthesis in situ. Limnology and Oceanography 38, 1646-1665.

- Konwar, K.M., Hanson, N.W., Bhatia, M.P., Kim, D., Wu, S.-J., Hahn, A.S., Morgan-Lang, C., Cheung, H.K., and Hallam, S.J. (2015). MetaPathways v2.5: quantitative functional, taxonomic and usability improvements. Bioinformatics 31, 3345-3347.
- Konwar, K.M., Hanson, N.W., Page, A.P., and Hallam, S.J. (2013). MetaPathways: a modular pipeline for constructing pathway/genome databases from environmental sequence information. BMC Bioinformatics 14, 202.
- Krause, S.M.B., Johnson, T., Samadhi Karunaratne, Y., Fu, Y., Beck, D.a.C., Chistoserdova, L., and Lidstrom, M.E. (2017). Lanthanide-dependent cross-feeding of methane-derived carbon is linked by microbial community interactions. Proceedings of the National Academy of Sciences 114, 358.
- Krzywinski, M., Birol, I., Jones, S.J., and Marra, M.A. (2012). Hive plots rational approach to visualizing networks. Briefings in Bioinformatics 13, 627-644.
- Lam, P., and Kuypers, M.M.M. (2011). Microbial Nitrogen Cycling Processes in Oxygen Minimum Zones. Annual Review of Marine Science 3, 317-345.
- Lam, P., and Kuypers, M.M. (2009). Revising the nitrogen cycle in the Peruvian oxygen minimum zone. Proc Natl Acad Sci U S A 106, 4752-4757.
- Lamontagne, M.G., and Holden, P.A. (2003). Comparison of free-living and particle-associated bacterial communities in a coastal Lagoon. Microbial Ecology 46, 228-237.
- Latora, V., and Marchiori, M. (2009). Is the Boston Subway a Small-World Network
- Lee, B.-S., Bullister, J.L., and Whitney, F.A. (1999). Chlorofluorocarbon CFC-11 and carbon tetrachloride removal in Saanich Inlet, an intermittently anoxic basin. Marine Chemistry 66, 171-185.
- Legendre, P., and Gallagher, E.D. (2001). Ecologically meaningful transformations for ordination of species data. Oecologia 129, 271-280.
- Lepp, P.W., and Schmidt, T.M. (1998). Nucleic acid content of Synechococcus spp. during growth in continuous light and light/dark cycles. Archives of Microbiology 170, 201-207.
- Lesniewski, R.A., Jain, S., Anantharaman, K., Schloss, P.D., and Dick, G.J. (2012). The metatranscriptome of a deep-sea hydrothermal plume is dominated by water column methanotrophs and lithotrophs. The ISME Journal 6, 2257-2268.
- Lewis, M.R., Warnock, R.E., and Platt, T. (1985). Absorption and photosynthetic action spectra for natural phytoplankton populations: Implications for production in the open ocean1. Limnology and Oceanography 30, 794-806.
- Li, M., Wang, J., and Chen, J. (Year). "A Fast Agglomerate Algorithm for Mining Functional Modules in Protein Interaction Networks", in: 2008 International Conference on BioMedical Engineering and Informatics), 3-7.

- Lilley, M.D., Baross, J.A., and Gordon, L.I. (1982). Dissolved hydrogen and methane in Saanich Inlet, British Columbia. Deep Sea Research Part A. Oceanographic Research Papers 29, 1471-1484.

- Lin, X., Scranton, M.I., Varela, R., Chistoserdov, A., and Taylor, G.T. (2007). Compositional responses of bacterial communities to redox gradients and grazing in the anoxic Cariaco Basin. Aquatic Microbial Ecology 47, 57-72.

- Llorens-Mares, T., Yooseph, S., Goll, J., Hoffman, J., Vila-Costa, M., Borrego, C.M., Dupont, C.L., and Casamayor, E.O. (2015). Connecting biodiversity and potential functional role in modern euxinic environments by microbial metagenomics. ISME J 9, 1648-1661.

- Lopez-Garcia, P., Rodriguez-Valera, F., Pedros-Alio, C., and Moreira, D. (2001). Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. Nature 409, 603-607.

- Louca, S., Hawley, A.K., Katsev, S., Torres-Beltrán, M., Bhatia, M.P., Kheirandish, S., Michiels, C.C., Capelle, D., Lavik, G., Doebeli, M., Crowe, S.A., and Hallam, S.J. (2016). Integrating biogeochemistry with multiomic sequence information in a model oxygen minimum zone. Proceedings of the National Academy of Sciences 113, E5925-E5933.

- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T., Steppi, S., Jobb, G., Frster, W., Brettske, I., Gerber, S., Ginhart, A.W., Gross, O., Grumann, S., Hermann, S., Jost, R., Knig, A., Liss, T., Lmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A., and Schleifer, K.H. (2004). ARB: a software environment for sequence data. Nucleic Acids Research 32, 1363-1371.

- Lke, C., Speth, D.R., Kox, M.a.R., Villanueva, L., and Jetten, M.S.M. (2016). Metagenomic analysis of nitrogen and methane cycling in the Arabian Sea oxygen minimum zone. PeerJ 4, e1924.

- Maclean, D., Jones, J.D., and Studholme, D.J. (2009). Application of 'next-generation' sequencing technologies to microbial genetics. Nature Reviews Microbiology 7, 260.

- Marcy, Y., Ouverney, C., Bik, E.M., Lsekann, T., Ivanova, N., Martin, H.G., Szeto, E., Platt, D., Hugenholtz, P., Relman, D.A., and Quake, S.R. (2007). Dissecting biological 'dark matter" with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. Proceedings of the National Academy of Sciences of the United States of America 104, 11889-11894.

- Massana, R., Balagu, V., Guillou, L., and Pedros-Alio, C. (2004a). Picoeukaryotic diversity in an oligotrophic coastal site studied by molecular and culturing approaches. FEMS microbiology ecology 50, 231-243.

- Massana, R., Castresana, J., Balagu, V., Guillou, L., Romari, K., Groisillier, A., Valentin, K., and Pedrs-Ali, C. (2004b). Phylogenetic and ecological analysis of novel marine stramenopiles. Applied and environmental microbiology 70, 3528-3534.

- Massana, R., Del Campo, J., Sieracki, M.E., Audic, S., and Logares, R. (2014). Exploring the uncultured microeukaryote majority in the oceans: reevaluation of ribogroups within stramenopiles. ISME J 8, 854-866.
- Massana, R., Murray, A.E., Preston, C.M., and Delong, E.F. (1997). Vertical distribution and phylogenetic characterization of marine planktonic Archaea in the Santa Barbara Channel. Applied and Environmental Microbiology 63, 50-56.
- Mau, S., Blees, J., Helmke, E., Niemann, H., and Damm, E. (2013). Vertical distribution of methane oxidation and methanotrophic response to elevated methane concentrations in stratified waters of the Arctic fjord Storfjorden (Svalbard, Norway). Biogeosciences 10, 6267-6278.
- Mcdonald, I.R., and Murrell, J.C. (1997). The particulate methane monooxygenase gene pmoA and its use as a functional gene probe for methanotrophs. FEMS Microbiology Letters 156, 205-210.
- Metcalf, W.W., Griffin, B.M., Cicchillo, R.M., Gao, J., Janga, S.C., Cooke, H.A., Circello, B.T., Evans, B.S., Martens-Habbena, W., Stahl, D.A., and Van Der Donk, W.A. (2012). Synthesis of methylphosphonic acid by marine microbes: a source for methane in the aerobic ocean. Science (New York, N.Y.) 337, 1104-1107.
- Milucka.G. Ferdelman, L. Polerecky, D. Franzke, G. Wegener, M. Schmid, et al. (2012).Zero-valent sulphur is a key intermediate in marine methane oxidation Nature, 491,541-546.
- Modin, O., Fukushi, K., and Yamamoto, K. (2007). Denitrification with methane as external carbon source. Water Research 41, 2726-2738.
- Moeseneder, M.M., Winter, C., and Herndl, G.J. (2001). Horizontal and vertical complexity of attached and free-living bacteria of the eastern Mediterranean Sea, determined by 16S rDNA and 16S rRNA fingerprints. Limnology and Oceanography 46, 95-107.
- Montagnes, D.J.S., Dower, J.F., and Figueiredo, G.M. (2010). The Protozooplankton-Ichthyoplankton Trophic Link: An Overlooked Aspect of Aquatic Food Webs1. Journal of Eukaryotic Microbiology 57, 223-228.
- Monteiro, P.M.S., Van Der Plas, A., Mohrholz, V., Mabille, E., Pascall, A., and Joubert, W. (2006). Variability of natural hypoxia and methane in a coastal upwelling system: Oceanic physics or shelf biology. Geophysical Research Letters 33,.
- Moon-Van Der Staay, S.Y., De Wachter, R., and Vaulot, D. (2001). Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. Nature 409, 607-610.
- Mou, X., Sun, S., Edwards, R.A., Hodson, R.E., and Moran, M.A. (2008). Bacterial carbon processing by generalist species in the coastal ocean. Nature 451, 708-711.
- Murphy, J., and Riley, J.P. (1962). A modified single solution method for the determination of phosphate in natural waters. Analytica Chimica Acta 27, 31-36.

- Murray, A.E., Blakis, A., Massana, R., Strawzewski, S., Passow, U., Alldredge, A., and Delong, E.F. (1999). A time series assessment of planktonic archaeal variability in the Santa Barbara Channel. Aquatic Microbial Ecology 20, 129-145.

- Murrell, J.C. (2010). "The Aerobic Methane Oxidizing Bacteria (Methanotrophs)," in Handbook of Hydrocarbon and Lipid Microbiology, ed. K. Timmis. Springer Berlin Heidelberg), 1953-1966.

- McGlynn G.L. Chadwick, C.P. Kempes, V.J. Orphan. (2015).Single cell activity reveals direct electron transfer in methanotrophic consortia. Nature, 526,531-535.

- Naqvi, S.W.A., Bange, H.W., Faras, L., Monteiro, P.M.S., Scranton, M.I., and Zhang, J. (2010). Marine hypoxia/anoxia as a source of CH4 and N2O. Biogeosciences 7, 2159-2190.

- Naqvi, S.W.A., Kumar, M.D., Narvekar, P.V., De Sousa, S.N., George, M.D., and D'silva, C. (1993). An intermediate nepheloid layer associated with high microbial metabolic rates and denitrification in the northwest Indian Ocean. Journal of Geophysical Research: Oceans 98, 16469-16479.

- Not, F., Campo, J.D., Balague, V., Vargas, C.D., and Massana, R. (2009). New insights into the diversity of marine picoeukaryotes. PLoS ONE 4.

- Not, F., Valentin, K., Romari, K., Lovejoy, C., Massana, R., Tbe, K., Vaulot, D., and Medlin, L.K. (2007). Picobiliphytes: a marine picoplanktonic algal group with unknown affinities to other eukaryotes. science 315, 253-255.

- Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., R.Minchin, P., O'hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H., and Wagner, H. (2015). vegan: Community Ecology Package. http://CRAN.R-project.org/package=vegan R package version 2.3-1.

- Orsi, W., Song, Y.C., Hallam, S., and Edgcomb, V. (2012). Effect of oxygen minimum zone formation on communities of marine protists. The ISME Journal 6, 1586-1601.

- Orsi, W.D., Smith, J.M., Wilcox, H.M., Swalwell, J.E., Carini, P., Worden, A.Z., and Santoro, A.E. (2015). Ecophysiology of uncultivated marine euryarchaea is linked to particulate organic matter. ISME J 9, 1747-1763.

- Oshkin, I.Y., Beck, D.a.C., Lamb, A.E., Tchesnokova, V., Benuska, G., Mctaggart, T.L., Kalyuzhnaya, M.G., Dedysh, S.N., Lidstrom, M.E., and Chistoserdova, L. (2015). Methane-fed microbial microcosms show differential community dynamics and pinpoint taxa involved in communal response. The ISME Journal 9, 1119-1129.

- Pack, M.A., Heintz, M.B., Reeburgh, W.S., Trumbore, S.E., Valentine, D.L., Xu, X., and Druffel, E.R.M. (2015). Methane oxidation in the eastern tropical North Pacific Ocean water column. Journal of Geophysical Research: Biogeosciences 120, 1078-1092.

- Padilla, C.C., Bertagnolli, A.D., Bristow, L.A., Sarode, N., Glass, J.B., Thamdrup, B., and Stewart, F.J. (2017). Metagenomic Binning Recovers a Transcriptionally Active Gammaproteobacterium Linking Methanotrophy to Partial Denitrification in an Anoxic Oxygen Minimum Zone. Frontiers in Marine Science 4, 23.

- Padilla, C.C., Bristow, L.A., Sarode, N., Garcia-Robledo, E., Gomez Ramirez, E., Benson, C.R., Bourbonnais, A., Altabet, M.A., Girguis, P.R., Thamdrup, B., and Stewart, F.J. (2016). NC10 bacteria in marine oxygen minimum zones. ISME J.

- Padilla, C.C., Ganesh, S., Gantt, S., Huhman, A., Parris, D.J., Sarode, N., and Stewart, F.J. (2015). Standard filtration practices may significantly distort planktonic microbial diversity estimates. Frontiers in Microbiology 6, 547.

- Park, S., Brown, K.W., and Thomas, J.C. (2002). The effect of various environmental and design parameters on methane oxidation in a model biofilter. Waste Management Research 20, 434-444.

- Parris, D.J., Ganesh, S., Edgcomb, V.P., Delong, E.F., and Stewart, F.J. (2014). Microbial eukaryote diversity in the marine oxygen minimum zone off northern Chile. Frontiers in Microbiology 5, 543.

- Paulmier, A., and Ruiz-Pino, D. (2009). Oxygen minimum zones (OMZs) in the modern ocean. Progress in Oceanography 80, 113-128.

- Pedrs-Ali, C., Caldern-Paz, J.I., and Gasol, J.M. (2000). Comparative analysis shows that bacterivory, not viral lysis, controls the abundance of heterotrophic prokaryotic plankton. FEMS Microbiology Ecology 32, 157-165.

- Perez, S. (2015). Exploring microbial community structure and resilience through visualization and analysis of microbial co-occurence networks. The University of British Columbia.

- Pesant, S., Not, F., Picheral, M., Kandels-Lewis, S., Le Bescot, N., Gorsky, G., Iudicone, D., Karsenti, E., Speich, S., Troubl, R., Dimier, C., and Searson, S. (2015). Open science resources for the discovery and analysis of Tara Oceans data. Scientific Data 2, 150023.

- Petersen, J.M., and Dubilier, N. (2009). Methanotrophic symbioses in marine invertebrates. Environmental Microbiology Reports 1, 319-335.

- Peura, S., Bertilsson, S., Jones, R.I., and Eiler, A. (2015). Resistant Microbial Cooccurrence Patterns Inferred by Network Topology. Applied and Environmental Microbiology 81, 2090.

- Proulx, S.R., Promislow, D.E.L., and Phillips, P.C. (2005). Network thinking in ecology and evolution. Trends in Ecology Evolution 20, 345-353.

- Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., and Glckner, F.O. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Research 35, 7188-7196.

- Rapp, M.S., and Giovannoni, S.J. (2003). The Uncultured Microbial Majority. Annual Review of Microbiology 57, 369-394.

- Rcoreteam (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

- Reeburgh, W.S. (2007). Oceanic Methane Biogeochemistry. Chemical Reviews 107, 486-513.

- Reeburgh, W.S., Ward, B.B., Whalen, S.C., Sandbeck, K.A., Kilpatrick, K.A., and Kerkhof, L.J. (1991). Black Sea methane geochemistry. Deep Sea Research Part A. Oceanographic Research Papers 38, Supplement 2, S1189-S1210.

- Richards, T.A., and Bass, D. (2005). Molecular screening of free-living microbial eukaryotes: diversity and distribution using a meta-analysis. Current Opinion in Microbiology 8, 240-252.

- Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.-F., Darling, A., Malfatti, S., Swan, B.K., Gies, E.A., Dodsworth, J.A., Hedlund, B.P., Tsiamis, G., Sievert, S.M., Liu, W.-T., Eisen, J.A., Hallam, S.J., Kyrpides, N.C., Stepanauskas, R., Rubin, E.M., Hugenholtz, P., and Woyke, T. (2013). Insights into the phylogeny and coding potential of microbial dark matter. Nature 499, 431-437.

- Rodriguez-Mora, M.J., Scranton, M.I., Taylor, G.T., and Chistoserdov, A.Y. (2013). Bacterial community composition in a large marine anoxic basin: a Cariaco Basin time-series survey. FEMS Microbiology Ecology 84, 625-639.

- Romari, K., and Vaulot, D. (2004). Composition and temporal variability of picoeukaryote communities at a coastal site of the English Channel from 18S rDNA sequences. Limnology and Oceanography 49, 784-798.

- Roslev, P., and King, G.M. (1995). Aerobic and Anaerobic Starvation Metabolism in Methanotrophic Bacteria. Applied and Environmental Microbiology 61, 1563-1570.

- Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S., Wu, D., Eisen, J.A., Hoffman, J.M., Remington, K., Beeson, K., Tran, B., Smith, H., Baden-Tillson, H., Stewart, C., Thorpe, J., Freeman, J., Andrews-Pfannkoch, C., Venter, J.E., Li, K., Kravitz, S., Heidelberg, J.F., Utterback, T., Rogers, Y.-H., Falcn, L.I., Souza, V., Bonilla-Rosso, G., Eguiarte, L.E., Karl, D.M., Sathyendranath, S., Platt, T., Bermingham, E., Gallardo, V., Tamayo-Castillo, G., Ferrari, M.R., Strausberg, R.L., Nealson, K., Friedman, R., Frazier, M., and Venter, J.C. (2007). The Sorcerer II Global Ocean Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific. PLoS Biology 5, e77.

- Salcher, M.M., Posch, T., and Pernthaler, J. (2013). In situ substrate preferences of abundant bacterioplankton populations in a prealpine freshwater lake. The ISME Journal 7, 896-907.

- Sansone, F.J., Popp, B.N., Gasc, A., Graham, A.W., and Rust, T.M. (2001). Highly elevated methane in the eastern tropical North Pacific and associated isotopically enriched fluxes to the atmosphere. Geophysical Research Letters 28, 4567-4570.

- Santelli, C.M., Orcutt, B.N., Banning, E., Bach, W., Moyer, C.L., Sogin, M.L., Staudigel, H., and Edwards, K.J. (2008). Abundance and diversity of microbial life in ocean crust. Nature 453, 653-656.

- Sauter, L.M., Latypova, E., Smalley, N.E., Lidstrom, M.E., Hallam, S., and Kalyuzhnaya, M.G. (2012). Methanotrophic communities of Saanich Inlet: A microcosm perspective. Systematic and Applied Microbiology 35, 198-203.

- Scheller,S., H. Yu, G.L. Chadwick, S.E. McGlynn, V.J. Orphan. (2016). Artificial electron acceptors decouple archaeal methane oxidation from sulfate reduction Science, 351,703-707.

- Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., and Weber, C.F. (2009). Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Applied and Environmental Microbiology 75, 7537-7541.

- Schmidtko, S., Stramma, L., and Visbeck, M. (2017). Decline in global oceanic oxygen content during the past five decades. Nature 542, 335-339.

- Semrau, J.D., Dispirito, A.A., and Yoon, S. (2010). Methanotrophs and copper. FEMS Microbiology Reviews 34, 496-531.

- Shanks, A.L., and Reeder, M.L. (1993). Reducing microzones and sulfide production in marine snow. Mar.Ecol.Prog.Ser 96, 43-47.

- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Research 13, 2498-2504.

- Shi, Y., Tyson, G.W., and Delong, E.F. (2009). Metatranscriptomics reveals unique microbial small RNAs in the ocean/'s water column. Nature 459, 266-269.

- Simon, M., Grossart, H.-P., Schweitzer, B., and Ploug, H. (2002). Microbial ecology of organic aggregates in aquatic ecosystems. Aquatic Microbial Ecology 28, 175-211.

- Skennerton, C.T. (2017) Methane-fueled syntrophy through extracellular electron transfer:uncovering the genomic traits conserved within diverse bacterial partners of anaerobic methanotrophic Archaea. mBio 8, e00530-17.

- Skovgaard, A., Massana, R., Balague, V., and Saiz, E. (2005). Phylogenetic position of the copepod-infesting parasite Syndinium turbo (Dinoflagellata, Syndinea). Protist 156, 413-423.

- Smith, D.C., Simon, M., Alldredge, A.L., and Azam, F. (1992). Intense hydrolytic enzyme activity on marine aggregates and implications for rapid particle dissolution. Nature 359, 139-142.

- Smith, M.W., Zeigler Allen, L., Allen, A.E., Herfort, L., and Simon, H.M. (2013). Contrasting genomic properties of free-living and particle-attached microbial assemblages within a coastal ecosystem. Frontiers in Microbiology 4, 120.

- Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., Arrieta, J.M., and Herndl, G.J. (2006). Microbial diversity in the deep sea and the underexplored 'rare biosphere'. Proceedings of the National Academy of Sciences 103, 12115-12120.

- Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., and Miller, H.L. (2007). Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

- Steele, J.A., Countway, P.D., Xia, L., Vigil, P.D., Beman, J.M., Kim, D.Y., Chow, C.-E.T., Sachdeva, R., Jones, A.C., Schwalbach, M.S., Rose, J.M., Hewson, I., Patel, A., Sun, F., Caron, D.A., and Fuhrman, J.A. (2011). Marine bacterial, archaeal and protistan association networks reveal ecological linkages. The Isme Journal 5, 1414.

- Stein, J.L., Marsh, T.L., Wu, K.Y., Shizuya, H., and Delong, E.F. (1996). Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. Journal of Bacteriology 178, 591-599.

- Stein, L.Y., and Klotz, M.G. (2011). Nitrifying and denitrifying pathways of methanotrophic bacteria. Biochemical Society Transactions 39, 1826.

- Steinberg, D.K., Carlson, C.A., Bates, N.R., Johnson, R.J., Michaels, A.F., and Knap, A.H. (2001). Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-scale look at ocean biology and biogeochemistry. Deep Sea Research Part II: Topical Studies in Oceanography 48, 1405-1447.

- Stevens, H., and Ulloa, O. (2008). Bacterial diversity in the oxygen minimum zone of the eastern tropical South Pacific. Environmental Microbiology 10, 1244-1259.

- Stewart, F.J., Dalsgaard, T., Young, C.R., Thamdrup, B., Revsbech, N.P., Ulloa, O., Canfield, D.E., and Delong, E.F. (2012a). Experimental Incubations Elicit Profound Changes in Community Transcription in OMZ Bacterioplankton. PLOS ONE 7, e37118.

- Stewart, F.J., Ottesen, E.A., and Delong, E.F. (2010). Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. ISME J 4, 896-907.

- Stewart, F.J., Ulloa, O., and Delong, E.F. (2012b). Microbial metatranscriptomics in a permanent marine oxygen minimum zone. Environmental Microbiology 14, 23-40.

- Stilwell, C.P. (2007). Microbial community dynamics in Saanich Inlet: Insights into population structure and methane cycling in a seasonally anoxic fjord. Thesis. Master of Science. The University of British Columbia.

- Stramma, L., Johnson, G.C., Sprintall, J., and Mohrholz, V. (2008). Expanding Oxygen-Minimum Zones in the Tropical Oceans. Science 320, 655.

- Stramma, L., Prince, E.D., Schmidtko, S., Luo, J., Hoolihan, J.P., Visbeck, M., Wallace, D.W.R., Brandt, P., and Krtzinger, A. (2011). Expansion of oxygen minimum zones may reduce available habitat for tropical pelagic fishes. Nature Climate Change 2, 33.

- Stramma, L., Schmidtko, S., Levin, L.A., and Johnson, G.C. (2010). Ocean oxygen minima expansions and their biological impacts. Deep Sea Research Part I: Oceanographic Research Papers 57, 587-595.

- Streams, E.M., Fisher, R.C., and Fiala-Mdioni, A. (1997). Methanotrophic symbiont location and fate of carbon incorporated from methane in a hydrocarbon seep mussel. Marine Biology 129, 465-476.

- Strong, P.J., Xie, S., and Clarke, W.P. (2015). Methane as a Resource: Can the Methanotrophs Add Value. Environmental Science Technology 49, 4001-4018.

- Strous, M., and Jetten, M.S.M. (2004). Anaerobic Oxidation of Methane and Ammonium. Annual Review of Microbiology 58, 99-117.

- Strous, M., Pelletier, E., Mangenot, S., Rattei, T., Lehner, A., Taylor, M.W., Horn, M., Daims, H., Bartol-Mavel, D., Wincker, P., Barbe, V., Fonknechten, N., Vallenet, D., Segurens, B., Schenowitz-Truong, C., Mdigue, C., Collingro, A., Snel, B., Dutilh, B.E., Op Den Camp, H.J.M., Van Der Drift, C., Cirpus, I., Van De Pas-Schoonen, K.T., Harhangi, H.R., Van Niftrik, L., Schmid, M., Keltjens, J., Van De Vossenberg, J., Kartal, B., Meier, H., Frishman, D., Huynen, M.A., Mewes, H.-W., Weissenbach, J., Jetten, M.S.M., Wagner, M., and Le Paslier, D. (2006). Deciphering the evolution and metabolism of an anammox bacterium from a community genome. Nature 440, 790.

- Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G., Mende, D.R., Alberti, A., Cornejo-Castillo, F.M., Costea, P.I., Cruaud, C., D'ovidio, F., Engelen, S., Ferrera, I., Gasol, J.M., Guidi, L., Hildebrand, F., Kokoszka, F., Lepoivre, C., Lima-Mendez, G., Poulain, J., Poulos, B.T., Royo-Llonch, M., Sarmento, H., Vieira-Silva, S., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Tara Oceans, C., Bowler, C., De Vargas, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Jaillon, O., Not, F., Ogata, H., Pesant, S., Speich, S., Stemmann, L., Sullivan, M.B., Weissenbach, J., Wincker, P., Karsenti, E., Raes, J., Acinas, S.G., and Bork, P. (2015). Structure and function of the global ocean microbiome. Science 348.

- Suter, E.A., Scranton, M.I., Chow, S., Stinton, D., Medina Faull, L., and Taylor, G.T. (2016). Niskin bottle sample collection aliases microbial community composition and biogeochemical interpretation. Limnology and Oceanography,.

- Suttle, C.A. (2005). Viruses in the sea. Nature 437, 356-361.
- Suzuki, R., and Shimodaira, H. (2015). pvclust: Hierarchical Clustering with P-Values via Multiscale Bootstrap Resampling. R package version 2.0-0. http://CRAN.R-project.org/package=pvclust.
- Swan, B.K., Martinez-Garcia, M., Preston, C.M., Sczyrba, A., Woyke, T., Lamy, D., Reinthaler, T., Poulton, N.J., Masland, E.D.P., Gomez, M.L., Sieracki, M.E., Delong, E.F., Herndl, G.J., and Stepanauskas, R. (2011). Potential for Chemolithoautotrophy Among Ubiquitous Bacteria Lineages in the Dark Ocean. Science 333, 1296-1300.
- Takeda, K., and Tanaka, K. (1980). Ultrastructure of intracytoplasmic membranes of Methanomonas margaritae cells grown under different conditions. Antonie van Leeuwenhoek 46, 15-25.
- Tavormina, P.L., Ussler, W., Joye, S.B., Harrison, B.K., and Orphan, V.J. (2010). Distributions of putative aerobic methanotrophs in diverse pelagic marine environments. ISME J 4, 700-710.
- Tavormina, P.L., Ussler, W., Steele, J.A., Connon, S.A., Klotz, M.G., and Orphan, V.J. (2013). Abundance and distribution of diverse membrane-bound monooxygenase (Cu-MMO) genes within the Costa Rica oxygen minimum zone. Environmental Microbiology Reports 5, 414-423.
- Taylor, C.D., and Doherty, K.W. (1990). Submersible Incubation Device (SID), autonomous instrumentation for the in situ measurement of primary production and other microbial rate processes. Deep Sea Research Part A. Oceanographic Research Papers 37, 343-358.
- Taylor, C.D., Edgcomb, V.P., Doherty, K.W., Engstrom, I., Shanahan, T., Pachiadaki, M.G., Molyneaux, S.J., and Honjo, S. (2015). Fixation filter, device for the rapid in situ preservation of particulate samples. Deep Sea Research Part I: Oceanographic Research Papers 96, 69-79.
- Taylor Gt, K.D., Pace Ml (1986). Impact of bacteria and zooflagellates on the composition of sinking particles: An in situ experiment. . Mar. Ecol. Prog. Ser. 29, 141-155.
- Thamdrup, B., Dalsgaard, T., and Revsbech, N.P. (2012). Widespread functional anoxia in the oxygen minimum zone of the Eastern South Pacific. Deep Sea Research Part I: Oceanographic Research Papers 65, 36-45.
- Tol, R.S.J., Heintz, R.J., and Lammers, P.E.M. (2003). Methane Emission Reduction: An Application of FUND. Climatic Change 57, 71-98.
- Torres-Beltrán, M., Hawley, A., Capelle, D., Bhatia, M., Durno, E., Tortell, P., and Hallam, S. (2016). Methanotrophic community dynamics in a seasonally anoxic fjord: Saanich Inlet, British Columbia. Frontiers in Marine Science 3, 268.
- Torres-Beltrán, M., Hawley, A.K., Capelle, D., Zaikova, E., Walsh, D.A., Mueller, A., Scofield, M., Payne, C., Pakhomova, L., Kheirandish, S., Finke, J., Bhatia, M., Shevchuk, O., Gies, E.A., Fairley, D., Michiels, C., Suttle, C.A., Whitney, F., Crowe, S.A., Tortell, P.D., and Hallam, S.J. (2017). A

compendium of geochemical information from the Saanich Inlet water column. Scientific Data 4, 170159.

- Tripp, H.J., Kitner, J.B., Schwalbach, M.S., Dacey, J.W.H., Wilhelm, L.J., and Giovannoni, S.J. (2008). SAR11 marine bacteria require exogenous reduced sulphur for growth. Nature 452, 741-744.

- Trosvik, P., De Muinck, E.J., and Stenseth, N.C. (2015). Biotic interactions and temporal dynamics of the human gastrointestinal microbiota. The ISME Journal 9, 533-541.

- Tsementzi, D., Wu, J., Deutsch, S., Nath, S., Rodriguez-R, L.M., Burns, A.S., Ranjan, P., Sarode, N., Malmstrom, R.R., Padilla, C.C., Stone, B.K., Bristow, L.A., Larsen, M., Glass, J.B., Thamdrup, B., Woyke, T., Konstantinidis, K.T., and Stewart, F.J. (2016). SAR11 bacteria linked to ocean anoxia and nitrogen loss. Nature 536, 179-183.

- Ulloa, O., Canfield, D.E., Delong, E.F., Letelier, R.M., and Stewart, F.J. (2012). Microbial oceanography of anoxic oxygen minimum zones. Proceedings of the National Academy of Sciences 109, 15996-16003.

- Ulloa, O., and Pantoja, S. (2009). The oxygen minimum zone of the eastern South Pacific. Deep Sea Research Part II: Topical Studies in Oceanography 56, 987-991.

- Van Den Bergh, M.R., Biesbroek, G., Rossen, J.W.A., De Steenhuijsen Piters, W.a.A., Bosch, A.a.T.M., Van Gils, E.J.M., Wang, X., Boonacker, C.W.B., Veenhoven, R.H., Bruin, J.P., Bogaert, D., and Sanders, E.a.M. (2012). Associations between Pathogens in the Upper Respiratory Tract of Young Children: Interplay between Viruses and Bacteria. PLOS ONE 7, e47711.

- Vaquer-Sunyer, R., and Duarte, C.M. (2008). Thresholds of hypoxia for marine biodiversity. Proceedings of the National Academy of Sciences 105, 15452.

- Vzquez-Domnguez, E., Vaqu, D., and Gasol, J.M. (2007). Ocean warming enhances respiration and carbon demand of coastal microbial plankton. Global Change Biology 13, 1327-1334.

- Velo-Surez, L., Brosnahan, M.L., Anderson, D.M., and Mcgillicuddy Jr, D.J. (2013). A quantitative assessment of the role of the parasite Amoebophrya in the termination of Alexandrium fundyense blooms within a small coastal embayment. PloS one 8, e81150.

- Venables, W.N., and Ripley, B.D. (2002). Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0.

- Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D., Paulsen, I., Nelson, K.E., Nelson, W., Fouts, D.E., Levy, S., Knap, A.H., Lomas, M.W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannkoch, C., Rogers, Y.-H., and Smith, H.O. (2004). Environmental Genome Shotgun Sequencing of the Sargasso Sea. Science 304, 66.

- Waidner, L.A., and Kirchman, D.L. (2007). Aerobic Anoxygenic Phototrophic Bacteria Attached to Particles in Turbid Waters of the Delaware and Chesapeake Estuaries. Applied and Environmental Microbiology 73, 3936-3944.

- Walsh, D.A., and Hallam, S.J. (2011). "Bacterial Community Strucuture and Dynamics ina a Seasonally Anoxic Fjord: Saanich Inlet, British Columbia," in Handbook of Molecular Microbial Ecology II: Metagenomics in Different Habitats, ed. F.J. De Brujin. Wiley-Blackwell), 253-267.

- Walsh, D.A., Zaikova, E., Howes, C.G., Song, Y.C., Wright, J.J., Tringe, S.G., Tortell, P.D., and Hallam, S.J. (2009). Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones. Science 326, 578-582.

- Wang, J., Zhong, J., Chen, G., Li, M., Wu, F.X., and Pan, Y. (2015). ClusterViz: A Cytoscape APP for Cluster Analysis of Biological Network. IEEE/ACM Transactions on Computational Biology and Bioinformatics 12, 815-822.

- Ward, B.B., Devol, A.H., Rich, J.J., Chang, B.X., Bulow, S.E., Naik, H., Pratihary, A., and Jayakumar, A. (2009). Denitrification as the dominant nitrogen loss process in the Arabian Sea. Nature 461, 78-81.

- Ward, B.B., and Kilpatrick, K.A. (1990). Relationship between substrate concentration and oxidation of ammonium and methane in a stratified water column. Continental Shelf Research 10, 1193-1208.

- Ward, B.B., and Kilpatrick, K.A. (1993). Methane oxidation associated with mid-depth methane maxima in the Southern California Bight. Continental Shelf Research 13, 1111-1122.

- Ward, B.B., Kilpatrick, K.A., Wopat, A.E., Minnich, E.C., and Lidstrom, M.E. (1989). Methane oxidation in Saanich inlet during summer stratification. Continental Shelf Research 9, 65-75.

- Wegener,G., V. Krukenberg, D. Riedel, H.E. Tegetmeyer, A. Boetius. (2015). Intercellular wiring enables electron transfer between methanotrophic archaea and bacteria Nature, 526,587-590.

- Weiss, R.F., and Price, B.A. (1980). Nitrous oxide solubility in water and seawater. Marine Chemistry 8, 347-359.

- Whitmire, A.L., Letelier, R.M., Villagrn, V., and Ulloa, O. (2009). Autonomous observations of in vivo fluorescence and particle backscattering in an oceanic oxygen minimum zone. Optics Express 17, 21992-22004.

- Whitney, F.A., Freeland, H.J., and Robert, M. (2007). Persistently declining oxygen levels in the interior waters of the eastern subarctic Pacific. Progress in Oceanography 75, 179-199.

- Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York http://ggplot2.org.

- Wickham, H., Francois, R., and (2015). dplyr: A Grammar of Data Manipulation. http://CRAN.R-project.org/package=dply R package version 0.4.3.

- Widder, S., Besemer, K., Singer, G.A., Ceola, S., Bertuzzo, E., Quince, C., Sloan, W.T., Rinaldo, A., and Battin, T.J. (2014). Fluvial network organization imprints on microbial co-occurrence networks. Proceedings of the National Academy of Sciences 111, 12799.

- Wiesenburg, D.A., and Guinasso, N.L. (1979). Equilibrium solubilities of methane, carbon monoxide, and hydrogen in water and sea water. Journal of Chemical Engineering Data 24, 356-360.

- Winkler, L.W. (1888). Die Bestimmung des im Wasser gelsten Sauerstoffes. Berichte der deutschen chemischen Gesellschaft 21, 2843-2854.

- Wishner, K.F., Gowing, M.M., and Gelfman, C. (1998). Mesozooplankton biomass in the upper 1000m in the Arabian Sea: overall seasonal and geographic patterns, and relationship to oxygen gradients. Deep Sea Research Part II: Topical Studies in Oceanography 45, 2405-2432.

- Wishner, K.F., Outram, D.M., Seibel, B.A., Daly, K.L., and Williams, R.L. (2013). Zooplankton in the eastern tropical north Pacific: Boundary effects of oxygen minimum zone expansion. Deep Sea Research Part I: Oceanographic Research Papers 79, 122-140.

- Woebken, D., Fuchs, B.M., Kuypers, M.M.M., and Amann, R. (2007). Potential Interactions of Particle-Associated Anammox Bacteria with Bacterial and Archaeal Partners in the Namibian Upwelling System. Applied and Environmental Microbiology 73, 4648-4657.

- Worm, B., Sandow, M., Oschlies, A., Lotze, H.K., and Myers, R.A. (2005). Global Patterns of Predator Diversity in the Open Oceans. Science 309, 1365.

- Wright, J.J., Konwar, K.M., and Hallam, S.J. (2012). Microbial ecology of expanding oxygen minimum zones. Nat Rev Microbiol 10, 381-394.

- Wright, J.J., Lee, S., Zaikova, E., Walsh, D.A., and Hallam, S.J. (2009). DNA Extraction from 0.22 mu;M Sterivex Filters and Cesium Chloride Density Gradient Centrifugation. e1352.

- Wrighton, K.C., Thomas, B.C., Sharon, I., Miller, C.S., Castelle, C.J., Verberkmoes, N.C., Wilkins, M.J., Hettich, R.L., Lipton, M.S., Williams, K.H., Long, P.E., and Banfield, J.F. (2012). Fermentation, Hydrogen, and Sulfur Metabolism in Multiple Uncultivated Bacterial Phyla. Science 337, 1661.

- Yu, Z., and Chistoserdova, L. (2017). Communal Metabolism of Methane and the Rare Earth Element Switch. Journal of Bacteriology 199, e00328-00317.

- Zaikova, E., Hawley, A., Walsh, D.A., and Hallam, S.J. (2009). Seawater Sampling and Collection. e1159.

- Zaikova, E., Walsh, D.A., Stilwell, C.P., Mohn, W.W., Tortell, P.D., and Hallam, S.J. (2010). Microbial community dynamics in a seasonally anoxic fjord: Saanich Inlet, British Columbia. Environ Microbiol 12, 172-191.
- Zeileis, A., Kleiber, C., and Jackman, S. (2008). Regression Models for Count Data in R. Journal of Statistical Software 27.
- Zhang, X., Hester, K.C., Ussler, W., Walz, P.M., Peltzer, E.T., and Brewer, P.G. (2011). In situ Raman-based measurements of high dissolved methane concentrations in hydrate-rich ocean sediments. Geophysical Research Letters 38,.

Appendix A

Sampling and processing methods impact microbial community structure and function

A.1 Supplementary results

A.1.1 Saanich Inlet microbial community monitoring time-series

Pyrotag data has been used to describe Saanich Inlet microbial community composition over time and throughout water column O_2 gradients. Results derived from analyzing 16S rDNA pyrotag data from samples taken over 2006 to 2011 revealed consistent microbial community partitioning into discrete clusters associated with an oxygen gradient throughout the oxic, dysoxic/suboxic and anoxic water column conditions (Supplementary Fig 1). This pattern is commonly observed in stratified ecosystems where O_2 deficiency is associated with redox-driven niche partitioning among and between microorganisms(Alldredge and Cohen, 1987; Shanks and Reeder, 1993; Wright *et al*., 2012). In addition, microbial community structure analysis have revealed among the abundant biosphere (relative abundance $>$1%)(Rapp and Giovannoni, 2003), operational taxonomic units (OTUs) affiliated with SAR11, SAR324, Nitrospina, SUP05, and Marine Group A within the bacterial domain, and Thaumarchaeota within the archaeal domain dominated most compositional profiles (Zaikova *et al*., 2010; Walsh and Hallam, 2011). Because of their abundance and distinctive metabolism these taxa have been previously described in marine O_2- deficient environments as community key groups (Field *et al*., 1997; Fuhrman and Davis, 1997; Rapp and Giovannoni, 2003; Brown and Donachie, 2007; Tripp *et al*., 2008; Lam *et al*., 2009; Walsh *et al*., 2009; Zaikova *et al*., 2010; Walsh and Hallam, 2011).

A.1.2 Filtering conditions effect on microbial community diversity

To investigate the effect of filtering conditions, both filter type and volume, on community's diversity, I evaluated the alpha and Shannon indices for 29 SSU rDNA pyrotag samples from 165 and 185m. Diversity

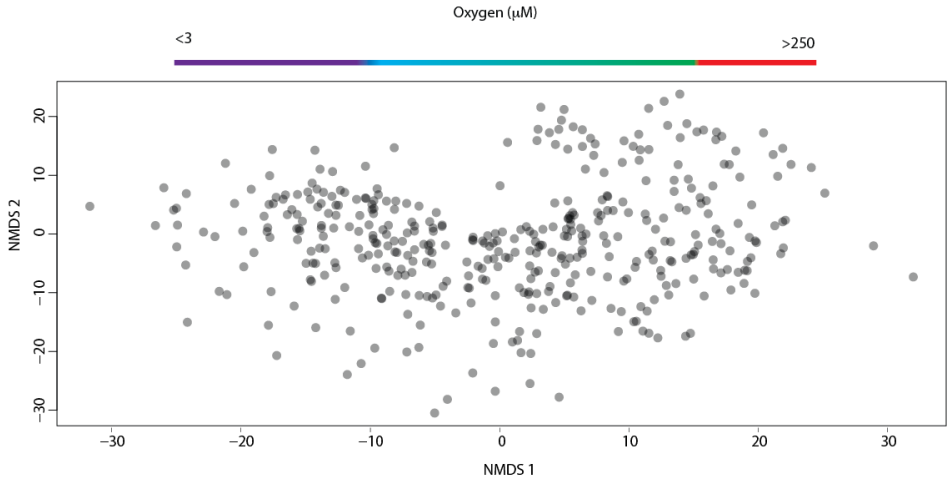

Figure A.1: Non-metric multidimensional scaling plot for Saanich Inlet Time-Series SSU rDNA pyrotag samples (2006 -2011) based on Manhattan distance (1000 iterations). Oxygen gradient from oxic (>250 μM; red) to anoxic (<3 μM; purple) is shown on top and pyrotag samples are depicted as gray dots.

Table A.1: Water column chemical properties during SCOR workshop on July14th, 2014. Oxygen (O_2), Phosphate (PO_4^{-3}), Silicic acid (SiO_2), Nitrate (NO_3^-), Nitrite (NO_2^-), Ammonium (NH_4^+) and Hydrogen Sulfide (H_2S) concentrations (μM) for 150, 165 and 185 m depth intervals. **Cadmium column closed

Depth (m)	O_2 (μM)	PO^{-3}_4 (μM)	SiO_2 (μM)	NO^-_3 (μM)	NO^-_2 (μM)
150	3.803	4.576	86.007	11.952	0.639
165	2.355	5.116	90.789	2.311	0
185	1.923	5.89	108.296	NaN**	0

values changed as function of filtering timing (*in situ* vs. on ship) and filter size. Overall 0.4 μm *in situ* samples showed higher alpha diversity values (>150) than those collected from bottles and filtered on ship, regardless the depth. However, Shannon diversity values were evenly distributed among samples, ranging from 4 to 5, for 0.4 and 2.7 μm pre-filtered samples regardless the depth and filtering timing (Supplementary Information Fig 2). Filtered water volume showed to have a different effect on pre-filtered 0.22 μm samples based on depth except. For instance, samples from 165m with <1.5 L water pre-filtered onto 0.4 μm showed higher diversity values than those with greater volume filtered, and samples from 185m with <2L water filtered onto 2.7 μm showed lower diversity than those with greater volume filtered.

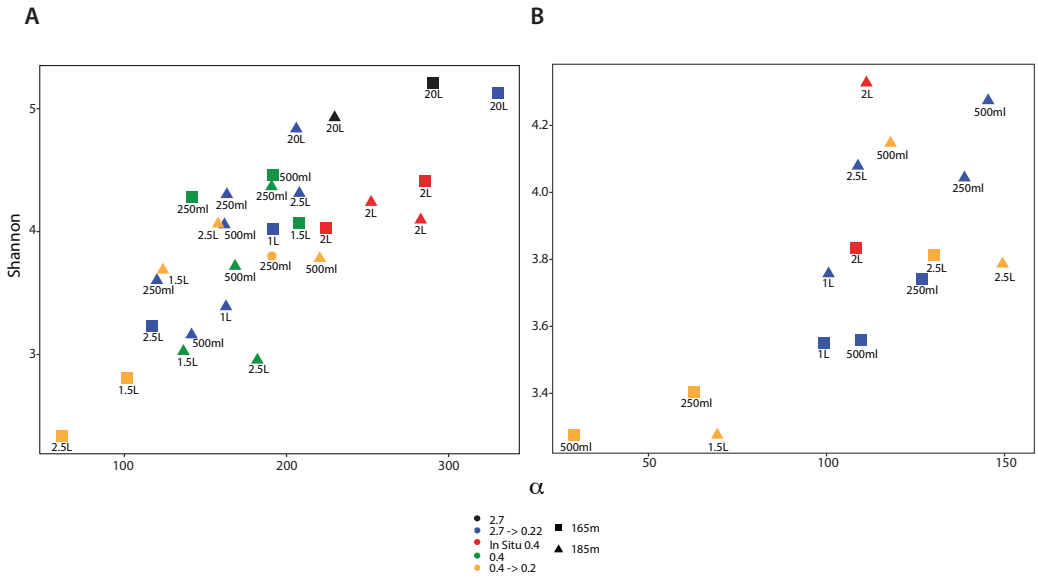

Figure A.2: Shannon and alpha (α) diversity indexes for SSU rDNA pyrotags SCOR samples. Filtering conditions used are depicted as shown in color key for *in situ* from 0.4μm *in situ* (red), on ship 0.4μm pre-filter (green) and pre-filtered (0.4 μm = yellow; 2.7 μm = blue) and in laboratory 2.7 μm (black) samples.

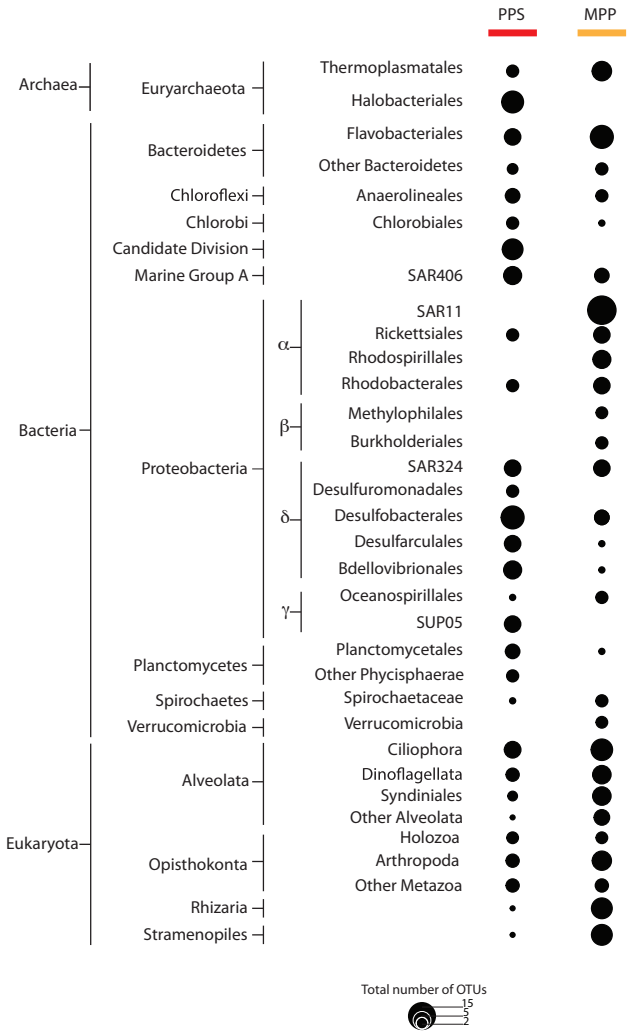

Figure A.3: Active microbial community composition based on rRNA: rDNA ratio (>1) for *in situ* 0.4μm filter (PPS; red) and on-ship 0.22μm pre-filtered onto 0.4mm filter (MPP; yellow). The size of dots depicts the total number of OTUs affiliated to specific taxa.

Appendix B

Protistan parasites along water column oxygen gradients: a network approach to assessing potential host-parasite interactions[1]

B.1 Supplementary results

B.1.1 Co-occurrence analysis and network

To determine potential interactions between OTUs throughout the water column over the stratification period (May-August) in Saanich Inlet, a co-occurrence network was constructed using both Bray-Curtis and Spearman correlation measures. All statistically significant correlations among OTUs resulting after permutations and bootstrap score distributions were included in the downstream analyses. The final edges matrix was visualized as a force directed network using Cytoscape 2.8.3 (Shannon *et al*., 2003). Each node represents an OTU and each edge a statistically significant positive correlation indicating co-occurrence. The resulting network contains 325 nodes, connected by 6,273 edges. Average node degree (mean edges per node) (Proulx *et al*., 2005) was 3, the average path length (the expected distance between two connected nodes) (Latora and Marchiori, 2009) was 2.73, and the network diameter (longest path between two nodes) (Cardoso *et al*., 2009) was equal to 8. The clustering coefficient (connectedness of a nodes neighbour) (Proulx *et al*., 2005) was 0.518, and connectance (proportion of all possible links realized) (Dunne *et al*., 2002) was 0.05. To describe the potential interactions occurring among Syndiniales OTUs and protists, OTUs were taxonomically identified in the Cytoscape network. Nodes corresponding to Syndiniales and their pairs were highlighted to summarize interactions. Based on nodes OTU number correlation data was extracted and exported for

[1] A version of this appendix appears as supplementary information in **Torres-Beltrán, M**., Sehein, T. et al. 2018. Protistan parasites along oxygen gradients in a seasonally anoxic fjord: a network approach to assessing potential host-parasite interactions. Deep Sea Res. II. doi.org/10.1016/j.dsr2.2017.12.026

co-occurrence description and visualization.

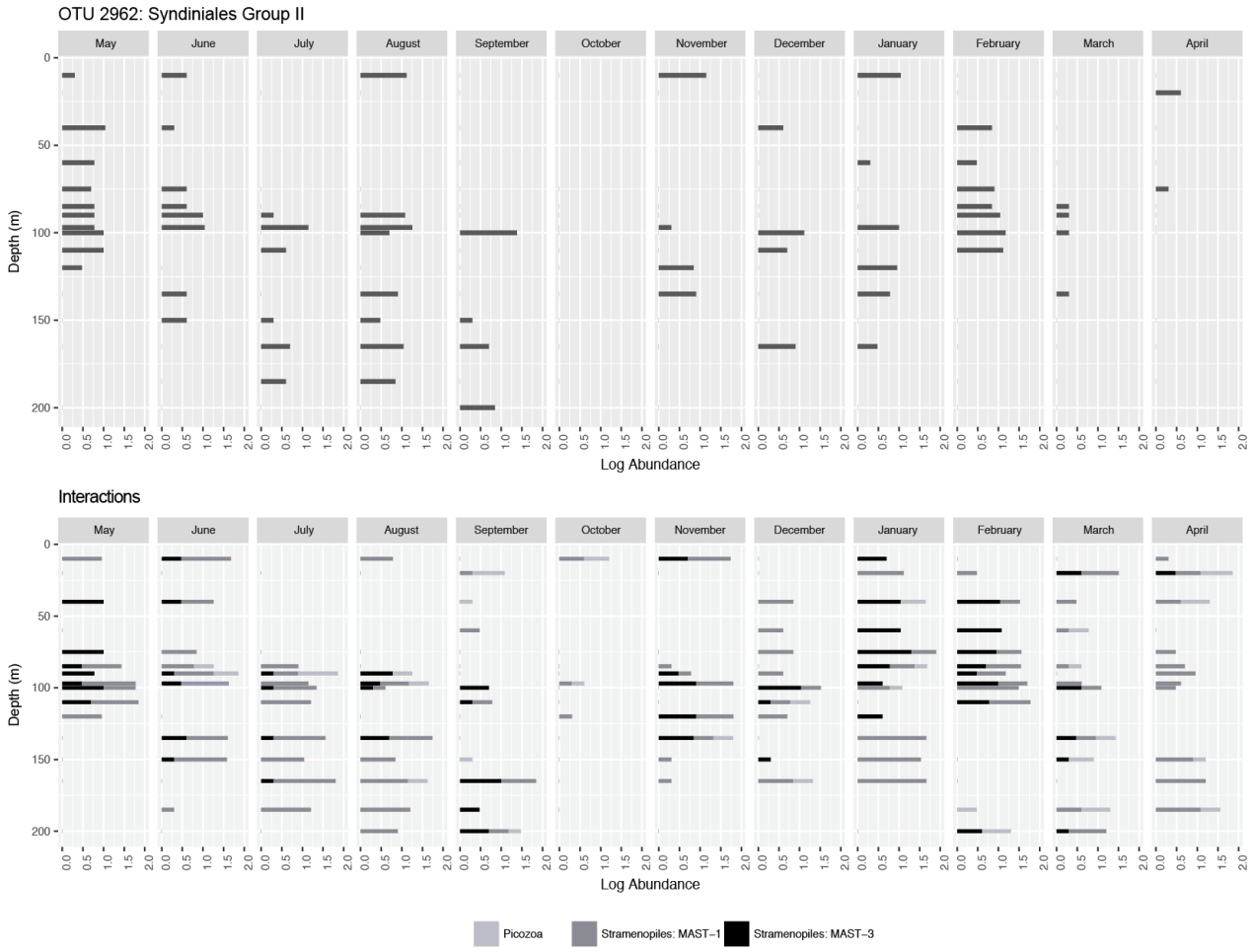

Figure B.1: Vertical distribution overtime for Syndiniales, Picobiliphyta, and Stramenopiles MAST1 and 3 OTUs. Top: Bars depict the abundance of Syndiniales OTUs found to correlate with Picobiliphyta, and Stramenopiles MAST1 and 3 throughout the water column from May 2008 to April 2009. Bottom: Stacked bars depict the abundance (log transformed) for the taxa OTUs found to correlate with Syndiniales throughout the water column from May 2008 to April 2009.

Table B.1: Indicator OTUs for summer stratification suboxic-dysoxic water column conditions from May to August 2008. The OTU Id number, and significant (p =0.05, $\alpha < 0.01$) indicator value are shown. Selected OTUs potentially showing a parasite-type interaction are highlighted.

OTU Id number	Indicator value	*p* value	α value	Taxonomy
15413	0.995	0.005	**	Eukaryota;Hacrobia;Haptophyta;Prymnesiophyceae;Prymnesiophyceae;Prymnesiophyceae;Phaeocystis;Phaeocystis+antarctica
4362	0.979	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-4;Amoebophrya;Amoebophrya+sp
13270	0.978	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-1;MAST-1A;MAST-1A;MAST-1A+sp
12385	0.969	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-1;Dino-Group-I-Clade-1;Dino-Group-I-Clade-1+sp
8298	0.965	0.005	**	Eukaryota;Alveolata;Alveolata;Ellobiopsidae;Ellobiopsidae;Ellobiopsidae;Thalassomyces;Thalassomyces+fagei
41856	0.962	0.005	**	Eukaryota;Opisthokonta;Choanoflagellida;Choanoflagellatea;Acanthoecida;Stephanoecidae_Group_D;Stephanoecidae_Group_D;Stephanoecidae_Group_D+sp
20554	0.948	0.005	**	Eukaryota;Rhizaria;Cercozoa;Filosa-Thecofilosea;Cryomonadida;Protaspa-lineage;Protaspa-lineage;Protaspa-lineage+sp
42840	0.945	0.005	**	Eukaryota;Hacrobia;Picobiliphyta;Picobiliphyta;Picobiliphyta;Picobiliphyta;Picobiliphyta;Picobiliphyta+sp
26650	0.94	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11+sp
2962	0.94	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-22;Dino-Group-II-Clade-22;Dino-Group-II-Clade-22+sp
41598	0.935	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-2;Dino-Group-I-Clade-2;Dino-Group-I-Clade-2+sp
35162	0.935	0.005	**	Eukaryota;Hacrobia;Picobiliphyta;Picobiliphyta;Picobiliphyta;Picobiliphyta;Picobiliphyta;Picobiliphyta+sp
8780	0.931	0.005	**	Eukaryota;Hacrobia;Haptophyta;Prymnesiophyceae;Prymnesiales;Chrysochromulinaceae;Chrysochromulina;Chrysochromulina+rotalis
35489	0.928	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-1;MAST-1B;MAST-1B;MAST-1B+sp
7561	0.927	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-7;Dino-Group-II-Clade-7;Dino-Group-II-Clade-7+sp
43589	0.924	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyta;Dinophyta;Dinophyta;Dinophyta;Dinophyta+sp
26582	0.923	0.005	**	Eukaryota;Opisthokonta;Choanoflagellida;Choanoflagellatea;Acanthoecida;Stephanoecidae_Group_D;Stephanoecidae_Group_D;Stephanoecidae_Group_D+sp
2274	0.916	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-4-6-7-8-9-10-11;MAST-7;MAST-7;MAST-7+sp
26380	0.915	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II;Dino-Group-II;Dino-Group-II+sp
27885	0.915	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-1;MAST-1C;MAST-1C;MAST-1C+sp
18459	0.905	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11+sp
38500	0.901	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-3-12;MAST-3;MAST-3;MAST-3+sp
10557	0.898	0.005	**	Eukaryota;Alveolata;Ciliophora;Spirotrichea;Choreotrichia;Choreotrichia;Choreotrichia;Choreotrichia+sp
50356	0.898	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-1;Dino-Group-I-Clade-1;Dino-Group-I-Clade-1+sp
42773	0.897	0.005	**	Eukaryota;Opisthokonta;Choanoflagellida;Choanoflagellatea;Acanthoecida;Stephanoecidae_Group_H;Stephanoecidae_Group_H;Stephanoecidae_Group_H+sp
10202	0.896	0.005	**	Eukaryota;Hacrobia;Telonemia;Telonemia;Telonemia;Telonemia-Group-2;Telonemia-Group-2;Telonemia-Group-2+sp
17737	0.892	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-8;Dino-Group-II-Clade-8;Dino-Group-II-Clade-8+sp
7847	0.881	0.005	**	Eukaryota;Archaeplastida;Chlorophyta;Mamiellophyceae;Mamiellales;Mamiellaceae;Micromonas;Micromonas+sp
23562	0.878	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
9437	0.878	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;Stramenopiles;Stramenopiles;Stramenopiles;StramenopilesX;StramenopilesX+sp
10460	0.87	0.005	**	Eukaryota;Hacrobia;Picobiliphyta;Picobiliphyta;Picobiliphyta;Picobiliphyta;Picobiliphyta;Picobiliphyta+sp
658	0.869	0.005	**	Eukaryota;Archaeplastida;Chlorophyta;Mamiellophyceae;Mamiellales;Bathycoccaceae;Bathycoccus;Bathycoccus+prasinos
25851	0.864	0.005	**	Eukaryota;Alveolata;Apicomplexa;Apicomplexa;Gregarines;Cephaloidophoroidea;Cephaloidophoroidea;Cephaloidophoroidea+sp
17547	0.862	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-V;Dino-Group-V;Dino-Group-V;Dino-Group-V+sp
31392	0.86	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;Stramenopiles;Stramenopiles;Stramenopiles;StramenopilesX;StramenopilesX+sp
22164	0.858	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
33553	0.851	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-20;Dino-Group-II-Clade-20;Dino-Group-II-Clade-20+sp
10761	0.845	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-4;Dino-Group-I-Clade-4;Dino-Group-I-Clade-4+sp
49179	0.843	0.005	**	Eukaryota;Hacrobia;Telonemia;Telonemia;Telonemia;Telonemia-Group-1;Telonemia-Group-1;Telonemia-Group-1+sp
30577	0.838	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-3-12;MAST-3;MAST-3;MAST-3+sp
27111	0.832	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-16;marine;marine+metagenome
18049	0.83	0.005	**	Eukaryota;Rhizaria;Cercozoa;Filosa-Chlorarachnea;Filosa-Chlorarachnea;LC104-lineage;LC104-lineage;LC104-lineage+sp
25511	0.827	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1+sp
16229	0.823	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1+sp
48960	0.822	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
40167	0.822	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-29;Dino-Group-II-Clade-29;Dino-Group-II-Clade-29+sp
4155	0.819	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-7;Dino-Group-II-Clade-7;Dino-Group-II-Clade-7+sp
28241	0.815	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-1;MAST-1C;MAST-1C;MAST-1C+sp
23310	0.805	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-4;Dino-Group-I-Clade-4;Dino-Group-I-Clade-4+sp
43345	0.801	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-4;Dino-Group-I-Clade-4;Dino-Group-I-Clade-4+sp
13554	0.791	0.005	**	Eukaryota;Opisthokonta;Choanoflagellida;Choanoflagellatea;Acanthoecida;Stephanoecidae_Group_D;Stephanoecidae_Group_D;Stephanoecidae_Group_D+sp
32070	0.789	0.025	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-3;Dino-Group-II-Clade-3;Dino-Group-II-Clade-3+sp
27690	0.788	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
44219	0.788	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-7;Dino-Group-II-Clade-7;Dino-Group-II-Clade-7+sp
39085	0.788	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II;Dino-Group-II;Dino-Group-II+sp
16672	0.784	0.005	**	Eukaryota;Alveolata;Ciliophora;Spirotrichea;Choreotrichia;Undellidae;Undella;Undella+marsupialis
40010	0.784	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
50251	0.783	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-5;Dino-Group-II-Clade-5;Dino-Group-II-Clade-5+sp
7286	0.782	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-2;Dino-Group-I-Clade-2;Dino-Group-I-Clade-2+sp
49329	0.778	0.01	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Gymnodinium;Gymnodinium+sp
36622	0.774	0.01	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
8102	0.77	0.005	**	Eukaryota;Alveolata;Ciliophora;Spirotrichea;Choreotrichia;Choreotrichia;Choreotrichia;Choreotrichia+sp
19736	0.769	0.005	**	Eukaryota;Opisthokonta;Choanoflagellida;Choanoflagellatea;Acanthoecida;Stephanoecidae_Group_D;Stephanoecidae_Group_D;Stephanoecidae_Group_D+sp
34794	0.768	0.02	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-12;Dino-Group-II-Clade-12;Dino-Group-II-Clade-12+sp
7259	0.766	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-7;Dino-Group-II-Clade-7;Dino-Group-II-Clade-7+sp
47404	0.764	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
13489	0.759	0.025	*	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dissodinium;Dissodinium+pseudolunula
10911	0.753	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-6;Dino-Group-I-Clade-6;Dino-Group-I-Clade-6+sp
50898	0.753	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-6;Dino-Group-II-Clade-6;Dino-Group-II-Clade-6+sp
50219	0.749	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5+sp
8473	0.746	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-30;Dino-Group-II-Clade-30;Dino-Group-II-Clade-30+sp
15504	0.744	0.005	**	Eukaryota;Excavata;Discoba;Euglenozoa;Diplonemea;Diplonemea;Diplonemea;Diplonemea+sp
31377	0.735	0.005	**	Eukaryota;Hacrobia;Katablepharidophyta;Katablepharidaceae;Katablepharidales;Katablepharidales;Katablepharidales;Katablepharidales+sp
38481	0.73	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11+sp
21594	0.726	0.005	**	Eukaryota;Archaeplastida;Chlorophyta;Trebouxiophyceae;Chlorellales;Chlorellales;Chlorellales;Chlorellales+sp
26361	0.725	0.01	**	Eukaryota;Alveolata;Ciliophora;Colpodea;Colpodea-1;Colpodea-1;Colpodea-1;Colpodea-1+sp
16728	0.721	0.02	*	Eukaryota;Alveolata;Ciliophora;Oligohymenophorea;Scuticociliatia;Scuticociliatia-1;Scuticociliatia-1;Scuticociliatia-1+sp
35326	0.719	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5+sp
1611	0.717	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-1;Dino-Group-I-Clade-1;Dino-Group-I-Clade-1+sp
14399	0.717	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-36;Dino-Group-II-Clade-36;Dino-Group-II-Clade-36+sp
10016	0.717	0.005	**	Eukaryota;Excavata;Discoba;Euglenozoa;Diplonemea;Diplonemea;Diplonemea;Diplonemea+sp
7059	0.714	0.015	*	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
45285	0.708	0.015	*	Eukaryota;Alveolata;Ciliophora;Colpodea;Colpodea-1;Colpodea-1;Colpodea-1;Colpodea-1+sp
24068	0.702	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-13;Dino-Group-II-Clade-13;Dino-Group-II-Clade-13+sp

Table B.1: Indicator OTUs for summer stratification suboxic-dysoxic water column conditions from May to August 2008. The OTU Id number, and significant (p =0.05, $\alpha < 0.01$) indicator value are shown. Selected OTUs potentially showing a parasite-type interaction are highlighted. (Continuation)

OTU Id number	Indicator value	*p* value	α value	Taxonomy
26382	0.697	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11+sp
14519	0.697	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11+sp
50042	0.678	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;Labyrinthulea;Labyrinthulales;Labyrinthulaceae;Labyrinthulaceae;Labyrinthulaceae+sp
5605	0.678	0.005	**	Organelle;nucleomorph-Archaeplastida;Cryptophyta-nucleomorph;Cryptophyta-nucleomorph;Cryptophyta-nucleomorph;Cryptophyta-nucleomorph;Teleaulax;Teleaulax+amphioxeia
2745	0.675	0.02	*	Eukaryota;Rhizaria;Cercozoa;Filosa-Chlorarachnea;Filosa-Chlorarachnea;NPK2-lineage;NPK2-lineage;NPK2-lineage+sp
43154	0.668	0.015	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5+sp
11111	0.665	0.01	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
17402	0.662	0.015	*	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
26536	0.661	0.035	*	Eukaryota;Hacrobia;Telonemia;Telonemia;Telonemia;Telonemia-Group-2;Telonemia-Group-2;Telonemia-Group-2+sp
13400	0.658	0.02	*	Eukaryota;Alveolata;Ciliophora;Oligohymenophorea;Scuticociliatia;Scuticociliatia-1;Scuticociliatia-1;Scuticociliatia-1+sp
1893	0.654	0.03	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II;Dino-Group-II;Dino-Group-II+sp
1973	0.65	0.03	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1+sp
27308	0.645	0.01	**	Eukaryota;Stramenopiles;Stramenopiles;Stramenopiles;Stramenopiles;Stramenopiles;StramenopilesX;StramenopilesX+sp
29105	0.641	0.025	*	Eukaryota;Opisthokonta;Fungi;Chytridiomycota;Chytridiomycotina;Chytridiomycetes;Spizellomycetales-and-Rhizophlyctidales;Spizellomycetales-and-Rhizophlyctidales+sp
14478	0.637	0.02	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-2;Dino-Group-I-Clade-2;Dino-Group-I-Clade-2+sp
503	0.637	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-27;Dino-Group-II-Clade-27;Dino-Group-II-Clade-27+sp
27014	0.637	0.015	*	Eukaryota;Hacrobia;Picobiliphyta;Picobiliphyta;Picobiliphyta;Picobiliphyta;Picobiliphyta;Picobiliphyta+sp
3056	0.637	0.005	**	Eukaryota;Opisthokonta;Choanoflagellida;Choanoflagellatea;Acanthoecida;Stephanoecidae_Group_D;Stephanoecidae_Group_D;Stephanoecidae_Group_D+sp
2939	0.637	0.015	*	Eukaryota;Stramenopiles;Stramenopiles;Bacillariophyta;Bacillariophyta;Radial-centric-basal-Coscinodiscophyceae;Corethron;Corethron+hystrix
21902	0.63	0.015	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-20;Dino-Group-II-Clade-20;Dino-Group-II-Clade-20+sp
4066	0.629	0.02	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1+sp
9442	0.629	0.03	*	Eukaryota;Hacrobia;Telonemia;Telonemia;Telonemia;Telonemia-Group-2;Telonemia-Group-2;Telonemia-Group-2+sp
19028	0.615	0.015	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-3;Amoebophrya;Amoebophrya+sp
6127	0.615	0.025	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-5;Dino-Group-II-Clade-5;Dino-Group-II-Clade-5+sp
42296	0.615	0.035	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-6;Dino-Group-II-Clade-6;Dino-Group-II-Clade-6+sp
13243	0.615	0.03	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-6;Dino-Group-II-Clade-6;Dino-Group-II-Clade-6+sp
41630	0.615	0.035	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-7;Dino-Group-II-Clade-7;Dino-Group-II-Clade-7+sp
6998	0.615	0.02	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-8;Dino-Group-II-Clade-8;Dino-Group-II-Clade-8+sp
26680	0.615	0.035	*	Eukaryota;Hacrobia;Cryptophyta;Cryptophyceae;Cryptophyceae;Cryptomonadales;Cryptomonadales;Cryptomonadales+sp
18473	0.615	0.015	*	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
5983	0.615	0.025	*	Eukaryota;Rhizaria;Cercozoa;Filosa-Imbricatea;Filosa-Imbricatea;Novel-clade-2;Novel-clade-2;Novel-clade-2+sp
40970	0.615	0.02	*	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-4-6-7-8-9-10-11;MAST-6;MAST-6;MAST-6+sp
29572	0.604	0.045	*	Eukaryota;Opisthokonta;Choanoflagellida;Choanoflagellatea;Craspedida;Monosigidae_Group_M;Monosigidae_Group_M;Monosigidae_Group_M+sp
47466	0.602	0.035	*	Eukaryota;Alveolata;Dinophyta;Dinophyta;Dinophyta;Dinophyta;Dinophyta;Dinophyta+sp
16021	0.593	0.03	*	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
9651	0.593	0.03	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11+sp
47939	0.593	0.03	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-14;Dino-Group-II-Clade-14;Dino-Group-II-Clade-14+sp
43319	0.593	0.015	*	Eukaryota;Opisthokonta;Choanoflagellida;Choanoflagellatea;Acanthoecida;Stephanoecidae_Group_D;Stephanoecidae_Group_D;Stephanoecidae_Group_D+sp
23147	0.593	0.02	*	Eukaryota;Rhizaria;Cercozoa;Filosa-Thecofilosea;Cryomonadida;Protaspa-lineage;Protaspa-lineage;Protaspa-lineage+sp
37245	0.593	0.04	*	Eukaryota;Stramenopiles;Stramenopiles;Dictyochophyceae;Dictyochophyceae;Florenciellales;Florenciellales;Florenciellales+sp
7889	0.593	0.035	*	Eukaryota;Stramenopiles;Stramenopiles;Stramenopiles-Group-9;Stramenopiles-Group-9;Stramenopiles-Group-9;Stramenopiles-Group-9;Stramenopiles-Group-9+sp
20006	0.593	0.02	*	Organelle;nucleomorph-Archaeplastida;Cryptophyta-nucleomorph;Cryptophyta-nucleomorph;Cryptophyta-nucleomorph;Cryptophyta-nucleomorph;Falcomonas;Falcomonas+daucoid
18789	0.58	0.035	*	Eukaryota;Hacrobia;Haptophyta;Prymnesiophyceae;Prymnesiophyceae;Prymnesiophyceae;Prymnesiophyceae;Prymnesiophyceae+sp
27024	0.569	0.025	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-26;Dino-Group-II-Clade-26;Dino-Group-II-Clade-26+sp
50929	0.569	0.035	*	Eukaryota;Opisthokonta;Choanoflagellida;Choanoflagellatea;Acanthoecida;Stephanoecidae_Group_D;Stephanoecidae_Group_D;Stephanoecidae_Group_D+sp

Table B.3: List of Spring indicator OTUs (May-June 2008). Table shows the OTU Id number, indicator value (IV > 0.6, p-value=0.05 and αvalue <0.01), and BLAST-based taxonomic assignment.

Surface indicator OTUs

OTU Id number	Indicator value	*p* value	α value	Taxonomy
38633	1	0.005	**	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
41739	0.933	0.01	**	Eukaryota;Stramenopiles;Stramenopiles;Bacillariophyta;Bacillariophyta;Polar-centric-Mediophyceae;Polar-centric-Mediophyceae;Polar-centric-Mediophyceae+sp
51135	0.894	0.01	**	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
47657	0.894	0.01	**	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
20157	0.894	0.01	**	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
17656	0.894	0.01	**	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
44881	0.894	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-14;Dino-Group-II-Clade-14;Dino-Group-II-Clade-14+sp
47889	0.894	0.01	**	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
45315	0.876	0.01	**	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
0	0.871	0.005	**	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
20422	0.859	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II;Dino-Group-II;Dino-Group-II+sp
4066	0.851	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1+sp
9651	0.814	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11+sp
23960	0.804	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-7;Dino-Group-II-Clade-7;Dino-Group-II-Clade-7+sp
50488	0.775	0.015	*	Eukaryota;Opisthokonta;Metazoa;Cnidaria;Cnidaria;Hydrozoa;Hydrozoa;Hydrozoa+sp
277	0.775	0.01	**	Eukaryota;Stramenopiles;Stramenopiles;Bacillariophyta;Bacillariophyta;Polar-centric-Mediophyceae;Chaetoceros;Chaetoceros+decipiens
50514	0.775	0.02	*	Eukaryota;Hacrobia;Telonemia;Telonemia;Telonemia;Telonemia-Group-2;Telonemia-Group-2;Telonemia-Group-2+sp
26360	0.775	0.015	*	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
1973	0.773	0.025	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1+sp
17143	0.759	0.01	**	Eukaryota;Stramenopiles;Stramenopiles;Labyrinthulea;Labyrinthulales;Labyrinthulaceae;Labyrinthulaceae;Labyrinthulaceae+sp
35395	0.757	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1+sp
19800	0.748	0.02	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-4;Dino-Group-I-Clade-4;Dino-Group-I-Clade-4+sp
25638	0.737	0.015	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1+sp
50019	0.737	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
44615	0.733	0.02	*	Eukaryota;Stramenopiles;Stramenopiles;Bacillariophyta;Bacillariophyta;Polar-centric-Mediophyceae;Skeletonema;Skeletonema+sp
33748	0.728	0.025	*	Eukaryota;Stramenopiles;Stramenopiles;Bacillariophyta;Bacillariophyta;Polar-centric-Mediophyceae;Polar-centric-Coscinodiscophyceae;Polar-centric-Coscinodiscophyceae+sp
27766	0.724	0.015	*	Eukaryota;Alveolata;Ciliophora;Oligohymenophorea;Scuticociliatia;Scuticociliatia;Scuticociliatia;Scuticociliatia+sp
33594	0.722	0.025	*	Eukaryota;Stramenopiles;Stramenopiles;Bacillariophyta;Bacillariophyta;Polar-centric-Mediophyceae;Polar-centric-Coscinodiscophyceae;Polar-centric-Coscinodiscophyceae+sp
16264	0.717	0.03	*	Eukaryota;Hacrobia;Katablepharidophyta;Katablepharidaceae;Katablepharidales;Katablepharidales;Katablepharis;Katablepharis+japonica
43348	0.697	0.025	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-3;Dino-Group-I-Clade-3;Dino-Group-I-Clade-3+sp
39795	0.689	0.04	*	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-3-12;MAST-12;MAST-12;MAST-12+sp
48230	0.679	0.035	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11;Dino-Group-II-Clade-10-and-11+sp
3694	0.663	0.03	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-22;Dino-Group-II-Clade-22;Dino-Group-II-Clade-22+sp
8473	0.661	0.045	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-30;Dino-Group-II-Clade-30;Dino-Group-II-Clade-30+sp
28694	0.632	0.03	*	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
28607	0.632	0.03	*	Eukaryota;Stramenopiles;Stramenopiles;Bacillariophyta;Bacillariophyta;Polar-centric-Mediophyceae;Polar-centric-Mediophyceae;Polar-centric-Mediophyceae+sp
15495	0.632	0.03	*	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
22207	0.632	0.03	*	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
33168	0.632	0.03	*	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
25560	0.632	0.04	*	Eukaryota;Stramenopiles;Stramenopiles;Stramenopiles-Group-7;Stramenopiles-Group-7;Stramenopiles-Group-7;Stramenopiles-Group-7;Stramenopiles-Group-7+sp

Sub-surface indicator OTUs

OTU Id number	Indicator value	*p* value	α value	Taxonomy
9442	0.869	0.005	**	Eukaryota;Hacrobia;Telonemia;Telonemia;Telonemia;Telonemia-Group-2;Telonemia-Group-2;Telonemia-Group-2+sp
16838	0.866	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1+sp
21162	0.829	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-27;Dino-Group-II-Clade-27;Dino-Group-II-Clade-27+sp
35088	0.777	0.01	**	Eukaryota;Excavata;Discoba;Euglenozoa;Diplonemea;Diplonemea;Diplonemea;Diplonemea+sp
18473	0.776	0.03	*	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
46903	0.744	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
37276	0.743	0.015	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-5;Dino-Group-II-Clade-5;Dino-Group-II-Clade-5+sp
43833	0.727	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-4;Dino-Group-I-Clade-4;Dino-Group-I-Clade-4+sp
24591	0.726	0.02	*	Eukaryota;Excavata;Discoba;Euglenozoa;Kinetoplastida;Neobodonid;Neobodo;Neobodo+designis
34542	0.702	0.05	*	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
2667	0.667	0.03	*	Eukaryota;Opisthokonta;Fungi;Ascomycota;Pezizomycotina;Eurotiomycetes;Aspergillus;Aspergillus+sp

Combined surface and sub-surface indicator OTUs

OTU Id number	Indicator value	*p* value	α value	Taxonomy
36180	0.925	0.015	*	Eukaryota;Opisthokonta;Metazoa;Cnidaria;Cnidaria;Hydrozoa;Hydrozoa;Hydrozoa+sp
46532	0.92	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-III;Dino-Group-III;Dino-Group-III;Dino-Group-III+sp
23562	0.919	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
4223	0.899	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-31;Dino-Group-II-Clade-31;Dino-Group-II-Clade-31+sp
503	0.894	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-27;Dino-Group-II-Clade-27;Dino-Group-II-Clade-27+sp
22164	0.854	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
36622	0.831	0.015	*	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
23147	0.796	0.005	**	Eukaryota;Rhizaria;Cercozoa;Filosa-Thecofilosea;Cryomonadida;Protaspa-lineage;Protaspa-lineage;Protaspa-lineage+sp
42296	0.786	0.035	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-6;Dino-Group-II-Clade-6;Dino-Group-II-Clade-6+sp
47404	0.784	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
38851	0.784	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-III;Dino-Group-III;Dino-Group-III;Dino-Group-III+sp
43154	0.782	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5+sp
25851	0.782	0.025	*	Eukaryota;Alveolata;Apicomplexa;Apicomplexa;Gregarines;Cephaloidophoroidea;Cephaloidophoroidea;Cephaloidophoroidea;Cephaloidophoroidea+sp
50251	0.776	0.025	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-5;Dino-Group-II-Clade-5;Dino-Group-II-Clade-5+sp
45843	0.759	0.015	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5+sp
50307	0.757	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-3-12;MAST-3;MAST-3;MAST-3+sp
6998	0.757	0.02	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-8;Dino-Group-II-Clade-8;Dino-Group-II-Clade-8+sp
13489	0.721	0.02	*	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dissodinium;Dissodinium+pseudolunula
39085	0.717	0.035	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II;Dino-Group-II;Dino-Group-II+sp
27690	0.696	0.05	*	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
1063	0.679	0.03	*	Eukaryota;Alveolata;Ciliophora;Spirotrichea;Choreotrichia;Choreotrichia;Choreotrichia;Choreotrichia+sp

Table B.4: List of Summer stratification indicator OTUs (July- August 2008). Table shows the OTU Id number, indicator value (IV > 0.6, p-value=0.05 and αvalue <0.01), and BLAST-based taxonomic assignment.

Upper oxycline indicator OTUs				
OTU Id number	**Indicator value**	***p* value**	**α value**	**Taxonomy**
46185	0.897	0.03	*	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Pseudocalanus;Pseudocalanus+sp
44660	0.753	0.02	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5+sp
38064	0.749	0.025	*	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Pseudocalanus;Pseudocalanus+sp
35395	0.707	0.025	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1+sp
39179	0.642	0.04	*	Eukaryota;Alveolata;Apicomplexa;Apicomplexa;Gregarines;Cephaloidophoroidea;Cephaloidophoroidea;Cephaloidophoroidea+sp

Upper oxycline indicator OTUs				
OTU Id number	**Indicator value**	***p* value**	**α value**	**Taxonomy**
24591	0.71	0.02	*	Eukaryota;Excavata;Discoba;Euglenozoa;Kinetoplastida;Neobodonid;Neobodo;Neobodo+designis
18185	0.7	0.025	*	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp

Complete oxycline indicator OTUs				
OTU Id number	**Indicator value**	***p* value**	**α value**	**Taxonomy**
35326	0.951	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5+sp
41598	0.937	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-2;Dino-Group-I-Clade-2;Dino-Group-I-Clade-2+sp
34794	0.886	0.02	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-12;Dino-Group-II-Clade-12;Dino-Group-II-Clade-12+sp
16229	0.883	0.03	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1+sp
51039	0.875	0.04	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-23;Dino-Group-II-Clade-23;Dino-Group-II-Clade-23+sp
10016	0.853	0.005	**	Eukaryota;Excavata;Discoba;Euglenozoa;Diplonemea;Diplonemea;Diplonemea;Diplonemea+sp
34620	0.84	0.01	**	Eukaryota;Excavata;Discoba;Euglenozoa;Diplonemea;Diplonemea;Diplonemea;Diplonemea+sp
13400	0.836	0.015	*	Eukaryota;Alveolata;Ciliophora;Oligohymenophorea;Scuticociliatia;Scuticociliatia-1;Scuticociliatia-1;Scuticociliatia-1+sp
35088	0.832	0.01	**	Eukaryota;Excavata;Discoba;Euglenozoa;Diplonemea;Diplonemea;Diplonemea;Diplonemea+sp
6567	0.832	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-14;Dino-Group-II-Clade-14;Dino-Group-II-Clade-14+sp
35577	0.827	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-III;Dino-Group-III;Dino-Group-III;Dino-Group-III+sp
17737	0.824	0.02	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-8;Dino-Group-II-Clade-8;Dino-Group-II-Clade-8+sp
3694	0.823	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-22;Dino-Group-II-Clade-22;Dino-Group-II-Clade-22+sp
1611	0.822	0.015	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-1;Dino-Group-I-Clade-1;Dino-Group-I-Clade-1+sp
7286	0.815	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-2;Dino-Group-I-Clade-2;Dino-Group-I-Clade-2+sp
40970	0.814	0.015	*	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-4-6-7-8-9-10-11;MAST-6;MAST-6;MAST-6+sp
49036	0.813	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-14;Dino-Group-II-Clade-14;Dino-Group-II-Clade-14+sp
44219	0.784	0.015	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-7;Dino-Group-II-Clade-7;Dino-Group-II-Clade-7+sp
42118	0.784	0.01	**	Eukaryota;Alveolata;Ciliophora;Oligohymenophorea;Apostomatia;Foettingeriidae;Hyalophysa;Hyalophysa+lwoffi
14478	0.784	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-2;Dino-Group-I-Clade-2;Dino-Group-I-Clade-2+sp
13243	0.773	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-6;Dino-Group-II-Clade-6;Dino-Group-II-Clade-6+sp
1063	0.773	0.01	**	Eukaryota;Alveolata;Ciliophora;Spirotrichea;Choreotrichia;Choreotrichia;Choreotrichia;Choreotrichia+sp
3845	0.767	0.01	**	Eukaryota;Alveolata;Ciliophora;Oligohymenophorea;Apostomatia;Foettingeriidae;Vampyrophrya;Vampyrophrya+pelagica
15376	0.758	0.015	*	Eukaryota;Excavata;Discoba;Euglenozoa;Kinetoplastida;Neobodonid;Neobodo;Neobodo+sp1
10911	0.756	0.03	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-6;Dino-Group-I-Clade-6;Dino-Group-I-Clade-6+sp
40167	0.754	0.03	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-29;Dino-Group-II-Clade-29;Dino-Group-II-Clade-29+sp
45285	0.753	0.015	*	Eukaryota;Alveolata;Ciliophora;Colpodea;Colpodea-1;Colpodea-1;Colpodea-1;Colpodea-1+sp
5605	0.701	0.05	*	Organelle;nucleomorph-Archaeplastida;Cryptophyta-nucleomorph;Cryptophyta-nucleomorph;Cryptophyta-nucleomorph;Cryptophyta-nucleomorph;Teleaulax;Teleaulax+amphioxeia
19479	0.679	0.035	*	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
1973	0.679	0.035	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1+sp
50349	0.679	0.05	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-16;marine;marine+metagenome

Table B.5: List of early Fall renewal indicator OTUs (September-October 2008). Table shows the OTU Id number, indicator value (IV > 0.6, p-value=0.05 and αvalue <0.01), and BLAST-based taxonomic assignment.

Mixed basin indicator OTUs				
OTU Id number	**Indicator value**	***p* value**	**α value**	**Taxonomy**
17371	0.943	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;Bicoecea;Bicoecea;Bicoecea;Bicoecea;Bicoecea+sp
36737	0.861	0.01	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
22836	0.728	0.02	*	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Herdmania;Herdmania+litoralis
47750	0.667	0.02	*	Eukaryota;Stramenopiles;Stramenopiles;Stramenopiles;Stramenopiles;Stramenopiles;StramenopilesX;StramenopilesX+sp
41334	0.638	0.025	*	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp

Mixed oxycline indicator OTUs				
OTU Id number	**Indicator value**	***p* value**	**α value**	**Taxonomy**
34620	0.931	0.005	**	Eukaryota;Excavata;Discoba;Euglenozoa;Diplonemea;Diplonemea;Diplonemea;Diplonemea+sp
16838	0.926	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1+sp
21162	0.898	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-27;Dino-Group-II-Clade-27;Dino-Group-II-Clade-27+sp
3938	0.862	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-25;Dino-Group-II-Clade-25;Dino-Group-II-Clade-25+sp
9442	0.853	0.01	**	Eukaryota;Hacrobia;Telonemia;Telonemia;Telonemia;Telonemia-Group-2;Telonemia-Group-2;Telonemia-Group-2+sp
35088	0.841	0.005	**	Eukaryota;Excavata;Discoba;Euglenozoa;Diplonemea;Diplonemea;Diplonemea;Diplonemea+sp
37276	0.8	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-5;Dino-Group-II-Clade-5;Dino-Group-II-Clade-5+sp
12917	0.786	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;Dictyochophyceae;Dictyochophyceae;Pedinellales;Pedinellales;Pedinellales+sp
24591	0.784	0.02	*	Eukaryota;Excavata;Discoba;Euglenozoa;Kinetoplastida;Neobodonid;Neobodo;Neobodo+designis
18473	0.761	0.03	*	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
45363	0.756	0.005	**	Eukaryota;Archaeplastida;Chlorophyta;Pyramimonadales;Pyramimonadales;Pyramimonadales;Pyramimonas;Pyramimonas+aurea
795	0.753	0.015	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5+sp
1063	0.738	0.025	*	Eukaryota;Alveolata;Ciliophora;Spirotrichea;Choreotrichia;Choreotrichia;Choreotrichia;Choreotrichia+sp
46903	0.73	0.035	*	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
41750	0.665	0.035	*	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
34132	0.641	0.045	*	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-4-6-7-8-9-10-11;MAST-4;MAST-4;MAST-4+sp

Combined mixed basin and oxycline indicator OTUs				
OTU Id number	**Indicator value**	***p* value**	**α value**	**Taxonomy**
27470	0.829	0.01	**	Eukaryota;Apusozoa;Hilomonadea;Planomonadida;Planomonadidae;Planomonadidae_Group-1;Planomonadidae_Group-1;Planomonadidae_Group-1+sp
35480	0.829	0.005	**	Eukaryota;Alveolata;Ciliophora;Colpodea;Colpodea-1;Colpodea-1;Colpodea-1;Colpodea-1+sp
43896	0.783	0.025	*	Eukaryota;Excavata;Discoba;Euglenozoa;Euglenida;Petalomonadales;Petalomonadales;Petalomonadales+sp
8780	0.951	0.005	**	Eukaryota;Hacrobia;Haptophyta;Prymnesiophyceae;Prymnesiales;Chrysochromulinaceae;Chrysochromulina;Chrysochromulina+rotalis
2274	0.945	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-4-6-7-8-9-10-11;MAST-7;MAST-7;MAST-7+sp
503	0.932	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-27;Dino-Group-II-Clade-27;Dino-Group-II-Clade-27+sp
46532	0.899	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-III;Dino-Group-III;Dino-Group-III;Dino-Group-III+sp
36180	0.898	0.01	**	Eukaryota;Opisthokonta;Metazoa;Cnidaria;Cnidaria;Hydrozoa;Hydrozoa;Hydrozoa+sp
22164	0.894	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
4223	0.882	0.005	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-31;Dino-Group-II-Clade-31;Dino-Group-II-Clade-31+sp
36622	0.868	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
0	0.863	0.01	**	Eukaryota;Opisthokonta;Metazoa;Arthropoda;Crustacea;Maxillopoda;Maxillopoda;Maxillopoda+sp
23147	0.832	0.01	**	Eukaryota;Rhizaria;Cercozoa;Filosa-Thecofilosea;Cryomonadida;Protaspa-lineage;Protaspa-lineage;Protaspa-lineage+sp
27308	0.819	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;Stramenopiles;Stramenopiles;Stramenopiles;StramenopilesX;StramenopilesX+sp
25851	0.818	0.005	**	Eukaryota;Alveolata;Apicomplexa;Apicomplexa;Gregarines;Cephaloidophoroidea;Cephaloidophoroidea;Cephaloidophoroidea+sp
47404	0.816	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae;Dinophyceae+sp
47523	0.811	0.04	*	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-4-6-7-8-9-10-11;MAST-4;MAST-4;MAST-4+sp
50251	0.81	0.01	**	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-5;Dino-Group-II-Clade-5;Dino-Group-II-Clade-5+sp
9437	0.809	0.015	*	Eukaryota;Stramenopiles;Stramenopiles;Stramenopiles;Stramenopiles;Stramenopiles;StramenopilesX;StramenopilesX+sp
50307	0.789	0.005	**	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-3-12;MAST-3;MAST-3;MAST-3+sp
28241	0.777	0.01	**	Eukaryota;Stramenopiles;Stramenopiles;MAST;MAST-1;MAST-1C;MAST-1C;MAST-1C+sp
1973	0.77	0.04	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1;Dino-Group-II-Clade-1+sp
43154	0.755	0.015	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-I;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5;Dino-Group-I-Clade-5+sp
13489	0.754	0.005	**	Eukaryota;Alveolata;Dinophyta;Dinophyceae;Dinophyceae;Dinophyceae;Dissodinium;Dissodinium+pseudolunula
38851	0.741	0.02	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-III;Dino-Group-III;Dino-Group-III;Dino-Group-III+sp
47466	0.741	0.035	*	Eukaryota;Alveolata;Dinophyta;Dinophyta;Dinophyta;Dinophyta;Dinophyta;Dinophyta+sp
6998	0.711	0.05	*	Eukaryota;Alveolata;Dinophyta;Syndiniales;Dino-Group-II;Dino-Group-II-Clade-8;Dino-Group-II-Clade-8;Dino-Group-II-Clade-8+sp
18963	0.693	0.025	*	Eukaryota;Hacrobia;Haptophyta;Prymnesiophyceae;Prymnesiales;Prymnesiaceae;Imantonia;Imantonia+sp
39173	0.577	0.05	*	Eukaryota;Alveolata;Ciliophora;Colpodea;Colpodea-1;Colpodea-1;Colpodea-1;Colpodea-1+sp

Appendix C

Methanotrophic community dynamics in Saanich Inlet[1]

C.1 Supplementary methods

C.1.1 Network description

To determine potential interactions between OTUs throughout the water column, a co-occurrence network was constructed using both Bray-Curtis and Spearman correlation measures. All statistically significant correlations among OTUs resulting after permutations and bootstrap score distributions were included in the network. Each node represents an OTU and each edge a statistically significant positive correlation indicating co-occurrence. The resulting network contains 520 nodes, connected by 9,499 edges. Average node degree (mean edges per node) (Proulx *et al*., 2005) was 4.4, the average path length (the expected distance between two connected nodes) (Latora and Marchiori, 2009) was 3.4, and the network diameter (longest path between two nodes) (Cardoso *et al*., 2009) was equal to 9. The clustering coefficient (connectedness of a node's neighbour) (Proulx *et al*., 2005) was 0.701, and connectance (proportion of all possible links realized) (Dunne *et al*., 2002) was 0.037.

To explore how O_2 influence network topology, modules (groups of highly clustered OTUs within the network) were identified with the ClusterViz (Wang *et al*., 2015) application on Cytoscape using the FAG-EC (Li *et al*., 2008) and MCODE (Bader and Hogue, 2003) algorithms. Three main modules were identified corresponding to O_2 water column compartments. Oxygen conditions were identified based on indicator OTUs within each module. Module 1, corresponding to the oxic (HCA group I) water column condition contained 140 OTUs, mostly affiliated with Marine Group II, Actinobacteria, SAR11, Rhodobacter, and indicator OTUs related to Bacteroidetes and Cyanobacteria. Module 2, corresponding to the dysoxic-suboxic (HCA group II) water column condition contained 197 OTUs, mostly affiliated with Bacteroidetes, Marine Group A, SAR11, Nitrospina, Methylophilales and indicator OTUs related to OPU_01, Planctomycetes,

[1] A version of this appendix appears as supplementary information in **Torres-Beltrán, M.** *et al*. 2016. Methanotrophic community dynamics in a seasonally anoxic fjord: Saanich Inlet, British Columbia. Front. Mar.Sci. doi.org/10.3389/fmars.2016.00268 .

and Verrucomicrobia. Module 3, corresponding to the anoxic (HCA group III) water column condition contained 140 OTUs, mostly affiliated to Thaumarchaeota, Marine SAR324, Nitrospina, Methylophaga, methanotrophic symbionts, and Verrucomicrobia, with indicator OTUs related to SUP05, Desulfobacteraceae, Marine Group A and Planctomycetes.

C.1.2 Methanotroph sub-networks

Hive plots (Krzywinski *et al*., 2012) were used to visualize positive correlations among methanotroph and microbial OTUs separately. Node axis assignment and positioning rules were chosen to reflect the number of potential interactions according to OTUs phylogeny and abundance. Nodes were placed on three axes according to their relative abundance, such that nodes closer to the center of the hive plot represent OTUs with low abundance values throughout the water column, and nodes located towards the outer perimeter of the hive plot represent OTUs with high abundance values throughout the water column.

Adjacent edges to Methanotroph OTUs and their connected nodes were selected in Cytoscape, and sub-network was built using a force directed layout. Sub-networks edges matrix was exported and visualized using the tool Hive Panel Explorer (https://github.com/hallamlab/HivePanelExplorer/wiki) (Perez, 2015). Sub-networks' composition is detailed in Supporting Information Table S4.

C.1.3 Methanotroph interactions in sub-networks

To support methanotroph interactions with bacterial OTUs in sub-networks, I followed a BLAST-comparison approach using as query the OTUs representative sequences derived from the pyrotag dataset against the SSU rRNA sequences derived from a CH_4 microcosm experiment carried out on Saanich waters (Sauter *et al*., 2012). Microcosm sequences were retrieved using the accession numbers JN172107-JN172588. I reported all blast hits with 80% identity or higher (e-value $<10^{-6}$) for all OTUs within methanotroph sub-networks that also resulted indicator OTUs (Supporting Information Table S4).

C.1.4 Statistical analyses considerations

To explore the correlation among methanotrophic OTUs distribution and to environmental variables, I conducted multivariate, negative binomial regression analyses. Multivariate regression allowed me to develop hypotheses on potential niche partitioning by inferring partial correlations between representative methanotroph OTUs while partialling out the effect of depth. Under this assumption, I considered a positive correlation as a concurrent fluctuation between OTUs abundance rather than an identical spatial distribution. Resulting correlations were not used in a deterministic fashion. I complement their interpretation for potential niche partitioning with further statistical models i.e negative binomial regression using environmental param. Negative binomial regression analyses allowed me to develop hypotheses on potential alternative electron acceptors (energy sources) by methanotroph OTUs. Negative binomial regression is commonly used for modeling count variables for over-dispersed count outcome variables (Venables and Ripley, 2002). Zero-inflated negative binomial regression is commonly used for modeling count variables with excessive zeros for over-dispersed count outcome variables (Zeileis *et al*., 2008), thus being accurate for rare taxa such as methanotroph OTUs particularly those affiliated with OPU1 and OPU3. In principle, regression

Table C.1: List of indicator OTUs for the oxic water column condition. Table shows the OUT ID number, indicator value (IV > 0.7 and p-value <0.001), and BLAST-based taxonomic assignment.

Oxic indicator OTUs		
OTU Id	**Indicator value**	**Taxonomy**
1374	0.864	Bacteria--Bacteroidetes--Flavobacteriales--F1CA7--Cc180
6332	0.859	Bacteria--Bacteroidetes--Flavobacteriales--Cytophaga
53435	0.845	Bacteria--Bacteroidetes--Flavobacteriales--Polaribacter
349	0.844	Bacteria--Bacteroidetes--Flavobacteriales--Cytophaga
4900	0.839	Bacteria--Proteobacteria--Alphaproteobacteria--PB21
1183	0.837	No blast hit
33882	0.834	Bacteria--Cyanobacteria--Chloroplasts--Haslea
59827	0.831	Bacteria--Cyanobacteria--Chloroplasts--Dinophysis
2103	0.83	Bacteria--Bacteroidetes--Saprospirales--Saprospiraceae
65004	0.826	No blast hit
43207	0.826	Bacteria--Proteobacteria--Gammaproteobacteria--HTCC2089--obligately_oligotrophic_bacteria_KI89C
16784	0.817	Bacteria--Verrucomicrobia--Opitutae
30319	0.815	Bacteria--Proteobacteria--Alphaproteobacteria--Rhodobacterales--Hyphomonadaceae--marine_bacterium_SCRIPPS_423
2749	0.81	Bacteria--Bacteroidetes--Saprospirales--Saprospiraceae
18569	0.804	Bacteria--Cyanobacteria--Chloroplasts--Haslea
50601	0.791	Bacteria--Proteobacteria--Gammaproteobacteria--Alteromonadales--Alteromonas
42684	0.789	Bacteria--Bacteroidetes--Flavobacteriales--Cytophaga--Psychroserpens_burtonensis--AEGEAN_179
56987	0.789	Bacteria--Proteobacteria--Alphaproteobacteria--Rhodobacterales--Rhodobacter--alpha_proteobacterium_HTCC2255
66645	0.781	No blast hit
6874	0.779	Bacteria--Proteobacteria--Alphaproteobacteria--Rhodobacterales--Rhodobacter--alpha_proteobacterium_HTCC2255
37460	0.778	Bacteria--Bacteroidetes--Flavobacteriales--Polaribacter
159	0.773	No blast hit
58808	0.766	Bacteria--Bacteroidetes--Flavobacteriales--Cytophaga
16176	0.757	No blast hit
26354	0.749	Bacteria--Verrucomicrobia--Opitutae--Fucophilus
52754	0.746	Bacteria--Bacteroidetes--Flavobacteriales
37852	0.737	Bacteria--Proteobacteria--Gammaproteobacteria
9156	0.734	Bacteria--Bacteroidetes--Flavobacteriales--Polaribacter
66354	0.732	No blast hit
28654	0.729	Bacteria--Bacteroidetes--Flavobacteriales--F1CA7--Cc180
2506	0.728	Bacteria--Proteobacteria--Gammaproteobacteria--Pseudomonadaceae--Dichelobacter--Dichelobacter_nodosus
50368	0.724	Bacteria--Bacteroidetes--Flavobacteriales--Polaribacter

analyses can identify predictive variables effect on the response variables, thus resulting negative regression coefficients between OTUs and environmental param may potentially indicate gas (O_2, CH_4 and H_2S) or nutrients (NO_3^- and NO_2^-) consumption by microbial agents. In addition, lack of significant correlation only indicates that a clear linear relation does not exist while statistical dependency can still exist. Although literature references provide evidence supporting our hypothesis to be permissive, correlations were not used in a deterministic fashion. I acknowledge and support the need for developing efforts to test derived statistical hypotheses by *in situ* activity measurements. To determine potential interactions between OTUs throughout the water column, a co-occurrence network was constructed using both Bray-Curtis and Spearman correlation measures. Resulting correlational edges were not considered direct metabolic links in this chapter. Observed OTUs interactions were used to hypothesize potential metabolic interactions, under the consideration that true metabolic interactions do require correlation. Similarly to regression analyses, literature references provide evidence supporting our hypothesis to be permissive. Nevertheless, I support the need for developing efforts to test derived statistical hypotheses by *in situ* microbial community activity measurements.

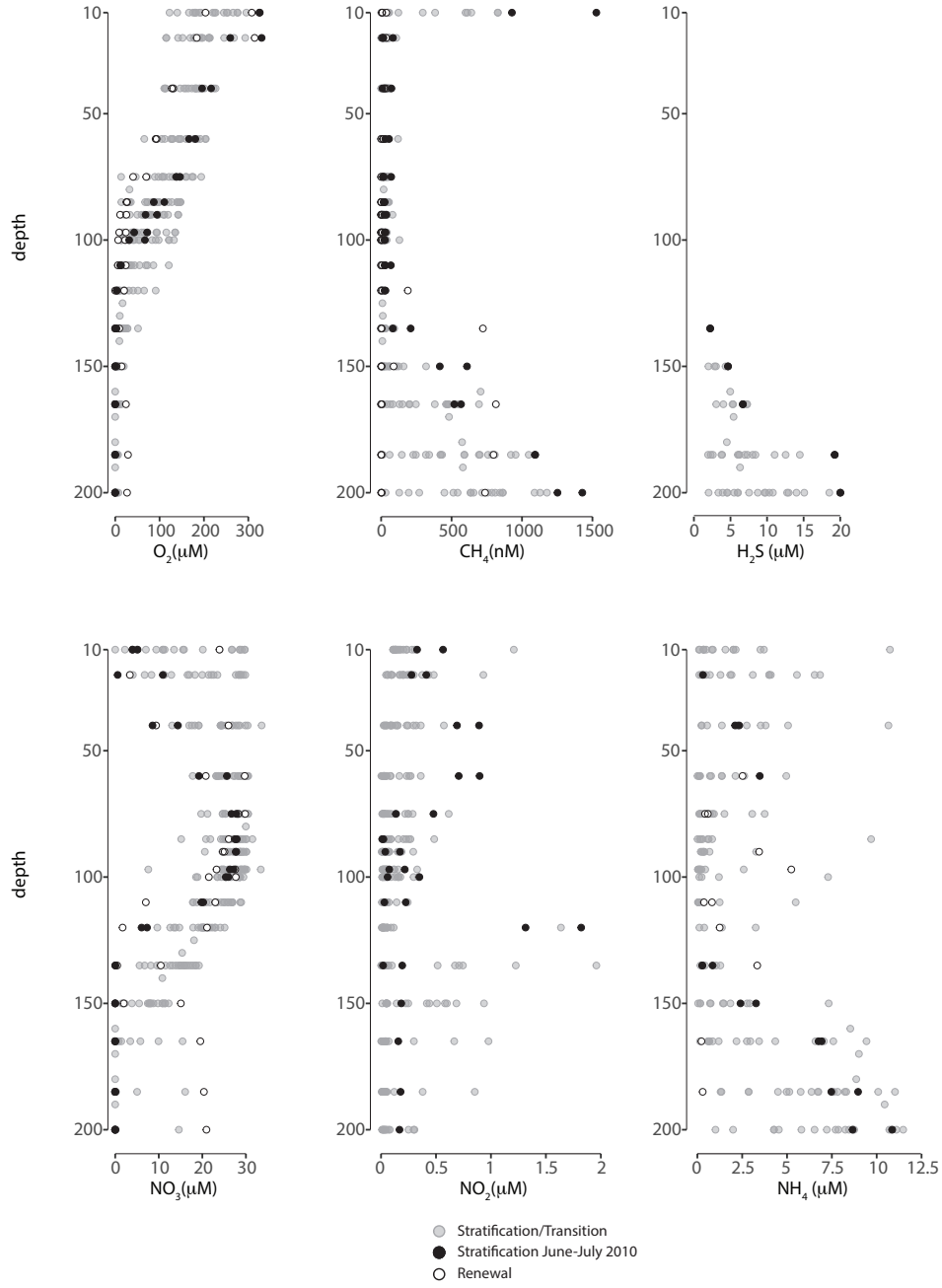

Figure C.1: Time-series chemical data shown as panel dot plots for oxygen (O_2), methane (CH_4), hydrogen sulphide (H_2S), nitrate (NO_3^-), nitrate (NO_2^-), and ammonium (NH_4^+) measurements along the depth profile for samples taken from May 2008 to July 2010 at Station S3 in Saanich Inlet.

Table C.2: List of indicator OTUs for the dysoxic-suboxic, and combined oxic-dysoxic-suboxic water column conditions. Table shows the OUT ID number, indicator value (IV > 0.7 and p-value <0.001), and BLAST-based taxonomic assignment.

Dysoxic-suboxic indicator OTUs			
OTU Id	**Indicator value**	**Taxonomy**	
47823	0.833	Archaea--Thermoplasmata_Eury--marine_group_II	
32415	0.83	Bacteria--Verrucomicrobia--Verrucomicrobia_subdivision_3--Verruco-3--CTD005-1B-02--ctg_NISA103	
24995	0.818	Bacteria--WS3	
66586	0.814	Bacteria--Proteobacteria--Deltaproteobacteria--NB1-j--NB1-i--JTB38	
48948	0.808	Bacteria--Proteobacteria--Deltaproteobacteria--PB19	
9662	0.805	Bacteria--Proteobacteria--Gammaproteobacteria	
14518	0.775	Bacteria--Proteobacteria--Deltaproteobacteria--NB1-j--NB1-i--JTB38	
53548	0.77	No blast hit	
53617	0.769	Bacteria--Proteobacteria--Deltaproteobacteria--Myxococcales--OM27--CTD005-73B-02	
50999	0.767	Bacteria--Verrucomicrobia--Verrucomicrobiae	
55333	**0.761**	**Bacteria--Proteobacteria--Gammaproteobacteria--Methylococcales**	**OPU1_01**
61434	0.76	Bacteria--Planctomycetes--WPS-1--BD2-16	
44989	0.746	No blast hit	
64275	0.741	Bacteria--Planctomycetes--agg27--CL500-15--B83	
1588	0.739	Archaea--Thermoplasmata_Eury--marine_group_II	
48598	0.732	Bacteria--Proteobacteria--Gammaproteobacteria--B2M28	
26820	0.703	Bacteria--Planctomycetes--agg27--OM190--ARKCH2Br2-76	
21744	0.702	Bacteria--Planctomycetes--Planctomycetacia	
34786	0.7	Bacteria--Chlamydiae--Parachlamydiaceae--	
54690	0.672	Bacteria--Planctomycetes--agg27--CL500-15--B83	
59977	0.656	Bacteria--Planctomycetes--Planctomycetacia	
37732	0.655	Bacteria--Proteobacteria--Deltaproteobacteria--PB19	
4737	0.635	Bacteria--Planctomycetes--agg27--OM190--ARKCH2Br2-76	
47114	0.591	Bacteria--Proteobacteria--Alphaproteobacteria--Consistiales--Pelagibacter--SAR11--Candidatus_Pelagibacter	

Combined oxic and dysoxic-suboxic indicator OTUs			
OTU Id	**Indicator value**	**Taxonomy**	
5987	0.983	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05	
64980	0.977	Bacteria--Proteobacteria--Alphaproteobacteria--Oleomonas	
65868	0.976	Bacteria--Marine_group_A--Arctic96B-7--	
7225	0.975	Bacteria--Marine_group_A--Arctic96B-7--	
63490	0.972	Archaea--Thermoplasmata_Eury--marine_group_II--SB95-72	
27691	0.97	Bacteria--Marine_group_A--SAR406--	
43578	0.969	Bacteria--Proteobacteria--Alphaproteobacteria--Sphingomonadales--Porphyrobacter	
5522	0.963	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum	
8199	0.963	Bacteria--Verrucomicrobia--Verrucomicrobiae	
26598	0.962	Bacteria--WS3	
26616	0.96	Bacteria--Proteobacteria--Deltaproteobacteria--Nitrospina	
10232	0.96	Bacteria--Bacteroidetes--Bacteroidales--VC21_Bac22	
1634	0.956	Bacteria--Marine_group_A--Arctic95A-2--	
18743	0.954	Bacteria--Proteobacteria--Alphaproteobacteria--Oleomonas--ctg_NISA150	
20751	**0.953**	**Bacteria--Proteobacteria--Gammaproteobacteria--Methylococcales**	**OPU3 cluster**
8461	0.942	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum	
30687	0.942	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c	
4524	0.939	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum	
30690	0.936	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum	
12901	0.935	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum	
69037	0.931	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum	
36890	0.928	Bacteria--Proteobacteria--Alphaproteobacteria	
13690	0.924	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c	
10699	0.922	Bacteria--Planctomycetes--agg27--OM190--ARKCH2Br2-76	
27647	0.921	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c	
47960	0.918	Bacteria--Proteobacteria--Alphaproteobacteria--Oleomonas	
34377	0.915	Bacteria--Proteobacteria--Alphaproteobacteria--S23_91	
66151	0.915	Bacteria--Proteobacteria--Alphaproteobacteria--Azospirillales	
35678	0.914	Archaea	
36491	0.914	Bacteria--Marine_group_A--SAR406	
17339	0.912	Archaea--Thermoplasmata_Eury--marine_group_II	
64913	0.907	No blast hit	
10140	0.903	Bacteria--Marine_group_A--SAR406	
22155	0.898	Bacteria--Planctomycetes--Planctomycetacia--Planctomycetales--H28--A24	
15539	0.894	Archaea--Thermoprotei_Cren--Thermoproteaceae--Vulcanisaeta	
22391	0.89	Bacteria--Chloroflexi--Chloroflexi-4--PAWS52f	
49225	0.889	Bacteria--ABY1_OD1--FW129	
3791	0.888	Bacteria--Proteobacteria--Alphaproteobacteria--Azospirillales	
32241	0.887	Bacteria--Proteobacteria--Deltaproteobacteria	
20039	0.886	Bacteria--Proteobacteria--Alphaproteobacteria--Sphingomonadales	
1796	0.88	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05--mussel_thioautotrophic_gill_symbiont_MAR1	
31622	0.876	Bacteria--Marine_group_A--A714017	
66501	0.874	Bacteria--Planctomycetes--Anammoxales	
39693	**0.872**	**Bacteria--Proteobacteria--Gammaproteobacteria--methanotrophic_gill_symbiont_of_mussel**	**Symbiont_01**
66119	0.867	Bacteria--Planctomycetes--agg27--OM190--ARKCH2Br2-76	

Table C.3: List of indicator OTUs for the anoxic, and combined dysoxic-suboxic and anoxic water column conditions. Table shows the OUT ID number, indicator value (IV > 0.7 and p-value <0.001), and BLAST-based taxonomic assignment.

Anoxic indicator OTUs		
OTU Id	**Indicator value**	**Taxonomy**
54827	0.975	Bacteria--Proteobacteria--Deltaproteobacteria--Desulfobacteraceae--Desulfobacter
348	0.974	Bacteria--Chloroflexi--Anaerolineae--Anaerolineales--envOPS12
19198	0.954	Bacteria--Proteobacteria--Gammaproteobacteria--BD7-8
43339	0.952	Bacteria--Proteobacteria--Deltaproteobacteria--OPB16
2834	0.94	Bacteria--Chloroflexi--Anaerolineae--Anaerolineales
64918	0.939	Bacteria--Planctomycetes--agg27--CL500-15--B83
35767	0.932	Bacteria--Proteobacteria--Deltaproteobacteria--Geobacter--Pelobacter_propionicus
57339	0.93	Archaea
9382	0.92	Archaea--pMC2A15
51517	0.915	Bacteria--Chloroflexi--Anaerolineae--Anaerolineales--PpalmC37
6820	0.914	Archaea
56521	0.911	Bacteria--Planctomycetes--Kueneniaceae--Scalindua
5588	0.908	Bacteria--Proteobacteria--Gammaproteobacteria--BD7-8
36238	0.904	Bacteria--Proteobacteria--Gammaproteobacteria--Alteromonadales
3942	0.902	Bacteria--Proteobacteria--Gammaproteobacteria--OM60
41410	0.9	Bacteria--ABY1_OD1--FW129--KNA6-NB12
3054	0.893	Archaea
49856	0.891	Archaea--DHVE3
68112	0.888	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c
57243	0.885	Bacteria--SR1--BD2-14--
37995	0.881	Archaea--pISA1
66746	0.878	Bacteria--Proteobacteria--Gammaproteobacteria--Alteromonadales
34039	0.876	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05
26172	0.873	Bacteria--Proteobacteria--Alphaproteobacteria--TSBb13
30361	0.873	Archaea--DHVE3
34904	0.872	Bacteria--Proteobacteria--Gammaproteobacteria
38431	0.872	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05--mussel_thioautotrophic_gill_symbiont_MAR1
58016	0.872	Bacteria--Planctomycetes
33942	0.871	Bacteria--ABY1_OD1--FW129--KNA6-NB12
8452	0.87	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum--pIVWA5
63742	0.869	Archaea--DHVE3
47479	0.869	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05
39916	0.868	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05
56355	0.865	Archaea
8870	0.862	Archaea--pISA1
27985	0.862	Bacteria--ABY1_OD1
2851	0.862	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05
39944	0.86	Archaea--DHVE3
44232	0.854	Bacteria--OP11--WCHB1-64--
31774	0.853	Archaea--BC07-2A-27
65422	0.852	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05--mussel_thioautotrophic_gill_symbiont_MAR1
24819	0.85	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05--mussel_thioautotrophic_gill_symbiont_MAR1
58861	0.847	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05
5015	0.845	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05
43793	0.842	Archaea--DHVE3
44286	0.836	Bacteria--ABY1_OD1--FW129--KNA6-NB12
48588	0.836	Bacteria--ZB2--BD5-13
39956	0.823	Archaea--DHVE3
11314	0.817	Bacteria--Marine_group_A--Arctic95A-2
37272	0.812	Bacteria--WS3
38607	0.807	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05--mussel_thioautotrophic_gill_symbiont_MAR1
8351	0.805	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05
674	0.805	Bacteria--Chloroflexi--TK17
20928	0.804	Bacteria--WS3
25929	0.804	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05

Table C.4: Complete list of interactions between methanotroph and bacterial OTUs based on Bray-Curtis and Spearman's rank correlations showed in sub-networks depicted as hive panels.

Source		Target
OPU1		OTU-34348
OPU1		OTU-39407
OPU1		OTU-49274
OTU-20592	Bacteria--Planctomycetes--Planctomycetacia--Unclassified--OTU_20592	OPU1
OTU-66501	Bacteria--Planctomycetes--Anammoxales--OTU_66501	OPU1
OPU3		OPU3
OTU-10140	Bacteria--Marine_group_A--SAR406--OTU_10140	OPU3
OTU-13046	Bacteria--VHS-B5-50--OTU_13046	OPU3
OTU-13690	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c--OTU_13690	OPU3
OTU-14007	Bacteria--Proteobacteria--Gammaproteobacteria--B2M28--OTU_14007	OPU3
OTU-20039	Bacteria--Proteobacteria--Alphaproteobacteria--Sphingomonadales--Unclassified--OTU_20039	OPU3
OTU-2119	Bacteria--Planctomycetes--Planctomycetacia--Planctomycetales--Unclassified--OTU_2119	OPU3
OTU-22391	Bacteria--Chloroflexi--Chloroflexi-4--PAWS52f--OTU_22391	OPU3
OTU-27647	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c--OTU_27647	OPU3
OTU-30687	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c--OTU_30687	OPU3
OTU-31873	Bacteria--Proteobacteria--Alphaproteobacteria--Oleomonas--ctg_NISA150--OTU_31873	OPU3
OTU-34377	Bacteria--Proteobacteria--Alphaproteobacteria--S23_91--OTU_34377	OPU3
OTU-36491	Bacteria--Marine_group_A--SAR406--OTU_36491	OPU3
OTU-39766	Bacteria--Proteobacteria--Gammaproteobacteria--Arctic96B-1--Gammaproteobacteria--OTU_39766	OPU3
OTU-4300	Bacteria--Verrucomicrobia--Verrucomicrobia_subdivision_3--Verruco-3--CTD005-1B-02--Unclassified--OTU_4300	OPU3
OTU-47823	Archaea--Thermoplasmata_Eury--marine_group_II--Unclassified--OTU_47823	OPU3
OTU-52929	Bacteria--Chloroflexi--Chloroflexi-4--PAWS52f--OTU_52929	OPU3
OTU-59979	Archaea--Thermoplasmata_Eury--marine_group_II--CTD005-13A--OTU_59979	OPU3
OTU-60749	Bacteria--Proteobacteria--Alphaproteobacteria--Sphingomonadales--Porphyrobacter--Unclassified--OTU_60749	OPU3
OTU-63234	Bacteria--Proteobacteria--Alphaproteobacteria--Unclassified--OTU_63234	OPU3
OTU-6504	Bacteria--Proteobacteria--Gammaproteobacteria--Unclassified--OTU_6504	OPU3
OTU-65093	Bacteria--Bacteroidetes--Flavobacteriales--Cytophaga--Psychroserpens_burtonensis--AEGEAN_129--OTU_65093	OPU3
OTU-66501	Bacteria--Planctomycetes--Anammoxales--OTU_66501	OPU3
OTU-66586	Bacteria--Proteobacteria--Deltaproteobacteria--NB1-j--NB1-i--JTB38--OTU_66586	OPU3
OPU3		OTU-10699
OPU3		OTU-17339
OPU3		OTU-23655
OPU3		OTU-26616
OPU3		OTU-36890
OPU3		OTU-39407
OPU3		OTU-43578
OPU3		OTU-44559
OPU3		OTU-45050
OPU3		OTU-59671
OPU3		OTU-63490
OPU3		OTU-7709
OPU3		OTU-8199
Symbiont		OTU-38607
Symbiont		OTU-26616
Symbiont		OTU-8461
Symbiont		OTU-8452
Symbiont		OTU-8351
Symbiont		OTU-34039
Symbiont		OTU-18743
Symbiont		OTU-8199
Symbiont		OTU-26172
Symbiont		OTU-3054
Symbiont		OTU-13337

Table C.5: BLAST-based comparison for methanotroph sub-networks OTUs with SSU rRNA sequences obtained from methane microcosm experiment on Saanich Inlet waters (Sauter *et al*., 2012), GreenGenes BLAST-based taxonomic assignment, and indicator value obtained from ISA.

Methanotroph group	OTU ID	Similarity (%)	Hit microcosm SSU rRNA sequences	Taxonomic affiliation	Indicator Value	Representative sequence in methanotrophs tree
Symbiont_01	39693	93	JN172291			FJ264570
OPU1_01	55333	90	JN172321			HQ673216
OPU3_01	6504	92	JN172144			HQ673315
	30361	-	-	Archaea--DHVE3	0.873	
	39944	-	-	Archaea--DHVE3	0.86	
	39956	-	-	Archaea--DHVE3	0.823	
	43793	-	-	Archaea--DHVE3	0.842	
	37995	-	-	Archaea--pISA1	0.881	
	9382	-	-	Archaea--pMC2A15	0.92	
	8452	-	-	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum--pIVWA5	0.87	
	8461	-	-	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum	0.942	
	12901	-	-	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum	0.935	
	30690	-	-	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum	0.936	
	5522	-	-	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum	0.963	
	69037	-	-	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum	0.931	
	17339	-	-	Archaea--Thermoplasmata_Eury--marine_group_II	0.912	
	47823	-	-	Archaea--Thermoplasmata_Eury--marine_group_II	0.833	
	59979	-	-	Archaea--Thermoplasmata_Eury--marine_group_II--CTD005-13A	0.829	
	63490	-	-	Archaea--Thermoplasmata_Eury--marine_group_II--SB95-72	0.972	
	63490	-	-	Archaea--Thermoplasmata_Eury--marine_group_II--SB95-72	0.972	
	24727	-	-	Archaea--Thermoprotei_Cren--Aeropyraceae	0.835	
	15539	-	-	Archaea--Thermoprotei_Cren--Thermoproteaceae--Vulcanisaeta	0.894	
	3054	-	-	Archaea	0.893	
	35678	-	-	Archaea	0.914	
	6820	-	-	Archaea	0.914	
	13337	-	-	Bacteria--ABY1_OD1--FW129	0.795	
	49225	-	-	Bacteria--ABY1_OD1--FW129	0.889	
	33942	-	-	Bacteria--ABY1_OD1--FW129--KNA6-NB12	0.871	
	41410	-	-	Bacteria--ABY1_OD1--FW129--KNA6-NB12	0.9	
	44286	-	-	Bacteria--ABY1_OD1--FW129--KNA6-NB12	0.836	
	27985	-	-	Bacteria--ABY1_OD1	0.862	
	10232	88	JN172357	Bacteria--Bacteroidetes--Bacteroidales--VC21_Bac22	0.96	
	34348	92	JN172437	Bacteria--Bacteroidetes--Flavobacteriales--Cytophaga--Chl112--OTU_34348	-	
	65093	94	JN172444	Bacteria--Bacteroidetes--Flavobacteriales--Cytophaga--Psychroserpens_burtonensis--AEGEAN_129--OTU_65093	-	
	3490	-	-	Bacteria--Chloroflexi--Anaerolineae--A4b--TK32	0.819	
	22391	-	-	Bacteria--Chloroflexi--Chloroflexi-4--PAWS52f	0.89	
	22391	-	-	Bacteria--Chloroflexi--Chloroflexi-4--PAWS52f	0.89	
	52929	-	-	Bacteria--Chloroflexi--Chloroflexi-4--PAWS52f--OTU_52929	-	
	674	-	-	Bacteria--Chloroflexi--TK17	0.805	
	31622	81	JN172126	Bacteria--Marine_group_A--A714017	0.876	
	11314	86	JN172337	Bacteria--Marine_group_A--Arctic95A-2	0.817	
	1634	81	JN172321	Bacteria--Marine_group_A--Arctic95A-2	0.956	
	65868	81	JN172337	Bacteria--Marine_group_A--Arctic96B-7	0.976	
	7225	82	JN172337	Bacteria--Marine_group_A--Arctic96B-7	0.975	
	27691	82	JN172126	Bacteria--Marine_group_A--SAR406	0.97	
	10140	80	JN172321	Bacteria--Marine_group_A--SAR406	0.903	
	36491	-	-	Bacteria--Marine_group_A--SAR406	0.914	
	46891	81	JN172337	Bacteria--Marine_group_A--SAR406--OTU_46891	-	
	37274	-	-	Bacteria--OP11	0.746	
	64918	81	JN172153	Bacteria--Planctomycetes--agg27--CL500-15--B83	0.939	
	10699	82	JN172152	Bacteria--Planctomycetes--agg27--OM190--ARKCH2Br2-76	0.922	
	10699	82	JN172152	Bacteria--Planctomycetes--agg27--OM190--ARKCH2Br2-76	0.922	
	59671	84	JN172152	Bacteria--Planctomycetes--agg27--OM190--ARKCH2Br2-76--OTU_59671	-	
	66501	83	JN172152	Bacteria--Planctomycetes--Anammoxales	0.874	
	66501	83	JN172152	Bacteria--Planctomycetes--Anammoxales	0.874	
	56521	82	JN172152	Bacteria--Planctomycetes--Kueneniaceae--Scalindua	0.911	
	2119	82	JN172152	Bacteria--Planctomycetes--Planctomycetacia--Planctomycetales	0.816	
	20592	82	JN172313	Bacteria--Planctomycetes--PlanctomycetaciaOTU_20592	-	
	58016	82	JN172153	Bacteria--Planctomycetes	0.872	
	25015	89	JN172210	Bacteria--Proteobacteria--Alphaproteobacteria	0.743	
	3945	92	JN172189	Bacteria--Proteobacteria--Alphaproteobacteria	0.802	
	36890	92	JN172130	Bacteria--Proteobacteria--Alphaproteobacteria	0.928	
	44559	90	JN172130	Bacteria--Proteobacteria--Alphaproteobacteria	0.849	
	63234	91	JN172189	Bacteria--Proteobacteria--Alphaproteobacteria	0.856	
	66151	91	JN172242	Bacteria--Proteobacteria--Alphaproteobacteria--Azospirillales	0.915	
	3791	89	JN172189	Bacteria--Proteobacteria--Alphaproteobacteria--Azospirillales	0.888	
	18743	91	JN172130	Bacteria--Proteobacteria--Alphaproteobacteria--Oleomonas--ctg_NISA150	0.954	
	31873	89	JN172145	Bacteria--Proteobacteria--Alphaproteobacteria--Oleomonas--ctg_NISA150	0.721	
	47960	92	JN172145	Bacteria--Proteobacteria--Alphaproteobacteria--Oleomonas	0.918	
	64980	91	JN172145	Bacteria--Proteobacteria--Alphaproteobacteria--Oleomonas	0.977	
	34377	92	JN172189	Bacteria--Proteobacteria--Alphaproteobacteria--S23_91	0.915	
	34377	92	JN172189	Bacteria--Proteobacteria--Alphaproteobacteria--S23_91	0.915	
	43578	95	JN172337	Bacteria--Proteobacteria--Alphaproteobacteria--Sphingomonadales--Porphyrobacter	0.969	
	43578	95	JN172337	Bacteria--Proteobacteria--Alphaproteobacteria--Sphingomonadales--Porphyrobacter	0.969	
	60749	94	JN172337	Bacteria--Proteobacteria--Alphaproteobacteria--Sphingomonadales--PorphyrobacterOTU_60749	-	
	20039	92	JN172145	Bacteria--Proteobacteria--Alphaproteobacteria--Sphingomonadales	0.886	
	26172	91	JN172189	Bacteria--Proteobacteria--Alphaproteobacteria--TSBb13	0.873	
	68112	91	JN172450	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c	0.888	
	13690	91	JN172145	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c	0.924	
	15060	92	JN172145	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c	0.756	
	30687	92	JN172145	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c	0.942	
	60274	92	JN172130	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c	0.737	
	9875	93	JN172145	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c	0.819	
	13690	91	JN172145	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c	0.924	
	27647	90	JN172130	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c	0.921	
	30687	92	JN172145	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c	0.942	
	23655	92	JN172130	Bacteria--Proteobacteria--Alphaproteobacteria--OTU_23655	-	
	32241	84	JN172404	Bacteria--Proteobacteria--Deltaproteobacteria	0.887	
	35767	-	-	Bacteria--Proteobacteria--Deltaproteobacteria--Geobacter--Pelobacter_propionicus	0.932	

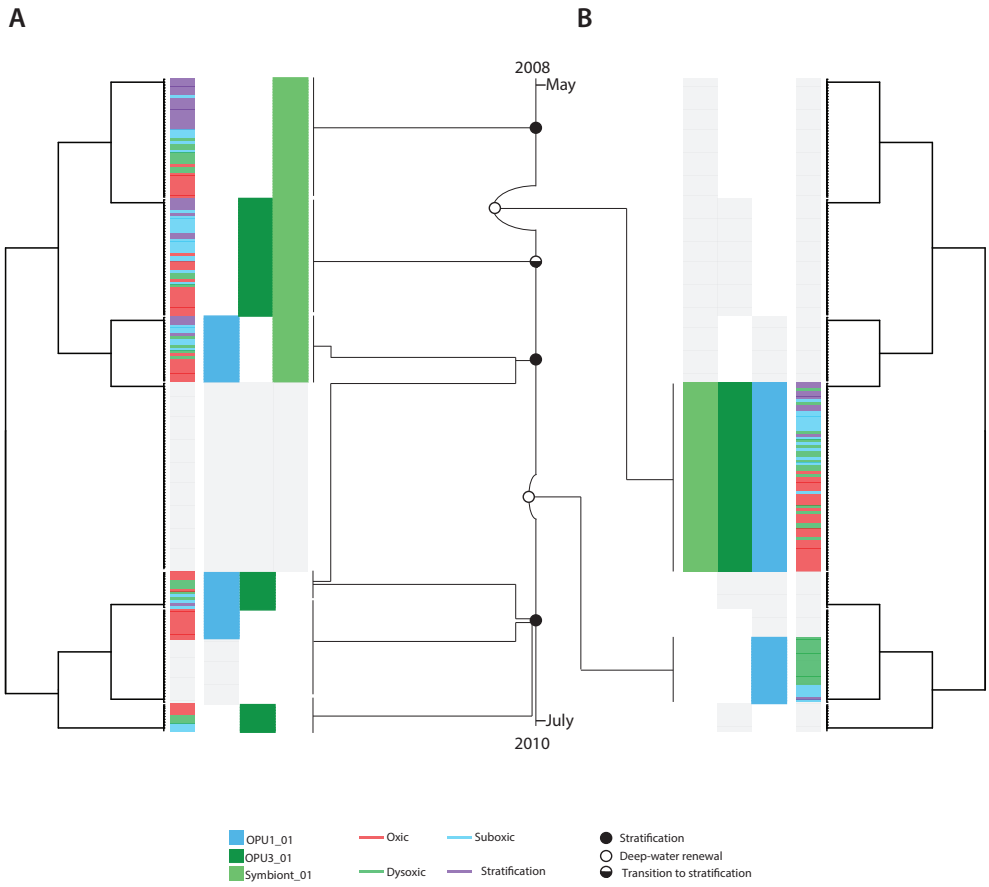

Figure C.2: Partitioning of methanotrophic OPU1, OPU3 and methanotrophic symbiont OTUs. Hierarchical clustering depict distribution patterns observed based on Manhattan distance associated with water column O_2 conditions (red= oxic, dysoxic= green, suboxic= light blue, and anoxic/sulfidic = purple) over: A) contrasting stratification periods and after deep-water renewal B) deep-water renewal perturbation OTUs presence in samples is depicted as colored rectangles according to phylogenetic affiliation as shown in key (OPU1_01 = blue, OPU3_01= dark green, and symbiont_01 = light green).

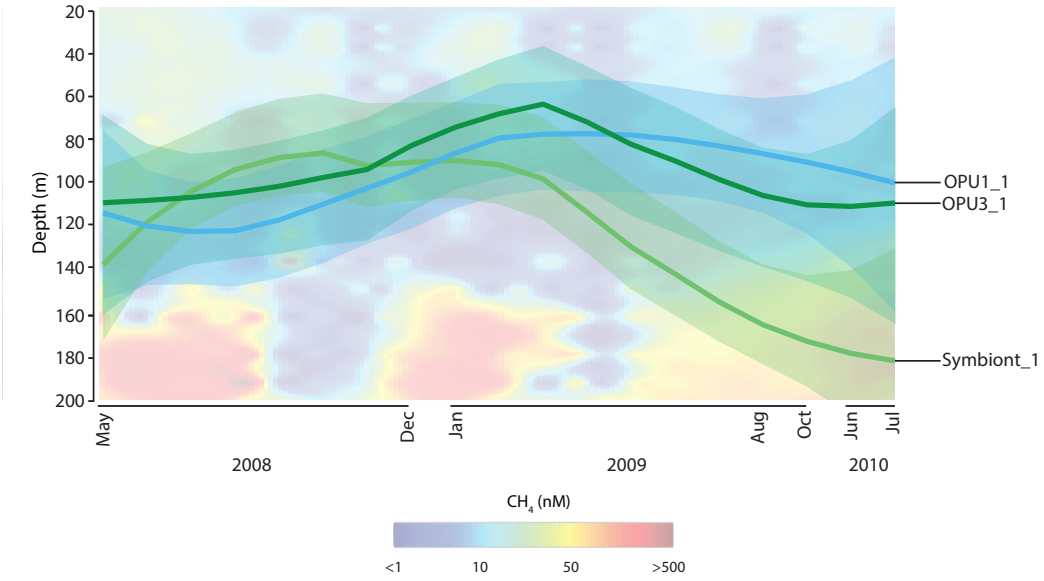

Figure C.3: Methane contour plot for gas concentration (nM) data throughout water column from May 2008 to July 2010. Overlapped is shown the methanotroph OTUs distribution mean trend and one-standard deviation throughout water column over time.

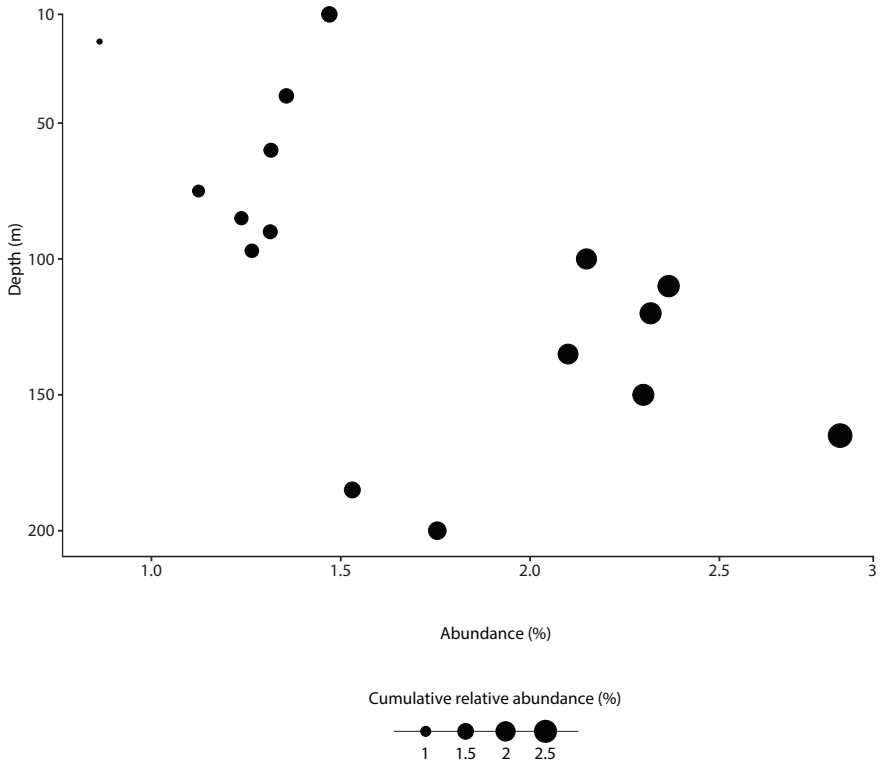

Figure C.4: Cumulative distribution of methanotroph OTUs throughout Saanich Inlet water column. Dots are sized based on the cumulative relative abundance (%) of OTUs affiliated with OPU1, OPU3 and mussel symbionts over time for a given depth.

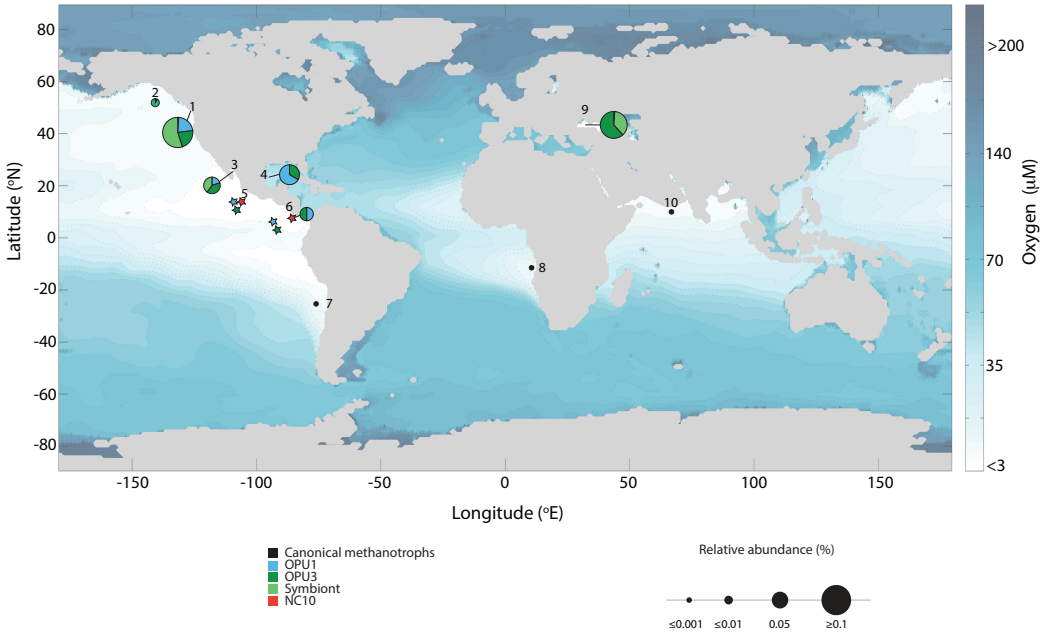

Figure C.5: Distribution and abundance of methanotrophic bacteria in OMZs. Open ocean and coastal OMZs, and enclosed basins from where methanotroph SSU rRNA full length or pyrotags sequences were obtained (reviewed in (Dick *et al*., 2013); (1) Saanich Inlet (Zaikova *et al*., 2010), (2) Northeastern Subarctic Pacific Ocean (Line P), (3) Guaymas Basin (Dick and Tebo, 2010), (4) Gulf of Mexico (Kessler *et al*., 2011), (6) Costa Rica (Tavormina *et al*., 2013), (7) Eastern Tropical South Pacific off Chile (Stevens and Ulloa, 2008), (8) Namibian upwelling (Lam *et al*., 2009), (9) Black Sea (Glaubitz *et al*., 2010; Fuchsman *et al*., 2011), and (10) Arabian Sea (Lke *et al*., 2016)) to be clustered with Saanich Inlet representative SSU rRNA and pyrotag sequences are indicated on the map. Pie charts size depicts the total relative abundance of the number of sequences clustering at 97% similarity with Canonical methanotrophs, OPU1, OPU3 and Methanotrophic symbionts. First environmental observation for OPU1 and OPU3 in the Eastern North Pacific OMZ (5) based on particulate monooxygenase (*pmoA*) gene (Hayashi *et al*., 2007), and for candidate NC10 Methylomirabilis oxyfera in the Eastern Tropical North Pacific OMZ off Mexico and Costa Rica (Padilla *et al*., 2016) are marked with a star.

Appendix D

Community-level interactions support methane oxidation in Saanicn Inlet oxygen-deficient water column

D.1 Supplementary Methods

D.1.1 Co-occurrence metabolic network properties

To determine potential metabolic interactions between microbial groups throughout the dysoxic-suboxic water column compartments, a co-occurrence network was constructed using both Bray-Curtis and Spearman correlation measures (Fig. 5). All statistically significant correlations among pathways resulting after permutations and bootstrap score distributions were included in the network. Each node represents a pathways and each edge a statistically significant positive correlation indicating co-occurrence. The resulting network contains 228 nodes, connected by 5,054 edges. Average node degree (mean edges per node) (Proulx *et al*.,2005) was 2.25, the average path length (the expected distance between two connected nodes) (Latora and Marchiori, 2009) was 2.725, and the network diameter (longest path between two nodes) (Cardoso *et al*., 2009) was equal to 9. The clustering coefficient (connectedness of a node's neighbour) (Proulx *et al*., 2005) was 0.725, and connectance (proportion of all possible links realized) (Dunne *et al*., 2002) was 0.833. Resulting edges were considered potential direct metabolic links under the consideration that true metabolic interactions are likely correlated. Metabolic interactions were supported by BLAST-compared amino acid sequences of ORFs within environmental pathways and incubation functional information. Combined functional and taxonomic information was used to determine microbial community metabolic interactions in relation to CH_4 oxidation, Carbon fixation, and Nitrogen metabolism.

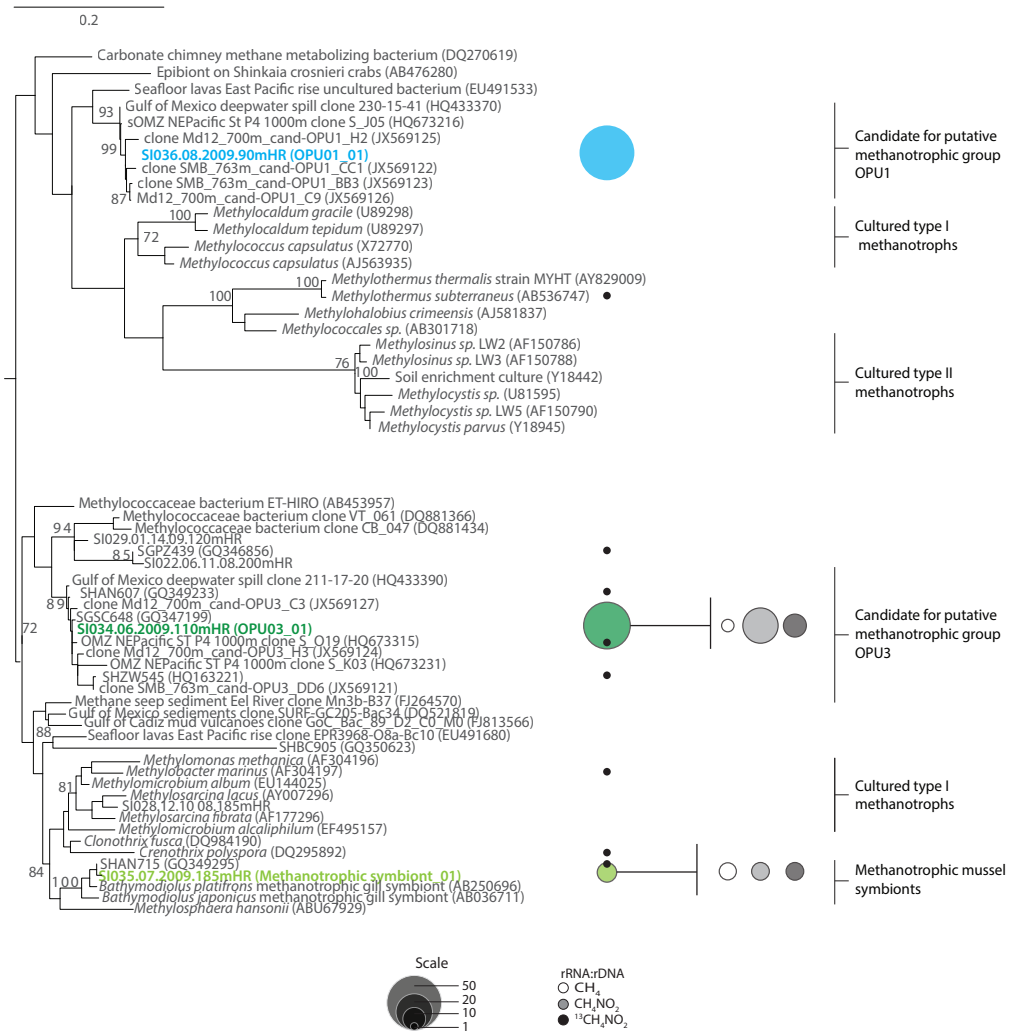

Figure D.1: Phylogenetic relationships between potentially active Methylococcaceae OTUs found in treatments. SSU tree inferred using maximum likelihood implemented in PHYML. The percentage (>70%) of replicates in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. Reference sequences for lineages are shown in black. Representative sequences obtained for pMMO-encoding groups OPU1, OPU3 and Symbiont are shown in light blue, dark green and light green, respectively. OTUs rRNA: rDNA is depicted as dots whose circumference represents cumulative ratio value among incubation treatments ($^{12}CH_4$ = open circle, $^{12}CH_4$ + NO_2^- = gray circle, and $^{13}CH_4$ + NO_2^- = solid circle).

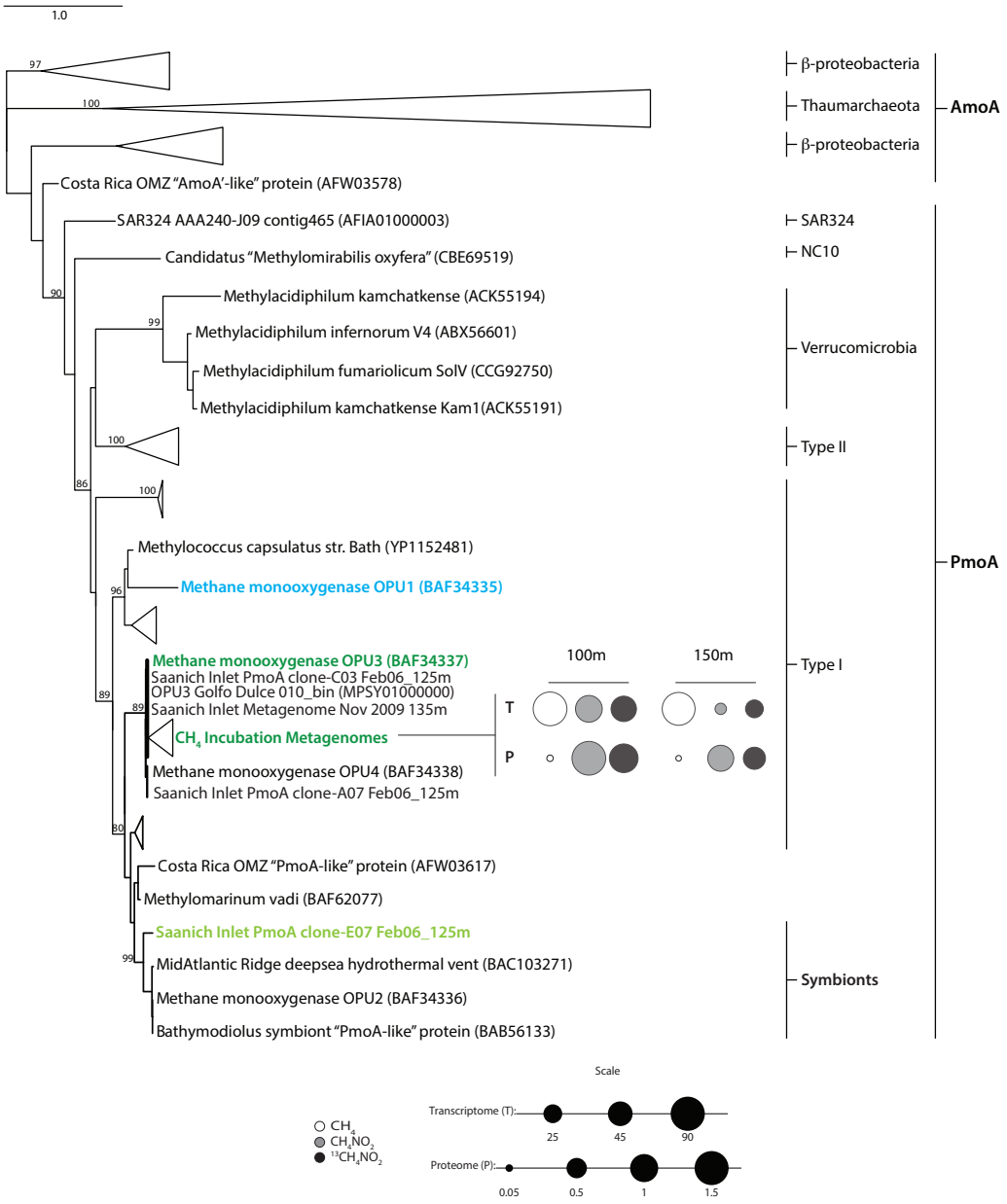

Figure D.2: Particulate methane monooxygenase subunit β (*pmoA*) phylogenetic tree. Topography was inferred using maximum likelihood on PmoA and AmoA amino acid sequences from reference sequences including pMMO-encoding groups OPU1 and OPU3, and symbionts. Bootstrap values (%) are based on 1000 replicates and are shown for branches with greater than 70% support. The scalebar represents 1.0 substitutions per site. PmoA abudance throughout incubation treatments per depth is depicted as dots whose size represents the RPKM or NSAF value for metatranscriptomic (T) and metaproteomic (P) datasets. Sample treatments are depicted as symbols: $^{12}CH_4$ = open circle, $^{12}CH_4$ + NO_2^-- = gray circle, and $^{13}CH_4$ + NO_2^-- =solid circle.

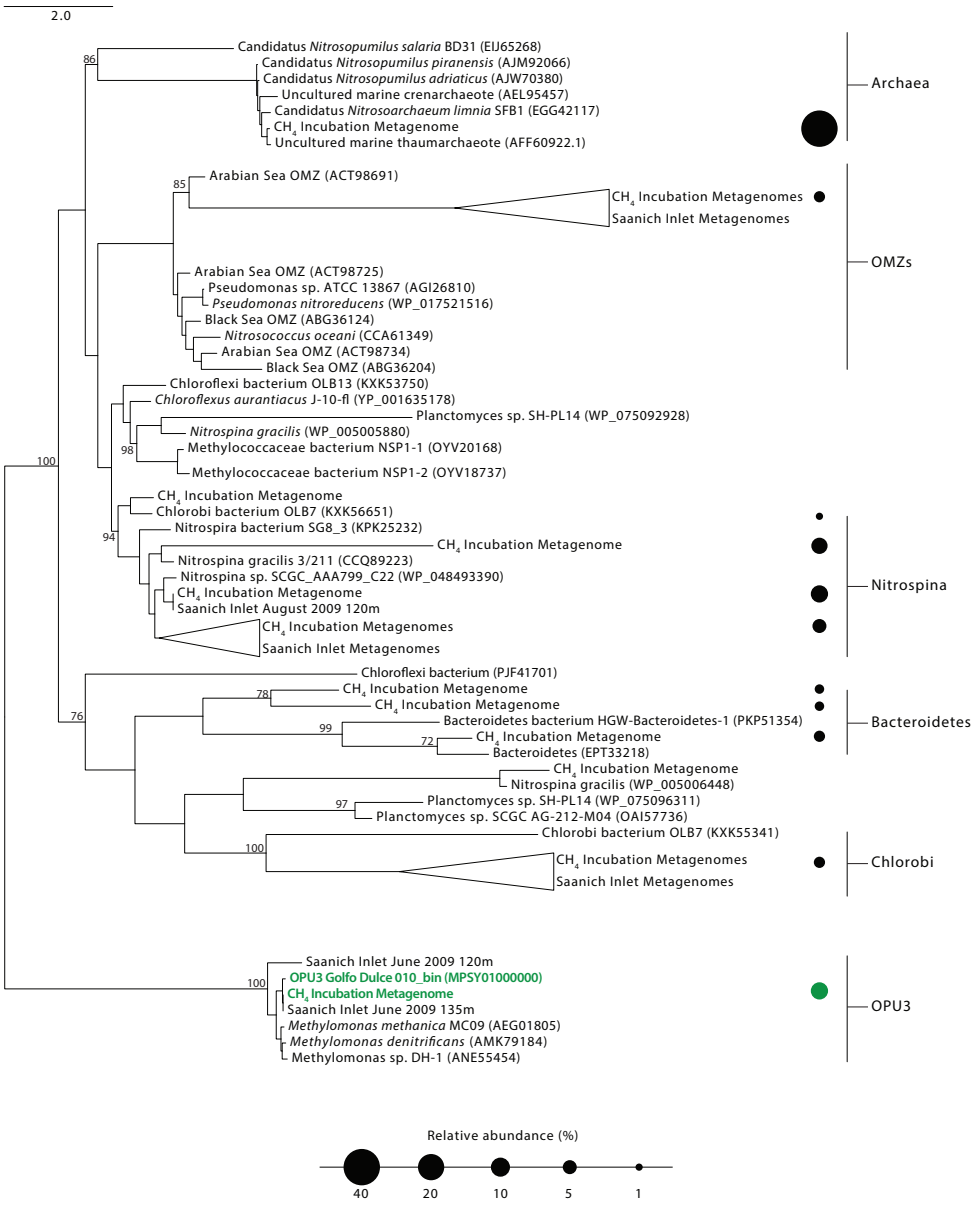

Figure D.3: Copper-containing nitrite reductase (nirK) phylogenetic tree. Topography was inferred using maximum likelihood on NirK amino acid sequences from reference sequences including pMMO-encoding group OPU3. Bootstrap values (%) are based on 1000 replicates and are shown for branches with greater than 70% support. The scalebar represents 2.0 substitutions per site. NirK abudance throughout incubation treatments is depicted as dots whose size represents the relative abundance (%) from the total number of annotated ORFs in the metagenomic datasets.

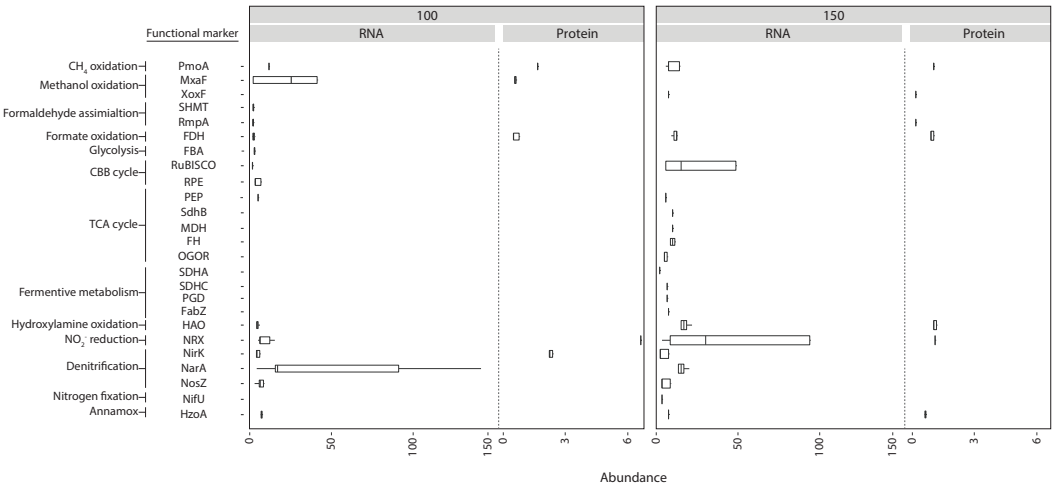

Figure D.4: Difference in transcript abundance between ^{12}C and ^{13}C incubation treatments. Abundance for transcripts related to key proteins in CH_4 oxidation, C1 metabolism and carbon incorporation, and Nitrogen metabolism pathways. The size of each box represents the relative abundance (RPKM) for each transcript in the ^{12}C and ^{13}C treatment sample of a given depth. Extended dashed lines (whiskers) represent at the base the lower and upper quartiles (25% and 75%) and at the end the minimum and maximum values encountered. The middle line represents the median.

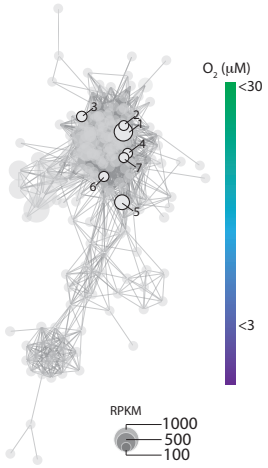

Figure D.5: Co-occurrence network of correlating pathways at dysoxic-suboxic water column conditions in the Saanich Inlet. Network was constructed based on Bray-Curtis and Spearman rank correlations between predicted pathways from time-resolved metagenomic data (June 2009 -August 2011). Nodes depict metabolic pathways, and edges positive significant (correlation value > 0.6, $p< 0.001$) correlations. Nodes are sized based on RPKM values. Nodes numbered according to functional annotation (1=CH_4 oxidation, 2= methanol oxidation to formaldehyde, 3= formaldehyde oxidation to formate via H4MPT pathway, 4=formate oxidation to CO_2, 5= CO_2 fixation via TCA pathway, 6= Ammonium oxidation, and 7= NO_2^--reduction).

Table D.1: BLAST-based comparison for incubation subnetworks OTUs with 16S rRNA sequences obtained time-resolved co-occurrence network anlaysis (Torres-Beltran *et al*., 2016).

Time-series reference OTU Id	Taxonomic affiliation	OTU Id
35678	Archaea--Unclassified--OTU_35678	1014
7709	Bacteria--Proteobacteria--Deltaproteobacteria--Nitrospina--OTU_7709	10840
3791	Bacteria--Proteobacteria--Alphaproteobacteria--Azospirillales--Unclassified--OTU_3791	10864
6504	**OPU3**	11160
66151	Bacteria--Proteobacteria--Alphaproteobacteria--Azospirillales--Unclassified--OTU_66151	11587
36491	Bacteria--Marine_group_A--SAR406--OTU_36491	11882
26616	Bacteria--Proteobacteria--Deltaproteobacteria--Nitrospina--OTU_26616	13887
22391	Bacteria--Chloroflexi--Chloroflexi-4--PAWS52f--OTU_22391	1422
39693	**Symbiont**	14587
65093	Bacteria--Bacteroidetes--Flavobacteriales--Cytophaga--Psychroserpens_burtonensis--AEGEAN_129--OTU_65093	14685
20039	Bacteria--Proteobacteria--Alphaproteobacteria--Sphingomonadales--Unclassified--OTU_20039	14806
25015	Bacteria--Proteobacteria--Alphaproteobacteria--Unclassified--OTU_25015	14923
13690	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c--OTU_13690	1526
23655	Bacteria--Proteobacteria--Alphaproteobacteria--Unclassified--OTU_23655	15879
52929	Bacteria--Chloroflexi--Chloroflexi-4--PAWS52f--OTU_52929	15906
31622	Bacteria--Marine_group_A--A714017--OTU_31622	15926
7225	Bacteria--Marine_group_A--Arctic96B-7--OTU_7225	16282
3942	Bacteria--Proteobacteria--Gammaproteobacteria--OM60--OTU_3942	16587
62650	Bacteria--Proteobacteria--Deltaproteobacteria--Sva0853--SAR324--OTU_62650	16664
36890	Bacteria--Proteobacteria--Alphaproteobacteria--Unclassified--OTU_36890	17036
63234	Bacteria--Proteobacteria--Alphaproteobacteria--Unclassified--OTU_63234	17128
27691	Bacteria--Marine_group_A--SAR406--OTU_27691	17249
32241	Bacteria--Proteobacteria--Deltaproteobacteria--Unclassified--OTU_32241	1759
47960	Bacteria--Proteobacteria--Alphaproteobacteria--Oleomonas--Unclassified--OTU_47960	17773
8199	Bacteria--Verrucomicrobia--Verrucomicrobiae--Unclassified--OTU_8199	18045
5522	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum--Unclassified--OTU_5522	18901
31873	Bacteria--Proteobacteria--Alphaproteobacteria--Oleomonas--ctg_NISA150--OTU_31873	2307
5987	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05--Unclassified--OTU_5987	2449
26598	Bacteria--WS3--Unclassified--OTU_26598	2525
30687	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c--OTU_30687	2642
35767	Bacteria--Proteobacteria--Deltaproteobacteria--Geobacter--Pelobacter_propionicus--Unclassified--OTU_35767	2688
3490	Bacteria--Chloroflexi--Anaerolineae--A4b--TK32--OTU_3490	2725
64918	Bacteria--Planctomycetes--agg27--CL500-15--B83--OTU_64918	2864
8461	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum--Unclassified--OTU_8461	2970
26172	Bacteria--Proteobacteria--Alphaproteobacteria--TSBb13--OTU_26172	4308
34377	Bacteria--Proteobacteria--Alphaproteobacteria--S23_91--OTU_34377	4430
10140	Bacteria--Marine_group_A--SAR406--OTU_10140	4935
8351	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05--Unclassified--OTU_8351	5013
58861	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05--Unclassified--OTU_58861	5013
20592	Bacteria--Planctomycetes--Planctomycetacia--Unclassified--OTU_20592	5303
9875	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c--OTU_9875	5411
15060	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c--OTU_15060	5439
5588	Bacteria--Proteobacteria--Gammaproteobacteria--BD7-8--OTU_5588	5696
66746	Bacteria--Proteobacteria--Gammaproteobacteria--Alteromonadales--Unclassified--OTU_66746	5962
8452	Archaea--Thaumarchaeota--Cenarchaeales--Cenarchaeum--pIVWA5--Unclassified--OTU_8452	6212
64980	Bacteria--Proteobacteria--Alphaproteobacteria--Oleomonas--Unclassified--OTU_64980	731
43578	Bacteria--Proteobacteria--Alphaproteobacteria--Sphingomonadales--Porphyrobacter--Unclassified--OTU_43578	7635
20928	Bacteria--WS3--Unclassified--OTU_20928	7764
27647	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c--OTU_27647	7855
3945	Bacteria--Proteobacteria--Alphaproteobacteria--Unclassified--OTU_3945	8188
60274	Bacteria--Proteobacteria--Alphaproteobacteria--ZA3420c--OTU_60274	8338
38431	Bacteria--Proteobacteria--Gammaproteobacteria--SUP05--mussel_thioautotrophic_gill_symbiont_MAR1--OTU_38431	8862
10232	Bacteria--Bacteroidetes--Bacteroidales--VC21_Bac22--OTU_10232	8902
4300	Bacteria--Verrucomicrobia--Verrucomicrobia_subdivision_3--Verruco-3--CTD005-1B-02--Unclassified--OTU_4300	9121
2119	Bacteria--Planctomycetes--Planctomycetacia--Planctomycetales--Unclassified--OTU_2119	9964

Table D.2: List of correlating pathways under dysoxic-suboxic water column conditions based on Bray-Curtis and Spearmans rank correlations (Correlation value > 0.6, q-values = 0.05). Pathway abudance (RPKM) in time-series and incubation metagnomes is shown.

Id	Pathway name	Time-series	CH_4	$CH_4 + NO_2$
12DICHLORETHDEG-PWY	1,2-dichloroethane degradation	5522	18.97	12.99
3-HYDROXYPHENYLACETATE-DEGRADATION-PWY	4-hydroxyphenylacetate degradation	3950	51.73	61.67
7ALPHADEHYDROX-PWY	cholate degradation (bacteria, anaerobic)	257	4.98	4.68
ACETOACETATE-DEG-PWY	acetoacetate degradation (to acetyl CoA)	768	NA	22.55
ALKANEMONOX-PWY	two-component alkanesulfonate monooxygenase	750	12.82	12.62
AMMOXID-PWY	ammonia oxidation I (aerobic)	6713	13.80	10.33
ANARESP1-PWY	respiration (anaerobic)	124901	219.96	224.25
AST-PWY	arginine degradation II (AST pathway)	1118	48.26	56.25
DHGLUCONATE-PYR-CAT-PWY	glucose degradation (oxidative)	3509	43.17	24.69
FERMENTATION-PWY	mixed acid fermentation	130535	275.13	285.46
GALACTARDEG-PWY	D-galactarate degradation I	11416	21.33	40.94
GALACTCAT-PWY	D-galactonate degradation	1019	6.76	7.89
GALACTUROCAT-PWY	D-galacturonate degradation I	2545	8.32	11.43
GLUCARDEG-PWY	D-glucarate degradation I	5289	14.75	30.46
GLUDEG-I-PWY	glutamate degradation III (via 4-aminobutyrate)	1507	50.50	36.40
GLYOXYLATE-BYPASS	glyoxylate cycle	64449	81.53	87.40
LACTOSECAT-PWY	lactose and galactose degradation I	668	5.72	5.83
MALATE-ASPARTATE-SHUTTLE-PWY	aspartate degradation II	4438	70.33	205.89
MALTOSECAT-PWY	maltose degradation	466	1.98	1.46
METHANOGENESIS-PWY	methanogenesis from CO2	4837	62.39	31.82
MGLDLCTANA-PWY	methylglyoxal degradation VI	5648	65.80	76.72
NONOXIPENT-PWY	pentose phosphate pathway (non-oxidative branch)	56695	136.57	64.36
OXIDATIVEPENT-PWY	pentose phosphate pathway (oxidative branch)	20512	21.92	31.91
P105-PWY	TCA cycle IV (2-oxoglutarate decarboxylase)	123621	233.81	295.69
P108-PWY	pyruvate fermentation to propionate I	53735	210.27	210.01
P124-PWY	Bifidobacterium shunt	143099	557.11	198.62
P141-PWY	atrazine degradation I (aerobic)	1525	4.51	2.57
P181-PWY	nicotine degradation I	1276	36.68	33.93
P184-PWY	protocatechuate degradation I (meta-cleavage pathway)	7068	51.31	33.76
P21-PWY	pentose phosphate pathway (partial)	30638	104.58	43.57
P221-PWY	octane oxidation	5916	107.53	129.85
P282-PWY	nitrite oxidation	1232	1.24	0.33
P483-PWY	phosphonoacetate degradation	1146	25.22	25.34
P561-PWY	stachydrine degradation	1991	17.10	8.13
P621-PWY	nylon-6 oligomer degradation	885	3.31	4.66
PHOSPHONOTASE-PWY	2-aminoethylphosphonate degradation I	2542	73.93	76.34
PROTOCATECHUATE-ORTHO-CLEAVAGE-PWY	protocatechuate degradation II (ortho-cleavage pathway)	4442	23.46	23.44
PWY-0	putrescine degradation III	713	50.96	73.10
PWY-101	photosynthesis light reactions	4915	40.05	48.14
PWY-1263	taurine degradation I	1288	8.23	0.81
PWY-1723	formaldehyde oxidation VI (H4MPT pathway)	657	16.03	8.25
PWY-1801	formaldehyde oxidation II (glutathione-dependent)	4291	34.41	19.71
PWY-181	photorespiration	41438	165.37	75.28
PWY-1881	formate oxidation to CO2	11367	64.58	53.93
PWY-241	C4 photosynthetic carbon assimilation cycle, NADP-ME type	23262	40.74	43.68
PWY-2503	benzoate degradation I (aerobic)	414	3.04	2.84
PWY-2721	trehalose degradation III	212	1.44	1.76
PWY-2722	trehalose degradation IV	193	2.08	5.12
PWY-31	canavanine degradation	3592	10.20	11.14
PWY-4841	UDP-α-D-glucuronate biosynthesis (from myo-inositol)	205	3.12	2.86
PWY-4983	citrulline-nitric oxide cycle	17120	14.19	20.99
PWY-4984	urea cycle	17076	93.71	102.55
PWY-5033	nicotinate degradation II	483	0.86	NA
PWY-5159	4-hydroxyproline degradation II	852	2.34	2.64
PWY-5209	methyl-coenzyme M oxidation to CO2	1827	43.02	10.64
PWY-5250	methanogenesis from trimethylamine	10083	11.43	3.93
PWY-5276	sulfite oxidation I (sulfite oxidoreductase)	612	7.38	5.77
PWY-5329	L-cysteine degradation III	530	1.34	2.21
PWY-5418	phenol degradation I (aerobic	64	1.28	1.69
PWY-5436	threonine degradation IV	4471	54.93	59.72
PWY-5451	acetone degradation I (to methylglyoxal)	208	2.53	1.97
PWY-5458	methylglyoxal degradation V	875	59.86	72.44
PWY-5461	betanidin degradation	1299	2.84	3.49
PWY-5482	pyruvate fermentation to acetate II	11224	41.57	61.34
PWY-5506	methanol oxidation to formaldehyde IV	705	6.52	2.93
PWY-5515	L-arabinose degradation II	332	24.65	23.42
PWY-5516	xylose degradation II	181	25.83	23.50
PWY-5519	D-arabinose degradation III	377	1.82	0.41
PWY-5525	D-glucuronate degradation I	330	2.01	1.27
PWY-5534	propylene degradation	1328	3.25	1.71
PWY-5652	2-amino-3-carboxymuconate semialdehyde degradation to glutaryl-CoA	1853	4.70	5.32
PWY-5654	2-amino-3-carboxymuconate semialdehyde degradation to 2-oxopentenoate	2684	8.76	10.85
PWY-5691	urate degradation to allantoin	2876	4.68	5.23
PWY-5697	allantoin degradation to ureidoglycolate I (urea producing)	4907	19.16	6.58
PWY-5913	TCA cycle VI (obligate autotrophs)	166539	542.80	715.20
PWY-5938	(R)-acetoin biosynthesis I	37003	96.77	NA
PWY-5939	(R)-acetoin biosynthesis II	38275	97.62	124.73
PWY-6044	methanesulfonate degradation	152	1.66	0.58
PWY-6079	anthranilate degradation I (aerobic)	113	5.68	1.25
PWY-6183	salicylate degradation I	3099	4.18	2.21
PWY-6210	2-aminophenol degradation	1803	6.07	4.36
PWY-6215	4-chlorobenzoate degradation	3873	15.12	7.67
PWY-6342	noradrenaline and adrenaline degradation	108	29.63	32.98
PWY-6348	phosphate acquisition	699	11.53	5.67

Appendix E

OPU3 genes distribution

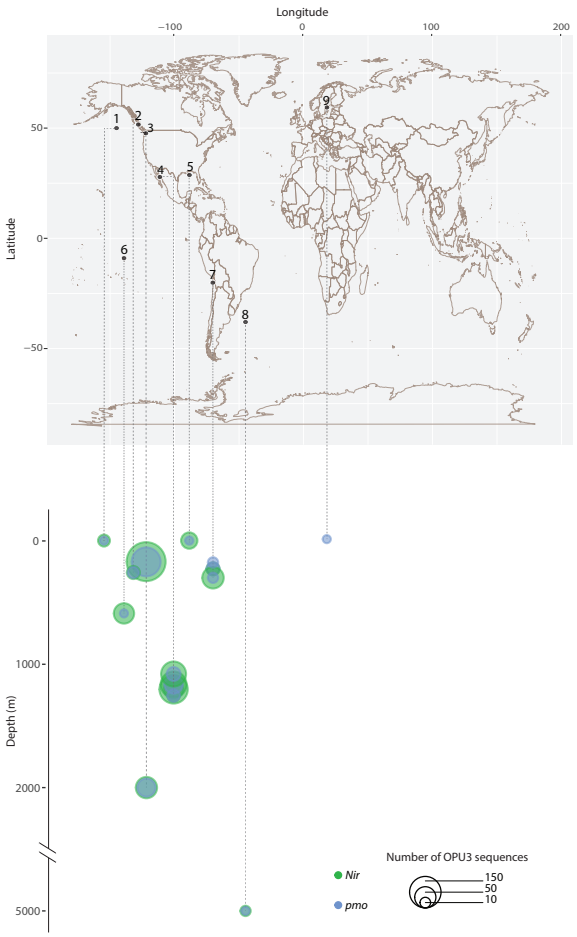

Figure E.1: Global distribution of OPU3 *pmoA* and *nirK* genes. Open ocean and coastal OMZs, and enclosed basins from where OPU3 *pmoA* and *nirK* genes sequences were obtained ((1) Northeastern Subarctic Pacific Ocean, (2) Saanich Inlet, (4) Guaymas Basin, (5) Gulf of Mexico,(6) Eastern Tropical South Pacific, (7) Eastern Tropical South Pacific off Chile, (8) South Atlantic Ocean and (9) Baltic Sea).

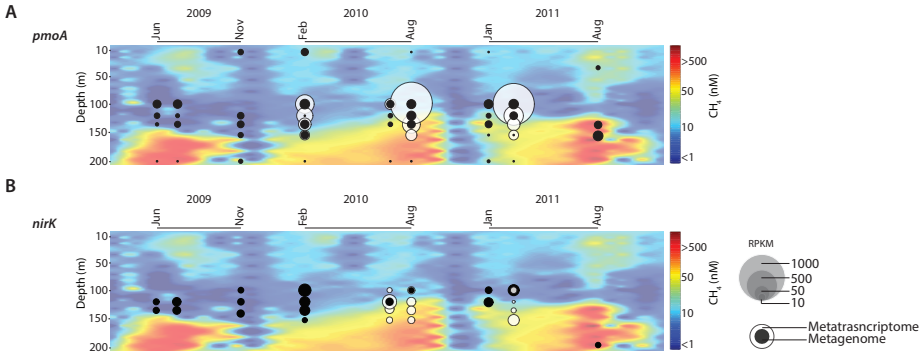

Figure E.2: Spatio-temporal distribution of OPU3 *pmoA* and *nirK* genes throughout Saanich Inlet water column. Time-series observations for OPU3 *pmoA* and *nirK* genes sequences from metagenomic datasets spanning 6 depths (10,100, 120, 135, 150 and 200 m) over 2-year period (June 2009 - August 2011).